Intercollegiate Faculty of Biotechnology
University of Gdańsk and Medical University of Gdańsk

Elucidating the role and mechanism of the regulatory network of genes encoding dioxygenases potentially involved in plant adaptation to terrestrial environments

Rola i mechanizm regulacji ekspresji genów kodujących dioksygenazy w adaptacji roślin do wzrostu i rozwoju w warunkach panujących na lądzie

Izabela Perkowska, MSc
Laboratory of Plant Protection and Biotechnology, IFB UG and MUG, Gdańsk

.....

Supervisor: **Prof. Ewa Łojkowska**
Laboratory of Plant Protection and Biotechnology, IFB UG and MUG, Gdańsk
Auxiliary supervisor: **Dr. Anna Ihnatowicz**
Laboratory of Plant Protection and Biotechnology, IFB UG and MUG, Gdańsk

Doctoral thesis submitted to the Discipline Council
of Biotechnology at the University of Gdańsk

Gdańsk, 2025

This doctoral thesis was carried out under the PRELUDIUM BIS 1 grant program:

- **2019/35/O/NZ1/02751 PRELUDIUM BIS 1 (E.L.)**

- **PPN/STA/2021/1/00074/U/00001 PRELUDIUM BIS 1 (I.P.)**

Additional financial support:

National Science Centre in Poland grants:

- 2022/47/B/NZ2/01835 OPUS 24 (A.I.)
- 2014/15/B/NZ2/01073 OPUS 8 (A.I.)

the Polish National Agency for Academic Exchange grants:

- BPN/BFR/2022/1/00039 POLONIUM (A.I.)
- PPN/BFR/2019/1/00050 POLONIUM (A.I.)

and the Erasmus+ Mobility Scholarship (European Union) (I.P.)

Many people have contributed to the completion of this dissertation, and I would like to express my sincere gratitude to them here:
Do powstania tej pracy przyczyniło się wiele osób, którym w tym miejscu chciałabym serdecznie podziękować:

To my supervisors, Prof. Ewa Łojkowska and Dr. Anna Ihnatowicz, for the trust they placed in me, the time they devoted, and everything I have learned from them,
Moim promotorkom, Prof. Ewie Łojkowskiej i Dr Annie Ihnatowicz, za zaufanie, którym mnie obdarzyły, poświęcony czas oraz wszystko, czego się od nich nauczyłam,

To everyone I had the pleasure of working with at the Laboratoire Agronomie et Environnement in France – especially Dr. Eng. Alexandre Olry and MSc Clément Charles – for their invaluable support during my research,
Wszystkim, których miałam okazję poznać w Laboratoire Agronomie et Environnement we Francji – w szczególności dr inż. Alexandre Olry i mgr Clément Charles za ich nieocenione wsparcie w prowadzeniu badań,

To all my colleagues at IFB UG & MUG, in particular those from the Department of Plant Protection and Biotechnology – Dr. Joanna Siwińska and MSc Alicja Dobek – for sharing experiences and fruitful discussions that helped shape this work,
Wszystkim koleżankom i kolegom z MWB UG i GUM, w szczególności z Zakładu Ochrony i Biotechnologii Roślin – dr Joannie Siwińskiej i mgr Alicji Dobek – za dzielenie się swoimi doświadczeniami i owocne dyskusje, które miały wpływ na kształt tej pracy,

And finally, to my family and friends, for their understanding, unwavering support, and constant encouragement at every stage of this journey.
Wreszcie mojej rodzinie i przyjaciołom – za wyrozumiałość, nieustające wsparcie i motywowanie mnie do działania na każdym etapie.

TABLE OF CONTENTS

KEYWORDS

abiotic and biotic stresses, adaptation to land conditions, *Arabidopsis thaliana*, coumarins, dioxygenases, glucosyltransferases, Natural Antisense Transcript, natural variation, phenylpropanoids, plant evolution, secondary metabolism

1. ABSTRACT

The colonization of land by plants represents one of the most groundbreaking evolutionary events in Earth's history, requiring profound transformations in their biochemical and regulatory pathways. In response to novel environmental pressures such as water deficiency and increased ultraviolet radiation, plants developed a wide spectrum of adaptive changes, including modifications in the biosynthesis and regulation of secondary metabolites.

This doctoral dissertation investigates the interactions between various *Arabidopsis thaliana* (thale cress) genotypes, selected abiotic and biotic environmental stresses, and the plant's secondary metabolism. Particular attention is given to coumarins – secondary metabolites derived from the phenylpropanoid pathway – which play an essential role in plant adaptation to unfavorable conditions. Coumarins are known for their ability to chelate iron (Fe), as well as their antioxidant and antimicrobial properties. By combining metabolomics, transcriptomics, phylogenetics, and reverse genetics approaches, this research provides new insights into plant responses to environmental stresses in the context of coumarin biosynthesis and phenylpropanoid pathway regulation.

The research is presented as a series of five thematically related publications. Three of these have been published in peer-reviewed scientific journals, while the remaining two are included as manuscripts in preparation. The first three publications (No. 1, 2, and 3) address the first research goal: to explore the correlation between coumarin accumulation, genotypic variation (natural accessions and mutants), and environmental factors, especially Fe availability and *Dickeya* spp. infection. The aim was to better understand how *A. thaliana* modulates coumarin metabolism in response to diverse environmental conditions, considering both genetic background and the nature of the applied stress.

The first study examined natural variation in the accumulation of simple coumarins (scopoletin, umbelliferone, esculetin) and their glycosylated forms across 28 *A. thaliana* accessions. UHPLC-MS metabolic profiling revealed significant intraspecific differences in coumarin profiles influenced by tissue type, growth conditions, and genotype (Perkowska *et al.,* 2021a). The second publication demonstrated that coumarin accumulation is modulated by Fe availability and correlates with resistance to *Dickeya* spp. – necrotrophic bacterial pathogens of agronomic importance, selected due to their virulence strategies and Fe-scavenging mechanisms. These findings suggest that coumarins may act as integral components in both nutrient acquisition and biotic stress defense (Perkowska *et al.*, 2021b). The third publication explored the relationship between gene expression dynamics and coumarin levels under specific environmental conditions. Based on previous coumarin content profiling, three genetically distinct accessions (Col-0, Est-1, and Tsu-1) were selected. To simulate contrasting conditions, plants were grown in liquid *in vitro* culture and in soil. Roots and shoots were harvested for parallel metabolite profiling (UHPLC-MS) and quantitative gene

expression analysis (qRT-PCR) targeting the phenylpropanoid pathway. Selected genes were re-sequenced to identify polymorphisms potentially contributing to variations in protein biosynthesis and ultimately to metabolic diversity (Ihnatowicz *et al.,* 2024).

The first three publications presented in this thesis highlight how genetic diversity and environmental conditions shape coumarin biosynthesis in *A. thaliana*, revealing links between metabolite levels, phenylpropanoid gene expression, and adaptive potential. The fourth manuscript focuses on the functional characterization of *DOXC21-A*, its natural antisense transcript (NAT), and the paralog *DOXC21-B*. Both genes encode 2OGD enzymes of previously uncharacterized biological function. Phylogenetic analysis confirmed that these genes are conserved across multiple land plant lineages, suggesting an evolutionarily maintained function and highlighting their potential biological significance. Wild type and mutant plants were grown hydroponically under low pH (4.5) and osmotic stress (3% PEG), and transcriptomic, metabolomic, and phenotypic analyses were conducted. The results support a model in which DOXC21-A and its NAT form a regulatory unit that may redirect phenylpropanoid flux from lignin biosynthesis toward flavonoid production in response to stress, potentially involving phytohormonal pathways. A conserved miRNA-homologous sequence found within NAT suggests possible post-transcriptional regulation. The distinct phenotypes observed in *doxc21-a* and *nat* mutants suggest that, while both are involved in abiotic stress responses, they may operate through partially overlapping yet functionally non-redundant molecular pathways.

The study described in the fifth manuscript addresses the functional characterization of UGT79B9, a UDP-glucosyltransferase potentially involved in stress-induced glycosylation. Experiments conducted on *ugt79b9* knockout, overexpression and complementation lines revealed that UGT79B9 affects root exudates profiles and glycosylation of stress-responsive metabolites under Fe deficiency, osmotic stress, and drought. Untargeted metabolomic profiling uncovered several yet-unidentified m/z features associated with UGT79B9 mutation and PEG-induced osmotic stress.

In conclusion, the five articles comprising the presented doctoral thesis broaden the understanding of how *A. thaliana* integrates environmental signals with the regulation of specialized metabolism, with particular emphasis on coumarin biosynthesis. The findings uncover a complex network of interactions involving 2OGDs, UGTs, and NATs that collectively participate in the adaptive reprogramming of metabolic pathways. This research highlights the significance of natural genetic variation and regulatory plasticity in enhancing plant resilience to abiotic and biotic stressors.

2. STRESZCZENIE (Polish version summary)

Kolonizacja lądu przez rośliny była jednym z najbardziej przełomowych wydarzeń w historii życia na Ziemi, wymagającym głębokich przemian w funkcjonowaniu ich szlaków biochemicznych i regulacyjnych. W odpowiedzi na nowe czynniki środowiskowe, takie jak deficyt wody czy zwiększone promieniowanie UV, w tkankach roślin zaszły liczne adaptacyjne przekształcenia, obejmujące m. in. regulację mechanizmów biosyntezy metabolitów wtórnych.

W ramach przedstawianej pracy doktorskiej przeanalizowałam wzajemne zależności pomiędzy genotypami rośliny modelowej *Arabidopsis thaliana* (rzodkiewnik pospolity), abiotycznymi i biotycznym stresami środowiskowymi, a metabolizmem wtórnym. Szczególną uwagę poświęciłam kumarynom – metabolitom wtórnym, syntetyzowanym w szlaku fenylopropanoidów, które odgrywają istotną rolę w strategiach adaptacyjnych roślin do niekorzystnych warunków środowiskowych. Kumaryny wykazują właściwości chelatujące jony żelaza (Fe), a także mają działanie antyoksydacyjne i przeciwdrobnoustrojowe. Łącząc metabolomikę, transkryptomikę, filogenetykę oraz odwrotną genetykę, przedstawiona praca przyczynia się do pogłębienia wiedzy na temat odpowiedzi roślin na niekorzystne warunki środowiskowe w kontekście biosyntezy kumaryn i regulacji szlaku biosyntezy fenylopropanoidów.

Badania przedstawiono w postaci cyklu pięciu tematycznie powiązanych publikacji, z których trzy zostały opublikowane w recenzowanych międzynarodowych czasopismach, natomiast dwie kolejne przedstawiono w postaci manuskryptów będących na etapie przygotowania. Pierwsze trzy prace, odpowiednio publikacje nr 1, 2 i 3, dotyczą pierwszego celu badawczego, którym było zbadanie korelacji między poziomem akumulacji kumaryn, zróżnicowaniem genotypowym roślin (naturalne akcesje i mutanty), a oddziaływaniem wybranych czynników stresowych – w szczególności dostępnością Fe, oraz infekcją *Dickeya* spp. Badania miały na celu zrozumienie, w jaki sposób *A. thaliana* dostosowuje metabolizm kumaryn w odpowiedzi na zmienne warunki środowiskowe, z uwzględnieniem różnic genotypowych roślin i rodzaju czynnika stresowego.

W pierwszej publikacji opisano zmienność naturalną w zakresie akumulacji wybranych prostych kumaryn (skopoletyny, umbeliferonu i eskuletyny) i ich glikozylowanych form w 28 naturalnych akcesjach *A. thaliana.* Profilowanie metaboliczne (UHPLC-MS) wykazało istotne zróżnicowanie wewnątrzgatunkowe zawartości kumaryn, na które wpływają genotyp i środowiskowe warunki wzrostu (Perkowska *et al.*, 2021a). W kolejnej publikacji wykazano, że akumulacja kumaryn koreluje z niedoborem Fe i jest powiązana z odpornością roślin na *Dickeya* spp. – nekrotroficzne patogeny bakteryjne o znaczeniu agronomicznym, wybrane ze względu na ich wysoką wirulencję i zdolność do wychwytywania Fe ze środowiska. Uzyskane wyniki wskazują, że kumaryny mogą ułatwiać pozyskiwanie Fe, jak i wspierać indukowane stresem biotycznym mechanizmy obronne (Perkowska *et al.*, 2021b). Badania opisane w trzeciej publikacji skupiały się na powiązaniu dynamiki ekspresji genów z akumulacją wybranych kumaryn w różnych tkankach roślinnych, w wybranych warunkach

środowiskowych. Bazując na wcześniejszej analizie akumulacji kumaryn wybrano trzy genetycznie zróżnicowane, naturalne akcesje *A. thaliana* (Col-0, Est-1 i Tsu-1), dla których wykazano zróżnicowaną akumulację tych związków. Aby symulować kontrastujące warunki środowiskowe, rośliny uprawiano w hodowli płynnej *in vitro*, jak i w glebie. Korzenie i pędy zbierano w celu równoległego profilowania metabolitów (UHPLC-MS) oraz analizy ekspresji genów (qRT-PCR) ze szlaku biosyntezy fenylopropanoidów. Re-sekwencjonowanie wybranych genów pozwoliło na identyfikację polimorfizmów, które mogą wyjaśniać zróżnicowanie zawartości kumaryn między badanymi akcesjami (Ihnatowicz *et al.*, 2024).

W trzech pierwszych publikacjach przedstawiono, jak zmienność genotypowa i warunki środowiskowe wpływają na biosyntezę kumaryn w tkankach *A. thaliana,* ujawniając zależności między poziomem metabolitów wtórnych, ekspresją genów szlaku fenylopropanoidów i potencjałem adaptacyjnym akcesji. Czwarta publikacja dotyczy charakterystyki genów *DOXC21-A*, jego antysensownego transkryptu (NAT) oraz paralogu *DOXC21-B*. *DOXC21-A* i *DOXC21-B* kodują dioksygenazy (2OGD) o nieznanej funkcji biologicznej. Analiza filogenetyczna potwierdziła ich wysoki stopień konserwacji u roślin lądowych, co sugeruje ewolucyjnie zachowaną rolę oraz podkreśla potencjalne znaczenie biologiczne. Rośliny hodowano w warunkach niskiego pH (4,5) i stresu osmotycznego (3% PEG), a następnie przeprowadzono analizy transkryptomiczne, metabolomiczne i fenotypowe. Uzyskane wyniki pozwoliły zaproponować model, w którym DOXC21-A i NAT tworzą jednostkę regulacyjną, mogącą koordynować z udziałem hormonów przekierowanie szlaku fenylopropanoidowego z lignin w stronę flawonoidów. Stwierdzono także obecność konserwowanej sekwencji homologicznej do regionów prekursorowych miRNA u roślin lądowych, co może sugerować udział w regulacji posttranskrypcyjnej. Odmienne fenotypy mutantów *doxc21-a* i *nat* sugerują, że chociaż obydwa elementy przyczyniają się do reakcji na stres abiotyczny, mogą działać poprzez nakładające się, ale nieredundantne szlaki molekularne.

Ostatnie badania skupiają się na charakterystyce funkcjonalnej UGT79B9 (AT3G53990), UDP-glukozyltransferazie potencjalnie zaangażowanej w glikozylację związków indukowanych stresem. Badania przeprowadzone na mutantach *ugt79b9* typu knockout i nadekspresorach *UGT79B9* pozwoliły na wykazanie, że UGT79B9 wpływa na profile eksudatów korzeniowych i poziom glikozylacji metabolitów, gromadzonych na skutek wystąpienia stresu środowiskowego (niedobór Fe, stres osmotyczny, susza). Nieukierunkowana analiza metabolomiczna ujawniła szereg dotąd niezidentyfikowanych sygnałów m/z związanych z występowaniem mutacji w badanym genie *UGT79B9* i odpowiedzią na stres osmotyczny indukowany PEG.

Podsumowując, wyniki badań przedstawionych w pięciu publikacjach składających się na niniejszą rozprawę doktorską poszerzają wiedzę o tym, jak *A. thaliana* integruje sygnały środowiskowe z regulacją metabolizmu wtórnego, ze szczególnym uwzględnieniem biosyntezy kumaryn. Uzyskane rezultaty, podkreślając znaczenie zmienności genetycznej i plastyczności regulacyjnej w odporności roślin na stres abiotyczny i biotyczny, wspierają tezę o złożonej sieci

interakcji pomiędzy 2OGD, UGT i NAT, które wspólnie uczestniczą w adaptacyjnym przeprogramowaniu szlaków metabolicznych.

3. LIST OF ABBREVIATIONS

2OGD: 2-oxoglutarate and iron (Fe)-dependent dioxygenase
4CL1, 4CL2, 4CL3: Coumaroyl-CoA ligases 1, 2, and 3
ABA: Abscisic acid
ACO: 1-aminocyclopropane-1-carboxylate oxidase
ANOVA: Analysis of variance
ANS: Anthocyanidin synthase
ATP: Adenosine triphosphate
ABC transporters: ATP synthase (ATP)-binding cassette transporters
C3'H: Coumaroyl shikimate 3-hydroxylase
CAS: Chrome Azurol S
CCoAOMT1: Caffeoyl-Coenzyme A O-methyltransferase 1
CCR1, CCR2: Cinnamoyl-CoA reductases 1 and 2
cDNA: Complementary DNA
CDS: coding sequence
CoA: Coenzyme A
Col-0: Columbia 0 A. thaliana accession
CYP82C2: Cytochrome P450 monooxygenase 82C2
DEGs: Differentially expressed genes
DNA: Deoxyribonucleic acid
DOX: Dioxygenase
DOXA: Dioxygenase clade A
DOXB: Dioxygenase clade B
DOXC: Dioxygenase clade C
Est-1: Estland-1 *A. thaliana* accession
F3H: Flavanone 3-hydroxylase
F6'H1, F6'H2: Feruloyl-CoA 6'-ortho-hydroxylase 1 and 2
Fe: iron
FLS: Flavonol synthase
FW: Fresh weight
GA20OX: Gibberellin 20-oxidase
GA3OX: Gibberellin 3-oxidase
GWAS: Genome-Wide Association Studies
HCT: Hydroxycinnamoyltransferase
HR: Hypersensitive response
HSP: Heat shock proteins
JA: Jasmonic acid
ko: knock-out
MS: Murashige and Skoog medium
NAT: Natural Antisense Transcript
NAT-siRNA: Small interfering RNA derived from NAT
mRNA: messenger RNA
PAMP: Pathogen-associated molecular pattern
PC: Principal component
PCA: Principal component analysis
PCR: Polymerase chain reaction
PDR9:
PEG: Polyethylene glycol
PDR9: Pleiotropic drug resistance 9
PR: Pathogenesis-related
PR1: Pathogenesis-related 1
qPCR: Quantitative PCR
QTL: Quantitative Trait Loci
RIL: Recombinant inbred line
RNA: Ribonucleic acid
ROS: Reactive oxygen species
S8H: Scopoletin 8-hydroxylase
SA: Salicylic acid
SAR: Systemic acquired resistance
SDS-PAGE: Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SOD1: Superoxide dismutase 1
TF: Transcription Factor
T-DNA: Transfer DNA
Tsu-1: Tsushima-1 *A. thaliana* accession
UHPLC: Ultra high-performance liquid Chromatography
UHPLC-MS: Ultra high-performance liquid chromatography coupled with mass spectrometry
UV-B: Ultraviolet B radiation
UVR8: UV-B resistance 8
WGD: Whole-genome duplication
WT: Wild type

4. INTRODUCTION

Plant evolution has remarkably shaped Earth's landscapes and ecosystems. One of the most transformative events in our planet's history was the colonization of terrestrial habitats by plants originating from aquatic environments. This foundational shift enabled the vast diversification of plant life into the countless species we see today [De Vries & Archibald, 2018]. A deeper understanding of biodiversity, mechanisms driving adaptation, ecological processes, and resilience, is crucial for addressing the pressing scientific and societal challenges we face nowadays. Such knowledge can help slow down or mitigate irreversible changes threatening natural habitats, which contribute to the loss of inter- and intraspecific variation and the degradation of ecosystems [Kumar *et al.,* 2023; Alekseeva, Doebeli & Ispolatov, 2020; Thompson & Fronhofer, 2019].

The significance of natural variation is particularly evident in the context of plant adaptation to terrestrial environments. Unlike aquatic habitats, land conditions are highly heterogeneous, exposing plants to a wide range of abiotic and biotic stresses. For example, water shortage, intense solar radiation, temperature fluctuations, nutrient limitations, and increased pressure from living organisms have a profound effect on plant gene expression, leading to significant phenotypic modifications. Due to their sessile nature, plants are especially vulnerable to environmental challenges and have developed diverse adaptive strategies [Cushman, Denby & Mittler, 2022; Foyer & Kranner, 2023].

Plants originated from charophyte green algae (charophytes), a group that diverged from other algae approximately over 1 billion years ago. Both charophytes and land plants (embryophytes) shared key features such as the ability to perform photosynthesis and cell wall formation composed of cellulose [Domozych, Poppe & Sørensen, 2016; Fürst-Jansen & de Vries, 2020]. Bryophytes, including mosses, represent some of the earliest diverging lineages of land plants and developed critical adaptations such as a waxy cuticle to prevent water loss and multicellular gametangia to enhance reproductive success in the desiccation-prone terrestrial habitat [Graham *et al.*, 2014].

The profound difference between aquatic and terrestrial environments is evident and has required from plants to develop physiological, biochemical, and molecular adaptations to survive under new conditions. Adaptation to terrestrial habitats required extensive modifications in plant development, comprising mechanisms to prevent water loss, provide structural support, UV and temperature protection, enhance nutrients acquisition, improve reproductive efficiency, and pathogen and pest resistance [Kumar *et al.*, 2018; Schreiber, Rensing & Gould, 2022].

One of the most critical challenges on land is periodic water shortage due to drought, which significantly affects plant growth, physiology and metabolism, often leading to reduced productivity and yield. To adapt to desiccation, plants have developed structural modifications such as deep root system to access water reservoirs, thicker and waxy cuticles to reduce water loss, and stomatal regulation to minimize transpiration [Kane *et al.*, 2020; Lewandowska, Keyl & Feussner, 2020].

At the physiological level, plants accumulate osmoprotectants (proline, sugars and polyamines) to maintain cell turgor and osmotic balance, and abscisic acid (ABA) which triggers stomatal closure and activates drought-responsive genes [Lawson & Blatt, 2014]. At the molecular level, plants undergo transcriptomic and epigenetic changes, upregulating stress-responsive genes, antioxidant defense mechanism, and secondary metabolites production such as phenylpropanoids to mitigate oxidative stress caused by water deficiency and UV radiation [Yang et al., 2021; Wahab et al., 2022].

Intracellular structural reinforcement was achieved through the development of rigid cell walls enriched with lignin, providing not only mechanical strength, but also resistance to desiccation, and protection against pathogens. The evolution of vascular tissues – xylem and phloem – enabled efficient, long-distance water and nutrient transport [Ménard *et al.*, 2022].

Terrestrial plants are exposed to higher doses of UV radiation than aquatic plants, especially those growing on high altitudes. Excessive exposure to UV and high-intensity visible light can lead to oxidative stress, DNA damage, membrane destabilization, and protein degradation. To mitigate these effects plants produce UV-absorbing secondary metabolites such as flavonoids or phenolics which reduce oxidative stress [Takshak & Agrawal, 2019; Nascimento & Tattini, 2022; Leonardelli *et al.*, 2024]. UV-induced DNA repair mechanism can be activated by the specialized receptors such as UV-B resistance 8 (UVR8), which can lead to the activation of transcription factors that increase the expression of UV-protective genes. On the other hand, high sunlight intensity can affect hormonal signaling pathways (jasmonic acid, JA; salicylic acid, SA) and thus leaf development, stomatal closure, and induce the production of heat shock proteins (HSP) [Yin & Ulm, 2017].

HSP also help plants to cope with fluctuating temperature *via* protein stabilization and prevention from denaturation under stress conditions. Proline, sugars and dehydrins may act as osmoprotectants, cryoprotectants, while unsaturated fatty acids are responsible for maintaining membrane fluidity and cellular functions. Temperature is also affecting hormonal signaling leading to ABA accumulation in cold and ethylene-mediated response in heat [Niu & Xiang, 2018; Shah *et al.,* 2024].

Another problem that had to be optimized for terrestrial plants was the uptake of nutrients from the soil, where conditions are completely different from those in water. The development of an extensive root system with root hairs led to an increased absorption surface and better anchoring in the soil. Symbiotic associations with mycorrhizal fungi and nitrogen-fixing bacterial improved the availability of phosphorus, nitrogen and other essential nutrients to plants. Regulation at the molecular, transcriptional, as well as post-transcriptional levels, ensured efficient nutrients uptake and redistribution of macro- and microelements through transporter proteins and secondary metabolites [Gruber *et al.*, 2013; Holz *et al.,* 2024].

Life on land also required the adaptation of reproductive processes to new, demanding environmental conditions, including the protection of embryos within multicellular structures, the development of pollen grains for aerial fertilization, and the production of desiccation-tolerant seeds for long-term survival. Specialized reproductive structures such as flowers and fruits

containing vast array of specialized compounds enhance plants productivity by the attraction of pollinators [Friedman & Barrett, 2009].

Physical barriers, such as a thickened cuticle, lignin-reinforced cell walls, and stomatal closure also contributed to plant resistance to pathogens by preventing microbial entry and infection [Ziv *et al.*, 2018]. The activation of immune receptor proteins for pathogen-associated molecular pattern (PAMP) recognition triggers systemic acquired resistance (SAR) and the hypersensitive response (HR), limiting infection spread. Finally, secondary metabolites with antimicrobial properties – such as flavonoids, coumarins, and phytoalexins – together with pathogenesis-related (PR) proteins, help inhibit the development of disease symptoms caused by pathogens and pests [Ramaroson *et al.*, 2022; Ali, Tyagi & Mir, 2024].

Secondary metabolism plays a crucial role in the interactions between plants and their environment. Plants are estimated to produce between 200,000 and 1,000,000 different compounds, with the exact number continually increasing due to advancements in metabolomics and analytical techniques. Examples of adaptive biosynthetic pathways include the production of phenylopropanoids (flavonoids, phenolics, lignins, coumarins, anthocyanins), terpenoids (monoterpenes, sesquiterpenes, diterpenes, carotenoids, and gibberellins), alkaloids, glucosinolates, oxylipins, and others, which provide diverse ecological benefits. They play key roles in 1) tolerance to abiotic stresses, including drought, UV, and extreme temperature; 2) nutrients acquisition; 3) attraction of pollinators; 4) defense against pathogens and pests; 4) microbe symbiosis; 5) allelopathy [Alseekh & Fernie, 2023; Butnariu & Tietel, 2023; Wu, Northen & Ding, 2023].

Such adaptive changes related to the synthesis of various secondary metabolites belonging to different classes of chemical compounds were possible due to the involvement of a broad range of enzymes involved in plant secondary metabolism. They can be divided into eight major classes due to their function: 1) oxidoreductases that catalyze oxidation-reduction reactions by electron transfer (peroxidases, monooxygenases); 2) lyases catalyzing the cleavage of bonds without hydrolysis or oxidation; 3) transferases that transfer functional groups from one molecule to another (glucosyltransferases, acyltransferases); 4) hydrolases catalyzing the hydrolysis of bonds (glucosidases, esterases); 5) ligases catalyzing the joining of two molecules using ATP; 6) methyltransferases that transfer methyl groups to substrates; 7) isomerases catalyzing structural rearrangements within molecules; and 8) dioxygenases which catalyze the incorporation of oxygen atoms into substrates [Pyne, Narcross & Martin, 2019].

One type of dioxygenases, more precisely 2-oxoglutarate and iron (Fe)-dependent dioxygenases (2OGDs), has been of special interest in our research team for many years. This superfamily is widely distributed in nature and catalyze various types of reactions. Enzymes belonging to this family play crucial roles in plant primary and secondary metabolism including hydroxylation, demethylation and epoxidation. They are involved in biosynthesis of hormones (GA20OX, GA3OX, ACO), flavonoids (F3H, FLS, ANS), alkaloids, glucosinolates, cell wall formation, epigenetic regulation

and general stress response [Farrow & Facchini, 2014]. According to Kawai *et al.* (2014) plant 2OGDs

can be classified into three major functional clades DOXA, DOXB and DOXC. DOXA clade consist of nucleic acid-modifying 2OGDs responsible for DNA repair, RNA stability, end epigenetic regulation. DOXB clade enzymes are involved in hormone biosynthesis and regulation. Finally, DOXC represents the largest and most functionally diverse clade, compromising enzymes involved in the biosynthesis of secondary metabolites and contributing to plant defense, pigmentation, UV protection and ecological interactions with biotic stimuli.

The vast diversity of secondary metabolites in plants is largely attributed to whole-genome duplication (WGD) events, which have occurred multiple times during plant evolution. In *Arabidopsis thaliana* approximately 60% of nuclear genes originated *via* WGD [Mukherjee *et al.*, 2022]. These events have increased gene content and genetic redundancy. Duplicated genes may acquire new functions through a process known as neofunctionalization or partition existing functions *via* subfunctionalization [Cusack *et al.*, 2021]. Horizontal gene transfer has also contributed to the evolution of secondary metabolism, allowing plants to acquire certain bacterial or fungal-derived biosynthetic traits, particularly in specialized pathways such as alkaloid, terpene, and carotenoid biosynthesis [Glasner, Truong & Morse, 2020; Haimlich *et al.,* 2024]. Most natural biochemicals have evolved from primary metabolic pathways. For example, flavonoids originate from the shikimate and phenylpropanoid pathways, where they are derived from phenylalanine [Maeda, 2019]. The remarkable diversity of plant secondary metabolites is partly due to enzyme promiscuity, in which some enzymes exhibit broad substrate specificity, facilitating the emergence of new compounds [Kreis & Munkert, 2019].

Understanding biodiversity by tracing its evolutionary origins allows us to uncover the strategies that plants have developed over time. Natural variation among plant populations and their phylogenetic relationships, shed light on the genetic and phenotypic diversity that occurs naturally. The remarkable ability of plants to adapt to unfavorable and changing environmental conditions results from the combined effect of mutation, genetic recombination, gene flow, and environmental pressures. These processes form the foundation of plant evolution, which can subsequently be utilized in plant biotechnology, sustainable agriculture, conservation, and ecology [Anderson *et al.*, 2011; Bomblies & Peichel, 2020; Lasky *et al.*, 2023].

A. thaliana is a model plant species that has become a great tool for studying natural variation among populations [Koornneef *et al.*, 2004]. Although the initial observations date back to the 16th century, systematic research on this plant began in the 1940s, with its recognition as a model organism gaining full appreciation in the 1980s. This development has positioned *A. thaliana* as a pivotal tool for unraveling the complexities of genetics, cellular biology, and plant adaptation. *A. thaliana* is an ideal model organism due to its several features like small, fully sequenced genome, short generation time and broad natural genetic diversity [Krämer, 2015]. Genome-Wide Association

Studies (GWAS) are especially useful for identifying genetic loci associated with specific traits by analyzing natural variation [Korte & Farlow, 2013]. This approach facilitates the discovery of genes potentially influencing observed plant phenotypes, such as stress tolerance, disease resistance, and flowering time. GWAS has provided valuable insights into the genetic basis of complex traits and has contributed to identifying genes involved in plant responses to adverse environmental conditions, thereby aiding in adaptation and evolution [Alseekh *et al.*, 2021]. In contrast, Quantitative Trait Loci (QTL) mapping identifies genomic regions associated with specific traits by analyzing genetic variation in controlled populations, such as recombinant inbred lines (RILs), rather than natural populations [Kover *et al.*, 2009]. Together, GWAS and QTL mapping are powerful tools in translational research for crop science, enabling the application of discoveries from model organisms to the development of strategies for crop species [Garrido-Cardenas *et al.*, 2018; Anderson & Song, 2020].

Despite decades of research on *A. thaliana*, many questions remain unanswered, particularly those related to secondary metabolism. While significant progress has been made in understanding primary metabolism and fundamental processes such as respiration, photosynthesis, growth, and reproduction, secondary metabolism – which involves the production of specialized compounds, and their glycosides is still not fully understood. These compounds help plants to survive in adverse environments and are not only important because of that, but they often serve as economically important compounds [Le Roy *et al.*, 2016]. Given that plant-derived natural products often serve as medicines and pharmaceuticals, it is crucial to discover new compounds, improve existing ones and engineer their biosynthesis to achieve higher compound yields. Additionally, plants are valuable source of biofuels, which can be optimized through biotechnology to become more efficient and sustainable [Dixon & Dickinson, 2024].

Understanding the mechanisms underlying the production and regulation of secondary metabolites has been the focus of extensive research for many years. However, only recently we have been able to integrate high-throughput methods like transcriptomics, metabolomics and proteomics to gain a comprehensive understanding of how specific pathways interact and influence each other within complex networks. One of the critical aspects of this complexity is the trade-off signaling between growth and defense [Ha *et al.,* 2021]. Since plants are using same metabolic precursor within different biosynthetic pathways, dynamic signaling networks are needed to govern allocation of their resources. It is often possible due to the phytohormones like JA and SA, which mediate the balance between growth and secondary metabolite production [Jan *et al.,* 2021]. Recently Cheaib *et al.* (2024) implicated that JA signaling is a master regulator of the growth defense balance, coordinating the accumulation of secondary metabolites in rice during stress response. Comprehending this mechanism is essential for optimizing plant productivity, particularly in stress-prone environments where metabolic flexibility is crucial.

In summary, plants exposed to multiple and simultaneous environmental stresses must rely on a complex network of signaling molecules and metabolites to adapt and survive. These stress conditions trigger a coordinated interplay among various compound classes that are not only genotype-specific but also act synergistically to regulate growth, stress responses, signal transduction, and metabolic reprogramming. High-throughput methods such as transcriptomics and metabolomics enable linking the gene expression with metabolomic profiles, integrating the plant's genotype and phenotype. In the era of constantly growing human population and worldwide global crises – including armed conflicts, climate change, economic challenges, and health concerns, it is especially important to advance in the crop improvement and develop sustainable agriculture. I also strongly believe that understanding plant secondary metabolism and its regulatory mechanisms might be source of inspiration to produce new biologically active compounds and drugs combating diseases with which we are fighting nowadays.

5. AIMS

The main aims of my PhD dissertation were as follows:

1) **To investigate the correlation between accumulation of coumarins, plant genotypes (natural accessions and mutants), and environmental stress factors (including Fe availability and pathogen defense) to gain deeper insights into the metabolic adaptations and stress response in *Arabidopsis thaliana.***

2) **Comprehensive analysis of the regulatory network controlling the switch between coumarin and lignin biosynthesis in *A. thaliana* under osmotic and low pH stress using integrative transcriptomic and metabolomic approaches.**

3) **Functional characterization of two *A. thaliana* dioxygenases DOXC21-A, its NAT, and DOXC21-B potentially involved in ortho-hydroxylation of coumarin and lignin precursors.**

4) **Characterization of *A. thaliana* UDP-glucosyltransferase 79B9 involved in Fe deficiency and osmotic stress response.**

6. RESULTS

<u>Aim 1.</u> Comprehensive analysis of the correlation between accumulation of coumarins, plant genotypes (natural accessions and mutants), and environmental stress factors (including Fe availability and pathogen defense) to gain deeper insights into the metabolic adaptations and stress response in *Arabidopsis thaliana.*

In three subsequent publications (No. 1, 2, 3), I contributed to the analysis of correlations between coumarin accumulation, plant genotypes (natural accessions and mutants), and stress factors (Fe availability and pathogen defense) to gain deeper insight into secondary metabolic adaptations and stress response mechanisms in *A. thaliana*. In the first publication (Perkowska *et al*., 2021a), I participated in a comprehensive study of the distribution of simple coumarins in 28 natural *A. thaliana* populations, shedding light on the complex interaction between plant metabolism, genetic variation, and environment. The second study (Perkowska *et al.*, 2021b) focused on the role of Fe availability and coumarin accumulation in *A. thaliana* defense against plant pathogens from *Dickeya* spp., emphasizing the link between coumarin biosynthesis and plant resistance to biotic stress. In the most recent publication (Ihnatowicz *et al.*, 2024), I participated in a comprehensive analysis of coumarin content in three *A. thaliana* accessions from contrasting environments, investigating how tissue type and growth conditions affect variability in coumarin levels and identifying some genetic determinants possibly underlying the observed variation.

6.1. Publication No. 1

Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. **Perkowska, I.**, Siwinska, J., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Ihnatowicz, A. (2021). *Molecules*, 26(6), 1804.

Despite the growing interest in coumarins and phenylpropanoids in recent years, little was known about the variation in coumarin and their glycoside accumulation among natural populations (accessions) of *A. thaliana*. To address this knowledge gap, we conducted a comprehensive metabolite profiling study across 28 *A. thaliana* accessions using ultra-high-performance liquid chromatography coupled with mass spectrometry (UHPLC-MS).

The presented study focused on the identification and quantification of simple coumarins – such as scopoletin, umbelliferon, esculetin – and their glycosylated derivatives: scopolin, skimmin, and esculin, respectively. These compounds, derived from the phenylpropanoid pathway, are known to facilitate Fe mobilization and play critical roles in plant defense and adaptation to environmental stress [Stringlis *et al.*, 2019]. By comparing metabolite profiles across genetically diverse *A. thaliana* accessions, we revealed substantial natural variation in coumarin levels and composition in roots and leaves. Importantly, our findings revealed genotype-specific metabolic signatures and suggested potential associations between coumarin accumulation patterns and the geographical origin of the accessions.

To achieve accurate metabolic profiling, plants were cultivated *in vitro* in liquid medium to promote root development, as coumarins are primarily synthetized and accumulated in roots [Kai *et al.*, 2006]. We also aimed to highlight the importance of analyzing both aglycone (non-glycosylated) and glycosylated forms of coumarins, as glycosylation affects their solubility, stability, and biological activity [Le Roy *et al.*, 2016]. Metabolites were extracted with 80% methanol to obtain crude extracts, half of which were subsequently subjected to enzymatic hydrolysis with β-glucosidase for comparative analysis of both forms (Fig.1A). Finally, Principal Component Analysis (PCA) was performed to explore potential correlations between coumarin content and environmental variables.

One of the key findings was the significant intraspecies variation in coumarin accumulation among the 28 *A. thaliana* accessions. Scopoletin and its glycoside scopolin were the most abundant coumarins detected, with highly variable concentrations between accessions (Fig. 1B). Notably, coumarin profiles differed significantly between roots and leaves, with roots generally exhibiting higher levels, consistent which their role in nutrient acquisition and interaction with soil microorganisms [Stringlis *et al.*, 2019]. Interestingly, the detection of umbelliferone in *A. thaliana*

root tissues was described for the first time, as previous studies had not reported its presence in this model species. Although esculetin and esculin were only detected in trace amounts, they were present in both roots and leaves of some accessions. The presence of coumarins in leaves may suggest roles beyond nutrient uptake, possibly in long-distance signaling or plant-microbe communication, as postulated by Stassen *et al.* (2021).

PCA of UHPLC-MS data showed that PC1 (27.1%) and PC2 (21.9%) together explained 49% of the total metabolic variance. It is worth noting that accessions separated along PC1 and PC2 based on altitude and precipitation of their native habitats. Scopoletin and umbelliferone were major contributors to PC1, while scopolin and skimmin influenced PC2. A positive correlation was observed between scopoletin levels and precipitation, and a negative one with temperature, suggesting that coumarin biosynthesis may be modulated by climatic factors.

The findings presented in Perkowska *et al.* (2021a) highlight *A. thaliana* as a valuable model for investigating how genetic variation and environmental factors jointly influence secondary metabolism, particularly coumarin biosynthesis. Coumarin accumulation patterns were shown to be shaped by both genotype and ecological selective pressures, suggesting their potential role in adaptation to different conditions. Understanding these interactions is crucial for enhancing plant resilience to abiotic and biotic stresses, especially in context of climate change and growing agricultural challenges. Moreover, elucidating the regulatory networks underlying coumarin biosynthesis may support the metabolic engineering of these bioactive compounds in other plant species or microbial systems for pharmaceutical or crop protection applications. This work provides a methodological and conceptual foundation for future high-throughput analyses of natural variation in specialized metabolism across model and crop plants.

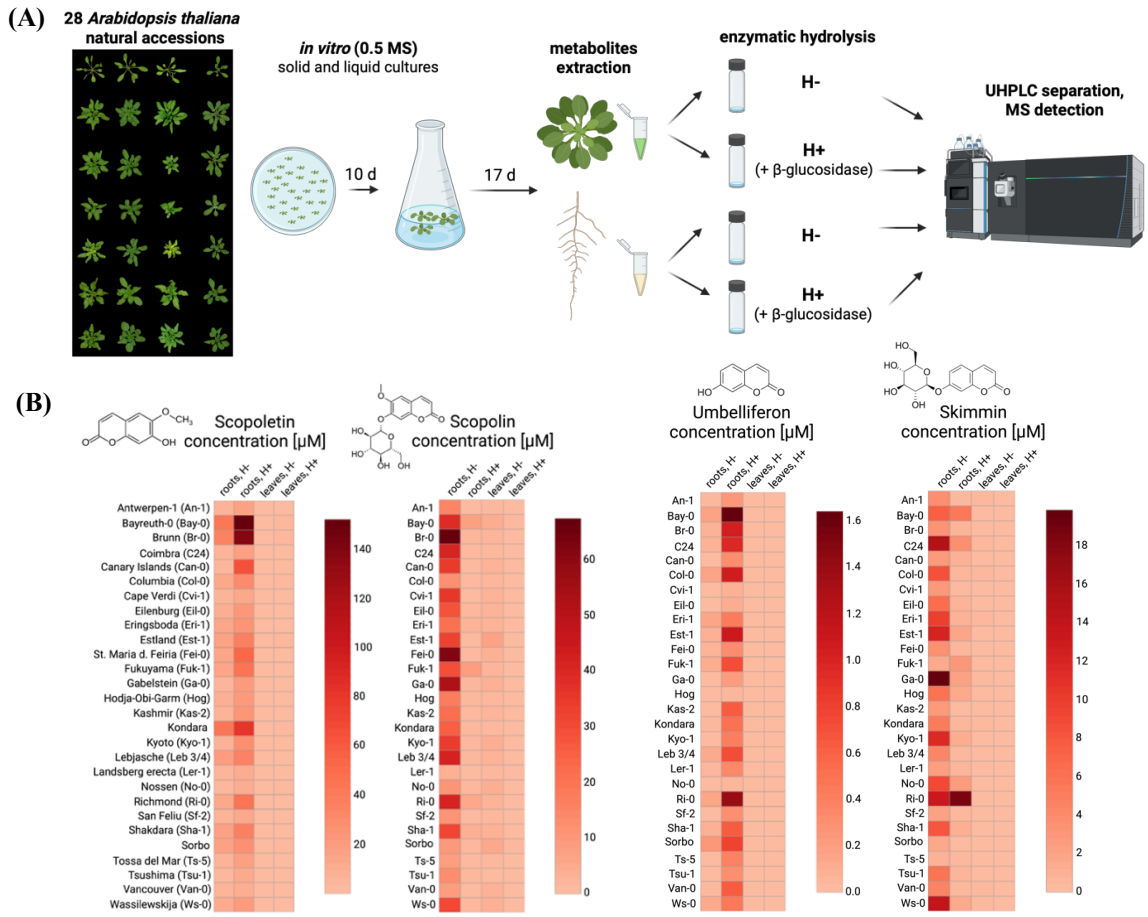


Figure 1.

(A) Schematic representation of the experimental procedure outlined in Publication No. 1, detailing the workflow used to determine selected coumarins in *A. thaliana* tissues.

(B) Quantification of scopoletin, scopolin (scopoletin glucoside), umbelliferone, and skimmin (umbelliferone glucoside) in roots and leaves of *A. thaliana* accessions, measured in extracts not subjected (H−) and subjected to enzymatic hydrolysis (H+).

6.2. Publication No. 2

Perkowska, I., Potrykus, M., Siwinska, J., Siudem, D., Lojkowska, E., & Ihnatowicz, A. (2021). Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *International Journal of Molecular Sciences*, 22(12), 6449.

Previous studies on coumarins accumulation within natural accessions of *A. thaliana* led us to the idea to explore their role more closely in biotic stress response. Given that coumarins play a key role in Fe acquisition by mobilizing it in the rhizosphere [Tsai & Schmidt, 2017], we investigated the interaction between *A. thaliana* interaction and plant pathogenic bacteria from the genus *Dickeya*. *Dickeya* spp. was chosenas a model pathogen since it is a model plant pathogen studied in our laboratory, which is known to produce Fe-chelating compounds called siderophores [Potrykus *et al.*, 2014; Potrykus *et al.*, 2020]. It was particularly interesting in the context of coumarin biosynthesis which have antimicrobial and Fe-chelating properties. However, the intricate mechanisms governing the interplay between coumarin biosynthesis, Fe homeostasis, and plant defense remained inadequately understood. Thus, our aim was to elucidate these relationships using *A. thaliana* and *Dickeya* spp. pathosystem as a model.

Primary objectives were as follow: I) assessment of the impact of Fe bioavailability on coumarin accumulation in *A. thaliana*, II) exploration of the role of coumarins in plant defense against *Dickeya* spp. under varying Fe conditions, III) analysis of gene expression patterns associated with coumarin biosynthesis and Fe homeostasis during pathogen infection. For this purpose, we employed a multifaceted experimental approach in which we used *A. thaliana* Col-0 wild-type (WT) and mutant lines deficient in the key enzymes of coumarin biosynthesis (*f6'h1 – feruloyl-CoA 6'-ortho-hydroxylase 1* and *s8h – scopoletin 8-hydroxylase*) and coumarin transport to rhizosphere (*pdr1 – pleiotropic drug resistance 1*). Plants were cultivated hydroponically under controlled environmental conditions with either sufficient Fe (40 μM Fe^{2+}) or Fe-deficient (0 μM Fe^{2+}) media to simulate varying Fe availability and in two commercial soil mixes with different chemical properties. Two strains of *Dickeya* spp. were selected for infection studies – the reference strain *D. dadantii* 3937 and the more virulent *D. solani* IFB0099. Inoculations were performed at the flowering stage and disease progression was assess for both culture types (hydroponics and soil). Disease severity was quantified using a developed in our laboratory, standardized scoring system ranging from 0 to 5, based on visible symptoms such as wilting, chlorosis, and tissue maceration. Quantitative PCR (qPCR) was conducted to monitor the expression levels of genes involved in coumarin biosynthesis (*F6'H1, S8H*) and stress response (*PR1, CYP82C2, SOD1*) during pathogen infection. To evaluate the impact of plant-derived compounds on bacterial siderophore production, we assessed the growth and halo formation of *D. dadantii* 3937 and *D. solani* IFB0099 on CAS-agar plates supplemented with leaf homogenates from *s8h* and *f6'h1* mutants grown in Fe-depleted hydroponic conditions (0 μM Fe^{2+}) (Fig. 2A).

Noteworthy finding described in this paper was consistent with prior results [Lan *et al.*, 2011; Rodríguez-Celma *et al.*, 2013], that Fe-deficient conditions significantly upregulated the expression of *F6'H1* and *S8H,* leading to increased accumulation of scopoletin and fraxetin [Siwinska *et al.*, 2018]. We showed that *f6'h1*and *s8h* mutant lines, impaired in coumarin production, displayed enhanced susceptibility to *D. solani* IFB0099 and *D. dadantii* 3937 under Fe-deficient conditions, as evidenced by more severe disease symptoms compared to WT plants (Fig. 2B). Chemical composition of soil also significantly impacted disease progression for both tested strains. While both *D. dadantii* 3937 and *D. solani* IFB0099 caused disease symptoms in *A. thaliana*, the latter strain induced more pronounced symptoms across all tested conditions. Infection with *D. solani* IFB0099 led to a marked induction of stress-responsive genes, including *SOD1* (antioxidant defense), *PR1* (SA-dependent SAR marker), and *CYP82C2* (JA-responsive gene) indicating an active defense response. It is worth noting that *D. solani* IFB0099 exhibited a significant increase in halo diameter on CAS-agar plates supplemented with *f6'h1* mutant leaf homogenate, in contrast to its typically lower Fe-chelation ability compared to *D. dadantii* 3937.

Our findings underscore the integral role of coumarins in mediating plant defense against *Dickeya* spp., particularly under Fe-limited conditions. The enhanced susceptibility of coumarin-deficient mutants to *D. solani* IFB0099 highlights their protective role, either by direct antimicrobial activity or by modulating defense signaling pathways. The differential responses observed between *D. dadantii* 3937 and *D. solani* IFB0099 infections suggest that pathogen-specific factors, such as virulence mechanisms and Fe acquisition strategies, influence the outcome of plant-pathogen interactions. This highlights the necessity for tailored approaches in managing diseases caused by distinct *Dickeya* species. From an applied perspective, this work has important implications for crop improvement. Enhancing coumarin biosynthesis or Fe homeostasis through breeding or genetic engineering represents a promising strategy to improve resistance against bacterial pathogens from *Dickeya* spp. in agriculturally important crops, as coumarins function both as Fe-mobilizing agents and antimicrobial compounds that shape beneficial rhizosphere interactions [Stringlis *et al.*, 2018]. In conclusion, our study contributes to a deeper understanding of the complex interplay between secondary metabolism, nutrient status, and immune regulation in plants.

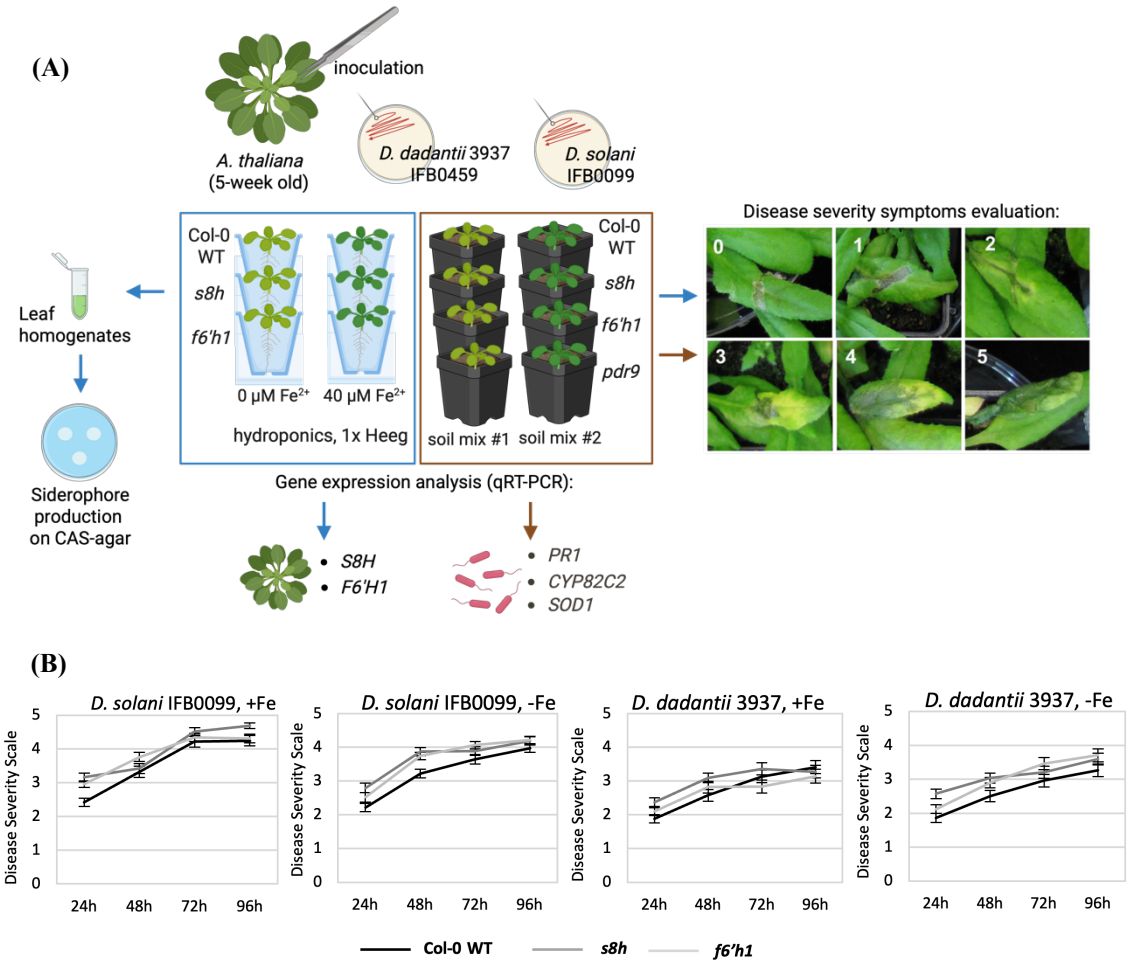


Figure 2.

(A) Schematic representation of the experimental procedure outlined in Publication No. 2.
(B) Disease progression caused by *D. solani* IFB0099 and *D. dadantii* 3937 on *A. thaliana* Col-0 WT (black line), and *s8h* (dark grey), *f6'h1* (light grey) mutant lines grown in +Fe (0 µM Fe^{2+}) and -Fe (0 µM Fe^{2+}) conditions. Disease symptoms were evaluated by visual symptom scoring (DSS, disease severity scale in which 0 – no signs of symptoms of the disease, 5 – visible maceration of the whole limb and the leaf) every 24 h for 4 days.

6.3. Publication No. 3

Ihnatowicz, A., Siwinska, J., **Perkowska, I**., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Olry, A. (2024). Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of *Arabidopsis thaliana*. *BMC Plant Biology*, 24(1), 806.

Our previous findings, detailed in Perkowska *et al.* (2021a), revealed natural variation in the accumulation of selected coumarins (scopoletin, umbelliferone, and their respective glycosides, scopolin and skimmin) across 28 natural *A. thaliana* accessions. Furthermore, their concentrations were tissue-dependent, with roots showing higher accumulation than leaves. In Perkowska *et al.* (2021b), we demonstrated that *A. thaliana* susceptibility to *Dickeya* spp. is influenced by coumarin content, as well as environmental factors, including iron availability and growth conditions (hydroponics *vs.* soil). To expand our understanding of gene-metabolite relationships governing coumarin biosynthesis, this study (Ihnatowicz *et al.*, 2024) integrated targeted metabolomic profiling, gene expression analysis, and genomic resequencing.

Our goal was to quantify the two selected coumarin compounds – scopoletin, which is the most abundant coumarin in *A. thaliana*, and umbelliferone, that has high pharmaceutical value, along with and their glycosylated derivatives (scopolin and skimmin, respectively) in root and leaf tissues. To achieve this, we selected three natural *A. thaliana* accessions – Col-0 (Columbia-0), Est-1 (Estland-1), and Tsu-1 (Tsushima-1), which exhibited significant differences in coumarin accumulation.

To evaluate the impact of environmental conditions on their accumulation, plants were cultivated in *in vitro* liquid cultures, which promote root growth and simulate stress conditions, and in soil, representing more physiological conditions. We analyzed the expression levels of key phenylpropanoid pathway genes, including *coumaroyl-CoA ligases* (*4Cl1, 4Cl2, 4Cl3*), *F6'H1*, *F6'H2*, *caffeoyl coenzyme A O-methyltransferase* (*CCoAOMT1*), *hydroxycinnamoyltransferase* (*HCT*), *coumaroyl shikimate 3-hydroxylase* (*C3'H*), and *cinnamoyl-CoA reductases* (*CCR1, CCR2*).

Additionally, to identify genetic polymorphisms within studied accessions, these genes were re-sequenced, allowing us to examine potential genotype-dependent differences in coumarin biosynthesis. Schematic representation of the experimental procedure is shown in the Fig. 3A. Plant material (roots and shoots of each accession) was divided into two portions – one used for the extraction and quantification of secondary metabolites, and the other for RNA isolation and qPCR analysis. To determine total scopoletin and umbelliferone content, half of the crude extracts were subjected to enzymatic hydrolysis with β-glucosidase, as coumarins are primarily stored in glycosylated, inactive forms under normal conditions [Duan *et al.*, 2021]. Detection and quantification of aglycones and glycosylated compounds were performed with UHPLC method.

Our findings confirm that coumarin accumulation varies significantly across genotypes and growth conditions. The highest scopoletin and umbelliferone accumulation was detected in roots grown *in vitro*, whereas soil-grown plants exhibited significantly lower concentrations (Fig. 3B). Leaves consistently accumulated lower levels of coumarin than roots, confirming previous observations [Kai *et al.*, 2006; Perkowska *et al.*, 2021a; Robe *et al.*, 2021] that coumarins are predominantly synthesized and stored in underground tissues (Fig. 3B). Additionally, glycosylated forms (scopolin and skimmin) were detected in all accessions, but their levels varied significantly between *in vitro* and soil-grown plants.

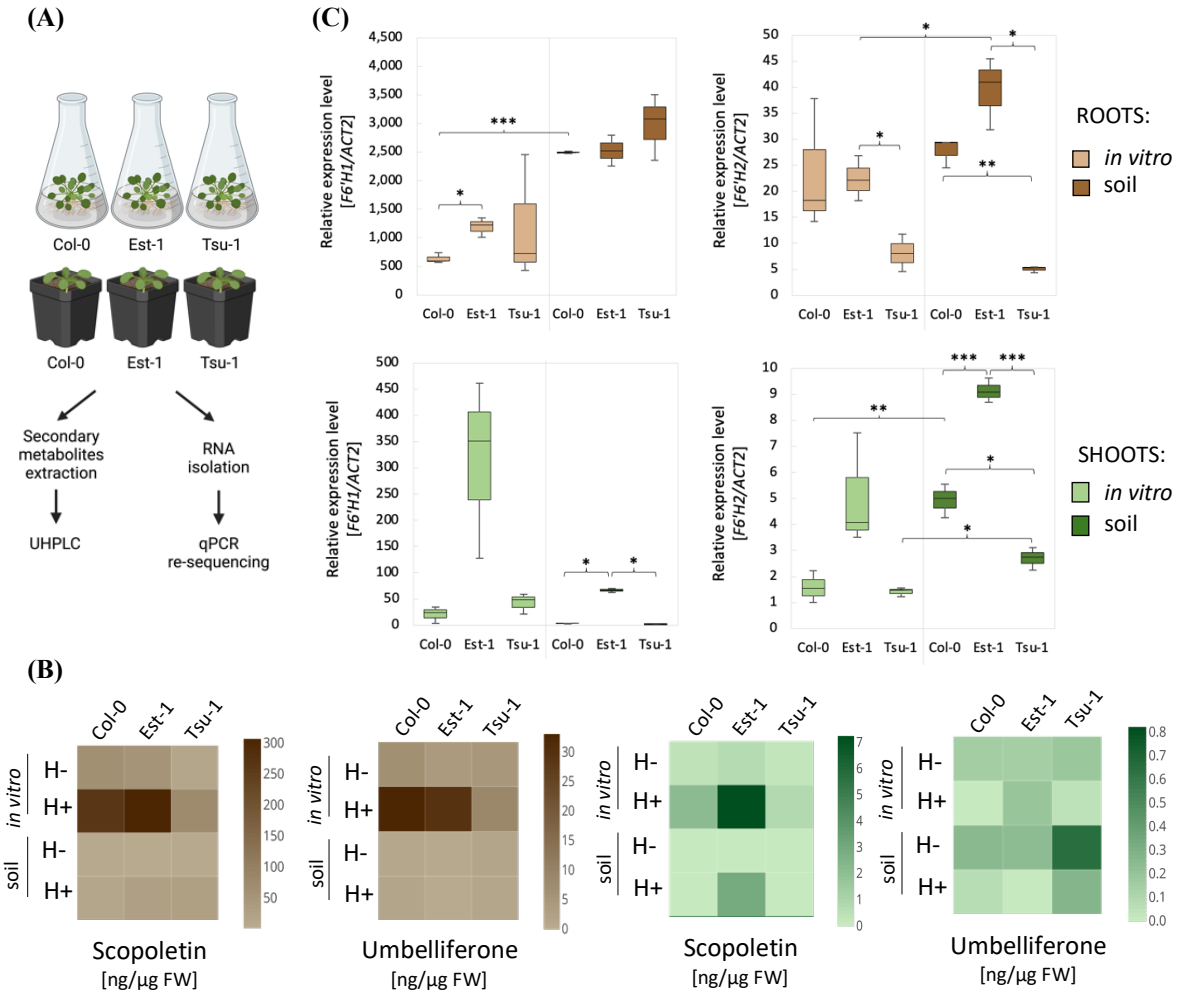


Figure 3.
(A) Schematic representation of the experimental workflow outlined in Publication No. 3.
(B) Quantification of umbelliferone and scopoletin in *A. thaliana* Col-0, Est-1 and Tsu-1 natural accessions, in roots (brown) and shoots (green) after cultivation *in vitro* or in soil.
(C) Relative expression level of *F6'H1* and *F6'H2* in *A. thaliana* Col-0, Est-1 and Tsu-1 natural accessions, in roots (brown) and shoots (green) after cultivation *in vitro* or in soil.

Gene expression analysis further confirmed the differential regulation of coumarin biosynthesis genes under distinct growth conditions. The *F6'H1* gene, a major contributor to scopoletin biosynthesis, was highly expressed in roots, particularly in soil-grown plants (Fig. 3C). Interestingly, *4CL1* and *4CL2*, acting upstream in coumarin biosynthesis, exhibited opposing expression trends between *in vitro* and soil conditions. Genes involved in lignin biosynthesis displayed tissue-specific expression patterns, *CCR1* has more highly expressed in leaves, while *CCR2* transcripts were more abundant in roots. Genomic re-sequencing identified polymorphisms in *F6'H2*, *4CL3*, and *CCR2*, suggesting that genotypic differences contribute to the variability in coumarin biosynthesis. Notably, SNPs in *4CL3* and *F6'H2* were found to alter amino acid sequences, potentially impacting enzyme function.

Presented study underscores the significant genetic and environmental control over coumarin biosynthesis in *A. thaliana*. The observed differences between *in vitro* and soil-grown plants suggest that stress conditions, such as nutrient availability and root architecture changes, influence coumarin production. Furthermore, the differential expression of key biosynthetic genes highlights the dynamic nature of the phenylpropanoid pathway, which is regulated by environmental factors [Dong & Lin, 2021]. Distinct metabolite profiles among Col-0, Est-1 and Tsu-1 likely reflect underlying genetic variation in biosynthetic pathway genes, reinforcing the idea that natural diversity plays a crucial role in shaping secondary metabolism. Given that coumarins contribute to Fe acquisition and pathogen defense (Perkowska *et al.*, 2021b), understanding their regulation provides insights into plant adaptation mechanisms in nutrient-limited and pathogen-rich environments.

To summarize the results presented in the three publications that collectively address the first objective of this PhD thesis, our findings enhance the understanding of how genetic and environmental factors influence coumarin biosynthesis and lay the foundation for future research into the molecular mechanisms underlying specialized metabolism in plants. By integrating metabolomics and genomic analyses, our study provides valuable insights into how genetic variation shapes plant secondary metabolism. Equally important to the integration of genetic, metabolic, and transcriptional data is the identification and functional characterization of genes encoding key enzymes in secondary metabolism pathways, along with elucidating the regulatory networks that govern plant responses to environmental stress factors. These insights may have important implications for a better understanding of plant adaptation, crop improvement, and biotechnological applications.

<u>Aim 2.</u> Comprehensive analysis of the regulatory network controlling the switch between coumarin and lignin biosynthesis in *A. thaliana* under osmotic and low pH stress using integrative transcriptomic and metabolomic approaches.

<u>Aim 3.</u> Functional characterization of two *A. thaliana* dioxygenases DOXC21-A, its NAT, and DOXC21-B potentially involved in ortho-hydroxylation of coumarin and lignin precursors.

To elucidate the molecular mechanisms underlying the balance between coumarin and lignin biosynthesis under abiotic stress, we investigated the regulatory networks responding to osmotic and low pH stress in *A. thaliana* using integrative transcriptomic and metabolomic approaches, alongside the functional characterization of two candidate dioxygenases (AT3G19000 and AT3G19010) potentially involved in ortho-hydroxylation of phenylpropanoid intermediates.

<u>6.4. Publication No. 4 (manuscript in preparation)</u>

Elucidating the role and regulatory mechanism of *Arabidopsis thaliana DOXC21-A*, its NAT and *DOXC21-B* in plant adaptation to terrestrial conditions. **Perkowska, I**., Olry, A., Barrit, T., Dobek, A., Magot, F., Charles, C., Siwinska, J., Lojkowska, E., & Ihnatowicz, A. (**2025a**)

Extending our previous studies of the phenylpropanoid pathway regulation under stress, this study investigates the functional roles of *A. thaliana* 2-oxoglutarate-dependent dioxygenase DOXC21-A (AT3G19000), its associated natural antisense transcript (NAT, AT3G19002), and DOXC21-B (AT3G19010) paralog. As reported by Kawai *et al.* (2014), DOXC21-A and DOXC21-B form distinct clade within DOXC subfamily, with orthologs conserved across major land plants lineages, from bryophytes to angiosperms (Fig. 4A). Phylogenetic analysis of *A. thaliana* 2OGD family places the DOXC21 clade in a separate lineage from well characterized dioxygenases (F6'H1, F6'H2, and S8H) involved in coumarin biosynthesis (Fig. 4B) (Siwinska *et al.*, 2018). Because of the close link to key enzymes in specialized metabolism, we decided to explore the role of DOXC21-A, its NAT and DOXC21-B in the context of stress adaptation.

To investigate the functional roles of DOXC21-A and its associated NAT, we employed an integrative approach combining *in silico* analysis, transcriptomic profiling, untargeted metabolomics, and reverse genetics. *A. thaliana* Col-0 WT and loss-of-function mutants deficient in either DOXC21-A (*doxc21-a*) or its NAT (*nat1, nat2*) were cultivated hydroponically under conditions designed to mimic environmentally relevant stresses. The baseline control condition was set at pH 4.5, a mildly acidic environment that simulates nutrient-poor soils and is known to influence nutrient solubility, root architecture, and secondary metabolism. While mildly stressful, this pH level provides a physiologically relevant context for evaluating additional stress responses. To ensure sufficient trace element availability under these conditions, a modified 10× Heeg nutrient solution was used, in which the micronutrient concentration was increased tenfold to avoid nutrient

limitations caused by reduced solubility at low pH [Neina, 2019]. Plants were subsequently subjected to osmotic stress by polyethylene glycol (PEG) supplementation to simulate water deficit conditions. Transcriptomic analysis (RNA-Seq) was performed for plants grown hydroponically in combined osmotic and low pH stresses, while metabolite profiling for plants grown in control and stress conditions (Fig. 5A). qRT-PCR was employed to confirm RNA-Seq data and our previous results. To broaden phenotypic assessment, additional soil cultures were conducted including *doxc21-b* mutant lines, under oxidative and cold stresses. For the functional characterization of DOXC21-A and DOXC21-B we overexpressed both proteins in heterologous systems, to test them for *in vitro* metabolization and *in vivo* transient expression in *Nicotiana benthamiana.*

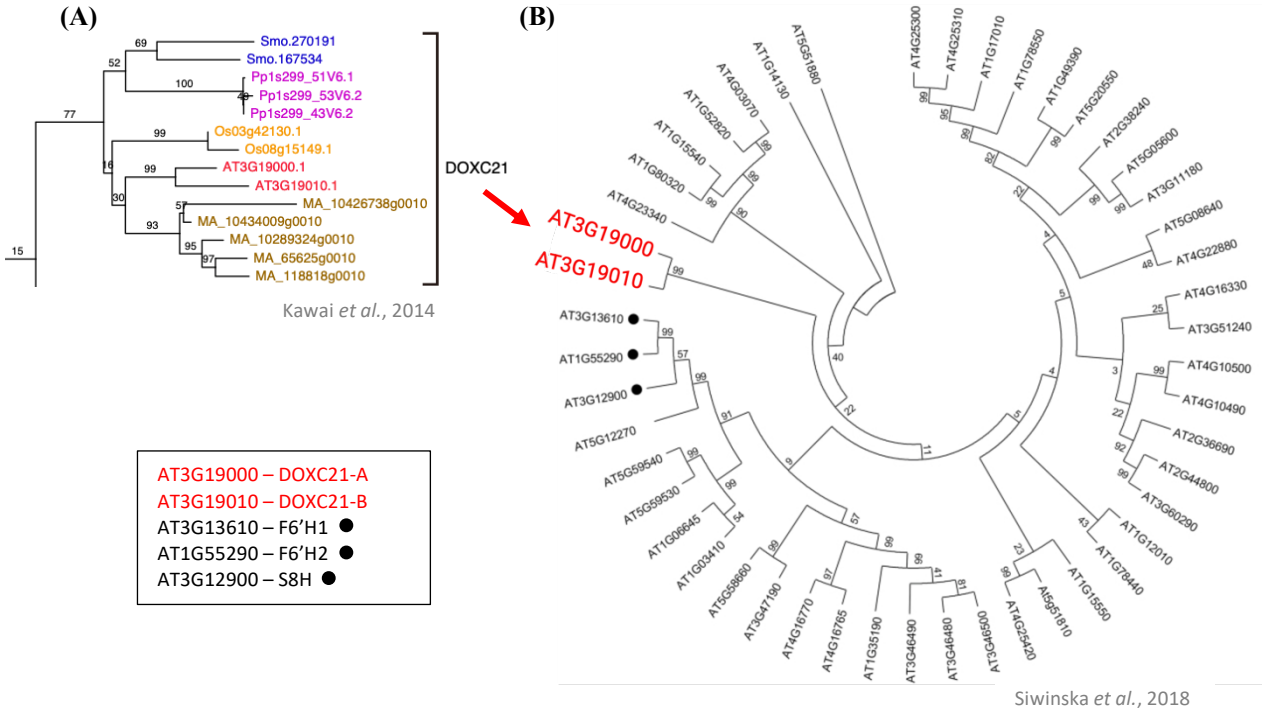


Figure 4. Phylogenetic placement of DOXC21-A and DOXC21-B.
(A) Phylogenetic tree showing the DOXC21 clade across representative land plant species, adapted from Kawai *et al.* (2014). *A. thaliana* DOXC21-A (AT3G19000) and DOXC21-B (AT3G19010) cluster together with homologs from diverse taxa, supporting evolutionary conservation across land plants.
(B) Circular tree of the *A. thaliana* 2OGD family, adapted from Siwinska *et al.* (2018). DOXC21-A and DOXC21-B are positioned near F6'H1, F6'H2, and S8H – well-characterized dioxygenases involved in coumarin biosynthesis – highlighting potential functional proximity within secondary metabolism.

Transcriptomic profiling revealed the distinct regulatory roles of DOXC21-A and its associated NAT under stress conditions. While both mutants showed significant transcriptional reprogramming under osmotic and low pH stresses, their differentially expressed genes (DEGs) profiles overlapped only partially, indicating non-redundant mechanisms. In *nat* mutants flavonoid biosynthesis genes were upregulated, suggesting a metabolic shift favoring antioxidant and protective secondary metabolites. In contrast, *doxc21-a* mutant exhibited broader transcriptional disruption, including

consistent downregulation of lignin biosynthesis genes and altered expression of multiple gene families involved in transport (notably ABC transporters), transcriptional regulation, and RNA metabolism. These results suggest that DOXC21-A and its NAT take part in abiotic stress responses, but they function through separate molecular pathways that meet at common metabolic and regulatory points.

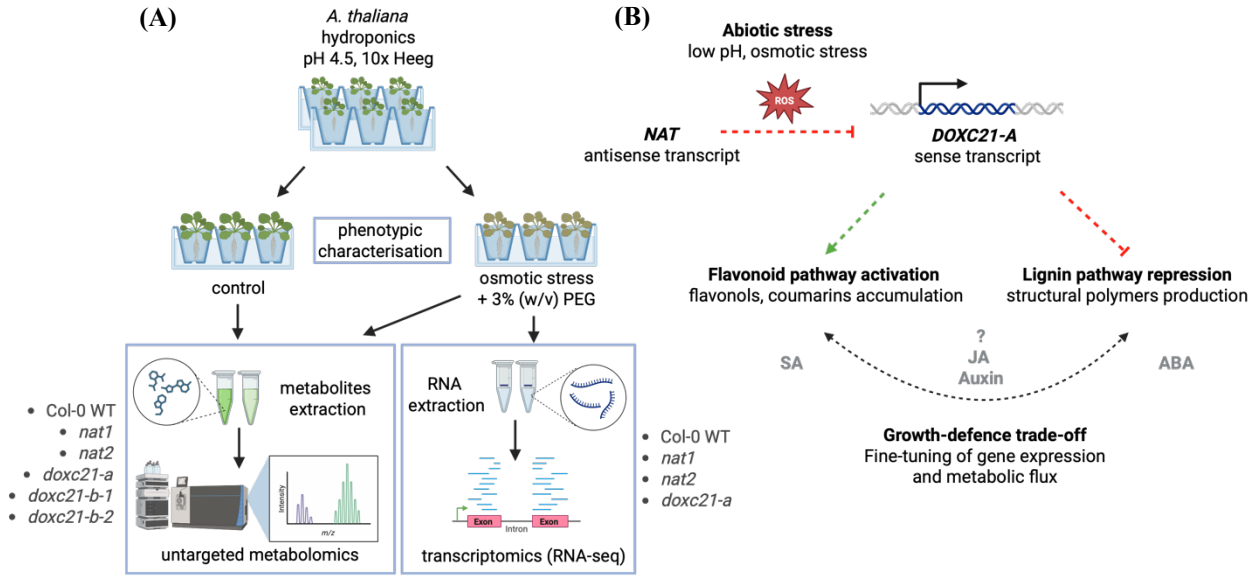


Figure 5. Experimental workflow and proposed model for DOXC21-A/NAT-mediated regulation under osmotic stress in *A. thaliana*.

(A) Schematic overview of the experimental setup. *A. thaliana* Col-0 WT and mutant lines (*doxc21-a, nat1, nat2, doxc21-b-1, doxc21-b-2*) were grown hydroponically in control conditions (pH 4.5, 10× Heeg medium) for three weeks, followed by either continued control or 3% PEG-induced osmotic stress treatment. RNA was extracted one week after stress initiation from leaves for transcriptome profiling *via* RNA-seq (Col-0 WT, *doxc21-a, nat1, nat2*), while metabolites were collected from leaves after two weeks for untargeted metabolomic analysis using Orbitrap mass spectrometry (Col-0 WT, *doxc21-a, nat1, nat2, doxc21-b-1, doxc21-b-2*).

(B) Proposed regulatory model. Under abiotic stress (e.g., low pH, osmotic), the *DOXC21-A* gene and its associated NAT form a dual regulatory module. The NAT is hypothesized to modulate *DOXC21-A* transcription post-transcriptionally *via* antisense-mediated repression. *DOXC21-A*, in turn, influences metabolic partitioning within the phenylpropanoid pathway by promoting flavonoid accumulation (e.g., flavonols, coumarins) and repressing lignin biosynthesis. These opposing outputs align with hormone signaling networks (SA, ABA, JA, auxin) and reflect a broader growth-defense trade-off strategy involving metabolic and transcriptional fine-tuning under stress [Lee et al., 2023].

Untargeted metabolomic profiling supported these transcriptomic trends, revealing increased accumulation of flavonoids in *nat* mutants, whereas *doxc21-a* plants exhibited a broader spectrum of non-canonical metabolites, including benzenoid volatiles and stress-associated peptides. These findings suggest that the NAT facilitates targeted metabolic reinforcement under stress, while DOXC21-A functions as a higher-order transcriptional regulator orchestrating more extensive metabolic reprogramming. In parallel, evolutionary analysis suggests that *DOXC21-A* and *DOXC21-B* originated from a gene duplication event and may have functionally diverged.

Despite their high sequence similarity, *DOXC21-B* lacks NAT-associated regulation and exhibits expression patterns linked to light signaling and redox balance, in contrast to stress-responsive profile of *DOXC21-A*.

We also explored potential molecular mechanisms underlying NAT-mediated regulation, including transcriptional interference, RNA splicing modulation, and chromatin remodeling [Ferdous *et al.*, 2015; Jha *et al.,* 2020, Fuchs *et al.*, 2021]. Our proposed model (Fig. 5B) places DOXC21-A/NAT at the center of growth-defense trade-offs, dynamically balancing repression of lignin biosynthesis with activation of flavonoid pathways, likely mediated *via* hormone signaling pathways such as JA, SA, and auxin [Zhao *et al.*, 2018; Mao *et al.*, 2021; Lee *et al.*, 2023]. This work highlights DOXC21-A and NAT as a regulatory module capable of integrating abiotic stress signals with transcriptional and metabolic reprogramming. Together, these findings expand our understanding of the adaptive plasticity of plant secondary metabolism and provide framework for future studies on the roles of 2OGDs in plant stress biology. Moreover, they suggest promising targets for regulatory engineering strategies aimed at improving crop resilience under climate variability and environmental stress conditions.

<u>Aim 4.</u> Characterization of *A. thaliana* UDP-glucosyltransferase 79B9 involved in Fe deficiency and osmotic stress response.

<u>6.5. Publication No. 5 (manuscript in preparation) (2025b)</u>

Arabidopsis UGT79B9-mediated crosstalk between secondary metabolism, root exudation and plant responses to osmotic stress and iron deficiency. **Perkowska, I**., Siwinska, J., Dobek, A, Magot, F., Grosjean, J., Hehn, A., Olry, A., Lojkowska, E., & Ihnatowicz, A. (2025b)

The last aim of my PhD thesis was to characterize *A. thaliana* UDP-glucosyltransferase 79B9 involved in Fe deficiency and osmotic stress response. Based on the previous functional characterization of 2OGDs and their role in stress-related transcriptional and metabolic reprogramming, the final part of this doctoral thesis focuses on characterizing a previously unstudied UDP-glycosyltransferase gene, *UGT79B9* (*AT5G53990*), in *A. thaliana. UGT79B9* was identified through QTL mapping as a candidate gene involved in coumarin biosynthesis, with expression localized primarily to roots (Siwinska *et al.*, 2014).

Transient expression of *UGT79B9* in *N. benthamiana* leaves led to elevated scopolin content, a glycosylated form of scopoletin, confirming *in planta* enzymatic activity. To gain further insight, *ugt79b9* knock-out mutants displayed disrupted scopolin-to-scopoletin ratios, and increased root and rosette growth under osmotic stress, suggesting altered stress responsiveness. No significant differences in coumarin content were found between Col-0 WT and *ugt79b9* under control or stress conditions, which necessitated in-depth studies using untargeted metabolomic profiling to explore unidentified metabolic changes.

To further explore the gene's function, *A. thaliana* overexpression lines were constructed by introducing the UGT79B9 coding sequence under the control of the constitutive 35S promoter. Untargeted metabolomics (UHPLC-MS) was performed for methanolic extracts from roots and exudates of plants grown under PEG-induced osmotic stress and Fe-deficiency. Metabolite profiling revealed clear metabolic shifts in *ugt79b9* and overexpression lines. Particularly, *ugt79b9* mutant lacked several m/z features that accumulated significantly in WT and overexpressing plants. Moreover, m/z 413.16 was completely absent in mutant exudates but strongly induced in overexpression lines, suggesting that UGT79B9 is essential for its synthesis or secretion. Metabolite clustering analyses grouped many of the UGT79B9-dependent signals with known glycosylated phenylpropanoids and coumarins, indicating biosynthetic and structural relationships. Soil-grown *ugt79b9* plants also showed lower electrolyte leakage and higher water content after drought in comparison to Col-0 WT, reinforcing UGT79B9's functional importance in abiotic stress mitigation.

In silico analyses and protein alignments placed UGT79B9 within a group of UGTs associated with coumarin glycosylation (Wang et al., 2023). Comparative analysis across *A. thaliana*

accessions, particularly Est-1, revealed functional divergence, suggesting that natural variation in UGT79B9 activity may contribute to adaptive metabolic traits.

Altogether, this work delivers the first biological evidence for UGT79B9's function in *A. thaliana* and suggests that it may act as a metabolic switch in specialized metabolism, linking coumarin biosynthesis to stress signaling. Its role appears particularly relevant under Fe-deficient and osmotic stress conditions – environments that are becoming increasingly common due to climate change. The potential involvement of UGT79B9 in managing the trade-off between secondary metabolism and stress resilience opens promising avenues for improving crop tolerance through metabolic engineering.

7. CONCLUSION

This doctoral dissertation provides a comprehensive insight into the complex interplay between genetic variation, environmental stress responses, and specialized metabolism in *A. thaliana*. By focusing on coumarins – secondary metabolites derived from the phenylpropanoid pathway – this work elucidates how different genotypes modulate metabolic outputs in response to both abiotic and biotic stimuli, such as iron deficiency, osmotic stress, acidic conditions, and bacterial *Dickeya* spp. infection. Through a series of five thematically linked studies, we systematically analyzed the accumulation patterns of coumarins and their glycosides across diverse accessions and pointed to key genetic factors and transcriptional regulators that may be involved in their biosynthesis.

The research suggests that DOXC21-A, its natural antisense transcript (NAT), and its paralog DOXC21-B form a unique regulatory module that may influence the balance between lignin and flavonoid biosynthesis during stress adaptation. Transcriptomic and metabolomic profiling of knockout mutants revealed distinct and overlapping functions, indicating that NATs can serve as important molecular switches fine-tuning gene expression in response to environmental cues. Additionally, the characterization of UGT79B9 as a stress-responsive glycosyltransferase revealed its contribution to the regulation of metabolite solubility and root exudation under stress, further highlighting the adaptive plasticity of the glycosylation machinery.

Overall, the results presented in this thesis expand our knowledge of the genetic and molecular frameworks controlling coumarin metabolism in plants. They also provide a conceptual model linking environmental perception to phenylpropanoid flux redirection through transcriptional and post-transcriptional regulation. These findings may contribute to future biotechnological strategies for engineering stress-tolerant plants by targeting key metabolic regulators, such as 2OGDs, UGTs, and NATs, offering potential applications in sustainable agriculture and crop improvement.

8. LIST OF REFERENCES:

Alekseeva, E., Doebeli, M., & Ispolatov, I. (2020). Evolutionary adaptation of high-diversity communities to changing environments. *Ecology and evolution*, 10(21), 11941-11953.

Ali, S., Tyagi, A., & Mir, Z. A. (2024). Plant immunity: at the crossroads of pathogen perception and defense response. *Plants*, 13(11), 1434.

Alseekh, S., Kostova, D., Bulut, M., & Fernie, A. R. (2021). Genome-wide association studies: assessing trait characteristics in model and crop plants. *Cellular and Molecular Life Sciences*, *78*, 5743-5754.

Alseekh, S., & Fernie, A. R. (2023). Expanding our coverage: strategies to detect a greater range of metabolites. *Current Opinion in Plant Biology*, 73, 102335.

Anderson, Jill T., and Bao-Hua Song. "Plant adaptation to climate change—Where are we?." *Journal of Systematics and Evolution* 58.5 (2020): 533-545.

Anderson, J. T., Willis, J. H., & Mitchell-Olds, T. (2011). Evolutionary genetics of plant adaptation. *Trends in Genetics*, *27*(7), 258-266.

Bomblies, K., & Peichel, C. L. (2022). Genetics of adaptation. *Proceedings of the National Academy of Sciences*, *119*(30), e2122152119.

Butnariu, M., & Tietel, Z. (2023). Biosynthesis, purification, characterization and uses of natural compounds in plants. *Frontiers in Plant Science*, 14, 1162676.

Cheaib, M., Nguyen, H. T., Couderc, M., Serret, J., & Soriano, A. (2024). Transcriptomic and metabolomic reveal OsCOI2 as the jasmonate-receptor master switch in rice root. *PLoS One*

Cusack, S. A., Wang, P., Lotreck, S. G., Moore, B. M., Meng, F., Conner, J. K., ... & Shiu, S. H. (2021). Predictive models of genetic redundancy in *Arabidopsis thaliana*. *Molecular Biology and Evolution,* 38(8), 3397-3414.

Cushman, J. C., Denby, K., & Mittler, R. (2022). Plant responses and adaptations to a changing climate. *The Plant Journal*, *109*(2).

De Vries, J., & Archibald, J. M. (2018). Plant evolution: landmarks on the path to terrestrial life. *New Phytologist*, 217(4), 1428-1434.

Delaux, P. M., & Schornack, S. (2021). Plant evolution driven by interactions with symbiotic and pathogenic microbes. *Science* 371: eaba6605.

Dixon, R. A., & Dickinson, A. J. (2024). A century of studying plant secondary metabolism—from “what?” to “where, how, and why?”. *Plant Physiology*, *195*(1), 48-66.

Domozych, D. S., Popper, Z. A., & Sørensen, I. (2016). Charophytes: evolutionary giants and emerging model organisms. *Frontiers in plant science*, 7, 1470.

Dong, N. Q., & Lin, H. X. (2021). Contribution of phenylpropanoid metabolism to plant development and plant–environment interactions. *Journal of integrative plant biology*, 63(1), 180-209.

Duan, Z., Yan, Q., Wu, F., Wang, Y., Wang, S., Zong, X., ... & Zhang, J. (2021). Genome-Wide analysis of the UDP-glycosyltransferase family reveals its roles in coumarin biosynthesis and abiotic stress in *Melilotus albus*. *International Journal of Molecular Sciences*, 22(19), 10826.

Farrow, S. C., & Facchini, P. J. (2014). Functional diversity of 2-oxoglutarate/Fe (II)-dependent dioxygenases in plant metabolism. *Frontiers in plant science*, 5, 524.

Ferdous, J., Hussain, S. S., & Shi, B. J. (2015). Role of micro RNA s in plant drought tolerance. *Plant biotechnology journal*, 13(3), 293-305.

Foyer, C. H., & Kranner, I. (2023). Plant adaptation to climate change. *Biochemical Journal*, *480*(22), 1865-1869.

Friedman, J., & Barrett, S. C. (2009). Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. *Annals of botany,* 103(9), 1515-1527.

Fuchs, A., Riegler, S., Ayatollahi, Z., Cavallari, N., Giono, L. E., Nimeth, B. A., ... & Kalyna, M. (2021). Targeting alternative splicing by RNAi: from the differential impact on splice variants to triggering artificial pre-mRNA splicing. *Nucleic Acids Research*, 49(2), 1133-1151.

Fürst-Jansen, J. M., de Vries, S., & de Vries, J. (2020). Evo-physio: on stress responses and the earliest land plants. Journal of Experimental Botany, 71(11), 3254-3269.

Garrido-Cardenas, J. A., Mesa-Valle, C., & Manzano-Agugliaro, F. (2018). Trends in plant research using molecular markers. *Planta*, *247*, 543-557.

Glasner, M. E., Truong, D. P., & Morse, B. C. (2020). How enzyme promiscuity and horizontal gene transfer contribute to metabolic innovation. *The FEBS journal*, 287(7), 1323-1342.

Graham, L., Lewis, L. A., Taylor, W., Wellman, C., & Cook, M. (2014). Early terrestrialization: transition from algal to bryophyte grade. *Photosynthesis in bryophytes and early land plants*, 9-28.

Gruber, B. D., Giehl, R. F., Friedel, S., & von Wirén, N. (2013). Plasticity of the Arabidopsis root system under nutrient deficiencies. *Plant physiology*, 163(1), 161-179.

Ha, C. M., Rao, X., Saxena, G., & Dixon, R. A. (2021). Growth–defense trade-offs and yield loss in plants with engineered cell walls. *New Phytologis*t, 231(1), 60-74.

Haimlich, S., Fridman, Y., Khandal, H., Savaldi-Goldstein, S., & Levy, A. (2024). Widespread horizontal gene transfer between plants and bacteria. *ISME communications*, 4(1), ycae073.

Holz, M., Zarebanadkouki, M., Benard, P., Hoffmann, M., & Dubbert, M. (2024). Root and rhizosphere traits for enhanced water and nutrients uptake efficiency in dynamic environments. *Frontiers in Plant Science*, 15, 1383373.

Ihnatowicz, A., Siwinska, J., **Perkowska, I**., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E. & Olry, A. (2024). Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of *Arabidopsis thaliana*. *BMC Plant Biology*, 24(1), 806.

Jan, R., Asaf, S., Numan, M., Lubna, & Kim, K. M. (2021). Plant secondary metabolite biosynthesis and transcriptional regulation in response to biotic and abiotic stress conditions. *Agronomy,* 11(5), 968.

Jha, U. C., Nayyar, H., Jha, R., Khurshid, M., Zhou, M., Mantri, N., & Siddique, K. H. (2020). Long non-coding RNAs: emerging players regulating plant abiotic stress response and adaptation. *BMC Plant Biology*, 20, 1-20.

Kai, K., Shimizu, B. I., Mizutani, M., Watanabe, K., & Sakata, K. (2006). Accumulation of coumarins in *Arabidopsis thaliana*. *Phytochemistry*, 67(4), 379-386.

Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., ... & Shimizu, B. I. (2008). Scopoletin is biosynthesized *via* ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in *Arabidopsis thaliana*. *The Plant Journal,* 55(6), 989-999.

Kane, C. N., Jordan, G. J., Jansen, S., & McAdam, S. A. (2020). A permeable cuticle, not open stomata, is the primary source of water loss from expanding leaves. *Frontiers in Plant Science*, 11, 774.

Kawai, Y., Ono, E., & Mizutani, M. (2014). Evolution and diversity of the 2–oxoglutarate-dependent dioxygenase superfamily in plants. *The Plant Journal*, 78(2), 328-343.

Koornneef, M., Alonso-Blanco, C., & Vreugdenhil, D. (2004). Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu. Rev. Plant Biol.*, 55(1), 141-172.

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant methods*, 9, 1-9.

Kover, P. X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I. M., Purugganan, M. D., ... & Mott, R. (2009). A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS genetics*, 5(7), e1000551.

Krämer, U. (2015). Planting molecular functions in an ecological context with *Arabidopsis thaliana. Elife*, *4*, e06100.

Kreis, W., & Munkert, J. (2019). Exploiting enzyme promiscuity to shape plant specialized metabolism. *Journal of experimental botany,* 70(5), 1435-1445.

Kumar, S., Sachdeva, S., Bhat, K. V., & Vats, S. (2018). Plant responses to drought stress: physiological, biochemical and molecular basis. *Biotic and abiotic stress tolerance in plants*, 1-25.

Kumar, P., Kumar, P., Verma, V., Irfan, M., Sharma, R., & Bhargava, B. (2023). How plants conquered land: evolution of terrestrial adaptation. *Journal of Evolutionary Biology*, *35*(5), 5-14.

Lan, P., Li, W., Wen, T. N., Shiau, J. Y., Wu, Y. C., Lin, W., & Schmidt, W. (2011). iTRAQ protein profile analysis of Arabidopsis roots reveals new aspects critical for iron homeostasis. *Plant physiology*, 155(2), 821-834.

Lasky, J. R., Josephs, E. B., & Morris, G. P. (2023). Genotype–environment associations to reveal the molecular basis of environmental adaptation. *The Plant Cell*, *35*(1), 125-138.

Lawson, T., & Blatt, M. R. (2014). Stomatal size, speed, and responsiveness impact on photosynthesis and water use efficiency. *Plant physiology*, 164(4), 1556-1570.

Le Roy, J., Huss, B., Creach, A., Hawkins, S., & Neutelings, G. (2016). Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Frontiers in plant science*, 7, 735.

Leonardelli, M., Tissot, N., Podolec, R., Ares-Orpel, F., Glauser, G., Ulm, R., & Demarsy, E. (2024). Photoreceptor-induced sinapate synthesis contributes to photoprotection in Arabidopsis. *Plant Physiology*, 196(2), 1518-1533.

Lewandowska, M., Keyl, A., & Feussner, I. (2020). Wax biosynthesis in response to danger: its regulation upon abiotic and biotic stress. *New Phytologist*, 227(3), 698-713.

Maeda, H. A. (2019). Evolutionary diversification of primary metabolism and its contribution to plant chemical diversity. *Frontiers in Plant Science*, 10, 881.

Mao, Y., Xu, J., Wang, Q., Li, G., Tang, X., Liu, T., ... & Lu, Y. (2021). A natural antisense transcript acts as a negative regulator for the maize drought stress response gene *ZmNAC48*. *Journal of Experimental Botany*, 72(7), 2790-2806.

Ménard, D., Blaschek, L., Kriechbaum, K., Lee, C. C., Serk, H., Zhu, C., ... & Pesquet, E. (2022). Plant biomechanics and resilience to environmental changes are controlled by specific lignin chemistries in each vascular cell type and morphotype. *The Plant Cell*, 34(12), 4877-4896.

Mukherjee, D., Saha, D., Acharya, D., Mukherjee, A., & Ghosh, T. C. (2022). Interplay between gene expression and gene architecture as a consequence of gene and genome duplications: evidence from metabolic genes of *Arabidopsis thaliana*. *Physiology and Molecular Biology of Plants*, 28(5), 1091-1108.

Nascimento, L. B. D. S., & Tattini, M. (2022). Beyond photoprotection: The multifarious roles of flavonoids in plant terrestrialization. *International Journal of Molecular Sciences*, 23(9), 5284.

Neina, D. (2019). The role of soil pH in plant nutrition and soil remediation. *Applied and environmental soil science*, 2019(1), 5794869.

Niu, Y., & Xiang, Y. (2018). An overview of biomembrane functions in plant responses to high-temperature stress. *Frontiers in plant science*, 9, 915.

Perkowska, I., Siwinska, J., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Ihnatowicz, A. (2021). Identification and Quantification of Coumarins by UHPLC-MS in *A. thaliana* Natural Populations. *Molecules*, 26(6), 1804.

Perkowska, I., Potrykus, M., Siwinska, J., Siudem, D., Lojkowska, E., & Ihnatowicz, A. (2021). Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *International Journal of Molecular Sciences*, 22(12), 6449.

Potrykus, M., Golanowska, M., Hugouvieux-Cotte-Pattat, N., & Lojkowska, E. (2014). Regulators involved in *Dickeya solani* virulence, genetic conservation, and functional variability. *Molecular Plant-Microbe Interactions*, 27(7), 700-711.

Potrykus, M., Decorosi, F., **Perkowska, I.,** Viti, C., Mengoni, A., Hugouvieux-Cotte-Pattat, N., & Lojkowska, E. (2020). The metabolic shift in highly and weakly virulent *Dickeya solani* strains is

more affected by temperature than by mutations in genes encoding global virulence regulators. *FEMS Microbiology Ecology*, 96(3), fiaa023.

Pyne, M. E., Narcross, L., & Martin, V. J. (2019). Engineering plant secondary metabolism in microbial systems. *Plant physiology*, 179(3), 844-861.

Ramaroson, M. L., Koutouan, C., Helesbeux, J. J., Le Clerc, V., Hamama, L., Geoffriau, E., & Briard, M. (2022). Role of phenylpropanoids and flavonoids in plant resistance to pests and diseases. *Molecules*, 27(23), 8371.

Robe, K., Izquierdo, E., Vignols, F., Rouached, H., & Dubos, C. (2021). The coumarins: secondary metabolites playing a primary role in plant nutrition and health. *Trends in Plant Science*, 26(3), 248-259.

Robe, K., Conejero, G., Gao, F., Lefebvre-Legendre, L., Sylvestre-Gonon, E., Rofidal, V., ... & Dubos, C. (2021). Coumarin accumulation and trafficking in *Arabidopsis thaliana*: a complex and dynamic process. *New Phytologist*, 229(4), 2062-2079.

Rodríguez-Celma, J., Pan, I. C., Li, W., Lan, P., Buckhout, T. J., & Schmidt, W. (2013). The transcriptional response of Arabidopsis leaves to Fe deficiency. *Frontiers in Plant Science*, 4, 276.

Schreiber, M., Rensing, S. A., & Gould, S. B. (2022). The greening ashore. *Trends in Plant Science*, 27(9), 847-857.

Shah, S. H., Carlson, J. E., Niklas, K. J., Benavides-Mendoza, A., & Ricachenevsky, F. K. (2024). Deciphering mechanisms of plant adaptation and resistance under cold temperature stress. *Frontiers in Plant Science*, 15, 1460573.

Siwinska, J., Kadzinski, L., Banasiuk, R., Gwizdek-Wisniewska, A., Olry, A., Banecki, B., ... & Ihnatowicz, A. (2014). Identification of QTLs affecting scopolin and scopoletin biosynthesis in *Arabidopsis thaliana*. *BMC plant biology*, 14, 1-14.

Siwinska, J., Siatkowska, K., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., ... & Ihnatowicz, A. (2018). Scopoletin 8-hydroxylase: a novel enzyme involved in coumarin biosynthesis and iron-deficiency responses in Arabidopsis. *Journal of Experimental Botany*, 69(7), 1735-1748.

Stringlis, I. A., Yu, K. E., Feussner, K., de Jonge, R., Van Bentum, S., Van Verk, M. C., ... & Pieterse, C. M. (2018). MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proceedings of the National Academy of Sciences*, 115(22), E5213-E5222.

Stringlis, I. A., De Jonge, R., & Pieterse, C. M. (2019). The age of coumarins in plant–microbe interactions. *Plant and Cell Physiology,* 60(7), 1405-1419.

Takshak, S., & Agrawal, S. B. (2019). Defense potential of secondary metabolites in medicinal plants under UV-B stress. *Journal of Photochemistry and Photobiology B: Biology*, 193, 51-88.

Thompson, P. L., & Fronhofer, E. A. (2019). The conflict between adaptation and dispersal for maintaining biodiversity in changing environments. *Proceedings of the National Academy of Sciences*, 116(42), 21061-21067.

Tsai, H. H., & Schmidt, W. (2017). Mobilization of iron by plant-borne coumarins. *Trends in Plant Science*, 22(6), 538-548.

Wahab, A., Abdi, G., Saleem, M. H., Ali, B., Ullah, S., Shah, W., ... & Marc, R. A. (2022). Plants' physio-biochemical and phyto-hormonal responses to alleviate the adverse effects of drought stress: A comprehensive review. *Plants*, 11(13), 1620.

Wu, M., Northen, T. R., & Ding, Y. (2023). Stressing the importance of plant specialized metabolites: omics-based approaches for discovering specialized metabolism in plant stress responses. *Frontiers in Plant Science*, 14, 1272363.

Yang, X., Lu, M., Wang, Y., Wang, Y., Liu, Z., & Chen, S. (2021). Response mechanism of plants to drought stress. *Horticulturae*, 7(3), 50.

Yin, R., & Ulm, R. (2017). How plants cope with UV-B: from perception to response. *Current Opinion in Plant Biology*, 37, 42-48.

Zhao, X., Li, J., Lian, B., Gu, H., Li, Y., & Qi, Y. (2018). Global identification of *Arabidopsis* lncRNAs reveals the regulation of *MAF4* by a natural antisense RNA. *Nature communications*, 9(1), 5056.

Ziv, C., Zhao, Z., Gao, Y. G., & Xia, Y. (2018). Multifunctional roles of plant cuticle during plant-pathogen interactions. *Frontiers in plant science*, 9, 1088.

9. APPENDIX (original publications and supplementary materials)

Publication No. 1

AIM

1) **Comprehensive analysis of the correlation between accumulation of coumarins, plant genotypes (natural accessions and mutants), and environmental stress factors (including Fe availability and pathogen defense) to gain deeper insights into the metabolic adaptations and stress response in *Arabidopsis thaliana.***

Perkowska Izabela, Siwinska Joanna, Olry Alexandre, Grosjean Jeremy, Hehn Alain, Bourgaud Frederic, Lojkowska Ewa, Ihnatowicz Anna.
Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. *Molecules*, 26(6), 1804 (2021)

Due to space limitations, supplementary data (including additional figures and tables) are not included in the printed version of this thesis but are accessible online at https://doi.org/10.3390/molecules26061804.

molecules

MDPI

Article

Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations

Izabela Perkowska [1], Joanna Siwinska [1], Alexandre Olry [2], Jérémy Grosjean [2], Alain Hehn [2], Frédéric Bourgaud [3], Ewa Lojkowska [1] and Anna Ihnatowicz [1,*]

[1] Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk, University of Gdansk, Abrahama 58, 80-307 Gdansk, Poland; izabela.perkowska@phdstud.ug.edu.pl (I.P.); siwinskajoanna@gmail.com (J.S.); ewa.lojkowska@biotech.ug.edu.pl (E.L.)
[2] Université de Lorraine-INRAE, LAE, 54000 Nancy, France; alexandre.olry@univ-lorraine.fr (A.O.); jeremy.grosjean@univ-lorraine.fr (J.G.); alain.hehn@univ-lorraine.fr (A.H.)
[3] Plant Advanced Technologies, 54500 Vandœuvre-lès-Nancy, France; frederic.bourgaud@plantadvanced.com
* Correspondence: anna.ihnatowicz@biotech.ug.edu.pl; Tel.: +48-58-5236330

Abstract: Coumarins are phytochemicals occurring in the plant kingdom, which biosynthesis is induced under various stress factors. They belong to the wide class of specialized metabolites well known for their beneficial properties. Due to their high and wide biological activities, coumarins are important not only for the survival of plants in changing environmental conditions, but are of great importance in the pharmaceutical industry and are an active source for drug development. The identification of coumarins from natural sources has been reported for different plant species including a model plant *Arabidopsis thaliana*. In our previous work, we demonstrated a presence of naturally occurring intraspecies variation in the concentrations of scopoletin and its glycoside, scopolin, the major coumarins accumulating in Arabidopsis roots. Here, we expanded this work by examining a larger group of 28 Arabidopsis natural populations (called accessions) and by extracting and analysing coumarins from two different types of tissues–roots and leaves. In the current work, by quantifying the coumarin content in plant extracts with ultra-high-performance liquid chromatography coupled with a mass spectrometry analysis (UHPLC-MS), we detected a significant natural variation in the content of simple coumarins like scopoletin, umbelliferone and esculetin together with their glycosides: scopolin, skimmin and esculin, respectively. Increasing our knowledge of coumarin accumulation in Arabidopsis natural populations, might be beneficial for the future discovery of physiological mechanisms of action of various alleles involved in their biosynthesis. A better understanding of biosynthetic pathways of biologically active compounds is the prerequisite step in undertaking a metabolic engineering research.

Keywords: analytical methods; model plant; natural genetic variation; natural products; simple coumarins

Citation: Perkowska, I.; Siwinska, J.; Olry, A.; Grosjean, J.; Hehn, A.; Bourgaud, F.; Lojkowska, E.; Ihnatowicz, A. Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. *Molecules* **2021**, *26*, 1804. https://doi.org/10.3390/molecules26061804

Academic Editor: Maria João Matos

Received: 27 February 2021
Accepted: 18 March 2021
Published: 23 March 2021

1. Introduction

Coumarins are secondary metabolites widely distributed throughout the plant kingdom. They are synthetized via the phenylpropanoid biosynthesis pathway. We can distinguish several simple coumarins like coumarin, scopoletin (7-hydroxy-6-methoxycoumarin), esculetin (6,7-dihydroxycoumarin), umbelliferone (7-hydroxycoumarin), fraxetin (7,8-dihydroxy-6-methoxycoumarin), sideretin (5,7,8-trihydroxy-6-methoxycoumarin) and their respective glycosylated forms–scopolin, esculetin, skimmin, fraxin and sideretin-glycoside, respectively [1]. Figure 1 presents the semi-developed formula of simple coumarins and their glycosides derivatives identified in this research.

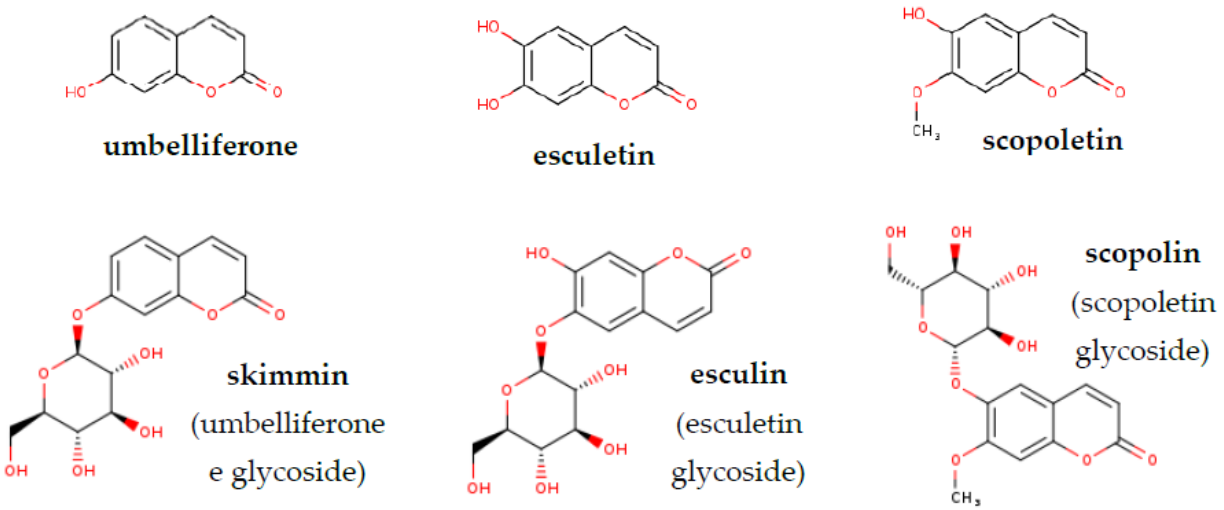


Figure 1. Chemical structures of simple coumarins and their glycosides analysed in this work (www.chem-space.com (accessed on 20 January 2021)).

Coumarins have been recognized for many years as an important class of pharmacologically active compounds. They have anticoagulant, anticancer, antiviral and anti-inflammatory properties [2]. In addition to the listed medicinal benefits, it was shown recently in numerous studies that coumarins play an important role in iron (Fe) homeostasis, oxidative stress response, plant-microbe interactions and that they can act as signalling molecules in plants [3–8]. In the last few years, an increasing number of reports concern the analysis of root extracts and root exudates that are rich in phenolic compounds, such as simple coumarins, which mediate multiple interactions in the rhizosphere. Coumarins were shown to have a strong impact on the plant interactions with microorganisms and play a crucial role in nutrient acquisition [6,9–18]. Moreover, the root-secreted scopoletin was proved to exert a selective antimicrobial action in the rhizosphere [8]. These numerous reports examining the biochemical and physiological functions of coumarins, make this class of specialized metabolites extremely interesting from a scientific and commercial point of view. The vast majority of these studies were performed using a reference accession, Col-0, of the model plant *Arabidopsis thaliana* (hereinafter Arabidopsis), and its mutants defective at various steps of coumarin biosynthesis.

Here, we conducted the qualitative and quantitative assessment of coumarin content in leaf and root tissue of a set of Arabidopsis natural populations (accessions). Numerous studies on primary and specialized metabolites profiling were conducted using the Arabidopsis model system [19–28]. Previously, due to the importance of coumarins for human health, most research on their metabolic profiling were carried out on plants of economic importance, such as e.g., sweet potato (*Ipomoea batatas* L.), rue (*Ruta graveolens* L.) or lettuce (*Lactuca sativa* L.) [29–34]. One of the first metabolic profiling of root exudates using Arabidopsis natural populations (Col-0, C24, Cvi-0, Ler) was made by Micallef et al. [35] who attempted to correlate them with the compositions of rhizobacterial communities. However, the authors of this work did not undertake the qualitative and quantitative evaluation of the isolated compounds. Consequently, they could only conclude that there are differences between Arabidopsis accessions in terms of the quality and quantity of released substances, which may have an impact on the composition of the rhizobacterial communities.

So far, only a few more studies focusing on the accumulation of coumarins in Arabidopsis natural populations have been published. As shown by our group [36], a significant natural variation in the accumulation of coumarins is present among the roots of Arabidopsis accessions. Using HPLC and GC-MS analytical methods, we identified and quantified coumarins in the roots of selected seven accessions-Antwerpen (An-1, Belgium),

Columbia (Col-0, Germany), Estland (Est-1, Estonia), Kashmir (Kas-2, India), Kondara (Kond, Tajikistan), Landsberg *erecta* (Ler, Poland) and Tsu (Tsu-1, Japan). Subsequently, we conducted a QTL mapping and identified new loci possibly underlying the observed variation in scopoletin and scopolin accumulation. Thereby, we demonstrated that Arabidopsis natural variation is an attractive tool for elucidating the basis of coumarin biosynthesis. Other studies focusing on differential accumulation of coumarins between Arabidopsis accessions were conducted by Mönchgesang et al. [14]. A non-targeted metabolite profiling of root exudates revealed the existence of distinct metabolic phenotypes for 19 Arabidopsis accessions. Scopoletin and its glycosides were among phenylpropanoids that differed in the exudates of tested accessions. This research group also focused on the plant-to-plant variability in root metabolite profiles of 19 Arabidopsis accessions [15]. In the current study, a larger set of 28 accessions was chosen, that represent a wide genetic variation existing in Arabidopsis. To increase the scope of this work, we extracted and quantified coumarins from two different types of tissue–roots and leaves. The latter one, in the light of our best knowledge, have never been tested for coumarin content using Arabidopsis natural variation. We believe our results will be beneficial for further studies focusing on a better understanding of coumarin physiological functions and the exact role of enzymes involved in their biosynthesis.

2. Results

2.1. UHPLC-MS Targeted Metabolite Profiling of Root and Leaf Tissues Reveals Distinct Metabolic Phenotypes for 28 Arabidopsis Accessions

The average content of each tested coumarins (Table 1) grouped by the 28 Arabidopsis accessions (Table 2), type of tissue (extracts from roots and leaves) and method of preparing extracts for analysis (without and after hydrolysis) were depicted through a general heatmap. For each compound, we quantified both the non-glycosylated coumarins—scopoletin (Figure 2A), umbelliferone (Figure 3A), esculetin (Figure 4A), and their respective glycosylated forms—scopolin (Figure 2B), skimmin (Figure 3B), and esculin (Figure 4B), respectively. The concentration µM was based on the fresh weight (FW).

Table 1. Coumarins and their glycosides identified in this study.

Peak Number	Retention Time tR (min)	Compound	Mass (*m*/*z* Ratio)	LOQ
(1)	14.5	Umbelliferone	163 ($M+H^+$)	0.2 µM
(2)	11.8	Esculetin	179 ($M+H^+$)	0.5 µM
(3)	14.8	Scopoletin	193 ($M+H^+$)	0.2 µM
(5)	10.1	Skimmin (glycosylated umbelliferone)	325 ($M+H^+$)	0.1 µM
(5)	11.8	Esculin (glycosylated esculetin)	341 ($M+H^+$)	0.1 µM
(6)	11	Scopolin (glycosylated scopoletin)	355 ($M+H^+$)	0.1 µM

Our analyses made evidence a significant variation in accumulation of all tested compounds between Arabidopsis accessions, both in roots and leaves. In accordance with the current state of knowledge [6,8,36–39], we identified the coumarin scopoletin and its glycoside, scopolin, to be the major metabolites that accumulate in Arabidopsis roots (Figure 2A,B).

Table 2. Basic information on the Arabidopsis accessions used in this study.

No.	Full Name	Abbreviation	Country of Origin
1	Antwerpen	An-1	Belgium
2	Bayreuth	Bay-0	Germany
3	Brunn	Br-0	Czech Republic
4	Coimbra	C24	Portugal
5	Canary Islands	Can-0	Spain
6	Columbia	Col-0	USA
7	Cape Verdi	Cvi-1	Cape Verde Islands
8	Eilenburg	Eil-0	Germany
9	Eringsboda	Eri-1	Sweden
10	Estland	Est-1	Russia
11	St. Maria d. Feiria	Fei-0	Portugal
12	Fukuyama	Fuk-1	Japan
13	Gabelstein	Ga-0	Germany
14	Hodja-Obi-Garm	Hog	Tajikistan
15	Kashmir	Kas-2	India
16	Kondara	Kondara	Tajikistan
17	Kyoto	Kyo-1	Japan
18	Lebjasche	Leb 3/4	Russia
19	Landsberg *erecta*	Ler-1	Germany
20	Nossen	No-0	Germany
21	Richmond	Ri-0	Canada
22	San Feliu	Sf-2	Spain
23	Shakdara	Sha-1	Tajikistan
24	Sorbo	Sorbo	Tajikistan
25	Tossa del Mar	Ts-5	Spain
26	Tsushima	Tsu-1	Japan
27	Vancouver	Van-0	Canada
28	Wassilewskija	Ws-0	Belarus

Scopoletin was the most abundant compound in each of the 28 Arabidopsis accessions studied, especially in the roots (from 2.61 to 151.90 μM), but interestingly this phytochemical was also detected in the leaf extracts (Table S1). The highest amount of scopoletin was detected in Bay-0, Br-0 and Kondara, respectively (Figure 2A), in samples prepared from the roots and subjected to hydrolysis. In non-hydrolyzed root samples, the highest content of scopoletin was detected for the same accessions–Kondara, Bay-0 and Br-0. As expected, scopoletin content in the leaf extracts was several dozen times smaller (from 0.03 to 2.6 μM) when compared with extracts prepared from the root tissue. Amount of scopoletin in the leaf sample was the highest in Br-0, Est-1 and Bay-0 when subjected to hydrolysis and in Est-1, Br-0, Col-0 and Bay-0 when not hydrolyzed.

Relatively large amounts of scopolin (from 2.94 to 67.26 μM) were found in almost all root extracts that were not subjected to hydrolysis (Figure 2B), the highest in Br-0, Fei-0, Ga-0, Leb 3/4, Ri-0, C24 and Bay-0 accessions (Table S1). We also identified some accessions (Fuk-1, Bay-0, Ri-0, Sha-1 and Eri-1) with relatively high content of scopolin in root extracts after hydrolysis (from 2.04 to 8.09 μM), most probably due to non-effective enzymatic reaction. Interestingly, another set of accessions (Est-1, Sorbo, Bay-0, Kyo-0, No-0, Ga-0, Fei-0, Ws-0) with relatively high scopolin concentration (from 2.58 to 6.48 μM) was also detected in leaf extracts not subjected to hydrolysis. As could be expected, in hydrolysed leaf samples in which sugar residues were cut off and most of scopolin was transformed into scopoletin, the amounts of scopolin were quite low or close to the LOQ.

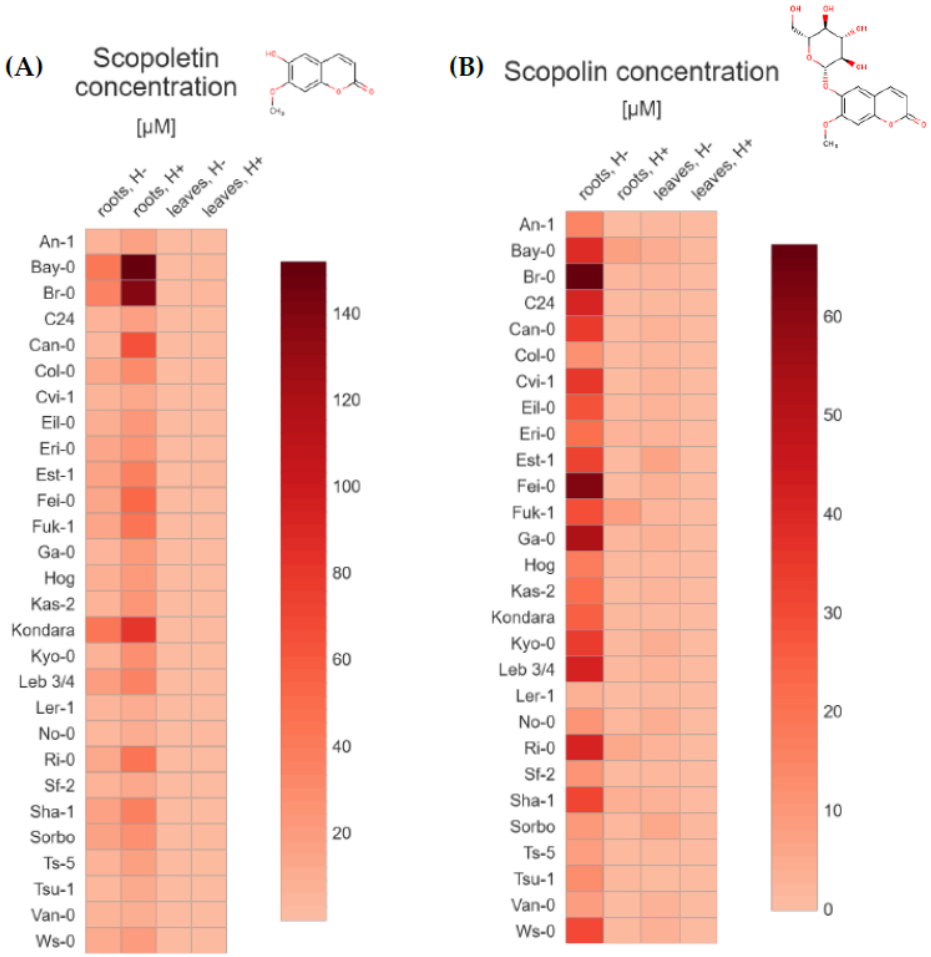


Figure 2. Heat maps based on the average (**A**) scopoletin and (**B**) scopolin concentration (µM/FW) in Arabidopsis tissue extracts from roots and leaves, without and after hydrolysis. The values used in the plots (https://app.displayr.com (accessed on 20 January 2021)) are the mean of 3 biological replicates. The mean values and standard deviations (±SD) are gathered in the supplementary materials (Table S1).

Interestingly, in this study, we identified small amounts of umbelliferone (from 0.02 to 1.64 µM) in Arabidopsis plants (Figure 3A). Importantly, we detected this phytochemical in all of the hydrolyzed root extracts (Table S2). The highest levels of umbelliferone were found in Bay-0, Ri-0, Est-1, Col-0, Br-0, C24, Sorbo, Fuk-1 and Leb 3/4, respectively. The quantity of umbelliferone in all leaf extract samples was below LOQ.

Skimmin, which is a glucoside of umbelliferone, was detected and quantified (from 0.69 to 19.80 µM) mostly in root samples of Arabidopsis accessions that were not subjected to enzymatic hydrolysis (Figure 3B). The highest levels were detected in extracts originating from Ga-0, C24, Ws-0, Ri-0, Est-1, Kyo-0 and Eri-0 accessions. It should be noted that skimmin could also be quantified (concentration from 0.04 to 18.76 µM) in all hydrolyzed root extract samples (Table S2), which needs further investigation.

Most of the results obtained for the leaf tissues were very low and near the LOQ, however in Eri-1, An-1, C24, Col-0, Van-0, Kondara, Ws-0, Ga-0, Fuk-1, Can-0 and Tsu-1 accessions, we observed values slightly above the limit.

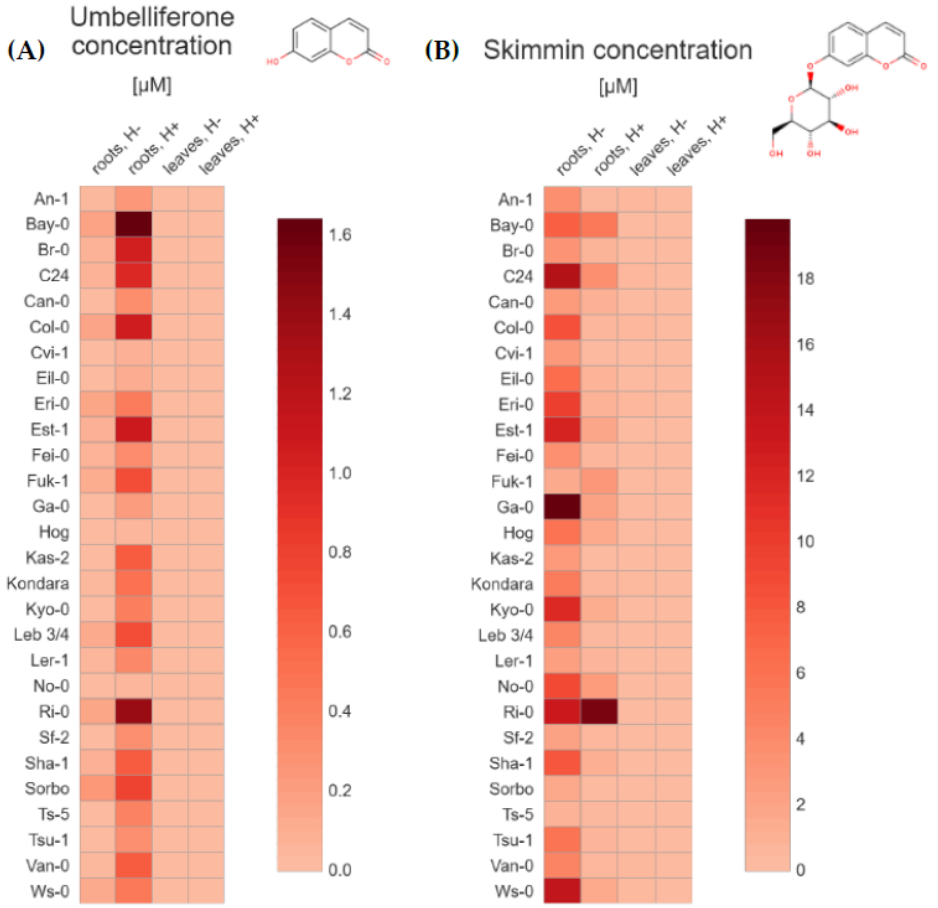


Figure 3. Heat maps based on the average (**A**) umbelliferone and (**B**) skimmin concentration (µM/FW) in Arabidopsis tissue extracts from roots and leaves, without and after hydrolysis. The values used in the plots (https://app.displayr.com (accessed on 20 January 2021)) are the mean of 3 biological replicates. The mean values and standard deviations (±SD) are gathered in the supplementary materials (Table S2).

Small amounts of esculetin were detected only in a few of root extracts (max. concentration 0.29 µM) and leaf samples (max. concentration 0.16 µM) (Figure 4A). In root non-hydrolysed samples, esculetin was present in Bay-0, Br-0 and Can-0 accessions, while in hydrolysed extracts it was detected in Can-0, Bay-0, Col-0, Ri-0 and Tsu-1 (Table S1). It may be puzzling that in some accessions, esculetin was only detected in samples which were not subjected to hydrolysis but not in the hydrolysed ones. This is the case for the root extract of Br-0 (0.07 µM), and leaf samples of C24, Br-0, An-1, Col-0, No-0, Ws-0 and Ri-0 accessions (from 0.01 to 0.16 µM, Table S3). In leaf samples after hydrolysis, only trace amount of esculetin was detected in Ws-0.

Esculin, which is a glycoside form of esculetin, was not found in any root extract (Figure 4B), except Col-0 sample with quantity near to LOQ (0.01 µM). Trace amounts of esculin were detected in some leaf extracts without hydrolysis (from 0.01 to 0.15 µM) with the highest content in Col-0 accession, and in the leaf samples subjected to hydrolysis (from 0.01 to 0.36 uM). Here, the highest esculin content was detected in Ws-0 accession (Table S3).

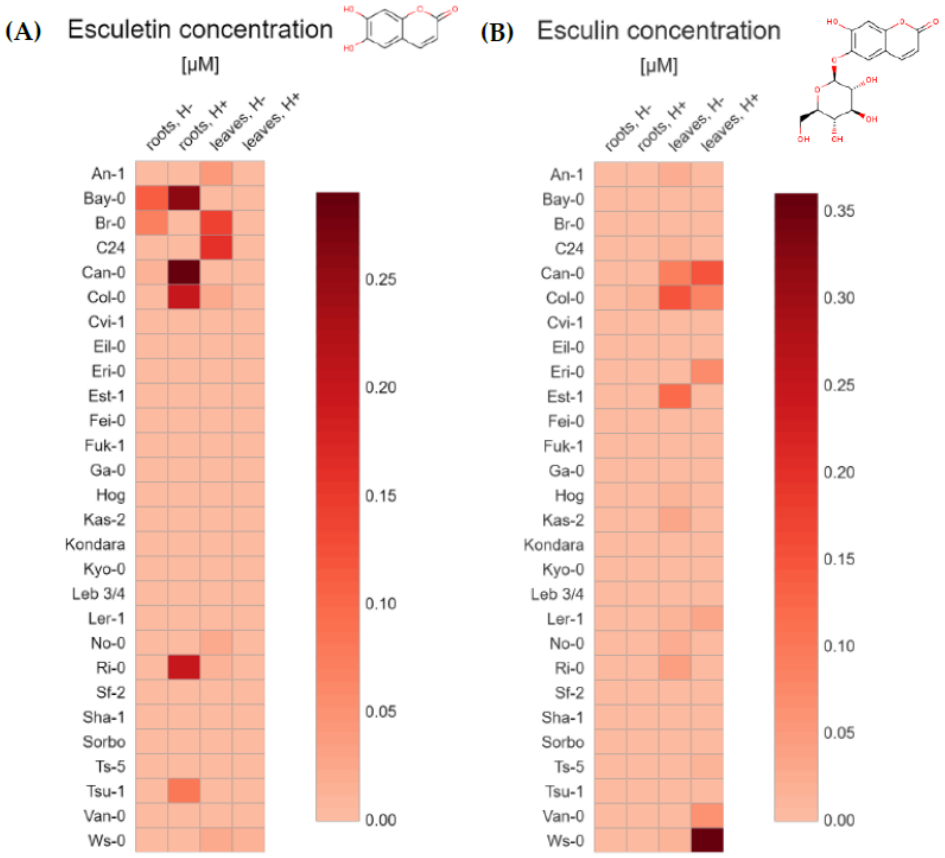


Figure 4. Heat maps based on the average (**A**) esculetin and (**B**) esculin concentration (μM/ FW) in Arabidopsis tissue extracts from roots and leaves, without and after hydrolysis. The values used in the plots (https://app.displayr.com (accessed on 20 January 2021)) are the mean of 3 biological replicates. The mean values and standard deviations (±SD) are gathered in the supplementary materials (Table S3).

2.2. Principal Component Analysis (PCA) for 28 Arabidopsis Accessions Using Coumarin Quantification by UHPLC-MS in Selected Geographic and under Diverse Climatic Factors

In order to compare and visualize the possible relationship between coumarin content variability present among 28 Arabidopsis accessions in selected geographic and in various climatic factors (maximal altitude [m], average winter minimal temperature [°C], average summer maximal temperature [°C] and average annual precipitation [mm]), we performed Principal Component Analysis (PCA). About half of the variance of used dataset was covered by the first two principal components, explaining 49% of the overall data variance (27.1% and 21.9% for PC1 and PC2, respectively) (Figure 5A).

According to the results presented on the Variables-PCA plot (Figure 5B), we assumed that there is a positive correlation between scopoletin, umbelliferone and scopolin concentration in root samples before hydrolysis. Despite the fact that scopolin content has relatively small contribution in explaining the variability between tested accessions, it can be also positively correlated with skimmin and umbelliferone concentration. Skimmin content is positively correlated with annual precipitation data.

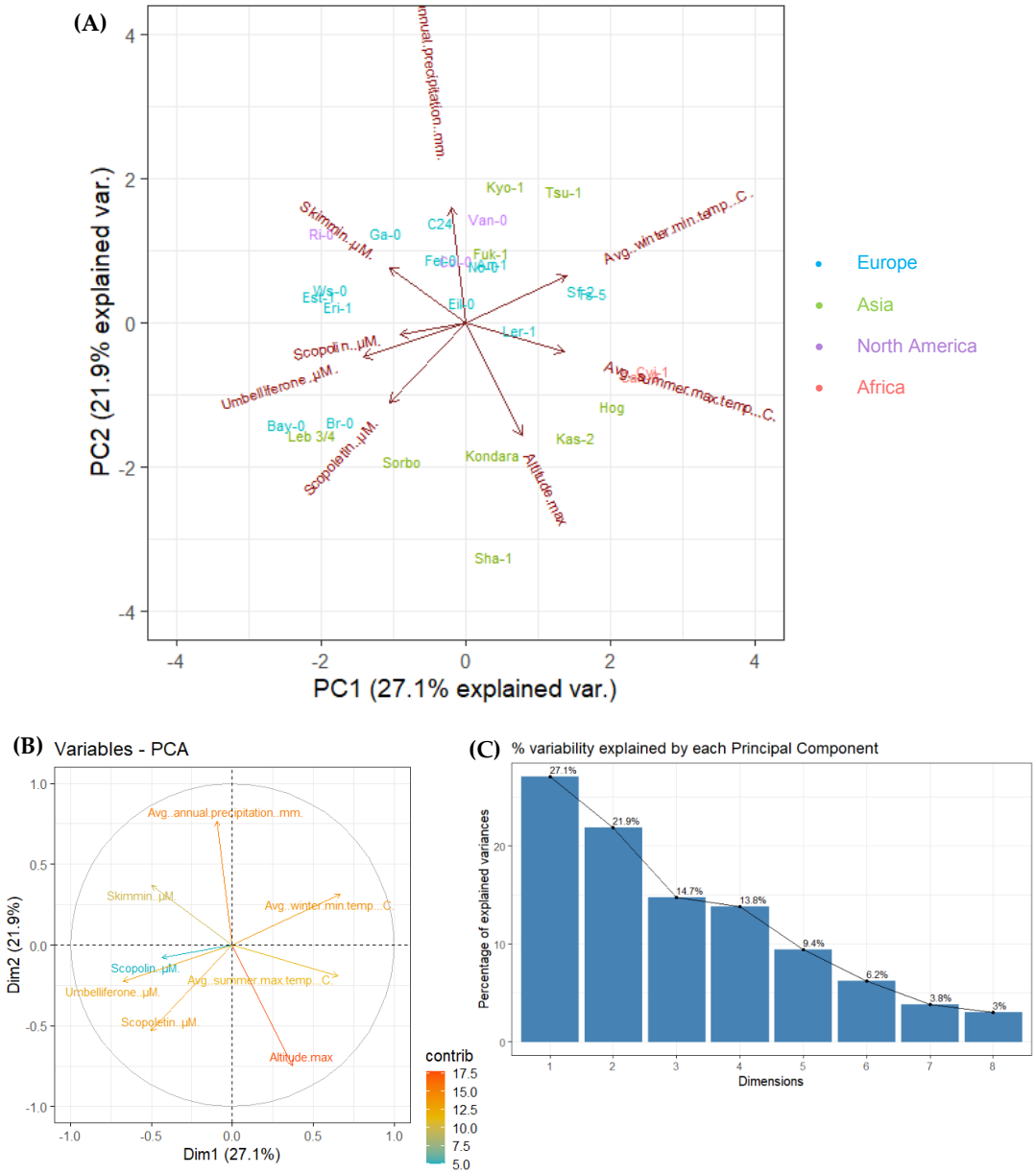


Figure 5. (**A**) Principal component analysis (PCA) for 28 Arabidopsis accessions using the concentration of umbelliferone, scopoletin and their corresponding glycosides (skimmin and scopolin, respectively) in root samples without hydrolysis, and four geographic and climatic factors (maximal altitude [m], average winter minimal temperature [°C], average summer maximal temperature [°C] and average annual precipitation [mm]; Table S4). Factor coordinates are marked with arrows. Observations indicated by blue accession names represent European locations (n = 14), green represent Asian locations (n = 9), violet represent North American locations (n = 3) and red represent African locations (n = 2). The abbreviations indicate the accessions according to Table 2. Component one and two explain 49% of the point variability. (**B**) The Variables-PCA contribution plot shows the correlation of the variables used in PCA with the respective contribution of each factor (contrib) indicated with a colour gradient. (**C**) The scree plot/graph of variables demonstrate the percentage of variability explained by each dimension (PC). Principal Component 1 and 2 explain 27.1% and 21.9% of the variance respectively.

A negative correlation is highlighted between the following variables: (1) umbelliferone concentration and temperatures (average winter minimal temperature and average summer maximal temperature); (2) scopolin concentration and temperatures (average

winter minimal temperature and average summer maximal temperature); (3) skimmin concentration and average summer maximal temperature, as well as skimmin concentration and maximal altitude; (4) scopoletin concentration and annual precipitation, as well as scopoletin concentration and average winter minimal temperature (Figure 5B).

On the Figure 5C we demonstrate the graph of variables (scree plot) which indicates the percentage of variability explained by each dimension (PC). Principal Component 1 and 2 explain 27.1% and 21.9% of the variance respectively, while the other 6 dimensions account for the total remaining variability between each accession (PC3 = 14.7%, PC4 = 13.8%, PC5 = 9.4%, PC6 = 6.2%, PC7 = 3.8% and PC8 = 3%).

3. Discussion

Our previous study strongly suggest that Arabidopsis is an excellent model for elucidating the basis of natural variation in coumarin accumulation [36]. Here, we identified and quantified a set of coumarin compounds in the root and leaf methanol extracts prepared from 28 Arabidopsis accessions. In the light of our best knowledge, it is the largest set of Arabidopsis natural populations used in the coumarin profiling analysis that should well represent a wide genetic variation existing in this model plant. It is assumed that these accessions reflect genetic adaptation to local environmental factors [40]. As a result of evolutionary pressure differentially acting on the studied accessions originating from various geographical locations, a large number of genetic polymorphisms is present that have led to different levels of expression of genes involved in the biosynthesis, transport and metabolism of coumarins, and ultimately to different levels of their accumulation. In the current work, we detected a significant natural variation in the content of simple coumarins present in the root and leaf extracts of 28 Arabidopsis accessions. Among tested compounds, scopoletin and its glycosylated form, scopolin, were the most abundant, which is in line with the current state of knowledge [6,14,15,36–38].

The previous study on differential accumulation of coumarins between 19 Arabidopsis accessions belonging to the MAGIC lines characterized by a high genetic variability [14], confirm the hypothesis that the composition of Arabidopsis accessions root exudates is genetically determined. They revealed the existence of distinct root metabolic phenotypes among tested natural populations, including variation in the accumulation of scopoletin and its glycoside. Another study focused on extensive profiling of specialized metabolites in root exudates of Arabidopsis reference accession, Col-0, by non-targeted metabolite profiling using reversed-phase UPLC/ESI-QTOFMS [16]. As many as 103 compounds were detected in exudates of hydroponically grown Col-0 plants. Among them, 42 were identified by authenticated standards, including the following coumarins: esculetin, scopoletin, and their glucosides esculin and scopolin. In addition to these coumarins, further esculetin and scopoletin conjugates were initially identified in the root exudates based on their mass spectral fragmentation pattern [16].

It has to be noted that among other coumarin compounds identified in our study, we detected umbelliferone for the first time in Arabidopsis model plant. Authors of the first publication on the accumulation of coumarins in Arabidopsis [37], in which various type of tissues (roots, shoots and callus) were tested, detected trace amounts of skimmin (umbelliferone glucoside) in the wild type plants and slightly increased skimmin level in mutants of CYP98A3. No umbelliferone was detected in that study, or in any other work with Arabidopsis to date, in the light of our best knowledge. It cannot be excluded that we were able to detect umbelliferone due to the sample types tested. We conducted coumarin profiling of extracts prepared from the plant tissues grown in in vitro liquid cultures. Moreover, umbelliferone was detected in root methanol extracts additionally subjected to enzymatic hydrolysis prior to quantification done by UHPLC-MS in order to hydrolyze the glycoside forms of coumarins, while its glycosylated form, skimmin, was detected in samples without enzymatic treatment. It should be highlighted that low amounts of umbelliferone were also detected in non-hydrolysed extracts.

Coumarins have become important players both in optimizing Fe uptake and shaping the root microbiome, thus affecting plant health [5,41]. The link between plant specialized metabolites, in particular coumarins, nutrient deficiencies and microbiome composition that was discovered in recent studies [7,8,42,43], could provide a new set of tools for rationally manipulating the plant microbiome [44]. The selection of underground tissue was an obvious choice in such analyses, considering that coumarins are essentially synthesized in roots where optimization of Fe uptake is coordinated with plant requirements and interaction with soil microorganisms. Therefore, most of the previous coumarin metabolic profiling analysis, including functional characterization of Arabidopsis mutants defective in genes encoding enzymes involved in coumarin biosynthesis or transport, were performed using the root exudates and root tissue [3,6,12,17,38,44,45]. It was also the case in research conducted on the effects of Fe, phosphorus (Pi) or both deficiencies on coumarin profiles in the root tissue of several T-DNA insertional mutants defective in genes involved in Pi or Fe homeostasis [11]. Importantly, in the current study we detected variation in accumulation of esculetin and esculin. The latter one was identified in Arabidopsis leaf extracts, both with and without enzymatic hydrolysis. This requires further research and is of particular interest in the light of recent research findings on coumarin cellular localization, trafficking and signalling [5,7]. Coumarins were found to be involved in the plant response to pathogens in aerial tissues [41,46] and proposed to play an important signalling role in bidirectional chemical communication along the microbiome-root-shoot axis [7].

The study of natural variation in coumarin content present among Arabidopsis accessions is a starting point in elucidating direct links between metabolic phenotypes and genotypes. In the presented research, we also checked whether the climatic and geographic data on the regions from which particular accessions originate, are correlated with the concentrations of tested coumarins. The conducted PCA showed a number of positive and negative correlations between climatic factors and coumarin content. Further investigation is needed to draw a more precise conclusion about possible relationship between the accumulation of coumarins and habitat data. Taking into account, the recent studies showing an important role of coumarins in plant interactions with soil microorganisms and nutrient acquisition, a more in-depth analysis, including data on soil parameters at the origin sites of a given accession, would explain the greater variance and give us more information on the potential correlations. It will be beneficial for the future discovery of physiological mechanisms of action of various alleles involved in the coumarin biosynthesis and can help to select biosynthetic enzymes for further metabolic engineering research.

4. Materials and Methods

4.1. Chemicals and Reagents

The coumarins standards umbelliferone (purity $\geq$ 99%), coumarin (>99% purity), esculin ($\geq$98% purity) were purchased from Sigma-Aldrich (St. Louis, MO, USA), scopoletin (>95% purity) and esculetin (>98% purity) from Extrasynthese (Genay, France), skimmin (98% purity) from Biopurify Phytochemicals (Chengdu, China), scopolin (>98% purity) from Chemicals Aktin Inc. (Chengdu, China). Stock solutions of each standard at a concentration of 10 mmol/L were prepared by diluting the powder in dimethyl sulfoxide (Fisher scientific, Illkirch, France) and kept at -18 °C until use. HPLC-grade methanol was purchased from CarloErba Reagents (Val de Reuil, France), formic acid was purchased from Fisher Scientific (Illkirch, France). Water was purified by a PURELAB Ultra system (Veolia Water S.T.I., Antony, France).

4.2. Plant Material

All seeds of the 28 Arabidopsis accessions (Table 2) from various habitats which were used in this study were obtained courtesy of prof. Maarten Koornneef.

4.3. In Vitro Plant Culture

All the Arabidopsis accessions seeds were surface sterilized with 70% ethanol for 2 min, 5% calcium hypochlorite solution for 8 min and then washed 3 times with sterile ultrapure water. The seeds were placed in Petri dishes with $\frac{1}{2}$ Murashige-Skoog (MS) medium solidified with agar (Sigma-Aldrich) for in vitro plant culture and incubated for 72 h in the dark at 4 °C. Then the plates were transferred to a growth chamber (daily cycle: 16 h light 35 µmol m^{-2} s^{-1} temperature 20 °C and 8 h dark temperature 18 °C) for 10 days. After that time, seedlings were transferred from agar plates into 200 mL flasks (three individuals per flask) containing 5 mL of $\frac{1}{2}$ MS liquid medium containing 1% sucrose, MS salts, 100 mg/L myo-inositol, 1 mg/L thiamine hydrochloride, 0.5 mg/L pyridoxine hydrochloride and 0.5 mg/L nicotinic acid (Sigma-Aldrich/Merck KGaA, Darmstadt, Germany). Plants were grown in the growth chamber on a rotary platform with shaking 120 rpm. After one week, 3 mL of fresh $\frac{1}{2}$ MS medium was added. Plants were grown for 17 supplementary days and after that time were rinsed with demineralized water, dried on paper towels. Roots and leaves samples were weighted (50 ± 2 mg fresh weight (FW)) and frozen in liquid nitrogen. The plant material was stored in a freezer at −80 °C until extraction process. All accessions were grown in three biological replicates (in three independent flasks, three seedlings per flask).

4.4. Metabolites Extraction

For the metabolites extraction, plant tissue frozen in liquid nitrogen was grinded by the usage of 5 mm diameter stainless steel beads (Qiagen, Hilden, Germany). To the 2 mL microtubes, 2 clean beads were added and samples were frozen in liquid nitrogen. Then, using vortex, samples were mixed. For the better performance, the freezing and vortexing procedure was repeated several times until all tissue was powdered. To the powdered tissues 0.5 mL of 80% methanol containing 5 µM 4-methylumbelliferone as an internal standard was added. After that, samples were sonicated for 30 min with ultrasonic cleaner (Proclean 3.0DSP, Ulsonix, Expondo, Berlin, Germany) (70% frequency, sweep function) and incubated in 4 °C in darkness for 24 h. Next day, all samples were vortexed, centrifuged at 13,000× *g* for 10 min and the supernatant was transferred into new microtubes. Centrifugation was repeated in order to get rid of any sediment. The extracts were firstly dried for 2 h in an incubator at 45 °C and then, for the next 2 h in a vacuum centrifuge (Savant SpeedVac vacuum concentrator, Thermo Fisher Scientific, Waltham, MA, USA). To the dried extracts 100 µL of 80% methanol was added to dissolve samples during the night at 4 °C. Then the extracts were vortexed for 10 min and separated by 50 µL. One of the replicates was subjected to enzymatic hydrolysis, and the second one was stored at −20 °C until UHPLC-MS analysis (Shimadzu Corp., Kyoto, Japan).

4.5. Enzymatic Hydrolysis

The enzymatic hydrolysis was performed according to Nguyen et al. [47]. Methanolic extracts were subjected to enzymatic hydrolysis with a β-glucosidase (Fluka Chemie GmbH, Buchs, Switzerland) in 0.1 M acetate buffer at a concentration of 0.5 mg/mL in order to determine the amounts of glycosylated compounds (o-glycosides). 50 µL of acetate buffer with β-glucosidase at pH 5.0 (0.1 M sodium acetate, 0.1 M acetic acid and 0.5 mg/mL β-glucosidase buffer) was added to 50 µL of the prepared extract and incubated for 22 h at 37 °C. The reaction was stopped by adding 100 µL of 96% ethanol to the reaction mixture. The extracts were dried in an incubator at 45 °C for 2 h and then for about 1 h in a vacuum centrifuge (Savant SpeedVac vacuum concentrator). The obtained extract was dissolved in 50 µL of 80% methanol overnight and stored at −20 °C until UHPLC-MS analysis.

4.6. UHPLC Separation

The coumarins analyses were performed using a NEXERA UPLC-MS system (Shimadzu Corp., Kyoto, Japan) equipped with two UHPLC pumps (LC-30AD), an automatic sampler (SIL-30AC), a photodiode array detector (PDA, SPDM-20A) and combined with

a mass spectrometer (single quadrupole, LCMS-2020). Coumarins separation was done on a C18 reversed phase column (ZORBAX Eclipse Plus), 150 × 2.1 mm, 1.8 µm (Agilent Technologies, Santa Clara, CA, USA) protected with an Agilent Technologies 1290 Infinity filter. The column was kept at 40 °C in a column oven (Shimadzu CTO-20AC). Mobile phase consisted of 0.1% formic acid in ultrapure water (buffer A) and 0.1% formic acid in methanol (buffer B) at a constant flow rate of 200 µL/min. The linear gradient solvent system was set as follows: 0 min, 10% B; 16 min, 70% B; 18 min, 99% B; 18.01 min, 10% B; 20 min, 10% B. The total analysis duration was 20 min. The injection volume was 5 µL.

4.7. MS Detection

The UHPLC system was connected to the MS by an electrospray ionization source (ESI), operating in positive mode (ESI+) and scanning in single ion monitoring mode (SIM). The inlet, desolvation line and heating block temperatures were set at 350 °C, 250 °C, and 400 °C, respectively. The capillary voltage was set at 4.5 kV. Dry gas flow was set at 15 L/min and nebulizing gas at 1.5 L/min. The instrument was operated and data were processed using LabSolution software version 5.52 sp2 (Shimadzu Corp., Kyoto, Japan).

4.8. Peak Identification and Quantitation

Each standard molecule was individually injected in the UHPLC-MS in full scan mode to determine retention time and *m/z* ratio for the analysis. The quantitation of each molecule (Table 2) was based on the signal obtained from the MS detection, using angelicin, as an analytical internal standard. Angelicin was added at the same concentration (5 µM) in all the samples before injection as well as in 7 calibration solutions. The calibration solutions contained all of the standard molecules at the same concentrations ranging from 0.1 to 10 µM (0.1, 0.2, 0.5, 1, 2, 5 and 10 µM). Calibration curves were drawn for each compound by linking its relative peak area (compound area divided by the angelicin area) and its concentration. Each curve fit type was linear. The limit of quantitation (LOQ) was calculated as the analyte concentration giving signal to signal to noise ratios (S/N) of 10. Three measurements were assessed per accession.

4.9. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) were performed using *prcomp()* package and visualize with the *factoextra* 1.0.7 version package in the R 4.0.4 environment developed by the R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (www.R-project.org (accessed on 20 January 2021)) and the RStudio Team (2019). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, USA (www.rstudio.com (accessed on 20 January 2021)). All the variables were standardized before analysis. Data used for the analysis are presented in Table S4.

5. Conclusions

Multi-pharmacological properties of coumarins that are widely used in medical applications, make the study of coumarin biosynthesis attractive from the commercial point of view. Considering that all medicinal plants currently used in studying the biosynthesis of coumarins are non-model organisms and many approaches are not available in those species, makes a model plant Arabidopsis, with its extensive genetic variation and numerous publicly accessible web-based databases, an excellent model to study accumulation of coumarins in natural populations. The presented results focusing on qualitative and quantitative characterization of natural resources provide a basis for further research on identification of genetic variants involved in coumarin biosynthesis in plants, which is the first step in metabolic engineering for the production of natural compounds. We identified scopoletin, and its glycosylated form, scopolin, to be the most abundant coumarins in Arabidopsis tissues. It should be emphasized that among other coumarin compounds identified in this study, we detected umbelliferone for the first time in Arabidopsis. In

view of the considerable importance of umbelliferone in synthesis and its pharmacological properties, this is a significant step in the study of biosynthesis of coumarins using this model plant.

Supplementary Materials: The following are available online, Table S1: The mean values (scopoletin and scopolin concentrations) and standard deviations (±SD) that were used in making heat maps shown in Figure 2, Table S2: The mean values (scopoletin and scopolin concentrations) and standard deviations (±SD) that were used in making heat maps shown in Figure 3, Table S3: The mean values (scopoletin and scopolin concentrations) and standard deviations (±SD) that were used in making heat maps shown in Figure 4, Table S4. Coumarin concentrations in root samples before hydrolysis and four geographic and climatic factors used in principal component analysis (PCA).

Author Contributions: Conceptualization, A.O., A.H. and A.I.; methodology, A.O., A.H. and A.I.; validation, I.P., J.S. and A.O.; investigation, I.P., J.S., A.O. and J.G.; writing—original draft preparation, I.P., E.L. and A.I.; writing—review and editing, E.L., A.O., A.H. and A.I.; supervision, E.L., F.B. and A.I.; funding acquisition, A.I. and A.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Narodowe Centrum Nauki (Polish National Science Centre) grant number 2014/15/B/NZ2/01073 and Narodowa Agencja Wymiany Akademickiej (Polish National Agency for Academic Exchange) grant number PPN/BFR/2019/1/00050 to A.I.; and by Campus France (PHC POLONIUM) grant number 45026ZE to A.O.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available within the article and supplementary materials [Tables S1–S4].

Acknowledgments: We thank Maarten Koornneef for providing all Arabidopsis seeds used in this study and Thibaut Duval for technical support.

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

Sample Availability: Coumarin standards used in the study and seeds corresponding to each accessions are available from the authors.

References

1. Matos, M.J.; Santana, L.; Uriarte, E.; Abreu, O.A.; Molina, E.; Yordi, E.G. Coumarins—An Important Class of Phytochemicals. *Phytochem. Isol. Charact. Role Hum. Health* **2015**, 113–140. [CrossRef]
2. Riveiro, M.E.; De Kimpe, N.; Moglioni, A.; Vazquez, R.; Monczor, F.; Shayo, C.; Davio, C. Coumarins: Old Compounds with Novel Promising Therapeutic Perspectives. *Curr. Med. Chem.* **2010**, *17*, 1325–1338. [CrossRef]
3. Rajniak, J.; Giehl, R.F.H.; Chang, E.; Murgia, I.; Von Wirén, N.; Sattely, E.S. Biosynthesis of redox-active metabolites in response to iron deficiency in plants. *Nat. Chem. Biol.* **2018**, *14*, 442–450. [CrossRef]
4. Reen, F.J.; Gutiérrez-Barranquero, J.A.; Parages, M.L.; O´gara, F. Coumarin: A novel player in microbial quorum sensing and biofilm formation inhibition. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 2063–2073. [CrossRef] [PubMed]
5. Robe, K.; Conejero, G.; Gao, F.; Lefebvre-Legendre, L.; Sylvestre-Gonon, E.; Rofidal, V.; Hem, S.; Rouhier, N.; Barberon, M.; Hecker, A.; et al. Coumarin accumulation and trafficking in *Arabidopsis thaliana*: A complex and dynamic process. *New Phytol.* **2021**, *229*, 2062–2079. [CrossRef]
6. Siwinska, J.; Siatkowska, K.; Olry, A.; Grosjean, J.; Hehn, A.; Bourgaud, F.; A Meharg, A.; Carey, M.; Lojkowska, E.; Ihnatowicz, A. Scopoletin 8-hydroxylase: A novel enzyme involved in coumarin biosynthesis and iron-deficiency responses in Arabidopsis. *J. Exp. Bot.* **2018**, *69*, 1735–1748. [CrossRef]
7. Stassen, M.J.; Hsu, S.-H.; Pieterse, C.M.; Stringlis, I.A. Coumarin Communication Along the Microbiome–Root–Shoot Axis. *Trends Plant Sci.* **2021**, *26*, 169–183. [CrossRef] [PubMed]
8. A Stringlis, I.; De Jonge, R. The Age of Coumarins in Plant–Microbe Interactions. *Plant Cell Physiol.* **2019**, *60*, 1405–1419. [CrossRef] [PubMed]
9. Ziegler, J.; Schmidt, S.; Chutia, R.; Müller, J.; Böttcher, C.; Strehmel, N.; Scheel, D.; Abel, S. Non-targeted profiling of semi-polar metabolites in Arabidopsis root exudates uncovers a role for coumarin secretion and lignification during the local response to phosphate limitation. *J. Exp. Bot.* **2016**, *67*, 1421–1432. [CrossRef] [PubMed]
10. Badri, D.V.; Loyola-Vargas, V.M.; Broeckling, C.D.; Vivanco, J.M. Root Secretion of Phytochemicals in Arabidopsis Is Predominantly Not Influenced by Diurnal Rhythms. *Mol. Plant* **2010**, *3*, 491–498. [CrossRef] [PubMed]

11. Chutia, R.; Abel, S.; Ziegler, J. Iron and Phosphate Deficiency Regulators Concertedly Control Coumarin Profiles in *Arabidopsis thaliana* Roots During Iron, Phosphate, and Combined Deficiencies. *Front. Plant Sci.* **2019**, *10*, 113. [CrossRef] [PubMed]
12. Fourcroy, P.; Sisó-Terraza, P.; Sudre, D.; Savirón, M.; Reyt, G.; Gaymard, F.; Abadía, A.; Abadia, J.; Álvarez-Fernández, A.; Briat, J.-F. Involvement of the ABCG37 transporter in secretion of scopoletin and derivatives by Arabidopsis roots in response to iron deficiency. *New Phytol.* **2014**, *201*, 155–167. [CrossRef] [PubMed]
13. Hildreth, S.B.; Foley, E.E.; Muday, G.K.; Helm, R.F.; Winkel, B.S.J. The dynamic response of the Arabidopsis root metabolome to auxin and ethylene is not predicted by changes in the transcriptome. *Sci. Rep.* **2020**, *10*, 1–15. [CrossRef]
14. Mönchgesang, S.; Strehmel, N.; Schmidt, S.; Westphal, L.; Taruttis, F.; Müller, E.; Herklotz, S.; Neumann, S.; Scheel, D. Natural variation of root exudates in *Arabidopsis thaliana*-linking metabolomic and genomic data. *Sci. Rep.* **2016**, *6*, 29033. [CrossRef]
15. Mönchgesang, S.; Strehmel, N.; Trutschel, D.; Westphal, L.; Neumann, S.; Scheel, D. Plant-to-Plant Variability in Root Metabolite Profiles of 19 *Arabidopsis thaliana* Accessions Is Substance-Class-Dependent. *Int. J. Mol. Sci.* **2016**, *17*, 1565. [CrossRef]
16. Strehmel, N.; Böttcher, C.; Schmidt, S.; Scheel, D. Profiling of secondary metabolites in root exudates of *Arabidopsis thaliana*. *Phytochemistry* **2014**, *108*, 35–46. [CrossRef] [PubMed]
17. Ziegler, J.; Schmidt, S.; Strehmel, N.; Scheel, D.; Abel, S. Arabidopsis Transporter ABCG37/PDR9 contributes primarily highly oxygenated Coumarins to Root Exudation. *Sci. Rep.* **2017**, *7*, 1–11. [CrossRef] [PubMed]
18. Clemens, S.; Weber, M. The essential role of coumarin secretion for Fe acquisition from alkaline soil. *Plant Signal. Behav.* **2016**, *11*, e1114197. [CrossRef]
19. Mitchell-Olds, T.; Pedersen, D. The molecular basis of quantitative genetic variation in central and secondary metabolism in Arabidopsis. *Genetics* **1998**, *149*, 739–747. [PubMed]
20. Hirai, M.Y.; Yano, M.; Goodenowe, D.B.; Kanaya, S.; Kimura, T.; Awazuhara, M.; Arita, M.; Fujiwara, T.; Saito, K. From the Cover: Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10205–10210. [CrossRef] [PubMed]
21. Von Roepenack-Lahaye, E.; Degenkolb, T.; Zerjeski, M.; Franz, M.; Roth, U.; Wessjohann, L.; Schmidt, J.; Scheel, D.; Clemens, S. Profiling of Arabidopsis Secondary Metabolites by Capillary Liquid Chromatography Coupled to Electrospray Ionization Quadrupole Time-of-Flight Mass Spectrometry. *Plant Physiol.* **2004**, *134*, 548–559. [CrossRef]
22. Keurentjes, J.J.; Fu, J.; De Vos, C.R.; Lommen, A.; Hall, R.D.; Bino, R.J.; Van Der Plas, L.H.; Jansen, R.C.; Vreugdenhil, D.; Koornneef, M. The genetics of plant metabolism. *Nat. Genet.* **2006**, *38*, 842–849. [CrossRef]
23. Fiehn, O. Metabolite Profiling in Arabidopsis. In *Arabidopsis Protocols*; Springer Science and Business Media LLC: Berlin, Germany, 2006; Volume 323, pp. 439–448.
24. Ebert, B.; Zöller, D.; Erban, A.; Fehrle, I.; Hartmann, J.; Niehl, A.; Kopka, J.; Fisahn, J. Metabolic profiling of *Arabidopsis thaliana* epidermal cells. *J. Exp. Bot.* **2010**, *61*, 1321–1335. [CrossRef] [PubMed]
25. Hannah, M.A.; Caldana, C.; Steinhauser, D.; Balbo, I.; Fernie, A.R.; Willmitzer, L. Combined Transcript and Metabolite Profiling of Arabidopsis Grown under Widely Variant Growth Conditions Facilitates the Identification of Novel Metabolite-Mediated Regulation of Gene Expression. *Plant Physiol.* **2010**, *152*, 2120–2129. [CrossRef] [PubMed]
26. Routaboul, J.-M.; Dubos, C.; Beck, G.; Marquis, C.; Bidzinski, P.; Loudet, O.; Lepiniec, L. Metabolite profiling and quantitative genetics of natural variation for flavonoids in Arabidopsis. *J. Exp. Bot.* **2012**, *63*, 3749–3764. [CrossRef] [PubMed]
27. Watanabe, M.; Balazadeh, S.; Tohge, T.; Erban, A.; Giavalisco, P.; Kopka, J.; Mueller-Roeber, B.; Fernie, A.R.; Hoefgen, R. Comprehensive Dissection of Spatiotemporal Metabolic Shifts in Primary, Secondary, and Lipid Metabolism during Developmental Senescence in Arabidopsis. *Plant Physiol.* **2013**, *162*, 1290–1310. [CrossRef]
28. Wu, S.; Alseekh, S.; Cuadros-Inostroza, Á.; Fusari, C.M.; Mutwil, M.; Kooke, R.; Keurentjes, J.B.; Fernie, A.R.; Willmitzer, L.; Brotman, Y. Combined Use of Genome-Wide Association Data and Correlation Networks Unravels Key Regulators of Primary Metabolism in *Arabidopsis thaliana*. *PLoS Genet.* **2016**, *12*, e1006363. [CrossRef]
29. Orlita, A.; Sidwa-Gorycka, M.; Paszkiewicz, M.; Malinski, E.; Kumirska, J.; Siedlecka, E.M.; Łojkowska, E.; Stepnowski, P. Application of chitin and chitosan as elicitors of coumarins and furoquinolone alkaloids in *Ruta graveolens* L. (common rue). *Biotechnol. Appl. Biochem.* **2008**, *51*, 91–96. [CrossRef]
30. Sidwa-Gorycka, M.; Krolicka, A.; Orlita, A.; Maliński, E.; Golebiowski, M.; Kumirska, J.; Chromik, A.; Biskup, E.; Stepnowski, P.; Lojkowska, E. Genetic transformation of *Ruta graveolens* L. by *Agrobacterium rhizogenes*: Hairy root cultures a promising approach for production of coumarins and furanocoumarins. *Plant Cell Tissue Organ Cult. (PCTOC)* **2009**, *97*, 59–69. [CrossRef]
31. Sidwa-Gorycka, M.; Królicka, A.; Kozyra, M.; Głowniak, K.; Bourgaud, F.; Łojkowska, E. Establishment of a co-culture of *Ammi majus* L. and *Ruta graveolens* L. for the synthesis of furanocoumarins. *Plant Sci.* **2003**, *165*, 1315–1319. [CrossRef]
32. Matsumoto, S.; Mizutani, M.; Sakata, K.; Shimizu, B.-I. Molecular cloning and functional analysis of the ortho-hydroxylases of p-coumaroyl coenzyme A/feruloyl coenzyme A involved in formation of umbelliferone and scopoletin in sweet potato, *Ipomoea batatas* (L.) Lam. *Phytochemistry* **2012**, *74*, 49–57. [CrossRef]
33. Vialart, G.; Hehn, A.; Olry, A.; Ito, K.; Krieger, C.; Larbat, R.; Paris, C.; Shimizu, B.-I.; Sugimoto, Y.; Mizutani, M.; et al. A 2-oxoglutarate-dependent dioxygenase from *Ruta graveolens* L. exhibits p-coumaroyl CoA 2′-hydroxylase activity (C2′H): A missing step in the synthesis of umbelliferone in plants. *Plant J.* **2012**, *70*, 460–470. [CrossRef]
34. Yang, X.; Wei, S.; Liu, B.; Guo, D.; Zheng, B.; Feng, L.; Liu, Y.; Tomás-Barberán, F.A.; Luo, L.; Huang, D. A novel integrated non-targeted metabolomic analysis reveals significant metabolite variations between different lettuce (*Lactuca sativa.* L.) varieties. *Hortic. Res.* **2018**, *5*, 1–14. [CrossRef] [PubMed]

35. Micallef, S.A.; Shiaris, M.P.; Colón-Carmona, A. Influence of *Arabidopsis thaliana* accessions on rhizobacterial communities and natural variation in root exudates. *J. Exp. Bot.* **2009**, *60*, 1729–1742. [CrossRef] [PubMed]
36. Siwinska, J.; Kadzinski, L.; Banasiuk, R.; Gwizdek-Wisniewska, A.; Olry, A.; Banecki, B.; Lojkowska, E.; Ihnatowicz, A. Identification of QTLs affecting scopolin and scopoletin biosynthesis in *Arabidopsis thaliana*. *BMC Plant Biol.* **2014**, *14*, 280. [CrossRef]
37. Kai, K.; Shimizu, B.-I.; Mizutani, M.; Watanabe, K.; Sakata, K. Accumulation of coumarins in *Arabidopsis thaliana*. *Phytochemistry* **2006**, *67*, 379–386. [CrossRef]
38. Kai, K.; Mizutani, M.; Kawamura, N.; Yamamoto, R.; Tamai, M.; Yamaguchi, H.; Sakata, K.; Shimizu, B.-I. Scopoletin is biosynthesized via ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in *Arabidopsis thaliana*. *Plant J.* **2008**, *55*, 989–999. [CrossRef] [PubMed]
39. Schmid, N.B.; Giehl, R.F.; Döll, S.; Mock, H.-P.; Strehmel, N.; Scheel, D.; Kong, X.; Hider, R.C.; Von Wirén, N. Feruloyl-CoA 6′-Hydroxylase1-Dependent Coumarins Mediate Iron Acquisition from Alkaline Substrates in Arabidopsis. *Plant Physiol.* **2014**, *164*, 160–172. [CrossRef] [PubMed]
40. Koornneef, M.; Alonso-Blanco, C.; Vreugdenhil, D. Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annu. Rev. Plant Biol.* **2004**, *55*, 141–172. [CrossRef]
41. Stringlis, I.A.; Yu, K.; Feussner, K.; De Jonge, R.; Van Bentum, S.; Van Verk, M.C.; Berendsen, R.L.; Bakker, P.A.H.M.; Feussner, I.; Pieterse, C.M.J. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E5213–E5222. [CrossRef]
42. Harbort, C.J.; Hashimoto, M.; Inoue, H.; Niu, Y.; Guan, R.; Rombolà, A.D.; Kopriva, S.; Voges, M.J.; Sattely, E.S.; Garrido-Oter, R.; et al. Root-Secreted Coumarins and the Microbiota Interact to Improve Iron Nutrition in Arabidopsis. *Cell Host Microbe* **2020**, *28*, 825–837.e6. [CrossRef] [PubMed]
43. Voges, M.J.E.E.E.; Bai, Y.; Schulze-Lefert, P.; Sattely, E.S. Plant-derived coumarins shape the composition of an Arabidopsis synthetic root microbiome. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 12558–12565. [CrossRef] [PubMed]
44. Jacoby, R.P.; Koprivova, A.; Kopriva, S. Pinpointing secondary metabolites that shape the composition and function of the plant microbiome. *J. Exp. Bot.* **2021**, *72*, 57–69. [CrossRef]
45. Schmidt, H.; Günther, C.; Weber, M.; Spörlein, C.; Loscher, S.; Böttcher, C.; Schobert, R.; Clemens, S. Metabolome Analysis of *Arabidopsis thaliana* Roots Identifies a Key Metabolic Pathway for Iron Acquisition. *PLoS ONE* **2014**, *9*, e102444. [CrossRef]
46. Döll, S.; Kuhlmann, M.; Rutten, T.; Mette, M.F.; Scharfenberg, S.; Petridis, A.; Berreth, D.-C.; Mock, H.-P. Accumulation of the coumarin scopolin under abiotic stress conditions is mediated by the *Arabidopsis thaliana* THO/TREX complex. *Plant J.* **2018**, *93*, 431–444. [CrossRef]
47. Nguyen, C.; Bouque, V.; Bourgaud, F.; Guckert, A. Quantification of Daidzein and Furanocoumarin Conjugates of *Psoralea cinerea* L. (Leguminosae). *Phytochem. Anal.* **1997**, *8*, 27–31. [CrossRef]

Publication No. 2

AIM

1) **Comprehensive analysis of the correlation between accumulation of coumarins, plant genotypes (natural accessions and mutants), and environmental stress factors (including Fe availability and pathogen defense) to gain deeper insights into the metabolic adaptations and stress response in *Arabidopsis thaliana.***

Perkowska Izabela, Potrykus Marta, Siwinska Joanna, Siudem Dominika, Lojkowska Ewa, Ihnatowicz Anna

Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *International Journal of Molecular Sciences*, 22(12), 6449 (2021)

Online access:
https://doi.org/10.3390/ijms22126449

International Journal of
Molecular Sciences

MDPI

Article

Interplay between Coumarin Accumulation, Iron Deficiency and Plant Resistance to *Dickeya* spp.

Izabela Perkowska [1], Marta Potrykus [1,2], Joanna Siwinska [1], Dominika Siudem [1], Ewa Lojkowska [1] and Anna Ihnatowicz [1,*]

[1] Laboratory of Plant Protection and Biotechnology, Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk, University of Gdansk, Abrahama 58, 80-307 Gdansk, Poland; izabela.perkowska@phdstud.ug.edu.pl (I.P.); marta.potrykus@gumed.edu.pl (M.P.); siwinskajoanna@gmail.com (J.S.); d.siudem@gmail.com (D.S.); ewa.lojkowska@biotech.ug.edu.pl (E.L.)
[2] Department of Environmental Toxicology, Faculty of Health Sciences with Institute of Maritime and Tropical Medicine, Medical University of Gdansk, Debowa 23 A, 80-204 Gdansk, Poland
* Correspondence: anna.ihnatowicz@biotech.ug.edu.pl; Tel.: +48-58-5236330

Abstract: Coumarins belong to a group of secondary metabolites well known for their high biological activities including antibacterial and antifungal properties. Recently, an important role of coumarins in plant resistance to pathogens and their release into the rhizosphere upon pathogen infection was discovered. It is also well documented that coumarins play a crucial role in the *Arabidopsis thaliana* growth under Fe-limited conditions. However, the mechanisms underlying interplay between plant resistance, accumulation of coumarins and Fe status, remain largely unknown. In this work, we investigated the effect of both mentioned factors on the disease severity using the model system of Arabidopsis/*Dickeya* spp. molecular interactions. We evaluated the disease symptoms in Arabidopsis plants, wild-type Col-0 and its mutants defective in coumarin accumulation, grown in hydroponic cultures with contrasting Fe regimes and in soil mixes. Under all tested conditions, Arabidopsis plants inoculated with *Dickeya solani* IFB0099 strain developed more severe disease symptoms compared to lines inoculated with *Dickeya dadantii* 3937. We also showed that the expression of genes encoding plant stress markers were strongly affected by *D. solani* IFB0099 infection. Interestingly, the response of plants to *D. dadantii* 3937 infection was genotype-dependent in Fe-deficient hydroponic solution.

Keywords: abiotic stress; biotic stress; fraxetin; iron deficiency; scopoletin; pathogen; plant–environment interactions; mineral nutrition

Citation: Perkowska, I.; Potrykus, M.; Siwinska, J.; Siudem, D.; Lojkowska, E.; Ihnatowicz, A. Interplay between Coumarin Accumulation, Iron Deficiency and Plant Resistance to *Dickeya* spp.. *Int. J. Mol. Sci.* **2021**, *22*, 6449. https://doi.org/10.3390/ijms22126449

Academic Editor: Marcello Iriti

Received: 15 April 2021
Accepted: 9 June 2021
Published: 16 June 2021

1. Introduction

The secretion of phenolic compounds from roots into the rhizosphere has long been recognised as a component of the acidification-reduction strategy to acquire iron (Fe), occurring in all plant species except grasses [1]. However, the molecular mechanisms underlying these processes remained elusive until recently, when several research groups including our team, independently demonstrated the important role of plant secondary metabolites called coumarins for the growth of a model plant *Arabidopsis thaliana* (hereafter Arabidopsis) under Fe-limited conditions [2–9]. It was proven that coumarins are involved in Fe chelation and that secretion of coumarins by Arabidopsis roots is induced under Fe-deficiency. The biological roles of novel enzymes involved in coumarin biosynthesis, which in parallel maintain Fe homeostasis in plants, were elucidated. A key enzyme for the biosynthesis of Arabidopsis major coumarin called scopoletin and its derivatives is Feruloyl-CoA 6′-Hydroxylase1 (F6′H1) that belongs to a large enzyme family of the 2-oxoglutarate and Fe(II)-dependent dioxygenases [3,5,10]. Our group elucidated the biological role of another member of this family, encoded by a strongly Fe-responsive gene (At3g12900), as a scopoletin 8-hydroxylase (S8H) involved in the last step of fraxetin biosynthesis [7]. Fraxetin is a coumarin derived from the scopoletin pathway, containing

two adjacent hydroxyl groups in the ortho-position that can efficiently solubilise Fe from the hydroxide precipitates [3,8]. We proved S8H to be involved in coumarin biosynthesis as part of the Fe acquisition machinery [7].

Fe is a crucial micronutrient for every kind of living organism. It plays an essential role in metabolic processes such as DNA synthesis, respiration, photosynthesis, and it is a cofactor of many enzymes. The role of Fe homeostasis in resistance to infections was also shown across all kingdoms of life—different types of pathogens are likely to compete with their hosts for the acquisition of Fe [11]. Mechanisms of Fe homeostasis in plants, pathogens, and beneficial microorganisms play key roles in plant-microbe interactions [12]. Moreover, one-third of the world's agricultural area is composed of calcareous soils, in which high pH leads to the precipitation of Fe that is finally not available and generate severe plant growth perturbation. Therefore, Fe deficiency is a widespread agricultural problem that reduces plant growth and crop yields, particularly in alkaline soils [13].

In addition to the important role in maintaining Fe homeostasis, coumarins can affect plant growth and fitness directly through their high biological activities including antibacterial and antifungal properties. Scopoletin that is one of the major Arabidopsis coumarins accumulating in roots [5,7,14], was shown to possess antimicrobial activity against various phytopathogens like *Ralstonia solanacearum* [15], *Alternaria alternata* [16], *Botrytis cinerea* [17], *Fusarium oxysporum* f. sp. *batatas* [18], *Sclerotinia sclerotiorum* [19], *Aspergillus flavus*, and *Aspergillus niger* [20], *Ceratocystis fimbriata* f. sp. *platani* [21] and acting against human pathogens [22] including multidrug-resistant *Pseudomonas aeruginosa* strains [23], *Salmonella typhi* [24] and clinical isolates of *Staphylococcus aureus* [25].

Recently, an important role for coumarins in microbiome modulation was demonstrated. It was shown that plant-derived coumarins shape the composition of Arabidopsis root microbial communities (rhizobiome) in Fe-starved plants, and possibly protect plants from pathogenic fungi [26]. Coumarins were shown to influence a reduced synthetic community (SynCom) of Arabidopsis root-isolated bacteria in synthetic media [27,28] and were proved to alter the rhizobiome and improve plant growth in Fe-limiting soil [29]. Even if, a role of root-exuded coumarins in structuring the rhizobiome was uncovered and their release into the rhizosphere upon pathogen infection was confirmed, the precise mechanisms underlying the above-described processes are only beginning to be discovered [26,28,30,31].

In the literature, there are examples of pathogens causing more severe disease symptoms on plants grown under high-Fe conditions when compared to Fe-deficient plants. One of the examples could be plant pathogenic bacteria *Dickeya dadantii* 3937, for which sufficient Fe uptake is essential to manifest full virulence on plants [32]. This bacterial pathogen causes soft rot and blackleg disease devastating potato and numerous other crops [33–35]. Bacteria from the *Dickeya* genus produce compounds called siderophores that form complexes with Fe and make it available to the microorganism. Taking into account that plants also produce Fe-chelating compounds, siderophore production is a part of the competitive relationships between plants and microorganisms that can promote infection [32]. It was demonstrated that Fe nutrition strongly affects the disease caused by *D. dadantii* 3937 also in a model plant Arabidopsis [36–38]. In Fe-starved Arabidopsis plants, authors observed a reduction in the expression of major bacterial virulence genes and finally a lower progression in disease symptoms on inoculated plants. The results obtained from a study of Arabidopsis response to *D. dadantii* 3937 infection highlight the major importance of the competition between plant and bacterial cells for Fe uptake during infection [36,38].

The above results reinforced the important role of coumarins in plant responses to disturbed Fe availability as well as their involvement in plant resistance and their release into the rhizosphere upon pathogen infection. The physiological functions of coumarins are strictly related to plant adaptation to various biotic and abiotic environmental stresses. Here, we investigated the Arabidopsis/*Dickeya* spp. pathosystem to better understand the relation between coumarins, plant Fe status, siderophores production and plant resistance to selected pathogenic bacteria.

2. Results

We used as a model, Arabidopsis wild-type plants (WT Col-0) and its mutants (*s8h* and *f6′h1*) defective in enzymes involved in coumarin biosynthesis (S8H and F6′H1, respectively) and *pdr9* mutant defective in coumarin transport to the rhizosphere (PDR9: Pleiotropic drug resistance 9 [2]). As plant pathogens, we included two *Dickeya* spp. strains: (1) a reference strain *D. dadantii* 3937 of medium virulence, (2) and *Dickeya solani* IFB0099 isolated from infected potato plant in Poland [39–41]. Interestingly, the selected strains differed in their ability to chelate Fe ions [40]. The ability to chelate Fe ions by *D. dadantii* 3937 strain was shown to be twice as high as the ability to chelate Fe ions by *D. solani* IFB0099 strain on CAS-agar medium [40].

2.1. Differential Susceptibility of Arabidopsis Plants Grown in Fe-Deficient Hydroponics to Tested Dickeya spp. Strains

To strictly control the growth conditions and maintain the nutrient composition of media, we conducted the hydroponic cultures (as described by [7] and [42]). Arabidopsis plants were cultivated in controlled conditions either under optimal (40 μM Fe^{2+}) or Fe-deficient conditions (0 μM Fe^{2+}) that induces coumarins accumulation [42] and subsequently were inoculated with *D. dadantii* 3937 and *D. solani* IFB0099 spp. It was reported previously that Arabidopsis roots release more phenolic-related compounds at later stages of life [43], therefore plants were inoculated at the flowering stage. We inoculated the Arabidopsis WT Col-0 and mutant lines defective in coumarin biosynthesis (*f6′h1*, *s8h*) with both *Dickeya* spp. strains, and evaluated disease progression according to the visual symptoms scoring with disease severity scale 0–5 (DSS) which allowed us to quantify the susceptibility of Arabidopsis plants to both bacterial strains (Figure 1).

Figure 1. Disease severity scale (DSS) on *Arabidopsis thaliana* wild-type (WT) Col-0 leaves inoculated with *D. dadantii* 3937 and/or *D. solani* IFB0099. Example images representing stages of DSS were taken 48 h after inoculation. The DSS was assigned to 0–5 scale and was defined as: (**0**) for no signs of symptoms of the disease and the wound has healed (observed for the mock control); (**1**) the necrotic tissue was observed in the inoculation site only; (**2**) the necrotic tissue observed in the inoculation site and max. 3 mm wide around it; (**3**) the maceration visible around the inoculation site spreading further with possible chlorosis of the leaf; (**4**) visible maceration of the whole leaf, possible chlorosis of other leaves, no maceration of the limb; (**5**) visible maceration of the whole limb and the leaf.

We observed that independently of genetic characteristic, all Arabidopsis genotypes inoculated with *D. solani* IFB0099 strain developed more severe disease symptoms (DSS up to 4.5) (Figure 2a,b) when compared to *D. dadantii* 3937 (DSS up to 3.5) (Figure 2c,d). It is

worth emphasizing that in case of all plant genotypes inoculated with *D. solani* IFB0099 strain, the symptoms of infection were more pronounced in conditions with optimal Fe availability (Figure 2a,b). Both mutants, in particular the *s8h* line, showed slightly higher disease symptoms after inoculation with *D. solani* IFB0099 when compared to WT Col-0 plants. Interestingly, when WT Col-0 plants were inoculated with *D. dadantii* 3937 strain, they developed the most severe disease symptoms at 96 h after inoculation when grown in optimal Fe solution (Figure 2c), while the opposite trend was detected for both Arabidopsis mutants with impaired biosynthesis of coumarins. During the *D. dadantii* 3937 infection, both Arabidopsis mutant lines showed a tendency to exhibit more severe disease symptoms in Fe-deficient hydroponics (Figure 2d) and not in the Fe-sufficient solution as WT Col-0 plants. This was particularly striking at 72 and 96 h after inoculation for the *f6′h1* mutants (Figure 2c,d), lacking the functional *F6′H1* gene, which does not synthesise scopoletin and its derivatives.

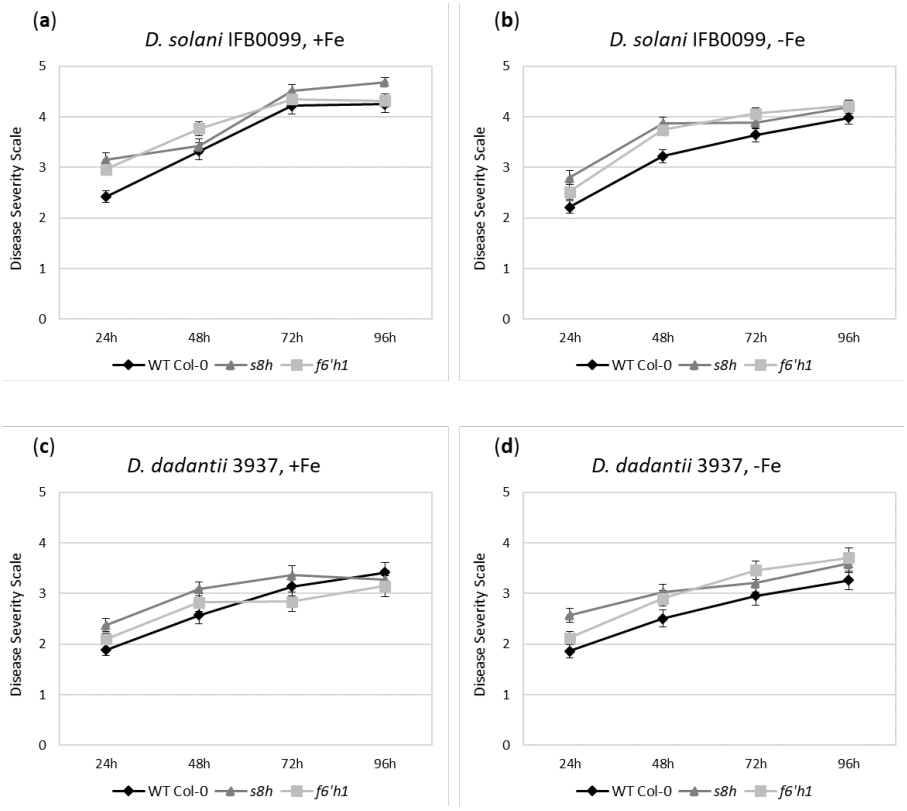


Figure 2. Disease progression caused by *Dickeya solani* IFB0099 (**a**,**b**) and *Dickeya dadantii* 3937 (**c**,**d**) strains on *Arabidopsis thaliana* wild-type (WT) Col-0 plants, *s8h* and *f6′h1* mutant lines grown in optimal (+Fe; 40 µM Fe^{2+}) (**a**,**c**) or Fe-deficient (−Fe, 0 µM Fe^{2+}) (**b**,**d**) hydroponic cultures by visual symptom scoring (Disease Severity Scale, DSS). The values represent the mean values of DSS originating from two independent experiments, in each experiment numerous individuals (n = 5–9) per plant genotypes (three leaves per plant) were inoculated for each time point. It is worth noting that the results averaged the DSS values obtained for two independent mutant lines for each tested gene. The mock-inoculated plants (with 0.85% NaCl) did not show the symptoms of the disease progression throughout the experiment. Error bars represent ± standard error (SE).

2.2. Inoculation of Arabidopsis WT Col-0 Grown in Fe-Deficient Hydroponics with Dickeya *spp. Cause Decrease in the Expression of S8H and F6′H1 Genes Involved in Coumarin Biosynthesis*

Next, to get insight into the expression levels of genes encoding key enzymes (S8H, F6′H1) involved in the biosynthesis of coumarins that are accumulated mostly in the

underground part of a plant, we performed the qPCR analysis with cDNA reverse transcribed using RNA isolated from the WT Col-0 roots grown under different Fe-regimes and inoculated with *D. dadantii* 3937 or *D. solani* IFB0099 strains. As expected, we observed up-regulation of both genes (*S8H* and *F6′H1*) in the Fe-deficient condition (Figure 3a,b). In particular, the *S8H* gene, which is known to be one of the most strongly Fe-responsive genes, was induced several hundred times in all treatments tested in Fe-deficient conditions compared to Fe-sufficient condition (Figure 3a).

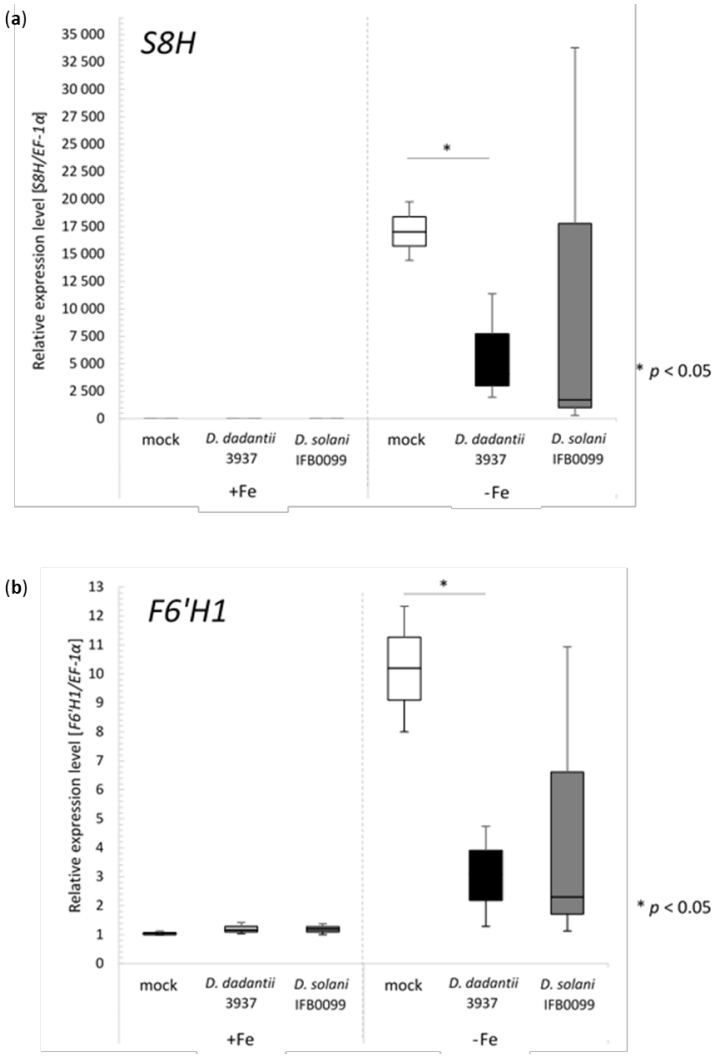


Figure 3. The relative expression levels of *S8H* (**a**) and *F6′H1* (**b**) genes were analysed in the roots of *Arabidopsis thaliana* wild-type (WT) Col-0 grown in hydroponics under optimal (40 μM Fe^{2+}) or Fe-deficient conditions (0 μM Fe^{2+}) and inoculated with *Dickeya dadantii* 3937 and *Dickeya solani* IFB0099 strains. Control plants were mock-inoculated with a 0.85% NaCl solution. As a reference, the *EF-1α* (ang. elongation factor-1α, At5g60390) gene was selected [44]. The pairwise t-test was used. Error bars, ±SD, from three biological replicates. * $p < 0.05$.

Interestingly, the expression of the *S8H* gene was significantly higher in the roots exposed only to abiotic stress (mock-inoculated WT Col-0 plants grown under Fe deficiency) when compared with those exposed to abiotic and biotic stress (*Dickeya* spp.- inoculated ones grown under Fe deficiency). It should be noted that the expression of *S8H* was relatively lower when combined environmental stress, composed of Fe-deficiency and bacterial

infection, was applied (Figure 3a: right panel). This relative decrease was particularly significant when *D. dadantii* 3937 was used to inoculate WT Col-0 plants ($p < 0.05$). A similar trend was observed for the *F6′H1* gene, for which a relative reduction in the expression levels was observed when the same two stress factors, Fe-deficiency and bacterial inoculation, were simultaneously applied. The *F6′H1* expression was approximately 2-fold lower in WT Col-0 plants subjected to biotic stress when compared to plant exposed only to abiotic stress under Fe-deficient condition (Figure 3b). In this case, the inoculation with *D. dadantii* 3937 also had a significant effect on the relative reduction of *F6′H1* expression in the WT Col-0 genetic background ($p < 0.05$). That is an interesting observation considering that this strain expressed a greater ability than *D. solani* IFB0099 to chelate Fe ions, as shown previously by the CAS-agar plate assay [40].

2.3. Fe-Chelation in CAS Agar Plate Assay

To test the potential of plant-produced compounds to affect the bacterial siderophore production, we observed the growth and halo production of *D. dadantii* 3937 and *D. solani* IFB0099 strains on CAS-agar plates supplemented with homogenates prepared from the leaves of *s8h* and *f6′h1* mutants (Figure 4a,b) that were grown in Fe-depleted hydroponic solution (0 μM Fe^{2+}).

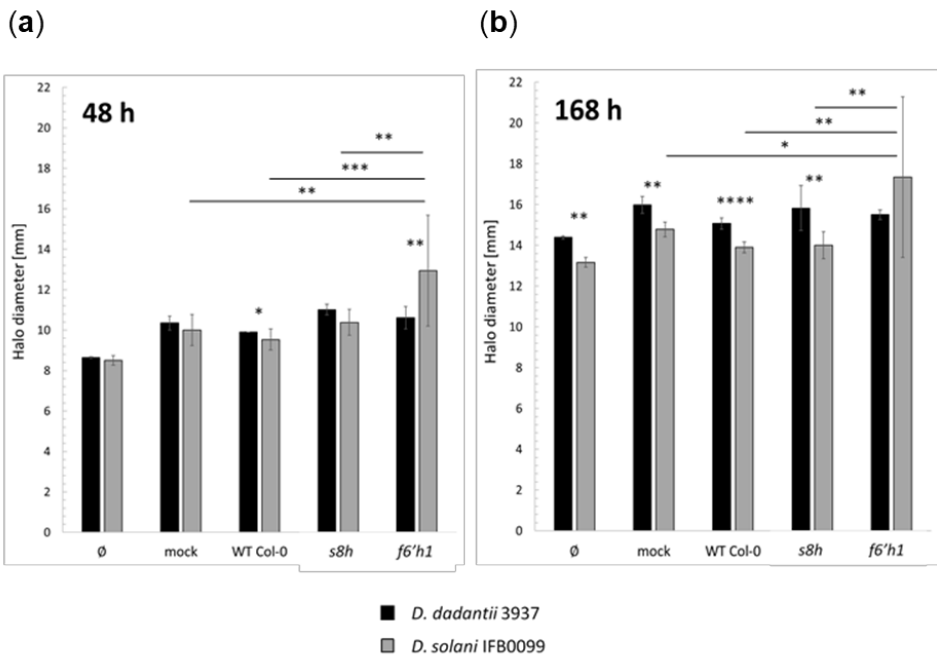


Figure 4. Siderophore production by *Dickeya dadantii* 3937 and *Dickeya solani* IFB0099 after incubation for (**a**) 48 and (**b**) 168 h on CAS-agar plates supplemented with leaf homogenates of *Arabidopsis thaliana* wild-type (WT) Col-0 and *s8h* or *f6′h1* mutant plants. The halo diameters were measured [mm] on control CAS-agar medium with no supplements (Ø), mock (phosphate buffer) and CAS-agar medium supplemented with leaf homogenates prepared from the Col-0 WT plants or *f6′h1, s8h* mutants. The experiment was performed twice, with 12 replicates (a set of pooled 3 plants was used for the biological replicates). The results are presented as average halo diameter. Error bars represent ± standard deviation (SD). Statistical significance is marked with asterisks * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$ based on student's *T*-test for independent samples.

Here, we selected leaves to prepare homogenates, as organs in which inoculation is conducted. No supplements or homogenates were added as a negative control (Ø), as a mock the phosphate buffer was used. It seems that both leaf homogenates and phosphate buffer itself can induce the production of bacterial siderophores. As could be expected,

in most cases *D. dadantii* 3937 strain that was previously characterised by a higher ability to chelate Fe ions [40], produced larger halos compared to *D. solani* IFB0099 (Figure 4a,b). But the most interestingly, we observed the opposite effect when homogenates originating from the leaves of *f6′h1* mutant were used as a supplement. We observed a significant increase in *D. solani* IFB0099 halos' diameter on agar plates supplemented with the *f6′h1* mutant leaf homogenates (Figure 4a,b).

2.4. Differential Susceptibility of Soil-Grown Arabidopsis Plants to Tested Dickeya spp. Strains

To shed light on the relationship between Fe homeostasis, coumarin accumulation and plant immunity in more physiological conditions, we grew a set of Arabidopsis mutants defective in coumarin accumulation (*f6′h1*, *s8h*) in the non-sterile soil environment. Here, we included in the experiment the *pdr9* mutant that is defective in coumarin transport to the rhizosphere [2]. Since soil mixture composition can alter Arabidopsis susceptibility to plant pathogens as was shown for the *Pseudomonas syringae* infection [45], we decided to estimate the disease symptoms caused by *D. dadantii* 3937 and *D. solani* IFB0099 on Arabidopsis plants grown in two different soil mixes (#1 and #2) derived from the commercial products that differ mainly in the level of salinity that can affect the availability of nutrients including Fe, and the content of some macro- and micronutrients like chlorides, phosphorus, potassium or calcium (Table 1).

Table 1. Chemical analysis of used soil mixes: (**a**) pH, salinity, macro- and (**b**) micronutrients.

(**a**)								
Soil Mix no.	**pH in H_2O**	**NaCl g/dm^3 Soil**	**NO_3**	**Cl**	**P**	**K**	**Ca**	**Mg**
		mg/dm^3 Soil						
#1	6.8	2.65	224	13.6	34.6	70.1	2960	>400 (548) [1]
#2	6.9	1.59	220	10.5	44.5	91.8	2509	>400 (498) [1]

(**b**)					
Soil Mix no.	**Cu**	**Zn**	**Mn**	**Fe**	**B**
	mg/dm^3 Soil				
#1	1.1	1.1	1.3	51.4	0.4
#2	1.0	1.3	2.3	49.2	0.4

[1] Results above upper limit of the method range for Mg = 400 mg/dm^3.

Arabidopsis plants of all genotypes (WT Col-0, *f6′h1*, *s8h* and *pdr9*) had better performance on soil mix #1, on which the plant rosettes were much bigger. In case of soil mix #2, the spontaneous plant wilt and die was also observed. This is an interesting observation since the conducted chemical analysis did not show any dramatic differences in the compositions of both soil mixes (Table 1). However, regardless of the soil in which the plants were grown, the Arabidopsis of all tested genotypes inoculated with *D. solani* IFB0099 strain developed more severe infection symptoms (DSS up to 4.5), compared to those challenged with *D. dadantii* 3937 reference strain (DSS up to 2.5) (Figure 5). However, most interestingly, we detected a variation in the disease symptoms between plant genotypes inoculated with the *D. solani* IFB0099 strain. In soil mix #1 characterised by a slightly higher salinity, the *pdr9* mutants inoculated with *D. solani* IFB0099 showed the mildest infection symptoms among all plant genotypes with the severity score up to 3.5 (Figure 5d) compared with the DSS up to 4.0 for WT Col-0 and *f6′h1* plants (Figure 5a,c) and 4.5 for *s8h* mutant line (Figure 5b). In soil mix #2, a slightly lower DSS (lower than 3) was observed for all mutant lines inoculated with *D. solani* IFB0099 (Figure 5b–d) compared to WT Col-0 plants (DSS up to 3) (Figure 5a). We did not observe such a variation in the infection symptoms on plants inoculated with the *D. dadantii* 3937 strain.

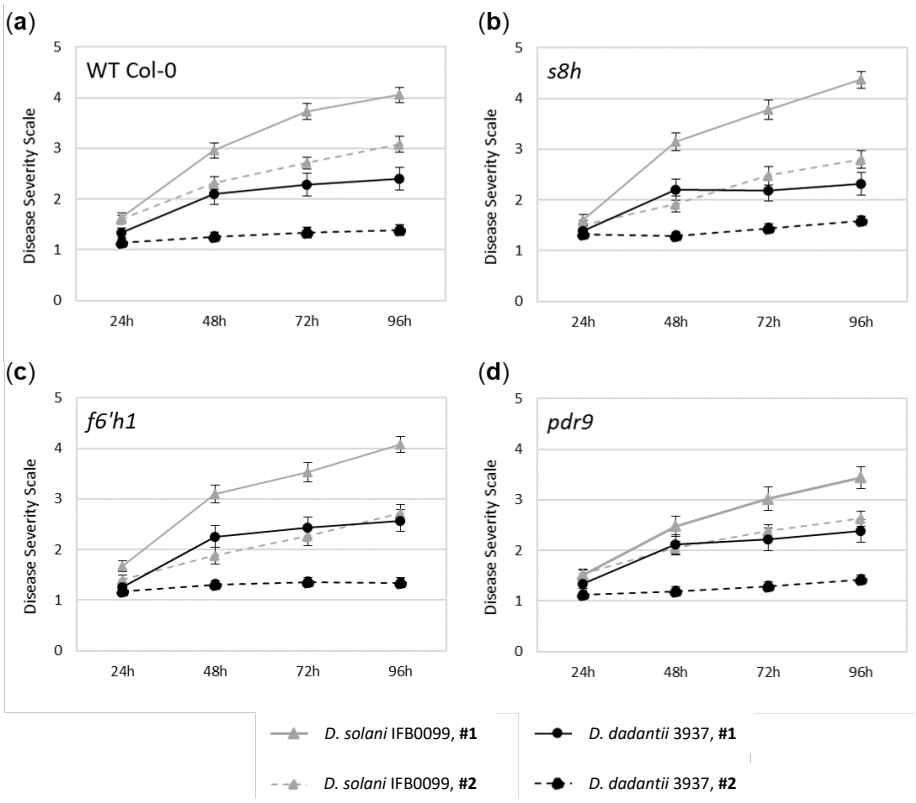


Figure 5. Disease progression caused by *Dickeya dadantii* 3937 and *Dickeya solani* IFB0099 strains on *Arabidopsis thaliana* (**a**) wild-type (WT) Col-0 plants, (**b**) *s8h*, (**c**) *f6′h1* and (**d**) *pdr9* mutant lines by visual symptom scoring (Disease Severity Scale, DSS). Plants were grown in two types of soil mixes (#1 and #2, see Table 1). The values represent the mean values of DSS originating from two independent experiments, in each experiment numerous individuals (n = 6–8) per plant genotypes (three leaves per plant) were inoculated for each time point. The mock-inoculated plants (with 0.85 % NaCl) did not show the symptoms of the disease progression throughout the experiment. Error bars represent ± standard error (SE).

2.5. Expression of Selected Plant Stress-Response Genes Is Strongly Induced in Arabidopsis Mutants Defective in Coumarin Accumulation Inoculated with D. solani IFB0099

Next, to measure the expression levels of selected genes being the plant stress markers, we analysed by qPCR the expression of genes related to oxidative stress (*SOD1*, AT1G08830), plant defence (pathogenesis-related *PR1*, At2g14610) and modulation of jasmonate-induced root growth inhibition (*CYP82C2*, At4g31970). We used RNA isolated from the leaves of Arabidopsis WT Col-0 and three mutant lines (*f6′h1*, *s8h*, *pdr9*) grown in soil mix #1, in which the disease symptoms after *D. solani* IFB0099 or *D. dadantii* 3937 strains infection were more prominent compared to soil mix #2 (Figure 5).

We observed a strong increase in expression of selected plant stress markers in plants inoculated with *D. solani* IFB0099 strain, which indicate that inoculation of Arabidopsis plants with this pathogen particularly induce the plant defence systems (Figure 6a–c). Interestingly, the expression levels of two out of three tested genes encoding plant stress markers (SOD1, PR1) were differentially induced among mutants and WT Col-0 plants, those inoculated with *Dickeya* spp. strains and mock-inoculated ones (Figure 6a,b). The level of the *SOD1* gene was visibly higher in all mutant lines compared to WT Col-0 plants (Figure 6a). While in the case of the *PR1* gene, we observed the opposite effect, where its expression was higher in WT Col-0 plants compared to mutants with disturbed coumarin biosynthesis (*f6′h1*, *s8h*) or transport (*pdr9*) (Figure 6b). The transcript levels

of *CYP82C2* were specifically induced by *D. solani* IFB0099 infection in all tested plant genotypes (Figure 6c).

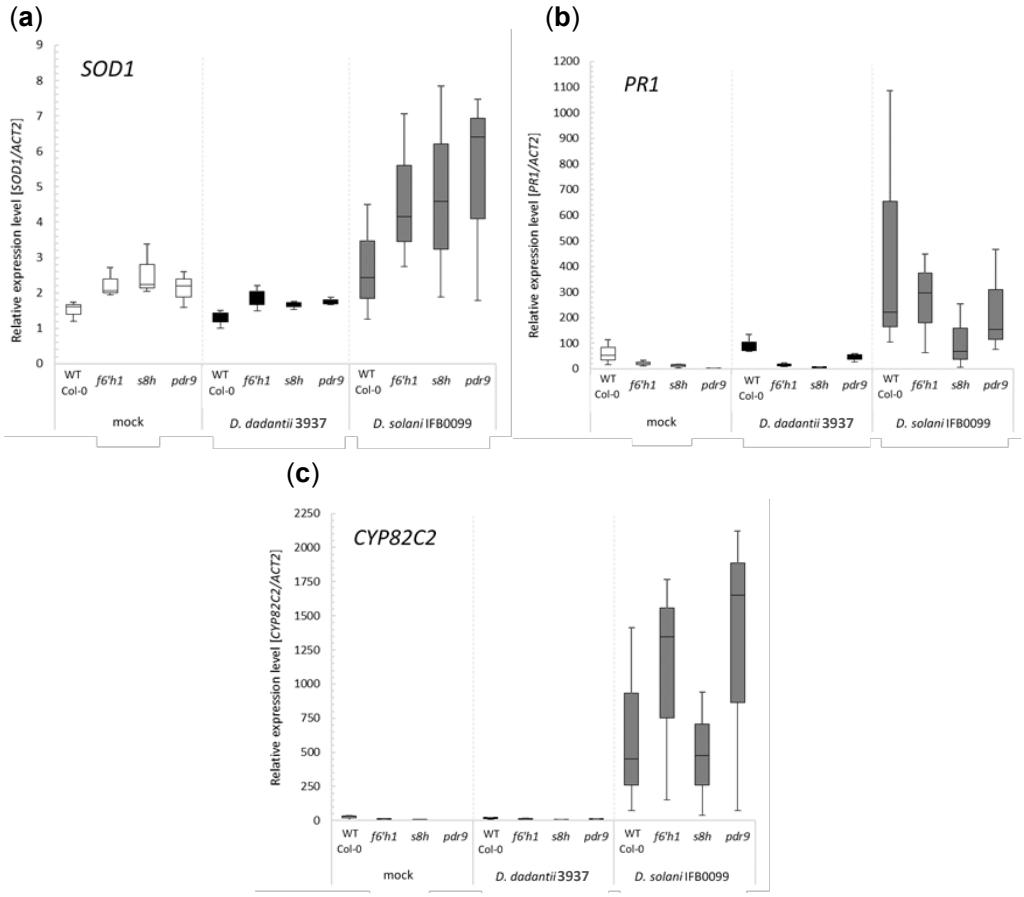


Figure 6. Relative expression levels of plant stress marker genes (**a**) *SOD1* (AT1G08830), (**b**) *PR1* (At2g14610) and (**c**) *CYP82C2* (At4g31970) analysed in the leaves of *Arabidopsis thaliana* wild-type (WT) Col-0 and mutant plants (*f6'h1, s8h, pdr9*) grown in soil mix #1 and inoculated with *Dickeya dadantii* 3937 and *Dickeya solani* IFB0099. 0.85% NaCl treated plants were used as a negative control. As a reference, the *ACT2* (At3g18780) gene was used [44]. Error bars, ±SD, from three biological replicates.

3. Discussion

Microorganisms that urgently need Fe for their growth, replication, metabolism and the infectious disease process, have evolved numerous strategies for Fe acquisition such as siderophore production. At the same time, plants are constantly subjected to various environmental stresses, including Fe-deficiency or pathogen attack during which Fe itself plays an important role. During a microbial infection, there is a competition between host and pathogen for the necessary nutritional resources. Numerous studies have shown that Fe ions play a key role in such competitive relationships [12,46–49]. It was also shown recently by several groups, including our research team, that the secretion of coumarins is essential for Fe acquisition under Fe-deficient conditions in a model plant *Arabidopsis thaliana* [7]. Plants and microorganisms have evolved a set of active strategies for Fe uptake from the soil that are based on acidification, chelation and reduction processes. Root exudation is one of such important processes determining the interaction of plants with the soil environment and microbiome. Coumarins that are secreted to the rhizosphere by roots

are involved in several processes determining plant interactions with the soil environment, both with biotic and abiotic factors.

In this study, we evaluated the disease symptoms caused by *Dickeya* spp. strains in Arabidopsis lines differing in coumarin accumulation that were grown under various growth conditions and Fe availability. The use of selected Arabidopsis mutants and bacterial strains of different origin enabled us to compare [1] the variation in disease symptoms among plant genotypes under numerous environmental scenarios, [2] and the expression of stress-related genes in plant genetic backgrounds with disturbed production and distribution of coumarins.

The presented analyses provided interesting insights into the differences in responses of the following plant genotypes: control plants (WT Col-0), coumarin-reach plants (WT Col-0 growing in Fe deficient environment), coumarin-deficient plants (*f6′h1*), fraxetin-deficient plants (*s8h*) and coumarin-hyperaccumulating plants (*pdr9*) to *D. dadantii* 3937 and *D. solani* IFB0099 strains. The developed model system of Arabidopsis/*Dickeya* spp. was applied to investigate the effect of two abiotic factors (Fe availability and coumarin content) on the disease severity. These studies are in line with the latest Top 10 Questions, selected by the International Congress on Molecular Plant-Microbe Interactions (IC-MPMI) community that met in Glasgow in 2019, covering the need to understand how the abiotic environment influence specific plant-microbe interactions [50]. It should be also highlighted that most of the previously published data describe the interaction between Arabidopsis plants and *D. dadantii* 3937. According to our best knowledge *D. solani*, which is an important plant pathogenic bacterium causing a loss in potato yield in Europe [34,51,52] was not tested before with a model plant Arabidopsis. Moreover, recent results show that *D. solani* strains cause severe disease symptoms in temperate climates, and are more aggressive than other blackleg-causing bacteria from genus *Dickeya* and *Pectobacterium* spp. [33,35,53,54].

To shed light on the strong relationship between Fe homeostasis (abiotic factor), coumarin accumulation and plant susceptibility to plant pathogenic bacteria, we grew WT Col-0 plants and two Arabidopsis mutants defective in coumarin accumulation (*s8h* and *f6′h1*) in the hydroponic cultures with strictly controlled Fe content and inoculated them with *Dickeya* spp. strains (biotic stress factor). We observed that all tested Arabidopsis genotypes (WT Col-0, *s8h* and *f6′h1*) inoculated with *D. solani* IFB0099 strain developed more severe disease symptoms than plants inoculated with *D. dadantii* 3937. The disease symptoms associated with *D. solani* IFB0099 infection were much more pronounced in Fe-sufficient hydroponics. A similar effect was observed for *D. dadantii* 3937-inoculated wild-type plants (WT Col-0). This is in line with the literature data showing that *D. dadantii* belongs to the pathogens causing more severe disease symptoms on plants grown under high-Fe conditions when compared to Fe-deficient environmental condition [32]. The most noticeable was a detection of the opposite effect for Arabidopsis mutants with impaired biosynthesis of coumarins. During *D. dadantii* 3937 infection of *s8h* and *f6′h1* mutant plants, the more severe disease symptoms were observed in Fe-deficient hydroponics, and not in the Fe-sufficient cultures, particularly at 72 h after inoculation. Both of these mutants (*s8h* and *f6′h1*) are defective in enzymes involved in the biosynthesis of coumarins, which are secondary metabolites important for Fe uptake in plants [4]. Consequently, the *Dickeya* spp. cells, which compete for Fe with the plant cells, can uptake Fe with higher efficiency and accumulate more of it than those infecting WT plants. It was shown before by other groups [32,37] that Fe uptake is important for bacteria ability to macerate plant tissue and the production of virulence factors. As a result of this, the *Dickeya* spp. strains infecting Arabidopsis mutants, which are defective in Fe acquisition, cause more severe disease symptoms in these plants. It is worth noticing that qPCR analysis of the corresponding genes (*S8H* and *F6′H1*) in the WT Col-0 genetic background, proved that their expression levels were relatively lower when a combined environmental stress, composed of Fe-deficiency and bacterial infection, was applied. Taking into account that *D. dadantii* 3937 has a higher ability to chelate Fe ions and the expression of several bacterial genes involved in siderophore-mediated Fe uptake is controlled by the Fe availability [32], it can be concluded

that (1) coumarins produced by plants influence more strongly pathogens for which siderophores production play a particularly important role in the pathogenesis process, (2) disorders of coumarin biosynthesis are more important for the disease symptoms under Fe-deficiency conditions.

Next, we explored the potential of coumarins and other factors possibly present in the selected plant homogenates to affect the bacterial siderophore production by measuring the halos' diameter produced by the *Dickeya* spp. strains on CAS-agar plates. We tested leaf homogenates prepared from a set of Arabidopsis plants (WT Col-0, *s8h* and *f6′h1*). In siderophore production assay, we observed greater halos produced by *D. dadantii* 3937 than that produced by *D. solani* IFB0099 for all tested Arabidopsis genotypes, except for the *f6′h1* mutants. It seems like coumarins and other possible factors present in leaf homogenates and phosphate buffer can induce, directly or indirectly, the production of bacterial siderophores. The existence of not-characterised yet interplay between coumarins and the bacterial siderophore production needs further investigation. It is important to continue and develop research on the role of coumarins as novel elements of chemical communication and to test if, in coumarin-deficient plants, the induction of other compensatory pathway occurs.

To better understand the responses of coumarin-deficient plants to combined environmental stimuli, we grew a set of Arabidopsis mutants defective in coumarin accumulation (*f6′h1*, *s8h*) and coumarin transport to the rhizosphere (*pdr9*) in more physiological conditions. The inoculation was conducted on plants grown in two soil mixes with some differences in chemical compositions (Table 1). It has to be highlighted that plants grew significantly better in soil mix #1, in which all rosettes were much larger. We can speculate that the lack of any of the elements of soil mix #2 is limiting plant growth. However, considering that the soil mix #2 consist of half of the peat moss, which is a natural product of organic origin, we can suspect the significant differences in the microbiomes of the tested soil mixes. These interesting questions should be clarified in the future. Importantly, regardless of the soil mix in which plants were grown, the Arabidopsis of all tested genotypes inoculated with the *D. solani* IFB0099 strain developed more severe infection symptoms compared to *D. dadantii* 3937 reference strain. For both bacterial strains, the symptoms of infection were more pronounced on plants with a better growth on soil mix #1 compared to plants grown in soil mix #2. Interestingly, we detected a variation in the disease symptoms between plant genotypes inoculated with *D. solani* IFB0099 strain, particularly on plants grown in the soil mix #1. The *pdr9* mutant plants that hyperaccumulate coumarins in their tissues, showed the mildest infection symptoms among all plant genotypes when inoculated with *D. solani* IFB0099 strain. While both Arabidopsis mutants (*f6′h1* and *s8h*) defective in coumarin biosynthetic genes, showed stronger symptoms of infection. The explanation of this phenomenon can be that coumarins are known for their antimicrobial activity, however, the observed genotype-specific mode of action needs further investigations. We did not observe such a clear variation in the infection symptoms on plants inoculated with *D. dadantii* 3937 strain that is characterised by a higher ability to chelate Fe ions (the CAS-plate assays presented in Figure 4 and [40]).

In this work, we also analysed the expression profiles of three plant genes (*PR1, SOD1, CYP82C2*) which products are involved in the plant tissue response to a wide range of stresses including biotic and abiotic factors. During *D. solani* IFB0099 infection of Arabidopsis, the expression of *PR1*, which is considered to be one of the markers for salicylic acid (SA)-dependent systemic acquired resistance (SAR) [55,56], was strongly induced in the leaves of all infected genotypes. This increase in the *PR1* gene expression was most pronounced in the WT Col-0 genetic background. A similar *PR1* expression profile was observed in *D. dadantii* 3937- and mock-inoculated plants, but the levels of *PR1* expression were much lower in these experimental setups. In parallel, we observed in *D. solani* IFB0099-inoculated plants an induction in the expression of the *SOD1* gene, which encodes a cytosolic copper/zinc superoxide dismutase CSD1 that can be regulated by biotic and abiotic stresses and detoxify superoxide radicals [57]. This indicates that plants infected with *D. solani* IFB0099 induce the defensive mechanism by increasing the

production of critical antioxidant enzymes protecting organisms from reactive oxygen species. Interestingly, plants with impaired biosynthesis or coumarin accumulation have a higher expression of *SOD1*, although the observed differences are not statistically significant. Furthermore, the expression of *CYP82C2* in *D. solani* IFB0099-inoculated Arabidopsis WT Col-0 and its mutants, was also clearly upregulated in comparison to *D. dadantii* 3937- and mock-inoculated plants. Since *CYP82C2* was shown to be involved in several aspects of jasmonic acid (JA) responses [58], it seems likely that infection with *D. solani* IFB0099 pathogen induces activation of the JA-dependent response.

The above results and data previously published obtained from a study of Arabidopsis response to *D. dadantii* 3937 infection highlight the major importance of the competition between plant and bacterial cells for Fe uptake during infection [36,38]. It was demonstrated that Fe nutrition strongly affects the disease caused by soft rot-causing plant pathogenic bacteria with a large plant host range including Arabidopsis. Plants have evolved various strategies to acquire Fe from their environment and mechanisms tightly regulating Fe uptake, transport and storage [11,13,59] including the production of Fe-mobilizing phenolic compounds like coumarins [1–5,7,8,60–62]. The production of exudates, which is dependent on the external environment, at the same time is genetically regulated in plants. It was shown by Micallef et al. [63] that natural populations of Arabidopsis originating from various geographical localization, called accessions, release a unique set of compounds into their rhizosphere. The authors detected that the rhizobacterial community composition and the relative abundance of particular ribotypes were also accession-dependent. They hypothesised that the observed natural variation in root exudation could partly explain the genotypic influence on bacterial communities in the rhizosphere [63]. Many studies of plant-microbe interactions revealed that plants are not only able to shape their rhizosphere microbiome, but also highlight this root-associated microbial community to be referred to as the second genome of the plant, which is crucial for plant health [64]. Our research group detected previously the existence of natural variation in the accumulation of antimicrobial coumarins, namely scopoletin and scopolin, among Arabidopsis accessions [14]. Lately, we also detected a significant variation in the content of other simple coumarins like umbelliferone and esculetin together with their glycosides: skimmin and esculin, respectively [65]. It was also shown recently that a natural variation exists in Arabidopsis tolerance to *Dickeya* spp. [66]. The significantly different susceptibility groups were uncovered within a small set of eight Arabidopsis accessions following inoculation with *D. dadantii* 3937, which suggested that tolerance associated loci might be present in this model plant. Even though *Dickeya* spp. are causative agents of severe diseases in a wide range of plant species and major economic losses, little data concerning potential resistance genes are available [67–69]. These data strongly suggest that Arabidopsis with its extensive genetic natural variation and a set of powerful genetic tools including web-accessible collections of mutants, provides an excellent model to study the interplay between secondary metabolites production, exudate profiles, Fe homeostasis and interaction with beneficial and plant pathogenic microbes.

4. Materials and Methods

4.1. Plant Material

The Arabidopsis thaliana accession Columbia was used as the wild type (Arabidopsis WT Col-0) together with a set of T-DNA insertional mutant lines in the Col-0 background: [1] *s8h-1* (SM_3.27151); *s8h-2* (SM_3.23443); [2] *f6′h1-1* (SALK_132418); *f6′h1-2* (SALK_050137C) and [3] *pdr9-1* (SALK_050885). Seeds of all lines are available at the Nottingham Arabidopsis Stock Centre (http://arabidopsis.info/, accessed on 15 April 2021).

4.2. Bacterial Strains, Media, Growth Conditions

The strains used in this study, *Dickeya dadantii* 3937 IFB0459 and *Dickeya solani* IFB0099, are available at the collection of bacterial pathogens located at the IFB UG & MUG, in

Poland. For the plant inoculation, the bacteria were grown overnight in lysogeny broth (LB) [70] liquid medium at 28 °C with agitation at 120 rpm. Then the bacterial cultures were centrifuged in Eppendorf tubes (5 min, 6500 rpm), washed in sterile 0.85% NaCl and centrifuged again. The bacterial suspension (at least 15 mL) was prepared in sterile 0.85% NaCl and adjusted to 1 MacFarland Unit (Densitometer DEN-1/DEN-1B, Buch & Holm Herlev, Denmark), approximately 10^8 cfu/mL.

4.3. Hydroponic Cultures

After a few days' stratification at 4 °C, Arabidopsis WT Col-0 plants were grown in a controlled environment (16 h light at 22 °C/~100 µmol m^{-2} s^{-1} and 8 h dark at 20 °C) in a modified 1 × Heeg solution [71], as described in details in [42] with the following modifications. Approximately 3-weeks-old plants in tube lids, filled with the solidified Heeg medium, were transferred from tip boxes with control solution (40 µM Fe^{2+}) to the modified 50 mL Falcon centrifuge tubes filled with optimal (40 µM Fe^{2+}) or Fe-deficient (0 µM Fe^{2+}) medium. The roots were passing through the 1-cm diameter hole drilled in Falcons' lid to support the seedling holder as proposed by Conn et al. [72]. Hydroponic solutions were replenished by the addition of a fresh medium.

4.4. Soil Cultures

Arabidopsis seeds were first stratified in Petri dishes on water-saturated Whatman paper followed by a cold treatment for 4 d at 4 °C and then planted into two different soil mixes derived from commercial products. Soil mix #1 consists of commercial soil 1 and vermiculite (3–6 mm in diameter) in a proportion of 3:1, respectively. Soil mix #2 consists of commercial de-acidified peat moss mix with commercial soil 2 and vermiculite in a proportion 2:1:1, respectively. Chemical analysis were conducted by the Regional Agro-Chemical Station in Gdansk, Poland (OSCh-R, http://www.oschrgdansk.pl/, accessed on 15 April 2021). Prior to sowing seeds, both soil mixes were soaked with general-purpose fertiliser (Substral, The Scotts Miracle-Gro Company, Marysville, OH USA). Arabidopsis plants were grown for five weeks under a photoperiod of 16 h light (120 µmol m^{-2} s^{-1}) at 22 °C and 8 h dark at 20 °C, before being inoculated.

4.5. Siderophore Production in the Presence of Plant Extracts

Homogenates from leaves of *Arabidopsis thaliana* were prepared from WT Col-0 plants, *s8h* and *f6′h1* mutants. Briefly, the pooled leaves (~300 ± 40 mg) stored at −80 °C were thawed and homogenised in Bioreba bags (BIOREBA AG, Reinach, Switzerland) with 3 mL of 50 mM phosphate buffer pH 7.0 with the use of hand homogeniser (BIOREBA AG, Reinach, Switzerland). Then, the homogenate was centrifuged twice in Eppendorf tubes (8500 rpm, 2 min). The supernatant was filtered with a syringe 0.22 µm filter into sterile Eppendorf tubes (in total 3.5 mL) and immediately used on plates. Siderophore production of bacterial strains was determined on chrome azurol S–agar (CAS-agar) plates [73] supplemented with 100 µL of each homogenate with sterile spreader 15 min before bacteria inoculation. The overnight cultures of bacterial strains were centrifuged and resuspended in sterile 0.85% NaCl and adjusted to 0.5 MacFarland (Densitometer DEN-1/DEN-1B, Buch & Holm). 10 µL of each bacterial suspension was put on the CAS-agar plates and incubated at 28 °C for up to 168 h. We measured the halo diameters developed on CAS-agar plates supplemented with plant homogenates every 24 h.

4.6. Plant Inoculation with Bacterial Strains

Approximately 5-weeks-old plants, grown in soil or hydroponically, were inoculated with the bacterial suspensions of either *D. dadanti* 3937 or *D. solani* IFB0099 with the use of laboratory pincers. The pincers were sterilised before use and approximately 1 cm of the pincer tip was dipped into the bacterial suspension (the final inoculum of 10^8 cfu/mL) and immediately the plant leaf was pinched with the pincers. At least 8 leaves were inoculated with each bacterial strain and mock. We pinched the middle parts of the

selected representative leaves (3 leaves per plant). The negative control (mock) were plants inoculated only with sterile 0.85% NaCl. For the mock-inoculated plants no symptoms development was observed throughout the experiment. The number of bacteria inoculated into the plant leaf with pincers was about 2×10^7 cfu/leaf and it was stable throughout experiments (data not shown). After inoculation, trays with plants were placed in the boxes filled with one litre of water, which were covered with transparent lids, to enable 100% humidity. Next, boxes were placed for 96 h in phytotron at 28 °C (16 h light at 28 °C/~100 µmol m^{-2} s^{-1} and 8 h dark at 26 °C), which is an optimal temperature for the development of disease symptoms by bacteria from the *Dickeya* spp. The whole rosettes were collected for each genotype grown in soil mix #1, frozen in liquid nitrogen and stored at −80 °C until further analysis. The roots of Col-0 plants grown hydroponically were gently removed from the agar droplets with tweezers and then rinsed in a beaker with distilled water. After drying on a paper towel, roots were frozen in liquid nitrogen and placed at −80 °C.

4.7. Quantification of Arabidopsis Plants Susceptibility to Dickeya *spp. Strains by Visual Symptom Scoring (Disease Severity Scale, DSS)*

The Arabidopsis plants were scored for soft rot/ blackleg symptoms development on leaves daily at 0 h, 24 h, 48 h, 72 h and 96 h post-inoculation. Developed symptoms of the disease were assigned to 0–5 scale (Figure 1), for which "0" meant no signs of symptoms of the disease and the wound has healed (it was observed for the negative control leaves); "1" the necrotic tissue was observed in the inoculation site only; "2" the necrotic tissue observed in the inoculation site and max. 3 mm wide around it; "3" the maceration visible around the inoculation site spreading further with possible chlorosis of the leaf; "4" visible maceration of the whole leaf, possible chlorosis of other leaves, no maceration of the limb; "5" visible maceration of the whole limb and the leaf.

4.8. RNA Extraction and Expression Analysis by qRT-PCR

Total RNA was extracted from plant material harvested 48 h after inoculation. The plant tissue was homogenised in liquid nitrogen using sterile mortars cleaned with isopropanol and baked for 4 h at 180 °C. A commercially available E.Z.N.A.® Plant RNA Kit (Omega Bio-tek, Inc., Norcross, GA USA) was used following the instructions of the manufacturer and including an additional step to remove the genomic DNA contamination from the mixture with RNase-Free DNase I Set (Omega Bio-tek, Inc., Norcross, GA USA). 500 ng RNA for RNA isolated from leaves or 200 ng RNA for RNA isolated from roots was used for reverse transcription by Maxima First Strand reverse transcriptase cDNA Synthesis Kit for RT-qPCR (Thermo Fisher Scientific, Waltham, MA USA). qPCR was performed using LightCycler® 480 Real-Time PCR System (Hoffmann-La Roche, Basel, Switzerland) and SYBR® Green Master Mix (Thermo Fisher Scientific, Waltham, MA USA), using the gene-specific primers shown in Table 2. Primers' specificities were confirmed by the analysis of the melting curves. Relative transcript levels (RLT) of the plant genes in leaf tissues were normalised to the transcript level of the house-keeping *ACTIN2* gene (At3g18780). As a reference for the root tissues, the *EF-1α* (ang. elongation factor-1α, At5g60390) gene was selected [44].

Table 2. Primer sequences for plant genes used in qPCR reactions.

Name	Sequence (5′-3′)	Description
3g18780For	CTTGCACCAAGCAGCATGAA	Primer for *ACT2* gene [1]
3g18780Rev	CCGATCCAGACACTGTACTTCCTT	Primer for *ACT2* gene [1]
AT2G14610_F	TTCTTCCCTCGAAAGCTCAAGA	Primer for *PR1* gene
AT2G14610_R	GTGCCTGGTTGTGAACCCTTA	Primer for *PR1* gene
AT4G31970_F	GATGGTGAGAATGGTGGCCG	Primer for *CYP82C2* gene
AT4G31970_R	GCCTCTTCGGCATCTTCAGG	Primer for *CYP82C2* gene
AT1G08830_F	TCAACCCCGATGGTAAAACAC	Primer for *SOD1* gene
AT1G08830_R	TCACCAGCATGTCGATTAGCA	Primer for *SOD1* gene
At5g60390_F	TGAGCACGCTCTTCTTGCTTTCA	Primer for *EF-1α* gene [1]
At5g60390_R	GGTGGTGGCATCCATCTTGTTACA	Primer for *EF-1α* gene [1]
S8HqPCR_F	GCCGAGACACTTGGCTTCTT	Primer for *S8H* gene
S8HqPCR_R	CAGCAGCTCCACCGAAACA	Primer for *S8H* gene
F6H1qPCRf	TGATGAGGACAGAGTCGCTGAA	Primer for *F6′H1* gene
F6H1qPCRr	CACTTGAAAGAACCCCCATTTC	Primer for *F6′H1* gene

[1] Reference [44].

5. Conclusions

We investigated here the possible interactions between plant resistance, coumarin content and Fe status by using the plant pathogenic *Dickeya* spp. strains and a set of selected Arabidopsis mutants defective in coumarin biosynthesis (*f6′h1*, *s8h*) and their transportation (*pdr9*). We studied the effect of disturbed coumarin accumulation and Fe deficiency on the disease severity using a model system of Arabidopsis/*Dickeya* spp. interactions. Arabidopsis plants grown in hydroponic cultures with different Fe regimes and two soil mixes were inoculated with *Dickeya* spp. strains or treated with NaCl as a control. Under all conditions tested, Arabidopsis plants inoculated with *D. solani* IFB0099 developed more severe disease symptoms compared to plants inoculated with *D. dadantii* 3937 strain. While the response of plants to *D. dadantii 3937* infection was genotype-dependent in Fe-deficient hydroponic solution. Subsequently, we showed that the expression of genes encoding plant stress markers was also strongly induced by *D. solani* IFB0099 infection. Interestingly, the inoculation of WT Col-0 plants grown in Fe-deficient hydroponics with both *Dickeya* spp. strains cause a decrease in the expression of *S8H* and *F6′H1* genes involved in coumarin biosynthesis.

Dickeya spp. was chosen as a plant pathogenic bacteria causing soft rot disease that can infect a broad spectrum of plants, whereas plant genotypes were selected due to their disturbed coumarin accumulation in roots and exudate profiles, which may have an impact on specific microbial consortia selection in the rhizosphere and influence plant response to pathogen attack. This may play a particularly important role for plant development and growth under Fe deficiency. The molecular mechanisms underlying these fascinating interactions are not yet well understood. We believe that Arabidopsis/*Dickeya* spp./pathosystem, together with a set of various Arabidopsis mutants defective in coumarin biosynthesis and its significant natural genetic variation, will be in future beneficial in uncovering a role of root-exuded coumarins in structuring the rhizobiome and plant resistance to pathogens.

Author Contributions: Conceptualization, M.P., E.L., and A.I.; methodology, M.P., J.S., and A.I.; validation, I.P., M.P., and J.S.; formal analysis, I.P.; investigation, I.P., M.P., J.S., and D.S.; writing—original draft preparation, M.P. and A.I.; writing—review and editing, E.L. and A.I.; supervision, E.L. and A.I.; project administration, A.I.; funding acquisition, A.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the NARODOWE CENTRUM NAUKI (Polish National Science Centre) grant number 2014/15/B/NZ2/01073 (OPUS to AI) and grant number 2019/35/O/NZ1/02751 (PRELUDIUM BIS to EL).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available within the article.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Clemens, S.; Weber, M. The Essential Role of Coumarin Secretion for Fe Acquisition from Alkaline Soil. *Plant Signal. Behav.* **2016**, *11*, e1114197. [CrossRef] [PubMed]
2. Fourcroy, P.; Sisó-Terraza, P.; Sudre, D.; Savirón, M.; Reyt, G.; Gaymard, F.; Abadía, A.; Abadia, J.; Alvarez-Fernández, A.; Briat, J.-F. Involvement of the ABCG37 Transporter in Secretion of Scopoletin and Derivatives by Arabidopsis Roots in Response to Iron Deficiency. *New Phytol.* **2014**, *201*, 155–167. [CrossRef] [PubMed]
3. Rajniak, J.; Giehl, R.F.H.; Chang, E.; Murgia, I.; von Wirén, N.; Sattely, E.S. Biosynthesis of Redox-Active Metabolites in Response to Iron Deficiency in Plants. *Nat. Chem. Biol.* **2018**, *14*, 442–450. [CrossRef] [PubMed]
4. Rodríguez-Celma, J.; Lin, W.-D.; Fu, G.-M.; Abadía, J.; López-Millán, A.-F.; Schmidt, W. Mutually Exclusive Alterations in Secondary Metabolism Are Critical for the Uptake of Insoluble Iron Compounds by Arabidopsis and Medicago Truncatula. *Plant Physiol.* **2013**, *162*, 1473–1485. [CrossRef]
5. Schmid, N.B.; Giehl, R.F.H.; Döll, S.; Mock, H.-P.; Strehmel, N.; Scheel, D.; Kong, X.; Hider, R.C.; von Wirén, N. Feruloyl-CoA 6′-Hydroxylase1-Dependent Coumarins Mediate Iron Acquisition from Alkaline Substrates in Arabidopsis. *Plant Physiol.* **2014**, *164*, 160–172. [CrossRef]
6. Sisó-Terraza, P.; Luis-Villarroya, A.; Fourcroy, P.; Briat, J.-F.; Abadía, A.; Gaymard, F.; Abadía, J.; Álvarez-Fernández, A. Accumulation and Secretion of Coumarinolignans and Other Coumarins in Arabidopsis Thaliana Roots in Response to Iron Deficiency at High PH. *Front. Plant Sci.* **2016**, *7*, 1711. [CrossRef]
7. Siwinska, J.; Siatkowska, K.; Olry, A.; Grosjean, J.; Hehn, A.; Bourgaud, F.; Meharg, A.A.; Carey, M.; Lojkowska, E.; Ihnatowicz, A. Scopoletin 8-Hydroxylase: A Novel Enzyme Involved in Coumarin Biosynthesis and Iron-Deficiency Responses in Arabidopsis. *J. Exp. Bot.* **2018**, *69*, 1735–1748. [CrossRef]
8. Tsai, H.-H.; Rodríguez-Celma, J.; Lan, P.; Wu, Y.-C.; Vélez-Bermúdez, I.C.; Schmidt, W. Scopoletin 8-Hydroxylase-Mediated Fraxetin Production Is Crucial for Iron Mobilization. *Plant Physiol.* **2018**, *177*, 194–207. [CrossRef]
9. Vanholme, R.; Sundin, L.; Seetso, K.C.; Kim, H.; Liu, X.; Li, J.; De Meester, B.; Hoengenaert, L.; Goeminne, G.; Morreel, K.; et al. COSY Catalyses Trans-Cis Isomerization and Lactonization in the Biosynthesis of Coumarins. *Nat. Plants* **2019**, *5*, 1066–1075. [CrossRef]
10. Kai, K.; Mizutani, M.; Kawamura, N.; Yamamoto, R.; Tamai, M.; Yamaguchi, H.; Sakata, K.; Shimizu, B. Scopoletin Is Biosynthesized via Ortho-Hydroxylation of Feruloyl CoA by a 2-Oxoglutarate-Dependent Dioxygenase in Arabidopsis Thaliana. *Plant J.* **2008**, *55*, 989–999. [CrossRef]
11. Curie, C.; Briat, J.-F. Iron Transport and Signaling in Plants. *Annu. Rev. Plant Biol.* **2003**, *54*, 183–206. [CrossRef]
12. Verbon, E.H.; Trapet, P.L.; Stringlis, I.A.; Kruijs, S.; Bakker, P.A.H.M.; Pieterse, C.M.J. Iron and Immunity. *Annu. Rev. Phytopathol.* **2017**, *55*, 355–375. [CrossRef]
13. Grillet, L.; Schmidt, W. Iron Acquisition Strategies in Land Plants: Not so Different after All. *New Phytol.* **2019**, *224*, 11–18. [CrossRef]
14. Siwinska, J.; Kadzinski, L.; Banasiuk, R.; Gwizdek-Wisniewska, A.; Olry, A.; Banecki, B.; Lojkowska, E.; Ihnatowicz, A. Identification of QTLs Affecting Scopolin and Scopoletin Biosynthesis in Arabidopsis Thaliana. *BMC Plant Biol.* **2014**, *14*, 280. [CrossRef]
15. Yang, L.; Ding, W.; Xu, Y.; Wu, D.; Li, S.; Chen, J.; Guo, B. New Insights into the Antibacterial Activity of Hydroxycoumarins against Ralstonia Solanacearum. *Molecules* **2016**, *21*, 468. [CrossRef]
16. Sun, H.; Wang, L.; Zhang, B.; Ma, J.; Hettenhausen, C.; Cao, G.; Sun, G.; Wu, J.; Wu, J. Scopoletin Is a Phytoalexin against Alternaria Alternata in Wild Tobacco Dependent on Jasmonate Signalling. *J. Exp. Bot.* **2014**, *65*, 4305–4315. [CrossRef]
17. El Oirdi, M.; Trapani, A.; Bouarab, K. The Nature of Tobacco Resistance against Botrytis Cinerea Depends on the Infection Structures of the Pathogen. *Environ. Microbiol.* **2010**, *12*, 239–253. [CrossRef]
18. Shimizu, B.; Miyagawa, H.; Ueno, T.; Sakata, K.; Watanabe, K.; Ogawa, K. Morning Glory Systemically Accumulates Scopoletin and Scopolin after Interaction with Fusarium Oxysporum. *Z. Naturforsch. C J. Biosci.* **2005**, *60*, 83–90. [CrossRef]
19. Prats, E.; Bazzalo, M.E.; León, A.; Jorrín, J.V. Fungitoxic Effect of Scopolin and Related Coumarins on Sclerotinia Sclerotiorum. A Way to Overcome Sunflower Head Rot. *Euphytica* **2006**, *147*, 451–460. [CrossRef]

20. Ndam, Y.N.; Nyegue, M.A.; Mounjouenpou, P.; Kansci, G.; Kenfack, M.J.; Eugène, E.E. LC-MS Quantification of Scopoletin in Cassava(Manihot Esculenta Crantz) Varieties, Local Derived Foods, and Activity on Some Food Spoilage Fungi. *J. Food Process. Preserv.* **2020**, *44*, e14387. [CrossRef]
21. Modafar, C.E.; Clérivet, A.; Vigouroux, A.; Macheix, J.J. Accumulation of Phytoalexins in Leaves of Plane Tree(Platanus Spp.) Expressing Susceptibility or Resistance ToCeratocystis Fimbriata f. Sp.Platani. *Eur. J. Plant Pathol.* **1995**, *101*, 503–509. [CrossRef]
22. Gnonlonfin, G.J.B.; Sanni, A.; Brimer, L. Review Scopoletin—A Coumarin Phytoalexin with Medicinal Properties. *Crit. Rev. Plant Sci.* **2012**, *31*, 47–56. [CrossRef]
23. Napiroon, T.; Bacher, M.; Balslev, H.; Tawaitakham, K.; Santimaleeworagun, W.; Vajrodaya, S. Scopoletin from Lasianthus Lucidus Blume(Rubiaceae): A Potential Antimicrobial against Multidrug-Resistant Pseudomonas Aeruginosa. *J. Appl. Pharm. Sci.* **2018**, *8*, 1–6. [CrossRef]
24. Acharya, D.; Bogati, B.; Risal, P. Scopoletin reduces intracellular survival of Salmonella typhi within U937 human macrophage cell line in vitro. *Afr. J. Microbiol. Res.* **2013**, *1*, 47–51.
25. Mogana, R.; Adhikari, A.; Tzar, M.N.; Ramliza, R.; Wiart, C. Antibacterial Activities of the Extracts, Fractions and Isolated Compounds from Canarium Patentinervium Miq. against Bacterial Clinical Isolates. *BMC Complement. Med. Ther.* **2020**, *20*, 55. [CrossRef] [PubMed]
26. Stringlis, I.A.; Yu, K.; Feussner, K.; de Jonge, R.; Van Bentum, S.; Van Verk, M.C.; Berendsen, R.L.; Bakker, P.A.H.M.; Feussner, I.; Pieterse, C.M.J. MYB72-Dependent Coumarin Exudation Shapes Root Microbiome Assembly to Promote Plant Health. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E5213–E5222. [CrossRef]
27. Bai, R.-R.; Wu, X.-M.; Xu, J.-Y. Current Natural Products with Antihypertensive Activity. *Chin. J. Nat. Med.* **2015**, *13*, 721–729. [CrossRef]
28. Voges, M.J.E.E.E.; Bai, Y.; Schulze-Lefert, P.; Sattely, E.S. Plant-Derived Coumarins Shape the Composition of an Arabidopsis Synthetic Root Microbiome. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 12558–12565. [CrossRef] [PubMed]
29. Harbort, C.J.; Hashimoto, M.; Inoue, H.; Niu, Y.; Guan, R.; Rombolà, A.D.; Kopriva, S.; Voges, M.J.E.E.E.; Sattely, E.S.; Garrido-Oter, R.; et al. Root-Secreted Coumarins and the Microbiota Interact to Improve Iron Nutrition in Arabidopsis. *Cell Host Microbe* **2020**, *28*, 825–837.e6. [CrossRef]
30. Stassen, M.J.J.; Hsu, S.-H.; Pieterse, C.M.J.; Stringlis, I.A. Coumarin Communication along the Microbiome-Root-Shoot Axis. *Trends Plant Sci.* **2021**, *26*, 169–183. [CrossRef]
31. Stringlis, I.A.; de Jonge, R.; Pieterse, C.M.J. The Age of Coumarins in Plant-Microbe Interactions. *Plant Cell Physiol.* **2019**, *60*, 1405–1419. [CrossRef]
32. Expert, D. WITHHOLDING AND EXCHANGING IRON: Interactions between Erwinia Spp. and Their Plant Hosts. *Annu. Rev. Phytopathol.* **1999**, *37*, 307–334. [CrossRef]
33. Motyka, A.; Zoledowska, S.; Sledz, W.; Lojkowska, E. Molecular Methods as Tools to Control Plant Diseases Caused by Dickeya and Pectobacterium Spp: A Minireview. *New Biotechnol.* **2017**, *39 Pt B*, 181–189. [CrossRef]
34. Potrykus, M.; Golanowska, M.; Sledz, W.; Zoledowska, S.; Motyka, A.; Kolodziejska, A.; Butrymowicz, J.; Lojkowska, E. Biodiversity of *Dickeya* Spp. Isolated from Potato Plants and Water Sources in Temperate Climate. *Plant Dis.* **2016**, *100*, 408–417. [CrossRef]
35. Toth, I.K.; van der Wolf, J.M.; Saddler, G.; Lojkowska, E.; Hélias, V.; Pirhonen, M.; Tsror (Lahkim), L.; Elphinstone, J.G. Dickeya Species: An Emerging Problem for Potato Production in Europe. *Plant Pathol.* **2011**, *60*, 385–399. [CrossRef]
36. Dellagi, A.; Rigault, M.; Segond, D.; Roux, C.; Kraepiel, Y.; Cellier, F.; Briat, J.-F.; Gaymard, F.; Expert, D. Siderophore-Mediated Upregulation of Arabidopsis Ferritin Expression in Response to Erwinia Chrysanthemi Infection. *Plant J.* **2005**, *43*, 262–272. [CrossRef]
37. Kieu, N.P.; Aznar, A.; Segond, D.; Rigault, M.; Simond-Côte, E.; Kunz, C.; Soulie, M.-C.; Expert, D.; Dellagi, A. Iron Deficiency Affects Plant Defence Responses and Confers Resistance to Dickeya Dadantii and Botrytis Cinerea. *Mol. Plant Pathol.* **2012**, *13*, 816–827. [CrossRef]
38. Segond, D.; Dellagi, A.; Lanquar, V.; Rigault, M.; Patrit, O.; Thomine, S.; Expert, D. NRAMP Genes Function in Arabidopsis Thaliana Resistance to Erwinia Chrysanthemi Infection. *Plant J.* **2009**, *58*, 195–207. [CrossRef]
39. Golanowska, M.; Potrykus, M.; Motyka-Pomagruk, A.; Kabza, M.; Bacci, G.; Galardini, M.; Bazzicalupo, M.; Makalowska, I.; Smalla, K.; Mengoni, A.; et al. Comparison of Highly and Weakly Virulent Dickeya Solani Strains, With a View on the Pangenome and Panregulon of This Species. *Front. Microbiol.* **2018**, *9*. [CrossRef]
40. Potrykus, M.; Golanowska, M.; Hugouvieux-Cotte-Pattat, N.; Lojkowska, E. Regulators Involved in Dickeya Solani Virulence, Genetic Conservation, and Functional Variability. *Mol. Plant Microbe Interact.* **2014**, *27*, 700–711. [CrossRef]
41. Sławiak, M.; Łojkowska, E.; Wolf, J.M.V.D. First Report of Bacterial Soft Rot on Potato Caused by Dickeya Sp. (Syn. Erwinia Chrysanthemi) in Poland. *Plant Pathol.* **2009**, *58*, 794. [CrossRef]
42. Ihnatowicz, A.; Siwinska, J.; Meharg, A.A.; Carey, M.; Koornneef, M.; Reymond, M. Conserved Histidine of Metal Transporter AtNRAMP1 Is Crucial for Optimal Plant Growth under Manganese Deficiency at Chilling Temperatures. *New Phytol.* **2014**, *202*, 1173–1183. [CrossRef]
43. Chaparro, J.M.; Badri, D.V.; Bakker, M.G.; Sugiyama, A.; Manter, D.K.; Vivanco, J.M. Root Exudation of Phytochemicals in Arabidopsis Follows Specific Patterns That Are Developmentally Programmed and Correlate with Soil Microbial Functions. *PLoS ONE* **2013**, *8*, e55731. [CrossRef]

44. Czechowski, T.; Stitt, M.; Altmann, T.; Udvardi, M.K.; Scheible, W.-R. Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis. *Plant Physiol.* **2005**, *139*, 5–17. [CrossRef]
45. Hassan, J.A.; de la Torre-Roche, R.; White, J.C.; Lewis, J.D. Soil Mixture Composition Alters Arabidopsis Susceptibility to Pseudomonas Syringae Infection. *Plant Direct* **2018**, *2*, e00044. [CrossRef]
46. Carvalhais, L.C.; Dennis, P.G.; Fan, B.; Fedoseyenko, D.; Kierul, K.; Becker, A.; von Wiren, N.; Borriss, R. Linking Plant Nutritional Status to Plant-Microbe Interactions. *PLoS ONE* **2013**, *8*, e68555. [CrossRef] [PubMed]
47. Lemanceau, P.; Expert, D.; Gaymard, F.; Bakker, P.A.H.M.; Briat, J.-F. Chapter 12 Role of Iron in Plant–Microbe Interactions. In *Advances in Botanical Research*; Academic Press: Cambridge, MA, USA, 2009; Volume 51, pp. 491–549. [CrossRef]
48. Marschner, P. Plant-Microbe Interactions in the Rhizosphere and Nutrient Cycling. In *Nutrient Cycling in Terrestrial Ecosystems. Soil Biology*; Marschner, P., Rengel, Z., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; Volume 10, pp. 159–182. [CrossRef]
49. Masclaux, C.; Expert, D. Signalling Potential of Iron in Plant—Microbe Interactions: The Pathogenic Switch of Iron Transport in Erwinia Chrysanthemi. *Plant J.* **1995**, *7*, 121–128. [CrossRef]
50. Harris, J.M.; Balint-Kurti, P.; Bede, J.C.; Day, B.; Gold, S.; Goss, E.M.; Grenville-Briggs, L.J.; Jones, K.M.; Wang, A.; Wang, Y.; et al. What Are the Top 10 Unanswered Questions in Molecular Plant-Microbe Interactions? *MPMI* **2020**, *33*, 1354–1365. [CrossRef]
51. Motyka-Pomagruk, A.; Zoledowska, S.; Sledz, W.; Lojkowska, E. The Occurrence of Bacteria from Different Species of Pectobacteriaceae on Seed Potato Plantations in Poland. *Eur. J. Plant Pathol.* **2021**, *159*, 309–325. [CrossRef]
52. Van der Wolf, J.M.; Nijhuis, E.H.; Kowalewska, M.J.; Saddler, G.S.; Parkinson, N.; Elphinstone, J.G.; Pritchard, L.; Toth, I.K.; Lojkowska, E.; Potrykus, M.; et al. Dickeya Solani Sp. Nov., a Pectinolytic Plant-Pathogenic Bacterium Isolated from Potato(Solanum Tuberosum). *Int. J. Syst. Evol. Microbiol.* **2014**, *64 Pt 3*, 768–774. [CrossRef]
53. Degefu, Y.; Potrykus, M.; Golanowska, M.; Virtanen, E.; Lojkowska, E. A New Clade of *Dickeya* Spp. Plays a Major Role in Potato Blackleg Outbreaks in North Finland. *Ann. Appl. Biol.* **2013**, *162*, 231–241. [CrossRef]
54. Golanowska, M.; Łojkowska, E. A Review on *Dickeya Solani*, a New Pathogenic Bacterium Causing Loss in Potato Yield in Europe. *BioTechnologia* **2016**, *97*, 109–127. [CrossRef]
55. Hamamouch, N.; Li, C.; Seo, P.J.; Park, C.-M.; Davis, E.L. Expression of Arabidopsis Pathogenesis-Related Genes during Nematode Infection. *Mol. Plant Pathol.* **2011**, *12*, 355–364. [CrossRef]
56. Van Loon, L.C.; Bakker, P.A.; Pieterse, C.M. Systemic Resistance Induced by Rhizosphere Bacteria. *Annu. Rev. Phytopathol.* **1998**, *36*, 453–483. [CrossRef]
57. Kliebenstein, D.J.; Monde, R.A.; Last, R.L. Superoxide Dismutase in Arabidopsis: An Eclectic Enzyme Family with Disparate Regulation and Protein Localization. *Plant Physiol.* **1998**, *118*, 637–650. [CrossRef]
58. Liu, F.; Jiang, H.; Ye, S.; Chen, W.-P.; Liang, W.; Xu, Y.; Sun, B.; Sun, J.; Wang, Q.; Cohen, J.D.; et al. The Arabidopsis P450 Protein CYP82C2 Modulates Jasmonate-Induced Root Growth Inhibition, Defense Gene Expression and Indole Glucosinolate Biosynthesis. *Cell Res.* **2010**, *20*, 539–552. [CrossRef]
59. Kobayashi, T.; Nishizawa, N.K. Iron Uptake, Translocation, and Regulation in Higher Plants. *Annu. Rev. Plant Biol.* **2012**, *63*, 131–152. [CrossRef]
60. Schmidt, H.; Günther, C.; Weber, M.; Spörlein, C.; Loscher, S.; Böttcher, C.; Schobert, R.; Clemens, S. Metabolome Analysis of Arabidopsis Thaliana Roots Identifies a Key Metabolic Pathway for Iron Acquisition. *PLoS ONE* **2014**, *9*, e102444. [CrossRef]
61. Tsai, H.-H.; Schmidt, W. One Way. Or Another? Iron Uptake in Plants. *New Phytol.* **2017**, *214*, 500–505. [CrossRef]
62. Zamioudis, C.; Hanson, J.; Pieterse, C.M.J. β-Glucosidase BGLU42 Is a MYB72-Dependent Key Regulator of Rhizobacteria-Induced Systemic Resistance and Modulates Iron Deficiency Responses in Arabidopsis Roots. *New Phytol.* **2014**, *204*, 368–379. [CrossRef]
63. Micallef, S.A.; Shiaris, M.P.; Colón-Carmona, A. Influence of Arabidopsis Thaliana Accessions on Rhizobacterial Communities and Natural Variation in Root Exudates. *J. Exp. Bot.* **2009**, *60*, 1729–1742. [CrossRef]
64. Berendsen, R.L.; Pieterse, C.M.J.; Bakker, P.A.H.M. The Rhizosphere Microbiome and Plant Health. *Trends Plant Sci.* **2012**, *17*, 478–486. [CrossRef]
65. Perkowska, I.; Siwinska, J.; Olry, A.; Grosjean, J.; Hehn, A.; Bourgaud, F.; Lojkowska, E.; Ihnatowicz, A. Identification and Quantification of Coumarins by UHPLC-MS in Arabidopsis Thaliana Natural Populations. *Molecules* **2021**, *26*, 1804. [CrossRef]
66. Rigault, M.; Buellet, A.; Masclaux-Daubresse, C.; Fagard, M.; Chardon, F.; Dellagi, A. Quantitative Methods to Assess Differential Susceptibility of Arabidopsis Thaliana Natural Accessions to Dickeya Dadantii. *Front. Plant Sci.* **2017**, *8*, 394. [CrossRef]
67. Austin, S.; Lojkowska, E.; Ehlenfeldt, K.; Kelman, A.; Helgeson, J.P. Fertile Interspecific Somatic Hybrids of Solanum: A Novel Source of Resistance to Erwinia Soft Rot. *Phytopathology* **1988**, *78*, 1216–1220. [CrossRef]
68. Pajerowska, K.M.; Parker, J.E.; Gebhardt, C. Potato Homologs of Arabidopsis Thaliana Genes Functional in Defense Signaling–Identification, Genetic Mapping, and Molecular Cloning. *Mol. Plant Microbe Interact.* **2005**, *18*, 1107–1119. [CrossRef]
69. Zimnoch-Guzowska, E.; Marczewski, W.; Lebecka, R.; Flis, B.; Schäfer-Pregl, R.; Salamini, F.; Gebhardt, C. QTL Analysis of New Sources of Resistance to Erwinia Carotovora Ssp. Atroseptica in Potato Done by AFLP, RFLP, and Resistance-Gene-Like Markers. *Crop Sci.* **2000**, *40*, 1156–1167. [CrossRef]
70. Lennox, E.S.; Luria, S.E.; Benzer, S. On the Mechanism of Photoreactivation of Ultraviolet-Inactivated Bacteriophage. *Biochim. Biophys. Acta* **1954**, *15*, 471–474. [CrossRef]

71. Heeg, C.; Kruse, C.; Jost, R.; Gutensohn, M.; Ruppert, T.; Wirtz, M.; Hell, R. Analysis of the Arabidopsis O-Acetylserine(Thiol)Lyase Gene Family Demonstrates Compartment-Specific Differences in the Regulation of Cysteine Synthesis. *Plant Cell* **2008**, *20*, 168–185. [CrossRef]
72. Conn, S.J.; Hocking, B.; Dayod, M.; Xu, B.; Athman, A.; Henderson, S.; Aukett, L.; Conn, V.; Shearer, M.K.; Fuentes, S.; et al. Protocol: Optimising Hydroponic Growth Systems for Nutritional and Physiological Analysis of Arabidopsis Thaliana and Other Plants. *Plant Methods* **2013**, *9*, 4. [CrossRef]
73. Schwyn, B.; Neilands, J.B. Universal Chemical Assay for the Detection and Determination of Siderophores. *Anal. Biochem.* **1987**, *160*, 47–56. [CrossRef]

Publication No. 3

AIM

1) **Comprehensive analysis of the correlation between accumulation of coumarins, plant genotypes (natural accessions and mutants), and environmental stress factors (including Fe availability and pathogen defense) to gain deeper insights into the metabolic adaptations and stress response in *Arabidopsis thaliana.***

Ihnatowicz Anna, Siwinska Joanna, **Perkowska Izabela,** Grosjean Jeremy, Hehn Alain, Bourgaud Frederic, Lojkowska Ewa and Olry Alexandre

Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of Arabidopsis thaliana. *BMC Plant Biology*, 24(1), 806. (2024)

Due to space limitations, supplementary data (including additional figures and tables) are not included in the printed version of this thesis but are accessible online at https://doi.org/10.1186/s12870-024-05491-w.

Ihnatowicz *et al. BMC Plant Biology* (2024) 24:806
https://doi.org/10.1186/s12870-024-05491-w

BMC Plant Biology

RESEARCH Open Access

Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of *Arabidopsis thaliana*

Anna Ihnatowicz[1*], Joanna Siwinska[1], Izabela Perkowska[1], Jeremy Grosjean[2], Alain Hehn[2], Frederic Bourgaud[2], Ewa Lojkowska[1] and Alexandre Olry[2*]

Abstract

Background Scopoletin and umbelliferone belong to coumarins, which are plant specialized metabolites with potent and wide biological activities, the accumulation of which is induced by various environmental stresses. Coumarins have been detected in various plant species, including medicinal plants and the model organism *Arabidopsis thaliana*. In recent years, key role of coumarins in maintaining iron (Fe) homeostasis in plants has been demonstrated, as well as their significant impact on the rhizosphere microbiome through exudates secreted into the soil environment. Several mechanisms underlying these processes require clarification. Previously, we demonstrated that Arabidopsis is an excellent model for studying genetic variation and molecular basis of coumarin accumulation in plants.

Results Here, through targeted metabolic profiling and gene expression analysis, the gene-metabolite network of scopoletin and umbelliferone accumulation was examined in more detail in selected Arabidopsis accessions (Col-0, Est-1, Tsu-1) undergoing different culture conditions and characterized by variation in coumarin content. The highest accumulation of coumarins was detected in roots grown in vitro liquid culture. The expression of 10 phenylpropanoid genes (*4CL1*, *4CL2*, *4CL3*, *CCoAOMT1*, *C3'H*, *HCT*, *F6'H1*, *F6'H2*, *CCR1* and *CCR2*) was assessed by qPCR in three genetic backgrounds, cultured in vitro and in soil, and in two types of tissues (leaves and roots). We not only detected the expected variability in gene expression and coumarin accumulation among Arabidopsis accessions, but also found interesting polymorphisms in the coding sequences of the selected genes through in silico analysis and resequencing.

Conclusions To the best of our knowledge, this is the first study comparing accumulation of simple coumarins and expression of phenylpropanoid-related genes in Arabidopsis accessions grown in soil and in liquid cultures. The large variations we detected in the content of coumarins and gene expression are genetically determined, but also

*Correspondence:
Anna Ihnatowicz
anna.ihnatowicz@biotech.ug.edu.pl
Alexandre Olry
alexandre.olry@univ-lorraine.fr

Full list of author information is available at the end of the article

tissue and culture dependent. It is particularly important considering that growing plants in liquid media is a widely used technology that provides a large amount of root tissue suitable for metabolomics. Research on differential accumulation of coumarins and related gene expression will be useful in future studies aimed at better understanding the physiological role of coumarins in roots and the surrounding environments.

Keywords Accessions, Coumarins, Genetic variation, Gene expression, Metabolic profiling, In vitro liquid culture, Scopolin, Skimmin, Soil environment

Background

Plants produce a wide range of specialized metabolites, among which phenylpropanoids constitute a large class. Their biosynthesis is a very complex and branched pathway, and their functions are, so far, not fully understood. The most studied phenylpropanoids are lignins, flavonoids, anthocyanins, chalcones and coumarins. Coumarins are widely produced by plants and were described as phytoalexins [1]. Scopoletin (7-hydroxy-6-methoxycoumarin), also reported as hydroxycoumarin is synthesized, among others, by diverse medicinal plants [2], cassava [3], sweet potato [4], sunflower [5], cotton [6], and the model plants tobacco (*Nicotiana tabacum*) and *Arabidopsis thaliana* [7]. In Arabidopsis, scopoletin and its glycosylated form scopolin (β-D-glucoside scopoletin) were firstly reported by Rohde et al. [8] and Bednarek et al. [9]. One year later, Kai et al. [10] showed additionally the presence of trace amounts of skimmin (glycosylated form of 7-hydroxycoumarin also known as umbelliferone) in the roots, as well as esculetin (6,7-dihydroxycoumarin) both in roots and shoots. Scopoletin and its glycoside, scopolin, are the major coumarins accumulating in Arabidopsis roots [10–16]. Our research team has demonstrated as the first one that Arabidopsis is an excellent model for studying the genetic basis of natural variation in coumarin biosynthesis by conducting a quantitative trait locus (QTL) mapping followed by identification of new potential candidate loci [13]. We then extended this work and showed that scopoletin was the most abundant coumarin compound in the roots of each of the 28 Arabidopsis accessions tested, but interestingly it was also detected in leaf extracts [16]. The latter is supported by the results of Robe et al. [17], which confirmed that scopoletin is synthesized in the roots, but can move throughout the plant body within the xylem sap and accumulate in the shoots. In the same study [16], we demonstrated for the first time in Arabidopsis the presence of small amounts of umbelliferone in hydrolyzed extracts prepared from the roots of all 28 tested accessions. Umbelliferone is an important intermediate for the biosynthesis of more complex coumarins – furanocoumarins and pyranocoumarins [18], which are of great importance in the pharmaceutical industry. Even if Arabidopsis does not produce furanocoumarins, our discovery of umbelliferone accumulation in Arabidopsis is a significant step in the study of coumarin biosynthesis using this model plant. However, we still do not know how umbelliferone is synthesized in Arabidopsis and it cannot be excluded to be only an intermediate in the synthesis of skimmin.

Recently, our research team and several other groups, demonstrated that coumarins, namely fraxetin, sideretin and scopoletin, play a crucial role in Fe chelation in Arabidopsis and secretion of coumarins by Arabidopsis roots was shown to be induced under Fe-deficiency [12, 14, 19–22]. Accumulation and secretion of coumarinolignans and other coumarins was also shown to be induced in Arabidopsis roots in response to Fe-deficiency at high pH [23]. Moreover, the excretion of an Fe-mobilizing scopoletin, which is regulated by the root-specific transcription factor MYB72, was revealed to have selective antimicrobial activity that shapes the root-associated microbial community [24]. Various studies have documented the antibacterial and antifungal effects of scopoletin and its derivatives. Scopoletin was described to be involved in the plant immune response in defense reactions to pathogens [25, 26] such as *Fusarium oxysporum*, *Fusarium solani*, *Rhizopus stolonifer*, and *Lasiodiplodia theobromae* [27], tobacco mosaic virus [28]. Scopoletin displays a higher growth inhibition effect on *F. oxysporum* than its β-D- glucoside scopolin [27] suggesting that the aglycon might be responsible for the defense reactions. We also recently investigated the mechanisms underlying the interplay between coumarin accumulation, Fe status, and plant pathogen resistance using the Arabidopsis/*Dickeya* spp. pathosystem. We observed that the response of different Arabidopsis lines (mutants defective in coumarin biosynthesis and transport) was dependent on the *Dickeya* species used and the genotype of plants grown in a Fe-deficient hydroponic culture [29].

Although coumarins are well known for their potent antibacterial and antifungal properties, they have recently received much attention as important factors influencing a number of processes that determine the interaction of plants with the soil environment, both biotic and abiotic factors [14, 15, 24, 28, 30–32]. Scopoletin was discovered as a new signal in the pre-penetration dialogue in plant-mycorrhizal associations that possibly have implications for chemical communication [33]. As shown by Cosme et al. [33], the coumarin scopoletin particularly stimulates pre-penetration development and metabolism in mycorrhizal fungi. The production of both scopoletin and fraxetin [described by 14, 21, 22] impact the root

microbiota as shown by Harbort et al. [31]. Their biosynthesis and secretion through PDR9 (plasma membrane-bound transporter described by [34] are determining root microbiota composition in a naturally Fe-limiting calcareous soil [31]. The important role of coumarins in communication on the microbiome-root-shoot axis, alongside strigolactones and flavonoids [32], is currently vigorously discussed.

Here we focused on selected simple coumarins: (1) scopoletin together with its glycoside scopolin, which are the main coumarins of Arabidopsis, (2) and umbelliferone that was recently detected for the first time by our group in this model plant [16] with its glycoside, skimmin. It should be remembered that other coumarin compounds like fraxetin, sideretin [21, 35] and esculetin are accumulated in Arabidopsis, whose biosynthesis and functions in plants also require further elucidation. The biosynthetic pathway leading to esculetin, which is postulated to be a strong Fe chelator due to the catechol functional group [20], is largely unclear in plants and remains a mystery in Arabidopsis. In this work, in addition to targeted metabolic profiling, we focused on the analysis of the expression of genes directly involved in scopoletin biosynthesis, but also of a number of genes encoding enzymes of the phenylpropanoid pathway located upstream and downstream to the biosynthesis of coumarins (Fig. 1). We detected significant variations in the content of coumarins and gene expression levels that were not only genetically determined but also tissue and culture dependent. The latter is particularly important considering that these are two different growing conditions, widely used by other authors and to the best of our knowledge, with no comparison between them in the context of coumarin accumulation until this article. Thus, by analyzing the differential expression patterns of selected genes in leaves and roots of three Arabidopsis genetic backgrounds and two contrasting environments, we can better understand the correlation between genetic variants (coding sequences) and phenotypic variation (coumarin content), especially in the context of the unknown biosynthesis of umbelliferone (Fig. 1).

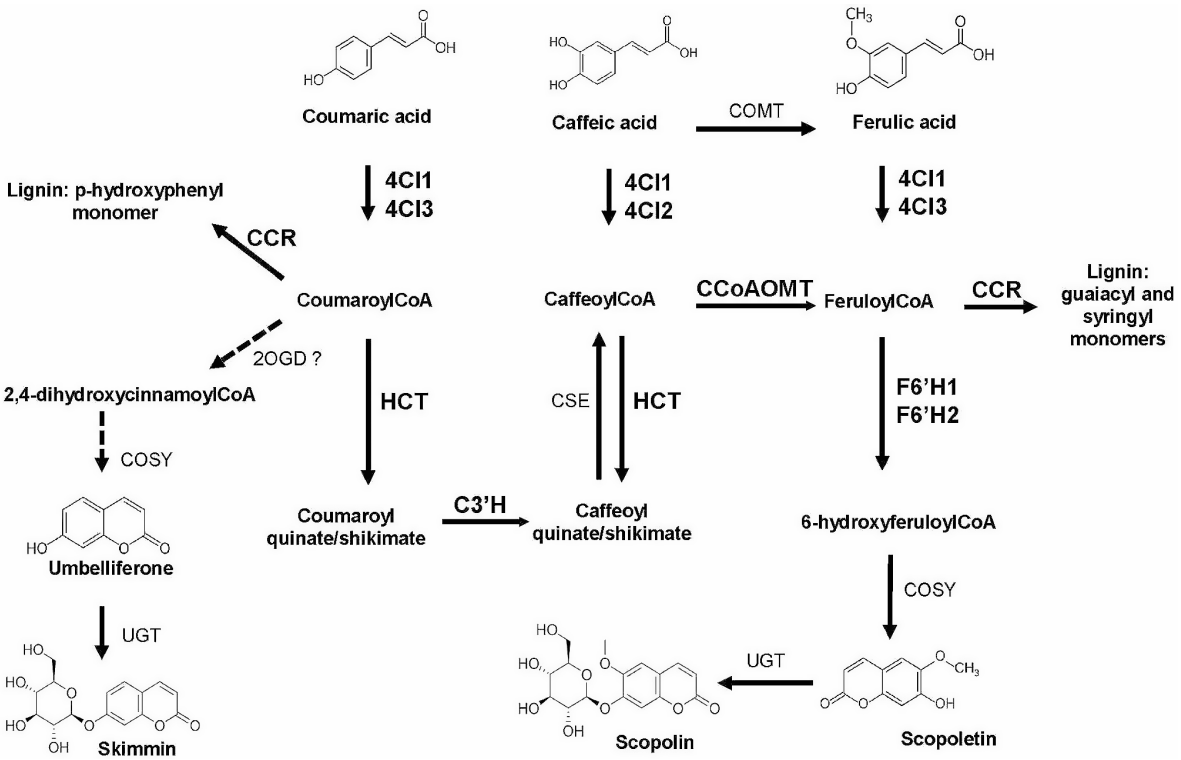


Fig. 1 Simplified schematic representation of the biosynthetic pathway of scopoletin, umbelliferone and their corresponding glycosides, scopolin and skimmin respectively, along with monolignols in *Arabidopsis thaliana*. The step of umbelliferone synthesis was hypothesized to occur via the same pathway as scopoletin synthesis. Genes whose expression was studied are highlighted: **HCT**, Hydroxycinnamoyl CoenzymeA shikimate: quinate hydroxycinnamoyltransferase; **C3'H**, Coumarate p-coumaroyl shikimate 3'-hydroxylase; **F6'H1**, **F6'H2**, Feruloyl-CoA 6'-Hydroxylase1 and 2; **4CL1**, **4CL2**, **4CL3**, 4-Coumarate CoenzymeA ligase 1–3; **CCoAOMT**, Caffeoyl CoenzymeA 3-o-methyltransferase; **CCR1**, **CCR2**, Cinnamoyl CoA reductase 1 and 2. Other enzymes shown in the diagram are: COSY, coumarin synthesis; putative UGT, UDP-glucosyltransferase. Adapted from [36] and [37]

Methods

Growth conditions in vitro culture

Arabidopsis seeds were surface sterilized and sown on 0.5x Murashige and Skoog (MS) medium as described by Siwinska et al. [13]. After stratification at 4 °C for 72 h in the dark, the plates were placed in a phytotron for another 10 days (light intensity 35 µmol m^{-2} s^{-1}, 20 °C day/18 °C night, photoperiod 16/8). Then, 10-day-old seedlings were transferred and grown in glass culture vessels with liquid medium according to Siwinska et al. [13]. On day 28 of cultivation (and after 17 days of growing Arabidopsis in in vitro liquid cultures on rotary platform shakers), the plants were harvested, leaves and roots weighed separately, frozen in liquid nitrogen and stored at -80 °C (in 2 ml microtubes). For all Arabidopsis accessions, three biological replicates were grown (in three independent glass culture vessels, at least three seedlings in each vessel).

Growth conditions in soil

Arabidopsis seeds were first stratified (at 4 °C for 4 days on water-soaked Whatman paper) and then sown into a soil mixture (commercially available Compo Sana soil with vermiculite in a 3:1 ratio). The plants were watered as it was required and once a week with the soil fertilizer (Substral). After 3 weeks, plants were collected (leaves and roots separately), weighed and frozen in 2 ml microtubes in liquid nitrogen and stored at -80 °C until further analysis. All Arabidopsis accessions were grown in three biological replicates (in independent pots). HOBO U12 data logger (Onset Computer Corporation, Bourne, MA) was used for monitoring plant growth conditions.

Preparation of root methanol extracts

The previously harvested plant material was divided into four sets according to the coumarins extraction method. The tissues in the first set (M-H-) were homogenized using steel beads and sonication, after that the homogenate was soaked with 80% methanol supplemented with 0.5 µM 4-methylumbelliferone and kept at 4 °C for 24 h. The methanol extracts were then centrifuged for 20 min at 13,000 x g. The second set (M+H+) was additionally incubated for 5 min in the microwave oven set at 700 W before the homogenization. After centrifugation, extracts were enzymatically hydrolyzed according to modified protocol of Nguyen et al. [38] as described in Siwinska et al. [13]. The third set of extracts was incubated only in a microwave oven (M+H-), while the fourth was only subjected to enzymatic hydrolysis (M-H+).

Chemicals

Coumarin standards umbelliferone (purity≥99%), esculin (glycosylated esculetin, >98% purity) were purchased from Sigma-Aldrich (St. Louis, USA), scopoletin (>95% purity) and esculetin (>98% purity) from Extrasynthese (Genay, France), skimmin (98% purity, glycosylated umbelliferone) from Biopurify Phytochemicals (Chengdu, China), scopolin (>98% purity, glycosylated scopoletin) from Aktin Chemicals Inc. (Chengdu, China). Stock solutions of each standard (at 10 mmol/L concentrations) were made by diluting the powder in dimethyl sulfoxide (Fisher scientific, Illkirch, France), which were subsequently kept at -18 °C. The methanol (HPLC-grade) was purchased from CarloErba Reagents (Val de Reuil, France), formic acid was purchased from Fisher Scientific (Illkirch, France). PURELAB Ultra system (Veolia Water S.T.I., Antony, France) was used for water purification.

Quantification of coumarins by UHPLC-MS targeted metabolite profiling

Targeted metabolite profiling of Arabidopsis methanol extracts prepared from roots and leaves was performed, namely quantification of selected coumarins using ultra-high-performance liquid chromatography combined with mass spectrometry analysis (UHPLC-MS), as described in the work of Perkowska et al. [16].

Quantitative real-time PCR analysis

Total RNA was isolated from Col-0, Est-1 and Tsu-1 Arabidopsis accessions grown in vitro in liquid culture and in soil according to Ihnatowicz et al. [39]. We conducted the quantitative real-time PCR analysis by using LightCycler® 480 Real-Time PCR System (Roche) with the Luminaris™ HiGreen qPCR Master Mix (Thermo Scientific) and primer sequences for genes summarized in Supplementary Table S1 (gene-specific) and for *ACTIN2* in [40]. Confirmation of primer specificity and normalization of relative transcript levels of the studied genes were performed as described in Ihnatowicz et al. [39]; the efficiencies of the PCR product amplification by the qPCR primers are provided in Supplementary Table S2.

Sequencing

PCR reactions were carried out in a 10 µl reaction mixture, which contained cDNA synthetized based on root RNA, 0.5 U of TaKaRa LA Taq® DNA polymerase, 200 µM dNTP, 1 µM primers, and 1 × LA PCR Buffer ll (Mg2+plus). After denaturation at 94 °C for 1 min, the reaction mixture was used in PCR amplification using 34 cycles of 98 °C for 10 s, 60 °C for 20 s, and 68 °C for 60 s in the Thermal Cycler C1000 Touch (Bio-Rad). Gene-specific primers used for gene amplification of CDS are summarized in Supplementary Table S3. PCR products were cloned into pGEM®-T vector. The *Escherichia coli* strain Gene Hogs (F- mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (StrR) endA1 nupG fhuA::IS2) was used for plasmid amplification and maintenance. The vector

specific primers M13pUCf and M13pUCr and BigDye® Terminator v3.1 (Life Technologies) were used for the sequencing of positive clones. The reaction products were sequenced by 3730xl DNA Analyzer, while sequence alignments were performed using CLUSTALW [41]. Multiple sequence alignment (MSA) indicating similar/different nucleotides/amino acids across the aligned sequences were visualized using BioEdit v5.0.9 software.

Statistical analysis

We processed data and conducted statistics (pairwise comparisons with Welch's t-test: two-sample assuming unequal variances) using R programming (https://app.displayr.com) and Microsoft Excel. Three biological replicates were included in all treatments. In the figures, means and error bars for absolute deviations are shown. The data points with significantly different mean values are indicated with asterisk(s), with the significance level of $p<0.001$ (***), $p<0.01$ (**) and $p<0.05$ (*).

Results

Methodology for the analysis of coumarin content in Arabidopsis extracts

The vast majority of coumarins present in plant cells are bound to sugars [42]. To obtain a global overview of the concentration of coumarins, we performed methanol/water (80:20) extraction of various tissues cultured in vitro and in soil under selected conditions. Half of the crude extracts were enzymatically hydrolyzed using β-glucosidase to analyze the total free scopoletin and umbelliferone. The quantification of their respective glycosylated counterparts, scopolin and skimmin, was performed on extracts that had not undergone enzymatic treatment. In parallel, as a control mode, all simple coumarins were quantified using the UHPLC method in the methanol extracts, both those subjected to enzymatic hydrolysis and those that were not hydrolyzed. Various data can be found in the literature regarding the effect of microwave oven on the activity of β-glucosidases naturally occurring in plant tissues and its use as a useful tool in the analysis of plant extracts [43, 44]. We decided to test it by using a microwave oven treatment in two sets of our plant extracts, both those subjected to enzymatic hydrolysis and those without it. Based on the results obtained, it cannot be concluded whether the microwave oven inhibited the β-glucosidase activity. However, it turned out that microwave treatment led to a reduction in the levels of all tested coumarins when compared to the untreated plant extracts, either through the direct energy distribution of the coumarin compounds or through the emitted heat (Figure S1, Figure S2). Additionally, the mean deviations were much higher in some microwave-treated samples for umbelliferone quantification (Figure S2). Therefore, we further quantified coumarins only in extracts that were not subjected to microwave oven treatment.

Quantification of free coumarins by UHPLC

To investigate the natural variability of scopoletin and umbelliferone, we focused on both roots and leaves of 3 Arabidopsis accessions (Col-0, Est-1, and Tsu-1) grown in different culture types (Fig. 2). We grew plants in vitro in liquid cultures - conditions that induce the growth of roots, which are the main tissue accumulating coumarins. Additionally, in this type of cultures, access to the roots is simple, which makes it easier to collect them for further analysis. This method of growing plants also mimics stressful conditions that might induce coumarin

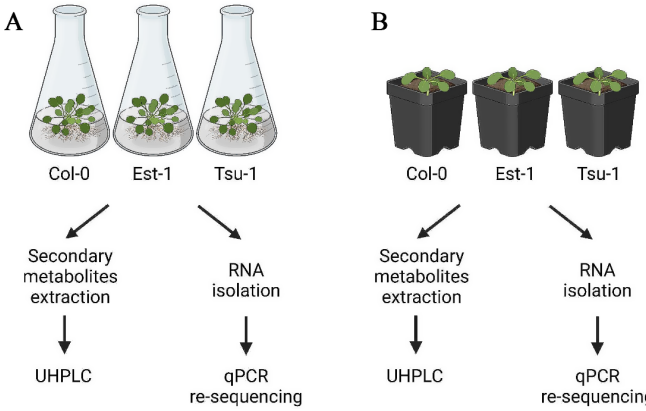


Fig. 2 Experimental scheme showing two types of culture used (**A**) in vitro liquid culture and (**B**) soil condition. Plants were grown in soil (optimal soil mix supplemented with fertilizer once per week) and in in vitro liquid cultures (light intensity 35 µmol m^{-2} s^{-1}, 20 °C day/18 °C night, photoperiod 16/8). Leaves and roots were harvested separately. All samples were divided for secondary metabolites extraction and RNA isolation. Plant extracts were divided in half and subsequent UHPLC analysis were performed in methanol extracts with and without enzymatic treatment. cDNA was synthetized for qPCR analysis. Created with BioRender.com

accumulation. In vitro liquid cultures were performed in light, which was reported to specifically increase phenylpropanoid production in roots [45]. In parallel, we cultivated plants in more physiological conditions, in soil enriched with fertilizer once a week, in accordance with the optimal growth conditions for Arabidopsis described in the literature [39].

Our results clearly showed that coumarin production was significantly higher in roots of Arabidopsis plants grown under stress conditions in in vitro liquid cultures when compared to soil-grown plants (Fig. 3). As expected, free coumarins, namely umbelliferone and scopoletin were mostly accumulated in MeOH roots extracts subjected to enzymatic hydrolysis (panels H+in Fig. 3A and B). These methanol extracts from Col-0 and Est-1 roots grown in vitro contain similar concentrations of scopoletin and umbelliferone, whereas Tsu-1 accumulates visibly lower levels of both molecules (Fig. 3AB). In the roots of all accessions grown in soil we detected significantly lower concentrations of the tested coumarins compared to the plants grown in vitro. The observed differences in coumarin content among plant genotypes grown in soil were minimal (Fig. 3AB).

Quantitative determination of coumarins in MeOH leaf extracts detected relatively high concentration of umbelliferon in the leaves of Tsu-1 genotype (compared to the reference Col-0 line) grown in soil (Fig. 4A). Unexpectedly, umbelliferone was mainly present in the MeOH leaf extracts without enzymatic hydrolysis (panel H- in Fig. 4A). In parallel, we detected relatively high levels of the major free Arabidopsis coumarin, scopoletin, in the MeOH leaf extracts prepared from the tissues of in vitro grown plants after enzymatic hydrolysis, but also in the soil-grown roots (panel H+in Fig. 4B). However, overall, higher levels of scopoletin were accumulated in leaves of plants grown in vitro compared to plants grown on soil (Fig. 4B).

In summary, under all tested conditions, both coumarin compounds were produced in small amounts in plant leaves (Fig. 4) and in larger quantities in roots (Fig. 3), which is consistent with the results of previous studies indicating that coumarins accumulate mainly in underground tissues [9, 16].

Quantification of glycosylated coumarins by UHPLC

As determined in our preliminary experiments, we performed the quantification of the glycosylated compounds in plant extracts without microwave treatment. The majority of glycosylated coumarins were expected to be detected in samples not subjected to enzymatic hydrolysis (H-), but as a control we quantified their levels also in samples treated with β-glucosidase (H+). Scopolin and skimmin were detected mostly in untreated root extracts (H-) in both in vitro and soil-grown plants (in Fig. 5AB), whereas some skimmin quantities were also identified in root samples subjected to enzymatic hydrolysis (H+) (Fig. 5A). Taking into account the scale presented in Fig. 5AB, the concentration of skimmin in the roots was

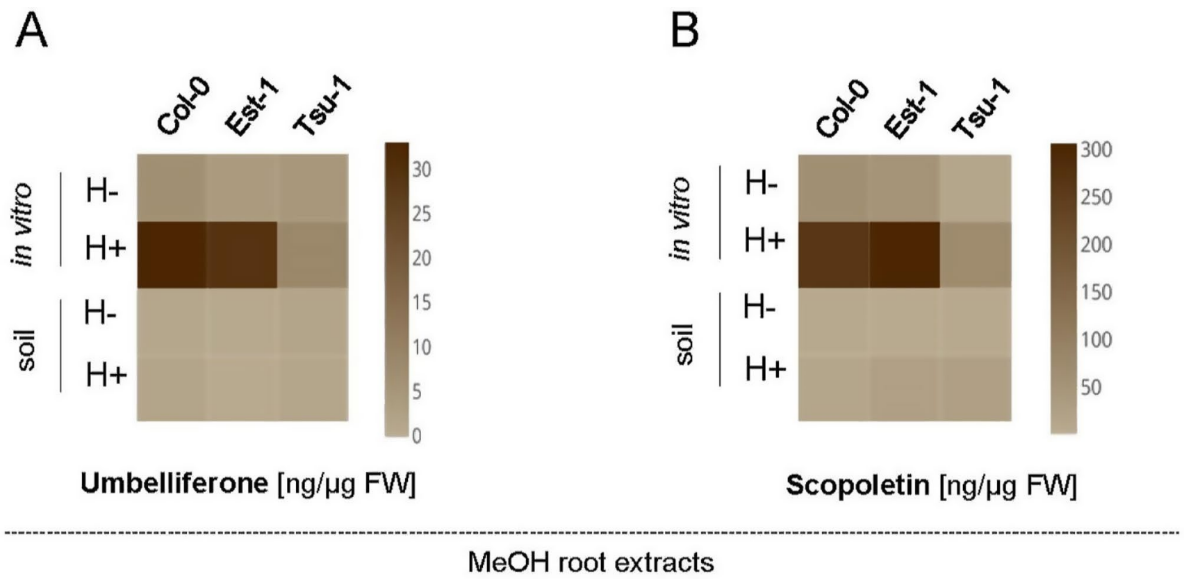


Fig. 3 Heat maps showing the quantification of umbelliferone (**A**) and scopoletin (**B**) in the Arabidopsis root extracts before (H-) and after (H+) enzymatic hydrolysis. Methanol extracts were prepared from Col-0, Est-1 and Tsu-1 Arabidopsis accessions grown in in vitro liquid culture and in soil, quantified by UHPLC. The color scale represents the compound concentration given in ng/µg FW (fresh weight); dark brown indicates high concentration and light brown denotes low concentration

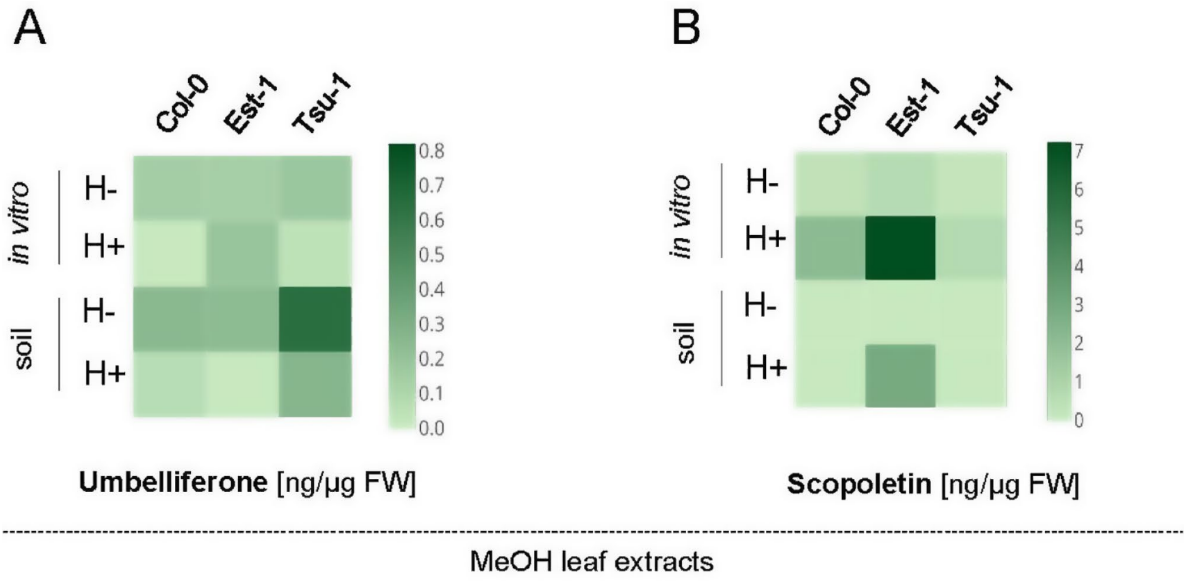


Fig. 4 Heat maps showing the quantification of umbelliferone (**A**) and scopoletin (**B**) in the Arabidopsis leaf extracts before (H-) and after (H+) enzymatic hydrolysis. Methanol extracts were prepared from Col-0, Est-1 and Tsu-1 Arabidopsis accessions grown in in vitro liquid culture and in soil, quantified by UHPLC. The color scale represents the compound concentration given in ng/µg FW (fresh weight); dark brown indicates high concentration and light brown denotes low concentration

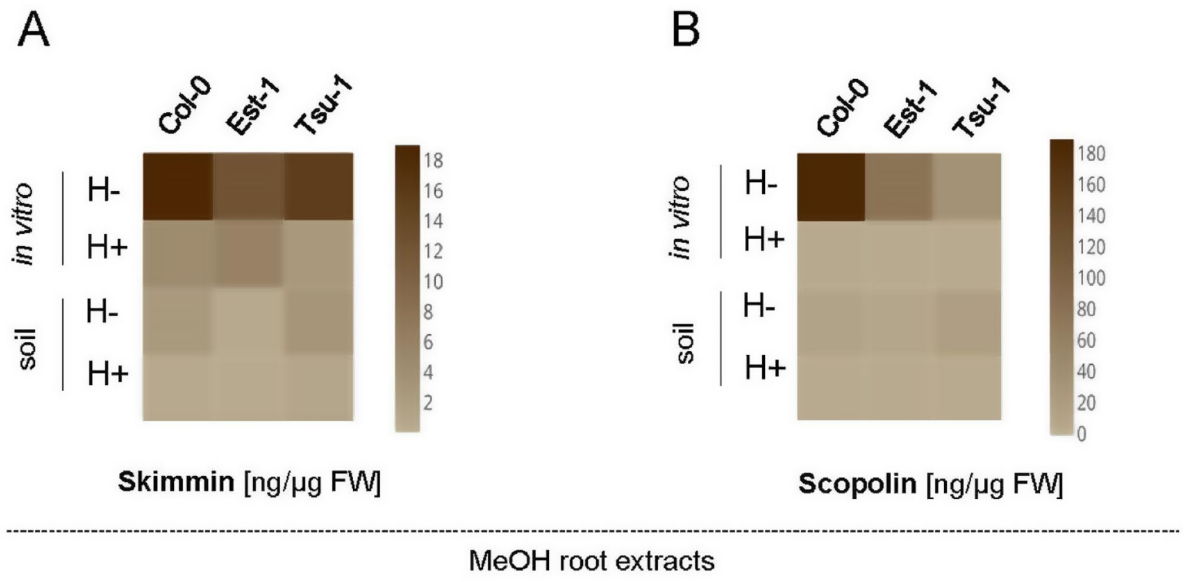


Fig. 5 Heat maps showing the quantification of skimmin (**A**) and scopolin (**B**) in the Arabidopsis root extracts before (H-) and after (H+) enzymatic hydrolysis. Methanol extracts were prepared from Col-0, Est-1 and Tsu-1 Arabidopsis accessions grown in in vitro liquid culture and in soil, quantified by UHPLC. The color scale represents the compound concentration given in ng/µg FW (fresh weight); dark brown indicates high concentration and light brown denotes low concentration

about ten times lower than concentration of scopolin. From a general point of view, it is worth noting that glycosylated coumarins in the roots of plants grown in soil were synthesized in much smaller amounts compared to plants grown in vitro. Regarding the genetic background specificity, scopolin was mostly detected in roots of Col-0 grown in vitro, while under this condition the lowest concentration was detected in Tsu-1 (Fig. 5B).

Concerning the quantification of coumarins in MeOH leaf extracts, the skimmin concentration was similarly

low for all the accessions cultivated in vitro but its content was strongly induced in soil-grown Est-1 and Tsu-1 leaves (both H+and H- in Fig. 6A).However, the main coumarin accumulated in the leaves was invariably scopolin, with the highest concentration detected in Est-1 accession both in vitro and in soil conditions (Fig. 6B).

Quantification of the expression level of genes involved in the synthesis of coumarins

To gain insight into coumarin compounds production at the molecular levels in different Arabidopsis genetic backgrounds, we assessed the expression levels of a number of genes encoding enzymes of the phenylpropanoid pathway located upstream and downstream of coumarin biosynthesis. We conducted real-time quantitative PCR (qPCR) experiments targeting 10 different genes: *4CL1*, *4CL2*, *4CL3*, *CCoAOMT1*, *C3'H*, *HCT*, *F6'H1*, *F6'H2*, *CCR1* and *CCR2* in three Arabidopsis accessions, grown in two conditions (liquid in vitro cultures and soil), and on both kind of tissues (leaves and roots). To be able to make robust comparisons between all samples, the expression was normalized based on the expression of *ACTIN2 (ACT2)* considered as a housekeeping gene [46].

Coumaroyl-CoA ligases (4CL) were the first enzymes we focused on. These enzymes are directly related to the synthesis of coumarins. Arabidopsis has four homologous genes encoding 4CLs. Here, we focused of three of them, since *4CL4* has a limited expression profile [37]. The transcription patterns of *4CL1* differ significantly between root samples isolated from soil and in vitro-grown plants for all three accessions (Fig. 7A). However, we did not observe significant differences among genotypes under the same growth conditions. The results were different when we investigated the *4CL1* expression profiles in leaf tissues. The transcription level of *4CL1* was much lower in Tsu-1 accession when compared to Col-0 and Est-1 cultured both in vitro and in soil. The opposite could be observed for the expression profile of *4CL2*. This gene is significantly more expressed in plants grown in vitro compared to soil conditions, in both tissue types - roots and leaves (Fig. 7B). The expression of *4CL3* drops down and is generally much lower in comparison to *4CL1* and *4CL2*, but interestingly it is the most variable between growth condition and among genotypes (Fig. 7C). Moreover, opposite trends in the relative induction of *4CL3* expression between tissues in both types of culture are visible. In leaves it is higher in plants grown in soil compared to in vitro, while in the roots it is the other way around - it is higher in vitro than in plants grown in soil (Fig. 7C).

The second gene family we focused on, are the oxoglutarate dependent dioxygenases responsible of the hydroxylation of feruloyl CoA. These enzymes are specific of the biosynthetic pathway of coumarins. As expected, our results confirmed that *F6'H1* and *F6'H2* are particularly more strongly expressed in root tissue, and the *F6'H1* gene is expressed at high level (Fig. 8). When grown in soil, the expression level of *F6'H1* in roots is comparable

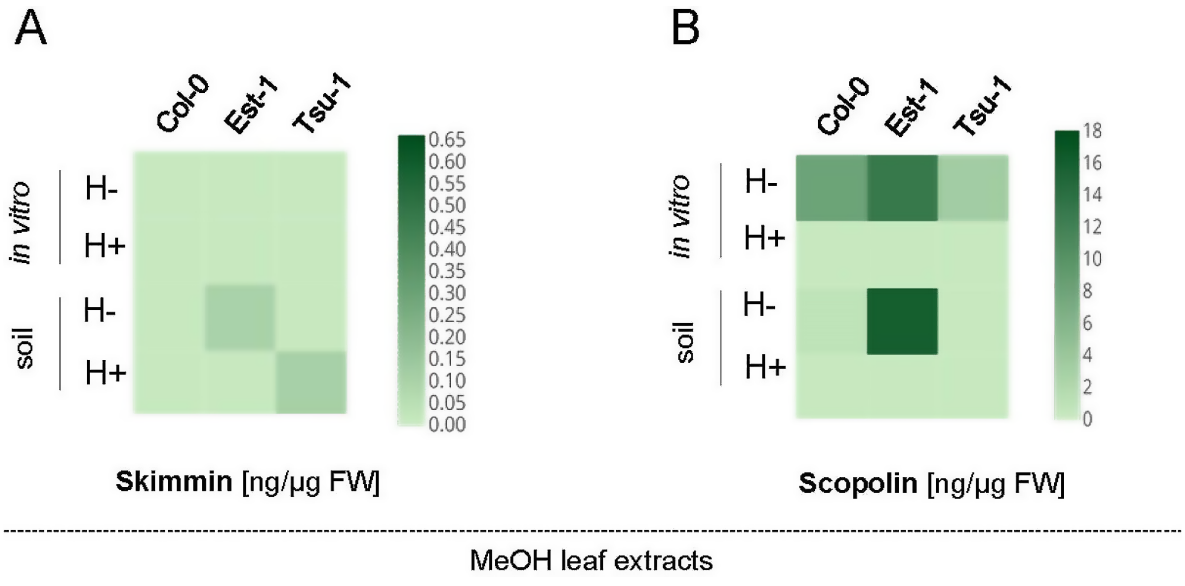


Fig. 6 Heat maps showing the quantification of skimmin (**A**) and scopolin (**B**) in the Arabidopsis leaf extracts before (H-) and after (H+) enzymatic hydrolysis. Methanol extracts were prepared from Col-0, Est-1 and Tsu-1 Arabidopsis accessions grown in in vitro liquid culture and in soil, quantified by UHPLC. The color scale represents the compound concentration given in ng/µg FW (fresh weight); dark brown indicates high concentration and light brown denotes low concentration

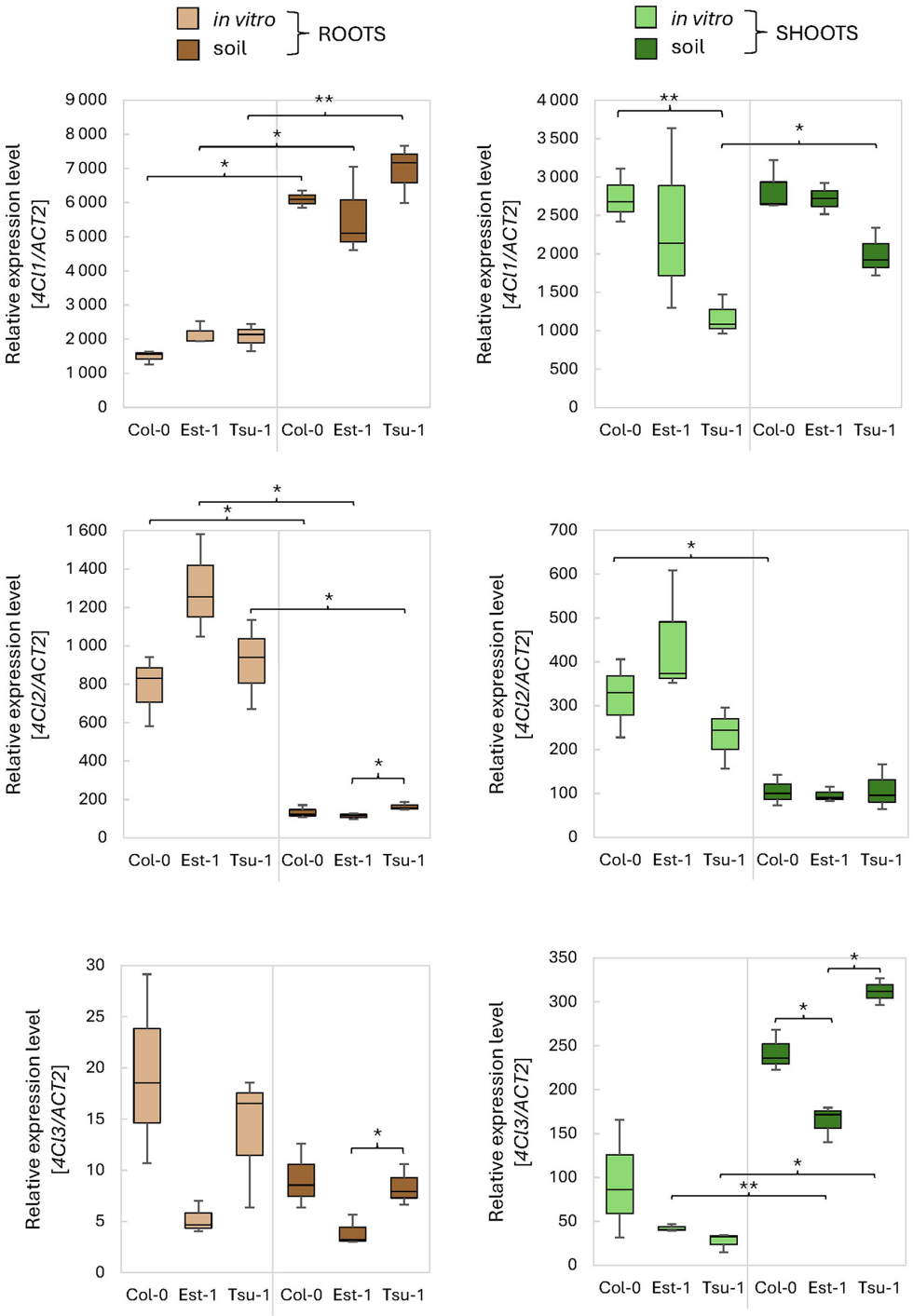


Fig. 7 Relative expression levels of the *4CL1-3* genes measured by qPCR. As a reference, the *ACT2* (At3g18780) gene was used. The expression levels were quantified in three Arabidopsis genetic backgrounds (Col-0, Est-1, Tsu-1) grown in in vitro liquid culture and in soil, in two types of tissues (leaves and roots). Means and error bars for absolute deviations are shown. Values: $p < 0.001$ (***), $p < 0.01$ (**) and $p < 0.05$ (*)

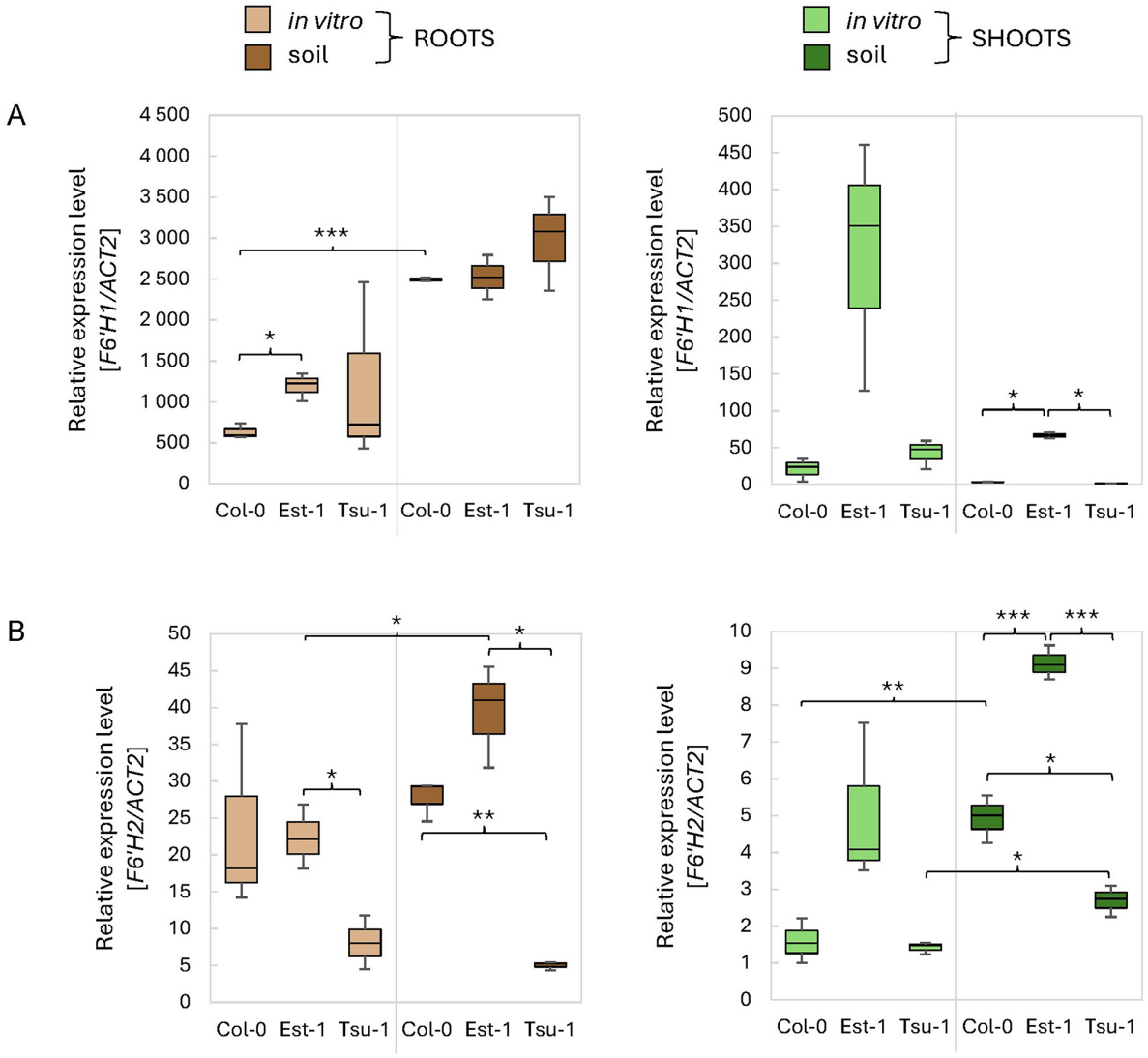


Fig. 8 Relative expression levels of (**A**) *F6'H1* and (**B**) *F6'H2* genes involved in the last step of the biosynthesis of the main coumarin (scopoletin) in Arabidopsis measured by qPCR. As a reference, the *ACT2* (At3g18780) gene was used. The expression levels were quantified in three Arabidopsis genetic backgrounds (Col-0, Est-1, Tsu-1) grown in in vitro liquid culture and in soil, in two types of tissues (leaves and roots). Means and error bars for absolute deviations are shown. Values: $p < 0.001$ (***), $p < 0.01$ (**) and $p < 0.05$ (*)

in all accessions (Fig. 8A), while in leaves it is significantly higher in Est-1 ($p<0.01$) than in Col-0 or Tsu-1. For plants grown in vitro, the expression of *F6'H1* was more variable. We observed significant differences between different tissues and accessions, but the *F6'H1* transcript was significantly induced in leaves of Est-1 when compared with other genetic backgrounds (Fig. 8A). The expression of *F6'H2* is highly significantly variable among accessions grown in soil. In general, its transcription is more efficient in soil conditions in both leaves and roots. Interestingly, the expression of *F6'H2* is the highest in Est-1 genetic background in all conditions and tissues tested (Fig. 8B). The relatively low expression level of *F6'H2* compared to *F6'H1* indicates its smaller contribution in the biosynthesis of scopoletin.

Finally, we assessed the expression level of several downstream genes including *CCoAOMT1*, *HCT*, *C3'H*, and *CCRs*. These enzymes are involved in the phenylpropanoid biosynthetic pathway but after the branching point leading the production of coumarins. Our results showed that *CCoAOMT1* and *HCT* have very similar expression patterns in roots, but *CCoAOMT1* is expressed at much higher level (Fig. 9AB). In soil-grown

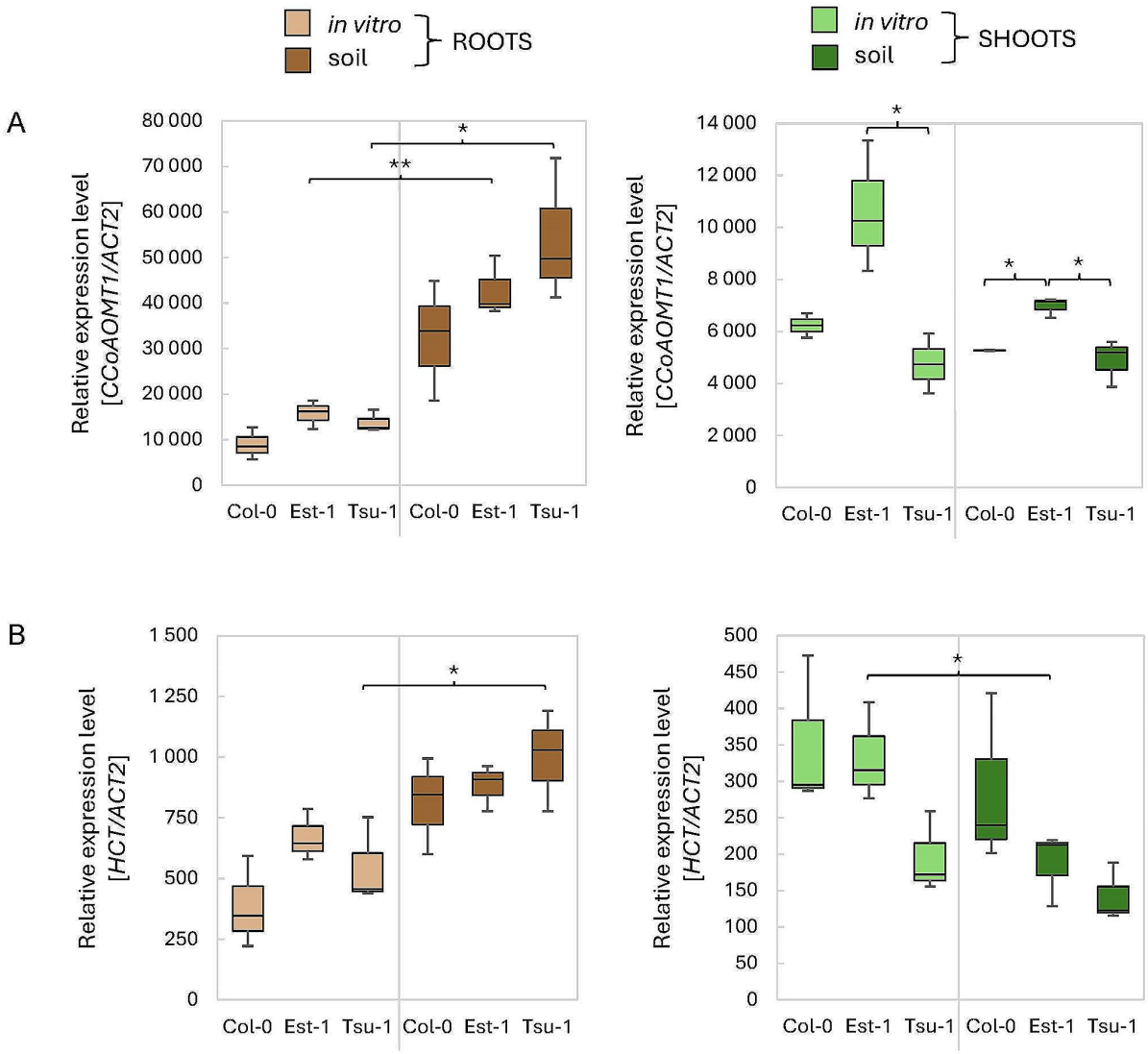


Fig. 9 Relative expression levels of the (**A**) *CCoAOMT1* and (**B**) *HCT* genes measured by qPCR. As a reference, the *ACT2* (At3g18780) gene was used. The expression levels were quantified in three Arabidopsis genetic backgrounds (Col-0, Est-1, Tsu-1) grown in in vitro liquid culture and in soil, in two types of tissues (leaves and roots). Means and error bars for absolute deviations are shown. Values: $p < 0.001$ (***), $p < 0.01$ (**) and $p < 0.05$ (*)

roots, both transcripts have significantly higher values in Tsu-1 compared to other accession ($p < 0.05$).

C3'H is specifically involved in the transformation of *p*-coumaroyl CoA into caffeoyl CoA (a precursor of feruloyl CoA). The expression level of *C3'H* varies significantly between the both conditions and tissues (Fig. 10A). Interestingly, transcription of this gene becomes significantly high in Est-1 leaves, especially under in vitro conditions ($p < 0.05$). Col-0 is relatively low accumulator of *C3'H* transcript in roots. A next set of genes, *CCRs*, is directly related to lignin synthesis (Fig. 10B). *CCR1* is more highly expressed compared to *CCR2*, especially in leaf tissue. The expression patterns of both *CCR* genes is very similar in in vitro cultured roots, with significantly variable levels of *CC1* and *CC2* transcript accumulation among accessions. However, in roots grown in soil, these differences are small. *CCR1* transcription is higher in roots than in leaves of Est-1 and Tsu-1 accessions grown in vitro, while in Col-0 the ratio is reversed.

Natural genomic variation

In order to further investigate the observed variations between Arabidopsis accessions and to better understand the possible correlation between the presence of coumarins/glycosylated coumarins and the expression level of various metabolic genes, we cloned and sequenced

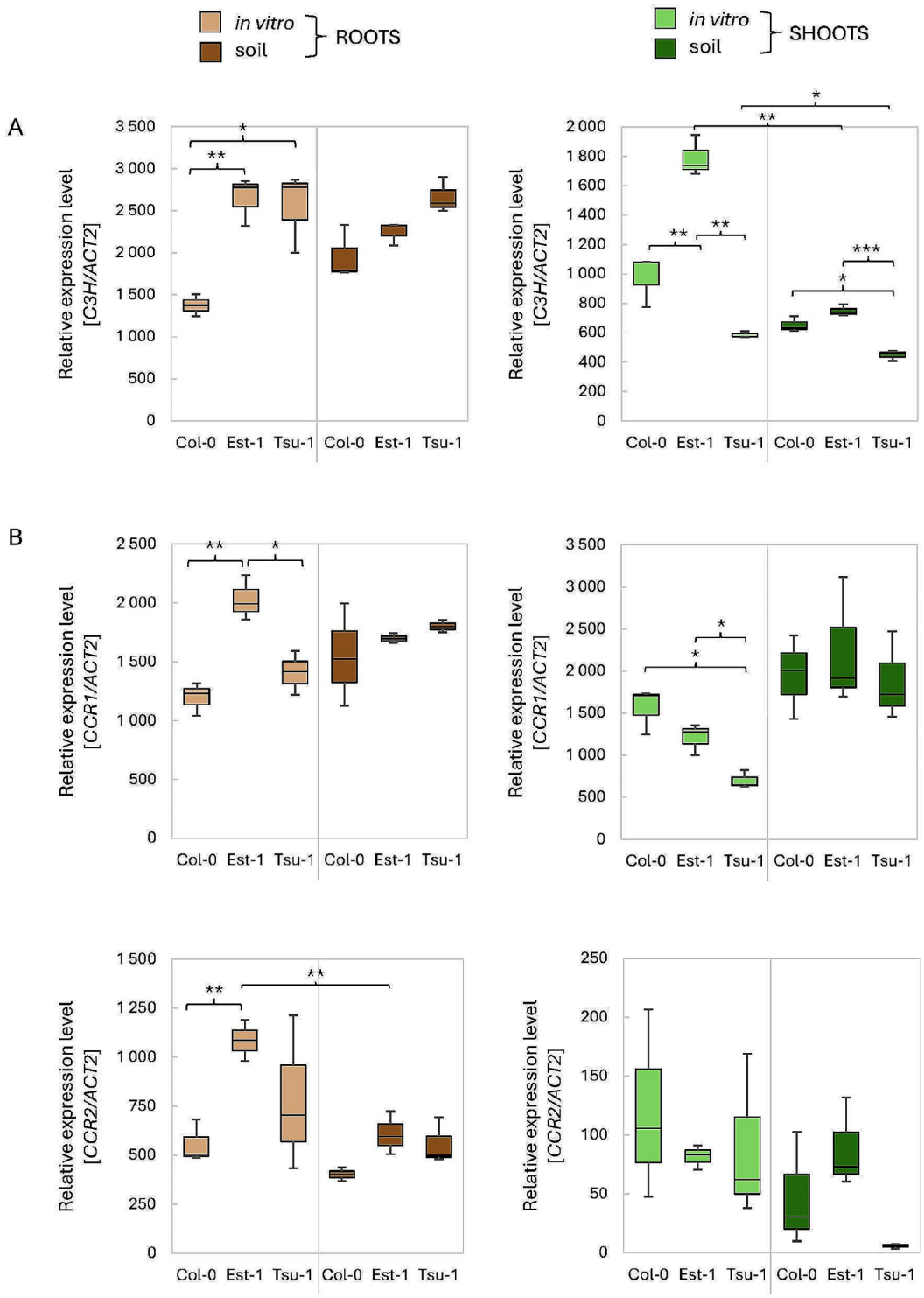


Fig. 10 Relative expression levels of the *C3'H* (**A**) and *CCRs* (**B**) genes measured by qPCR. As a reference, the *ACT2* (At3g18780) gene was used. The expression levels were quantified in three Arabidopsis genetic backgrounds (Col-0, Est-1, Tsu-1) grown in in vitro liquid culture and in soil, in two types of tissues (leaves and roots). Means and error bars for absolute deviations are shown. Values: $p < 0.001$ (***), $p < 0.01$ (**) and $p < 0.05$ (*)

a set of coding sequences involved in the biosynthesis of phenylpropanoid molecules. For each gene (*F6'H1*, *F6'H2*, *4CL*, *4CL2*, *4CL3*, *CCoAOMT1*, *C3'H* and *HCT*), except both *CCR* genes, two independent PCR products were cloned and sequenced separately. This sequencing task confirmed the presence of several SNPs that could be highlighted based on data available in the 1001 genomes database. Interestingly, we could also reveal some new SNPs, which are described below. Additionally, all detected SNPs of the tested Arabidopsis accessions along with possible resulting changes in amino acid sequences are presented in the Supplementary Information as Multiple sequence alignment (MSA) files from CLUSTALW (Supplementary Figure S3 and Figure S4). *4CL1* nucleotide sequence is highly conserved among the 3 studied accessions, Est-1 does not contain any SNPs, while Tsu-1 contains only one synonymous. *4CL2* does not show any polymorphisms in Tsu-1 whereas many SNPs could be detected in Est-1, especially on 5' end of the gene although the sequence of the protein resulting from it is well conserved. Only one SNP led to a change in amino acids (Arg26Lys) but exchanged a positively charged side chain amino acid by a similar one. This SNP is popular in accessions from all regions from the POLYMORPH database (https://tools.1001genomes.org/polymorph/). We could not detect any polymorphism for *4CL3* in Tsu-1, but, again, numerous SNPs on the 5'end of the corresponding gene in Est-1. In this case, the SNPs led to amino acids substitutions (Thr15Ser, Asp22Gly, His27Pro, Asp43Asn, Tyr93Cys and Arg170Val). An additional significant insertion at the N-terminus of the protein sequence could also be highlighted leading to His27_Ser28insPro_Pro_Pro. Concerning the dioxygenases, Est-1 does not contain any SNPs in *F6'H1* coding sequence, but Tsu-1 has one which causes the amino acid substitution Pro70Lys, that can affect the secondary structure of the synthetized protein. This SNP is present in 1001 genomes database for Tsu-1 (http://1001genomes.org/data/MPI/MPIOssowski2008/releases/current/strains/Tsu-1/), but also in POLYMORPH database for one accession, named HKT2.4, from Tübingen region. Est-1 *F6'H2* sequencing revealed two SNPs in the coding region, which were not present in 1001 genome project database, leading to an arginine to proline substitution in Est-1. For *F6'H2* in Tsu-1 15 SNPs were found, all of which were described in 1001 genome project database. Interestingly, as many as six SNPs broadly existing in accessions from all regions represented in POLYMORPH database (from Central Asia, Caucasus, Europe and North Africa) changed amino acid sequence of Tsu-1 *F6'H2* (Ile7Met, Lys199Thr, Phe208Leu, Gly299Ser, Ser311Asn and Lys334Arg). Finally, *CCoAOMT1* both in Est-1 and Tsu-1 does not contain any SNPs. *C3'H* coding sequence in Tsu-1 contains one synonymous SNP. *HCT* in Est-1 has three polymorphisms while Tsu-1 has 5, but in both accessions two SNPs cause the same amino acid change from hydrophobic to polar group: Ala125Thr.

Discussion

Here, additionally to scopolin, scopoletin and skimmin, umbelliferone was detected. We could not detect esculin in the investigated accessions as was detected by Kai et al. [10] and Perkowska et al. [16]. We detected a natural variation in accumulation of coumarins between Tsu-1, Col-0 and Est-1 accessions grown in vitro and in soil, together with variation in the gene expression between tested accessions. We verified SNPs present in the 1001 Genome database for Tsu-1 accession and found new SNPs and insertions in the Est-1 genetic background by re-sequencing.

Under non-stress conditions, coumarins are synthesized to a low extent. We observed only few significant differences in the production of coumarins in roots of tested accessions when grown in optimal soil mix (skimmin in Est-1 versus Tsu-1, $p<0.01$). Biosynthesis of coumarins is induced by various biotic and abiotic elicitors as well as by stress factors such as in vitro liquid culture. This is consistent with the work of Hemm et al. [45], who have shown that the phenylpropanoid production increases in roots exposed to light, similarly to the liquid culture system we used. This may be caused by changes in gene expression due to activation of photoreceptors that monitor different wavelengths of light [47]. Coumarins act as UV screens, so the overproduction of these compounds in Arabidopsis roots in in vitro liquid culture may constitute a defense mechanism against radiation. Experimental conditions of in vitro liquid culture also promote plant vitrification - hyperhydric malformations that affect the physiological state of plants [48] by reducing the content of chlorophylls, carotenoids and lignin [49]. Vitrification and light exposure may alter the coumarin biosynthetic profile because it is tightly connected to lignin biosynthesis and may be genotype (accession)- dependent. In the roots of plants grown in vitro, we observed a decrease in the expression of the *4CL1* gene, which has the biggest contribution into lignin biosynthesis, and an increase in the *4CL2* expression. These changes may lead to an imbalance in the accumulation of lignin in the cell wall and an increase in the level of coumarins. The Col-0 accessions was characterized by the lowest level of *4CL1* and *4CL2* expression in the roots of in vitro cultured plants compared to Est-1 and Tsu-1. Genes with different expression levels between accessions may have polymorphisms in their promoter regions.

We found some SNPs in the eight coding sequences (CDS) of the investigated accessions by resequencing. In parallel, SNPs in the *CCR1* and *CCR2* genes in Tsu-1

were checked in the 1001 Genomes database. The *CCR1* gene in Tsu-1 genetic background does not contain any SNPs, while *CCR2* has four, of which two at each end of the protein are nonsynonymous - Leu2Pro and Ser-332Pro. From the data obtained, it can be concluded that most of the nonsynonymous SNPs that might cause large changes in the enzyme structure occur in the CDS of enzyme homologues, which seem to be less important for homeostasis in non-stress conditions −4CL3, F6'H2 and CCR2, but are crucial for the survival during stress. The coding sequences for 4CL1, CCoAOMT1 and C3'H are highly conserved. This is consistent with the definition of natural variability, i.e. adaptation to local habitats that may experience different weather or geographic conditions. Evolution is driven by the need of adaptation. Could some genomic sequences evolve faster? An intraspecific race for survival?

Natural variation for various traits among Arabidopsis accessions have been investigated in many studies [13, 50–52]. It has been reported that Arabidopsis display a great natural variation for the accumulation of secondary metabolites [53]. This is logical in the context of the evolutionary history of plants, which had to adapt to various biotic and abiotic stress factors to survive. In addition to developing mechanical barriers such as wood, cuticle, and thorns, plants have developed complex biochemical machinery to produce and release a huge variety of compounds displaying antimicrobial or antifungal properties. As numerous studies carried out on various Arabidopsis accessions collected in areas with different environmental conditions, altitudes, humidity, and salinity have shown, there is an enormous variation at the level of the genome [54, 55], transcriptome [56] and metabolome [13, 16, 57]. Importantly, natural variation of root exudates of 19 Arabidopsis accession was detected and a direct link between metabolic phenotypes and genotypes were shown without using segregating populations [58]. Our study suggests that the observed variability in metabolic phenotypes may be genetically determined, and the integration of genomics and metabolomics data along with the gene expression analysis might be useful in elucidating the biosynthetic pathway.

Conclusion

Coumarins are secondary metabolites that have a range of important functions and biological activities valuable for both plants and humans. The unique structures of coumarins, make them useful in medicinal chemistry and pharmaceutical industry. In plants, they are involved in vital processes including adaptation to environmental stress factors, interactions with soil microorganisms and nutrient acquisition. Previously, we demonstrated for the first time that Arabidopsis with its extensive genetic variation and numerous publicly accessible web-based databases, is an exceptional model for studying molecular basis of natural variability underlying accumulation of coumarins in plants. Here, through targeted metabolic profiling and expression analysis of a set of phenylpropanoid genes, the gene-metabolite network was examined in more detail in the roots and leaves of selected three Arabidopsis accessions (Col-0, Est-1, Tsu-1) characterized by various levels of coumarin accumulation, which were grown in different types of cultures. We focused on two coumarin compounds, scopoletin and umbelliferone, along with their glycosides. This choice was dictated by the fact that scopoletin and scopolin are the main coumarins in Arabidopsis, and the biosynthesis of umbelliferone, recently discovered in this model plant by our research group, is completely unknown. We not only detected the expected variability in gene expression and coumarin accumulation among Arabidopsis grown in soil and in vitro cultures, but also found interesting polymorphisms in the coding sequences of the studied genes through in silico analysis and resequencing. Studying the natural variation in coumarin content present among Arabidopsis accessions followed by the analysis of various alleles possibly underlying the detected variation, may be useful in the future discovery of the physiological mechanisms of action of different alleles and better understanding the correlation between genetic and metabolic variants.

Abbreviations

4CL1	4-Coumarate: CoA Ligase 1
4CL2	4-Coumarate: CoA Ligase 2
4CL3	4-Coumarate: CoA Ligase 3
4CL4	4-Coumarate: CoA Ligase 4
ACT2	ACTIN 2 *ACTIN 2*
C2'H	p-Coumaroyl CoA 2'-Hydroxylase
C3'H	p-Coumaroyl 3'-Hydroxylase
C4H	Trans-Cinnamate 4-monooxygenase
CCoAOMT1	Caffeoyl Coenzyme A dependent O-Methyltransferase 1
CCR1	Cinnamoyl CoA Reductase 1
CCR2	Cinnamoyl CoA Reductase 2
CoA	Coenzyme A
Col-0	Columbia
Est-1	Estland
F6'H1	Feruloyl-CoA 6'-Hydroxylase 1
F6'H2	Feruloyl-CoA 6'-Hydroxylase 2
HCT	Shikimate O-Hydroxycinnamoyltransferase
MD	Mean Deviation
N0	Initial copy number
PAL1	Phenylalanine Ammonia-Lyase 1
PAL2	Phenylalanine Ammonia-Lyase 2
PDR9	ABC transporter G family member 37
qPCR	Quantitative PCR
SNP	Single Nucloitide Polymorphism
Tsu-1	Tsushima
UHPLC	Ultra High Performance Liquid Chromatography
UV	Ultraviolet
ΔCp	Crossing Point

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-024-05491-w.

Additional file 1_Table S1. The gene-specific primers used for qPCRs

Additional file 2_Table S2. Efficiencies of the PCR product amplification by the qPCR primers used in this study

Additional file 3_Table S3. Primers sequences used for re-sequencing (gene-specific primers used for CDS amplification)

Additional file 4_ Figure S1. Skimmin and scopolin quantification in Arabidopsis roots grown in vitro

Additional file 5_ Figure S2. Umbelliferone and scopoletin quantification in Arabidopsis roots grown in vitro

Additional file 6_Figure S3. Nucleotide sequence variants shown as a Multiple sequence alignment (MSA)

Additional file 7_Figure S4. Amino acid sequence variants shown as a Multiple sequence alignment (MSA)

Acknowledgements
We would like to thank Maarten Koornneef for providing seeds of Arabidopsis accessions used in this study and Thibault Duval for his valuable technical assistance.

Author contributions
AI designed the research, supervised the experiments and coordinated the project, analyzed data, was a major contributor in writing the manuscript, received funding for the project. JS performed experiments, analyzed data, contributed to the writing of manuscript draft. IP analyzed data, contributed to the manuscript writing, was a major contributor in figures preparation. JG performed the experiments. AH critically revised manuscript. FB critically revised manuscript. EL critically revised manuscript, received funding for the project. AO designed and performed the experiments, analyzed data, coordinated the project and contributed to the manuscript writing. All authors read and approved the final manuscript.

Funding
This work was supported by the Polish National Science Centre grants: UMO-2022/47/B/NZ2/01835 (AI) and 2019/35/O/NZ1/02751 (EL), and the Polish National Agency for Academic Exchange grants BPN/BFR/2022/1/00039 (AI) and PPN/BFR/2019/1/00050 (AI).

Data availability
All data supporting the findings of this study are available within the paper and its Supplementary Information. Gene-specific primer sequences used for qPCRs are provided in Supplementary Table S1, along with original reference describing the *ACTIN2* primer sequences used in this study. Efficiencies of the PCR product amplification by the qPCR primers are provided in Supplementary Table S2. Primer sequences used for re-sequencing are provided in Supplementary Table S3. Nucleotide sequences of the selected eight genes (*F6'H1*, *F6'H2*, *4CL1*, *4CL2*, *4CL3*, *CCoAOMT1*, *C3'H* and *HCT*) of the studied Arabidopsis accessions are deposited in the GenBank ® repository (NIH genetic sequence database; https://www.ncbi.nlm.nih.gov/genbank/) with the relevant accession numbers from PP507124 to PP507144. Additionally, all sequence variants are provided as Multiple sequence alignment (MSA) files in Supplementary Figure S3 and S4. The effect of microwave treatment on the content of coumarins in MeOH extracts is provided in Supplementary Figures S1 and S2.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Laboratory of Plant Protection and Biotechnology, Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk, University of Gdansk, Abrahama 58, Gdansk 80-307, Poland
[2]Université de Lorraine-INRAE, LAE, Nancy F-54000, France

Received: 17 April 2024 / Accepted: 6 August 2024
Published online: 27 August 2024

References

1. Matern U. Coumarins and other phenylpropanoid compounds in the defense response of plant cells. Planta Med. 1991;57(1):S15–20. https://doi.org/10.1055/s-2006-960224.
2. Gao XY, Li XY. Scopoletin: a review of its pharmacology, pharmacokinetics, and toxicity. Front Pharmacol. 2024;15:1268464. https://doi.org/10.3389/fphar.2024.1268464.
3. Bayoumi SA, Rowan MG, Blagbrough IS, Beeching JR. Biosynthesis of scopoletin and scopolin in cassava roots during post-harvest physiological deterioration: the EZ-isomerisation stage. Phytochemistry. 2008;69(17):2928–36. https://doi.org/10.1016/j.phytochem.2008.09.023.
4. Matsumoto S, Mizutani M, Sakata K, Shimizu BI. Molecular cloning and functional analysis of the Ortho-hydroxylases of p-coumaroyl coenzyme A/feruloyl coenzyme A involved in formation of umbelliferone and scopoletin in sweet potato, Ipomoea batatas (L.) Lam. Phytochemistry. 2012;74:49–57. https://doi.org/10.1016/j.phytochem.2011.11.009.
5. Rial C, Varela RM, Molinillo JM, Peralta S, Macías FA. Sunflower metabolites involved in resistance mechanisms against broomrape. Agronomy. 2021;11(3):501. https://doi.org/10.3390/agronomy11030501.
6. Alissandrakis E, Kibaris AC, Tarantilis PA, Harizanis PC, Polissiou M. Flavour compounds of Greek cotton honey. J Sci Food Agric. 2005;85(9):1444–52. https://doi.org/10.1002/jsfa.2124.
7. Lerat S, Babana AH, Oirdi E, El Hadrami M, Daayf A, Beaudoin F, Beaulieu N, C. Streptomyces scabiei and its toxin thaxtomin A induce scopoletin biosynthesis in tobacco and Arabidopsis thaliana. Plant Cell Rep. 2009;28:1895–903.
8. Rohde A, Morreel K, Ralph J, Goeminne G, Hostyn V, De Rycke R, Boerjan W. Molecular phenotyping of the pal1 and pal2 mutants of Arabidopsis thaliana reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. Plant Cell. 2004;16(10):2749–71. https://doi.org/10.1105/tpc.104.023705.
9. Bednarek P, Schneider B, Svatos A, Oldham NJ, Hahlbrock K. Structural complexity, differential response to infection, and tissue specificity of indolic and phenylpropanoid secondary metabolism in Arabidopsis roots. Plant Physiol. 2005;138(2):1058–70. https://doi.org/10.1104/pp.104.057794.
10. Kai K, Shimizu BI, Mizutani M, Watanabe K, Sakata K. Accumulation of coumarins in Arabidopsis thaliana. Phytochemistry. 2006;67(4):379–86. https://doi.org/10.1016/j.phytochem.2005.11.006.
11. Kai K, Mizutani M, Kawamura N, Yamamoto R, Tamai M, Yamaguchi H, Shimizu BI. Scopoletin is biosynthesized via Ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in Arabidopsis thaliana. Plant J. 2008;55(6):989–99. https://doi.org/10.1111/j.1365-313X.2008.03568.x.
12. Schmid NB, Giehl RF, Döll S, Mock HP, Strehmel N, Scheel D, von Wirén N. Feruloyl-CoA 6'-Hydroxylase1-dependent coumarins mediate iron acquisition from alkaline substrates in Arabidopsis. Plant Physiol. 2014;164(1):160–72. https://doi.org/10.1104/pp.113.228544.
13. Siwinska J, Kadzinski L, Banasiuk R, Gwizdek-Wisniewska A, Olry A, Banecki B, Ihnatowicz A. Identification of QTLs affecting scopolin and scopoletin biosynthesis in Arabidopsis thaliana. BMC Plant Biol. 2014;14:1–14. https://doi.org/10.1186/s12870-014-0280-9.
14. Siwinska J, Siatkowska K, Olry A, Grosjean J, Hehn A, Bourgaud F, Ihnatowicz A. Scopoletin 8-hydroxylase: a novel enzyme involved in coumarin biosynthesis and iron-deficiency responses in Arabidopsis. J Exp Bot. 2018;69(7):1735–48. https://doi.org/10.1093/jxb/ery005.
15. Stringlis IA, De Jonge R, Pieterse CM. The age of coumarins in plant–microbe interactions. Plant Cell Physiol. 2019;60(7):1405–19. https://doi.org/10.1093/pcp/pcz076.
16. Perkowska I, Siwinska J, Olry A, Grosjean J, Hehn A, Bourgaud F, Ihnatowicz A. Identification and quantification of coumarins by UHPLC-MS in Arabidopsis thaliana natural populations. Molecules. 2021a;26(6):1804. https://doi.org/10.3390/molecules26061804.

17. Robe K, Conejero G, Gao F, Lefebvre-Legendre L, Sylvestre-Gonon E, Rofidal V, Dubos C. Coumarin accumulation and trafficking in Arabidopsis thaliana: a complex and dynamic process. New Phytol. 2021;229(4):2062–79. https://doi.org/10.1111/nph.17090.
18. Bourgaud F, Hehn A, Larbat R, Doerper S, Gontier E, Kellner S, Matern U. Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. Phytochem Rev. 2006;5:293–308. https://doi.org/10.1007/s11101-006-9040-2.
19. Fourcroy P, Sisó-Terraza P, Sudre D, Savirón M, Reyt G, Gaymard F, Briat JF. Involvement of the ABCG 37 transporter in secretion of scopoletin and derivatives by Arabidopsis roots in response to iron deficiency. New Phytol. 2014;201(1):155–67. https://doi.org/10.1111/nph.12471.
20. Schmidt H, Günther C, Weber M, Spörlein C, Loscher S, Böttcher C, Clemens S. Metabolome analysis of Arabidopsis thaliana roots identifies a key metabolic pathway for iron acquisition. PLoS ONE. 2014;9(7):e102444. https://doi.org/10.1371/journal.pone.0102444.
21. Rajniak J, Giehl RF, Chang E, Murgia I, von Wirén N, Sattely ES. Biosynthesis of redox-active metabolites in response to iron deficiency in plants. Nat Chem Biol. 2018;14(5):442–50. https://doi.org/10.1038/s41589-018-0019-2.
22. Tsai HH, Rodríguez-Celma J, Lan P, Wu YC, Vélez-Bermúdez IC, Schmidt W. Scopoletin 8-hydroxylase-mediated fraxetin production is crucial for iron mobilization. Plant Physiol. 2018;177(1):194–207. https://doi.org/10.1104/pp.18.00178.
23. Sisó-Terraza P, Luis-Villarroya A, Fourcroy P, Briat JF, Abadía A, Gaymard F, Abadía J, Álvarez-Fernández A. Accumulation and Secretion of coumarinolignans and other coumarins in Arabidopsis thaliana roots in response to Iron Deficiency at High pH. Front Plant Sci. 2016;7. https://doi.org/10.3389/fpls.2016.01711.
24. Stringlis IA, Yu K, Feussner K, de Jonge R, Van Bentum S, Van Verk MC, Pieterse CM. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. Proc Natl Acad Sci. 2018;115(22):E5213–22. https://doi.org/10.1073/pnas.1722335115.
25. Jeandet P, Clément C, Courot E, Cordelier S. Modulation of phytoalexin biosynthesis in engineered plants for disease resistance. Int J Mol Sci. 2013;14(7):14136–70. https://doi.org/10.3390/ijms140714136.
26. Tohge T, Watanabe M, Hoefgen R, Fernie AR. The evolution of phenylpropanoid metabolism in the green lineage. Crit Rev Biochem Mol Biol. 2013;48(2):123–52. https://doi.org/10.3390/ijms140714136.
27. Peterson JK, Harrison HF, Jackson DM, Snook ME. Biological activities and contents of scopolin and scopoletin in sweetpotato clones. HortScience. 2003;38(6):1129–33. https://doi.org/10.21273/HORTSCI.38.6.1129.
28. Costet L, Fritig B, Kauffmann S. Scopoletin expression in elicitor-treated and tobacco mosaic virus-infected tobacco plants. Physiol Plant. 2002;115(2):228–35. https://doi.org/10.1034/j.1399-3054.2002.1150208.x.
29. Perkowska I, Potrykus M, Siwinska J, Siudem D, Lojkowska E, Ihnatowicz A. Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. Int J Mol Sci. 2021b;22(12):6449. https://doi.org/10.3390/ijms22126449.
30. Voges MJ, Bai Y, Schulze-Lefert P, Sattely ES. (2019). Plant-derived coumarins shape the composition of an Arabidopsis synthetic root microbiome. Proc Natl Acad Sci USA. 2019;116(25):12558-65 https://doi.org/10.1073/pnas.1820691116
31. Harbort CJ, Hashimoto M, Inoue H, Niu Y, Guan R, Rombolà AD, Schulze-Lefert P. Root-secreted coumarins and the microbiota interact to improve iron nutrition in Arabidopsis. Cell Host Microbe. 2020;28(6):825–37. https://doi.org/10.1016/j.chom.2020.09.006.
32. Stassen MJ, Hsu SH, Pieterse CM, Stringlis IA. Coumarin communication along the microbiome–root–shoot axis. Trends Plant Sci. 2021;26(2):169–83. https://doi.org/10.1016/j.tplants.2020.09.008.
33. Cosme M, Fernández I, Declerck S, van der Heijden MG, Pieterse CM. A coumarin exudation pathway mitigates arbuscular mycorrhizal incompatibility in Arabidopsis thaliana. Plant Mol Biol. 2021;106:319–34. https://doi.org/10.1007/s11103-021-01143-x.
34. Rodríguez-Celma J, Lin WD, Fu GM, Abadía J, López-Millán AF, Schmidt W. Mutually exclusive alterations in secondary metabolism are critical for the uptake of insoluble iron compounds by Arabidopsis and Medicago truncatula. Plant Physiol. 2013;162(3):1473–85. https://doi.org/10.1104/pp.113.220426.
35. Paffrath V, Tandron Moya YA, Weber G, von Wirén N, Giehl RF. A major role of coumarin-dependent ferric iron reduction in strategy I-type iron acquisition in Arabidopsis. Plant Cell. 2024;36(3):642–64. https://doi.org/10.1093/plcell/koad279.
36. Biała W, Jasinski M. The Phenylpropanoid case – it is transport that matters. Front Plant Sci. 2018;9:1610. https://doi.org/10.3389/fpls.2018.01610.
37. Li Y, Kim JI, Pysh L, Chapple C. Four isoforms of Arabidopsis 4-Coumarate:CoA ligase have overlapping yet distinct roles in Phenylpropanoid Metabolism. Plant Physiol. 2015;169:2409–21. https://doi.org/10.1104/pp.15.00838.
38. Nguyen C, Bouque V, Bourgaud F, Guckert A. Quantification of daidzein and furanocoumarin conjugates of Psoralea cinerea L.(Leguminosae). Phytochemical Analysis: Int J Plant Chem Biochem Techniques. 1997;8(1):27–31. https://doi.org/10.1002/(SICI)1099-1565(199701).
39. Ihnatowicz A, Siwinska J, Meharg AA, Carey M, Koornneef M, Reymond M. Conserved histidine of metal transporter at NRAMP 1 is crucial for optimal plant growth under manganese deficiency at chilling temperatures. New Phytol. 2014;202(4):1173–83. https://doi.org/10.1111/nph.12737.
40. Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible WR. Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. Plant Physiol. 2005;139(1):5–17. https://doi.org/10.1104/pp.105.063743.
41. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673–80. https://doi.org/10.1093/nar/22.22.4673.
42. Rataboul P, Alibert G, Boller T, Boudet AM. Intracellular transport and vacuolar accumulation of o-coumaric acid glucoside in Melitolus Alba mesophyll cell protoplasts. Biochim et Biophys Acta (BBA) - Biomembr. 1985;816(1):25–36. https://doi.org/10.1016/0005-2736(85)90389-X.
43. Xiping Chen J, Zhang Y. Rapid microwave-assisted hydrolysis for determination of Ginkgo flavonol glycosides in extracts of Ginkgo biloba leaves. J Chromatogr Sci. 2008;46(2):117–21. https://doi.org/10.1093/chromsci/46.2.117.
44. Wang H, Li Y, Ren Z, Cong Z, Chen M, Shi L, Han X, Pei J. Optimization of the microwave-assisted enzymatic extraction of *Rosa roxburghii* Tratt. Polysaccharides using response surface methodology and its antioxidant and α-d-glucosidase inhibitory activity. Int J Biol Macromol. 2018;112:473–82. https://doi.org/10.1016/j.ijbiomac.2018.02.003.
45. Hemm MR, Rider SD, Ogas J, Murry DJ, Chapple C. Light induces phenylpropanoid metabolism in Arabidopsis roots. Plant J. 2004;38(5):765–78. https://doi.org/10.1111/j.1365-313X.2004.02089.x.
46. Guénin S, Mauriat M, Pelloux J, van Wuytswinkel O, Bellini C, Gutierrez L. Normalization of RT-qPCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. J Exp Bot. 2009;60:487–93. https://doi.org/10.1093/jxb/ern305.
47. Kendrick RE, Kronenberg GH, editors. Photomorphogenesis in plants. Springer Science & Business Media; 2012. https://doi.org/10.1007/978-94-011-1884-2.
48. Hakkaart FA, Versluijs JM. Some factors affecting glassiness in carnation meristem tip cultures. Neth J Plant Pathol. 1983;89:47–53. https://doi.org/10.1007/BF01974443.
49. Lapeña L, Pérez-Bermúdez P, Segura J. Factors affecting shoot proliferation and vitrification in Digitalis obscura cultures. Vitro–Plant. 1992;28:121–4. https://doi.org/10.1007/BF02823059.
50. Sharma S, Lin W, Villamor JG, Verslues PE. Divergent low water potential response in Arabidopsis thaliana accessions Landsberg erecta and Shahdara. Plant Cell Environ. 2013;36(5):994–1008. https://doi.org/10.1111/pce.12032.
51. Fu SF, Chen PY, Nguyen QTT, Huang LY, Zeng GR, Huang TL, Huang HJ. Transcriptome profiling of genes and pathways associated with arsenic toxicity and tolerance in Arabidopsis. BMC Plant Biol. 2014;14(1):1–16. https://doi.org/10.1186/1471-2229-14-94.
52. Gawarecka K, Siwinska J, Poznanski J, Onysk A, Surowiecki P, Sztompka K, Ihnatowicz A. cis-prenyltransferase 3 and α/β-hydrolase are new determinants of dolichol accumulation in Arabidopsis. Plant Cell Environ. 2022;45(2):479–95. https://doi.org/10.1111/pce.14223.
53. Koornneef M, Alonso-Blanco C, Vreugdenhil D. Naturally occurring genetic variation in Arabidopsis thaliana. Annu Rev Plant Biol. 2004;55:141–72. https://doi.org/10.1146/annurev.arplant.55.031903.141605.
54. Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Weigel D. Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nat Genet. 2011;43(10):956–63. https://doi.org/10.1038/ng.911.
55. Hollister JD. Genomic variation in Arabidopsis: tools and insights from next-generation sequencing. Chromosome Res. 2014;22(2):103–15. https://doi.org/10.1007/s10577-014-9420-1.
56. Günther T, Lampei C, Schmid KJ. Mutational bias and gene conversion affect the intraspecific nitrogen stoichiometry of the Arabidopsis thaliana

transcriptome. Mol Biol Evol. 2013;30(3):561–8. https://doi.org/10.1093/molbev/mss249.
57. Routaboul JM, Dubos C, Beck G, Marquis C, Bidzinski P, Loudet O, Lepiniec L. Metabolite profiling and quantitative genetics of natural variation for flavonoids in Arabidopsis. J Exp Bot. 2012;63(10):3749–64. https://doi.org/10.1093/jxb/ers067.
58. Mönchgesang S, Strehmel N, Schmidt S, et al. Natural variation of root exudates in Arabidopsis thaliana-linking metabolomic and genomic data. Sci Rep. 2016;6:29033. https://doi.org/10.1038/srep29033.

Publisher's Note

Publication No. 4

AIM

2) **Comprehensive analysis of the regulatory network controlling the switch between coumarin and lignin biosynthesis in *A. thaliana* under osmotic and low pH stress using integrative transcriptomic and metabolomic approaches.**
3) **Functional characterization of two *A. thaliana* dioxygenases DOXC21-A, its NAT, and DOXC21-B potentially involved in ortho-hydroxylation of coumarin and lignin precursors.**

Perkowska, I., Olry, A., Barrit, T., Dobek, A., Magot, F., Charles, C., Januchowska, A., Siwinska, J., Lojkowska, E., & Ihnatowicz, A. **(2025a)**

Elucidating the role and regulatory mechanism of *Arabidopsis thaliana DOXC21-A*, its NAT and *DOXC21-B* in plant adaptation to terrestrial conditions

(manuscript in preparation)

Elucidating the role and regulatory mechanism of *Arabidopsis thaliana DOXC21-A,* its NAT and *DOXC21-B* in plant adaptation to terrestrial conditions

Izabela Perkowska[1], Alexandre Olry[2], Thibault Barrit[1], Alicja Dobek[1], Clement Charles[2], Florent Magot[3], Joanna Siwinska[1], Ewa Lojkowska[1] and Anna Ihnatowicz[1,*]

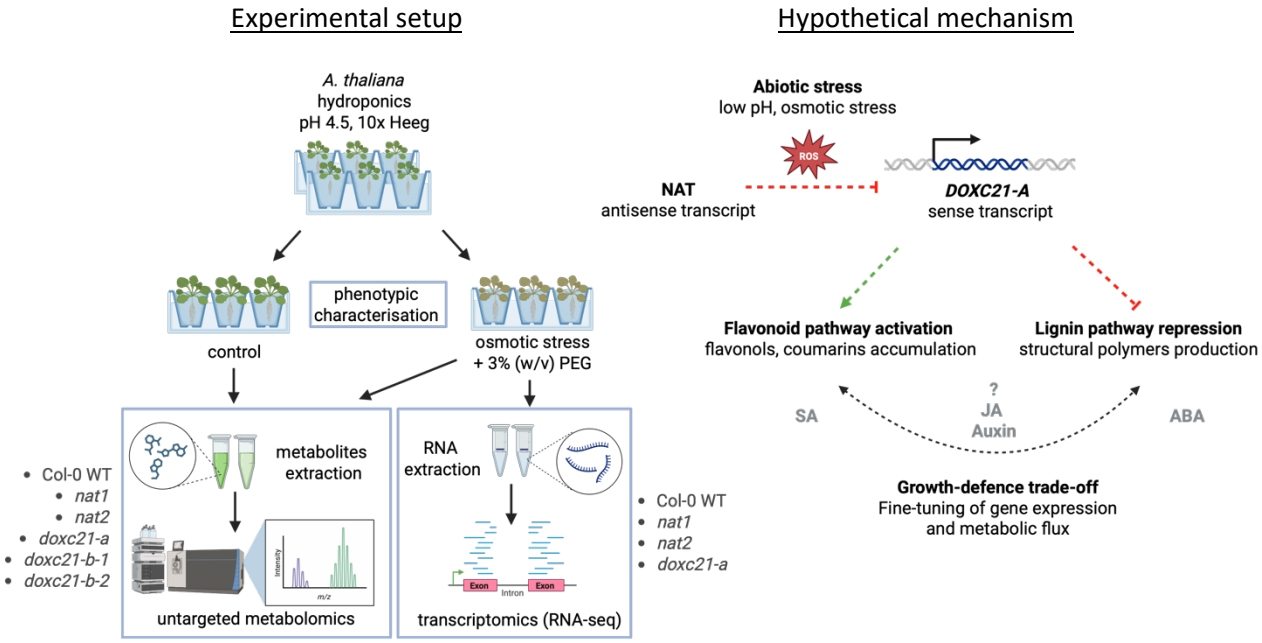
[1]Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk, University of Gdansk, Abrahama 58, 80-307 Gdansk, Poland

[2]Université de Lorraine, INRAE, LAE, 54000 Nancy, France

[3]Université de Tours, Biomolécules et Biotechnologies Végétales, EA2106 F-37200 Tours, France

***Corresponding author details**: Anna Ihnatowicz (Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk), anna.ihnatowicz@biotech.ug.edu.pl

GRAPHICAL ABSTRACT



BACKGROUND

The colonization of land by early plants represents one of the most consequential evolutionary transitions in Earth's history. This shift enabled the diversification of plant life into the complex terrestrial ecosystems we observe today [1]. Terrestrial environments exposed plants to a wide range of abiotic stresses, including drought, fluctuating temperatures, high UV-B radiation, osmotic imbalance, and nutrient limitations [2, 3]. These environmental factors induce profound morphological, physiological, and biochemical changes, including reinforced cell walls, vascular

tissues, and novel regulatory mechanisms for developmental plasticity and stress mitigation [4, 5]. Understanding the genetic and metabolic mechanisms that enabled terrestrial adaptation offers insights into current challenges such as drought tolerance and climate resilience.

Early land plant lineages, such as bryophytes, developed cuticles and desiccation-resistant reproductive structures to survive in water-limited conditions [6]. Land colonization was further supported by molecular strategies such as stomatal control, antioxidant production, and the synthesis of protective secondary metabolites [7, 8]. Thus, adaptation to land required the evolution of regulatory pathways capable of integrating internal cues with external stress signals in a coordinated manner.

Among the major evolutionary developments supporting terrestrial adaptation, plant secondary metabolism has emerged as a critical layer of environmental responsiveness. Specialized metabolites such as flavonoids, coumarins, terpenoids, alkaloids, and glucosinolates play roles in nutrient uptake, UV protection, osmotic adjustment, and pathogen defense [9, 10]. The phenylpropanoid pathway, in particular, serves as a metabolic hub for synthesizing lignin, flavonoids, and coumarins – compounds involved in structural integrity and stress resilience [11]. These biosynthetic routes are controlled by a network of stress-responsive transcription factors (TF) (e.g., MYB, bHLH) and enzymatic players, including cytochrome P450s (CYP) and dioxygenases [12].

Hormonal signaling pathways, including abscisic acid (ABA) and jasmonic acid (JA), orchestrate trade-offs between growth and defense, modulating resource allocation between primary and secondary metabolism [13]. While the major regulatory components of stress responses are relatively well characterized, the expression dynamics and biological functions of many enzymes – particularly under simultaneous or multifactorial stress conditions – remain not fully understood.

2-oxoglutarate-dependent dioxygenases (2OGDs) are a large superfamily of non-heme iron (Fe)-dependent enzymes that catalyze diverse oxidative reactions, including hydroxylation, demethylation, and epoxide formation [14, 15]. These enzymes function in primary metabolism, including the biosynthesis of hormones such as gibberellins and ethylene, and also contribute to secondary metabolism, notably in the production of coumarins, as well as flavonoids, glucosinolates, and alkaloids [16].

Genome of *Arabidopsis thaliana* encodes over 130 2OGDs, grouped into three functional clades: DOXA (histone demethylation), DOXB (proline hydroxylation), and DOXC (secondary metabolism). The DOXC clade, the most functionally diverse, includes enzymes involved in stress response, UV protection, and cell wall modification. However, despite broad interest, many DOXC genes remain poorly characterized in terms of stress responsiveness, enzymatic activity, and regulatory control [15].

We selected two previously uncharacterized *A. thaliana* genes (*AT3G19000* and *AT3G19010*), both members of the DOXC clade classified as clade 21 by Kawai *et al.* [15], which are conserved across all land plants and exhibit transcriptional responsiveness to abiotic stressors such as UV-B

and osmotic stress [17]. Based on their sequence similarity and genomic context, we hereby named them *DOXC21-A* and *DOXC21-B*, respectively. Additionally, *DOXC21-A* is overlapped by a natural antisense transcript (NAT), suggesting a possible layer of post-transcriptional regulation. NATs are increasingly recognized for their role in fine-tuning gene expression under environmental stress. They can modulate gene expression through several mechanisms, including direct inhibition of translation, alteration of mRNA stability, and formation of RNA duplexes that impact transcript accessibility and regulatory protein binding [18].

Given the rapid diversification of genes encoding DOXC enzymes and the presence of NAT regulation, *DOXC21-A* in particular presents an opportunity to explore gene function evolution and NAT-mediated regulation, which has emerged as a key layer of transcriptional control in plants. Exploring the transcriptional and metabolic consequence of *DOXC21-A* disruption may provide valuable insights into the broader functional diversity of the 2OGD protein superfamily.

While canonical 2OGDs such as FLS (flavonol synthase), F3'H (flavanone 3-hydroxylase), and F6'H1 (feruloyl-CoA 6'-ortho-hydroxylase 1) have well-defined roles in phenylpropanoid metabolism [19, 20], the physiological functions of DOXC21-A and DOXC21-B remain uncharacterized. Their deep evolutionary conservation, stress-responsive expression, and the presence of NAT overlapping *DOXC21-A* suggest a potential role in specialized metabolism and possibly in gene regulation through RNA-based mechanisms.

We hypothesize that DOXC21-A contributes to transcriptional and metabolic reprogramming during abiotic stress adaptation. We further propose that its expression and potential functional divergence from DOXC21-B may be shaped by evolutionary events such as gene duplication and subfunctionalization. The presence of a cis-NAT suggests possible regulatory complexity beyond enzymatic function alone.

This study aims to functionally characterize DOXC21-A and its associated NAT in the context of abiotic stress adaptation in *A. thaliana*, alongside its closely related paralog, *DOXC21-B* in order to investigate potential functional divergence or redundancy in stress-related pathways. Using a multi-omics approach integrating targeted phenotyping, transcriptomics, and metabolomics, we specifically aimed to: (I) Examine the evolutionary conservation, expression patterns, and stress-associated expression of *DOXC21-A* and *DOXC21-B*; (II) Assess the potential impact of NAT regulation and *DOXC21-A* disruption on phenotype appearance under abiotic stress; (III) Identify transcriptomic and metabolic changes associated with loss-of-function mutans of *DOXC21*-A, its associated NAT, and *DOXC21-B* under osmotic and low pH stress conditions.

METHODS

1. *In silico* analysis

1.1. Gene model annotation, NAT identification, SNP distribution analysis, and miRNA homology

Gene structure, exon–intron organization, and alternative transcript variants were retrieved from the ThaleMine database (https://bar.utoronto.ca/thalemine/) and the ePlant visualization tool (http://bar.utoronto.ca/eplant/). Exonic overlap and NAT presence for *DOXC21-A* were assessed based on annotated transcript models. Additional confirmation was obtained by BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) alignment of the antisense transcript sequence against the TAIR10 *A. thaliana* genome to validate strand-specific overlap and orientation.

Single Nucleotide Polymorphism (SNP) data for *DOXC21-A* and *DOXC21-B* were extracted from the 2029-line *A. thaliana* population dataset using a custom R script available at https://github.com/arthurkorte/SNP_extractor. The analysis includes all detected SNP types, ranging from UTR and synonymous variants to splice region and missense variants, providing insight into the degree of polymorphism and potential functional divergence of each gene [21].

To investigate sequence conservation between the *AT3G19002* NAT and members of the 2OGD gene family in *A. thaliana*, a 35-nucleotide region from the NAT transcript was selected for comparative alignment. This region included a 29-bp highly conserved core implicated in regulatory similarity. Homologous regions were first identified by querying the NAT sequence against the *A. thaliana* genome using the NCBI BLASTN tool (default parameters; TAIR10 genome; https://blast.ncbi.nlm.nih.gov/). High-confidence hits were retained from six 2OGD genes: *DOXC21-A* (*AT3G19000*), *DOXC21-B* (*AT3G19010*), *GA2OX8* (*AT4G21200*), *JID1* (*AT1G05260*), *F6'H2* (*AT1G55290*), and *F6'H1* (*AT3G13610*), based on ≥25/29 nucleotide identity within the core region.

Multiple sequence alignment of these homologous regions was performed using ClustalW with default settings as implemented in BioEdit v7.2.5. The alignment was manually curated, and the conserved 29-bp segment was highlighted to visualize nucleotide conservation and mismatches. Conserved positions were shaded in black, while mismatches were shown in white to emphasize sequence variation across homologs.

To investigate the evolutionary conservation and potential regulatory mimicry of a conserved segment within the *AT3G19002* NAT transcript, we compared its sequence to annotated or predicted plant microRNAs (miRNAs). A 29-nucleotide region was selected and queried against the miRBase database (https://www.mirbase.org) to identify homologous miRNA precursor sequences. Eleven representative plant species were included in the comparative analysis: *Gossypium raimondii* (gra-miR8697), *Physcomitrium patens* (ppt-miR1055), *A. thaliana* (pre-miR844, pre-miR395a), *Medicago truncatula* (pre-miR2670a), *Arabidopsis lyrata* (pre-miR4240), *Vitis vinifera*

(pre-miR395f), *Populus trichocarpa* (pre-miR6452), *Lotus japonicus* (pre-miR11111), *Glycine max* (pre-miR5042), and *Amborella trichopoda* (pre-miR8565h). The conserved NAT segment and the aligned portions of the miRNA precursors were subjected to multiple sequence alignment (MSA) using the MUSCLE algorithm implemented in Jalview v2.11.3.10. Alignment was manually curated and trimmed to positions 5–20 to focus on the most conserved subregion. A sequence logo was generated directly in Jalview using its built-in WebLogo visualization tool to represent nucleotide frequency and conservation across this region.

1.2. Phylogenetic tree construction

To investigate the evolutionary relationships within the DOXC21 clade of 2OGDs, we curated a dataset comprising 100 non-redundant protein sequences representing a wide phylogenetic range of plant species. Homologs of DOXC21-A and DOXC21-B were identified using BLASTp searches against publicly available proteomic databases, including Gramene database (https://ensembl.gramene.org/) [22] and NCBI RefSeq (https://www.ncbi.nlm.nih.gov/refseq/). Where multiple accessions were available per species, up to five high-confidence representatives were retained based on annotation quality and sequence completeness. The final dataset comprises sequences from bryophytes, lycophytes, gymnosperms, monocots, dicots, and basal angiosperms, ensuring comprehensive phylogenetic representation. All sequences were downloaded in FASTA format and assigned consistent naming conventions that include species and gene identifiers.

The protein sequences (n=100) were aligned using the MUSCLE algorithm [23] in MEGA11, applying default gap penalties and the UPGMA clustering method. Phylogenetic trees were constructed using the Neighbor-Joining (NJ) method with 500 bootstrap replicates to estimate branch support. The Jones–Taylor–Thornton (JTT) model of amino acid substitution was applied with gamma-distributed rates and invariant sites (+G+I). The resulting tree was rooted using *Marchantia polymorpha* subsp. *ruderalis*, an early-diverging liverwort, to establish evolutionary polarity and anchor the DOXC21 lineage within a broader land plant context. The phylogenetic tree was visualized and edited using MEGA11 (https://www.megasoftware.net/).

To assess residue-level conservation across the DOXC21-related sequences used for phylogenetic analysis, we performed a conservation profile analysis based on the MSA. Conservation scores were computed in Jalview (v2.11.4.1) [24] using the AAConWS plugin [25] with the Shenkin entropy-based method [26]. The conservation profile was normalized to allow relative comparison across positions, and the results were used to evaluate the evolutionary constraint of DOXC21 homologs across land plant species.

To assess the conservation of the canonical 2OGD Fe-binding motif, protein sequences were analyzed using the ScanProsite tool (https://prosite.expasy.org/scanprosite/). The conserved His–X–Asp/Glu–(X)n–His motif was used to identify potential Fe-binding sites characteristic of 2OGD enzymatic function. FASTA-formatted protein sequences were uploaded to ScanProsite, and matches were evaluated for motif presence, position, and conservation across sequences.

This analysis enabled the confirmation of functional residues required for Fe(II) coordination in the active site of DOXC21-related proteins [15].

1.3. Promoter region analysis, abiotic stress expression, protein-protein interactions

The 2 kilobase (kb) upstream promoter regions of *DOXC21-A* and *DOXC21-B* were analyzed for *cis*-regulatory elements using the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Known *cis*-elements associated with abiotic stress responses, hormone signaling, and TF binding (e.g., ABRE, MYB, MYC, W-box, GC-box, CCAAT-box) were identified and annotated. To assess stress-responsive expression profiles, transcript abundance under abiotic stress conditions (e.g., drought, salt, cold, oxidative stress) was obtained from publicly available Affymetrix ATH1 microarray datasets integrated in the eFP Browser (https://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) [17]. Putative protein–protein interaction (PPI) networks for DOXC21-A and DOXC21-B were obtained from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) v11.5 database (https://string-db.org/). PPI categories included experimental data, co-expression, gene neighborhood, and database-curated interactions.

2. Plant material and growth conditions

2.1. *Arabidopsis thaliana* lines and used genotypes

A. thaliana Columbia-0 (Col-0) wild type (WT) and T-DNA insertional mutant lines in Col-0 genetic background were used in this study. The following mutant alleles were selected: *nat1* (SALK_073389) as the first insertional line within the NAT sequence, *nat2* (SALK_201039) as second insertional line within the NAT sequence, *doxc21-a* (Wiscseq_DsLox387A07.2) for *DOXC21-A* gene, *doxc21-b-1* (SALK_070568) for the *DOXC21-B* gene, and *doxc21-b-2* (SALK_127289) as an independent line for *DOXC21-B*. Seeds were obtained from the Nottingham Arabidopsis Stock Centre (https://arabidopsis.info/). Genotyping validation is shown in Supplementary Fig. S1, and primer sequences used for PCR-based genotyping are provided in Supplementary Tab. S1. T-DNA flanking sequences were verified by Illumina sequencing (Genomed S.A., Warsaw, Poland).

2.2. Soil cultures and abiotic stress treatments

Prior to sowing, *A. thaliana* seeds of Col-0 WT, *nat1*, *nat2*, *doxc21-a*, *doxc21-b-1*, and *doxc21-b-2* were stratified for 4 days in 4°C on water-saturated Whatman filter paper in sealed Petri dishes to synchronize germination. Plants were grown in a soil mix consisting of a peat-free commercial substrate for seedlings and herbs (COMPO SANA®, COMPO, Germany) combined with vermiculite (3–6 mm granules) at 3:1 (v/v) ratio. For drought stress experiments, each pot filled with soil was weighted and adjusted to ensure uniform mass (120±0.5 g) across biological replicates before sowing, thereby standardizing initial water content (WC) and ensuring consistent drought stress conditions. Pots were arranged in collective trays and placed in growth chambers under a 16 h light / 8 h dark photoperiod, with a light intensity of 120 µmol m^{-2} s^{-1}, and temperatures maintained

at 22°C during the light phase and 20°C during the dark phase, respectively. Twenty-eight pots per tray were randomized within genotypes to minimize position effect on plant phenotype.

Drought stress was initiated at third week of growth by withholding irrigation for the following three weeks. After three weeks of dehydration, plants were rehydrated with 3 L of water per tray. Recovery rate was scored visually, one day after re-watering, based on the presence of green, viable rosette leaves. A minimum of 18 plants per genotype was evaluated. Fresh weight (FW) of aerial tissue was measured 3 days post-rehydration using analytical balance. Dry weight (DW) was obtained after drying FW samples at 37°C for one week until a constant weight was reached. WC was calculated using the formula: WC [%] = (FW-DW/FW) x100 (n=8). In parallel, electrolyte leakage was measured to assess membrane integrity. Freshly harvested leaves were weighed and incubated overnight in 20 mL of Milli-Q water on a rotary shaker (114 rpm) at room temperature. Conductivity of 6 mL of the incubation solution was determined using a conductivity meter (HANNA EDGE™, USA), and values were normalized to leaf FW. Each measurement was performed on at least eight biological replicates.

2.3. *In vitro* cultures

For *in vitro* cultivation, *A. thaliana* seeds of Col-0 WT, *nat1*, *nat2*, *doxc21-a*, *doxc21-b-1*, and *doxc21-b-2* were surface sterilized by soaking in 70% ethanol for 2 min, followed by treatment with 5% calcium hypochlorite for 8 min. Seeds were subsequently rinsed three times with autoclaved deionized water. To facilitate even distribution during sowing, sterilized seeds were suspended in 0.1% (w/v) agar in deionized water. Sterilized seeds were sown on Petri dishes containing half-strength Murashige and Skoog (½ MS, pH 5.7) basal medium (Sigma-Aldrich, USA) supplemented with 1% (w/v) sucrose and solidified with 0.7% (w/v) plant culture-grade agar (Sigma-Aldrich, USA). Seeds were stratified on plates for 4 days at 4°C in darkness, then transferred to a controlled growth chamber (16 h light/ 8 h dark; light intensity 70 µmol m^{-2} s^{-1}; 22°C day/ 20°C night) and incubated for 3 weeks. Osmotic stress *in vitro* was induced by supplementing the culture medium with 200 mM mannitol from the beginning of the cultivation period.

2.4. Hydroponic cultures

A. thaliana seeds of Col-0 WT, *nat1*, *nat2*, and *doxc21-a* were germinated on tip-cut microcentrifuge tubes filled with 0.65% agar-solidified 10× Heeg solution and placed into square Petri dishes containing liquid 10× Heeg for germination. Here and throughout the hydroponic experiments, we used a modified 10× Heeg nutrient solution [27], with modifications adapted from Siwinska *et al.* [28]. In this formulation, the concentration of micronutrients was increased tenfold compared to the standard 1x Heeg solution (see Supplementary Tab. S2). To enhance the phenotypic responses and simulate environmentally relevant conditions, plants were grown at pH 4.5 - a condition known to affect nutrient availability, root architecture, hormone signaling, and secondary metabolism [29]. pH stability was maintained using 2-(N-morpholino)ethanesulfonic acid (MES) buffer. Seeds were stratified at 4°C in darkness

for 4 days, then transferred to controlled environment chambers set to a 16 h light / 8 h dark photoperiod (100 µmol m^{-2} s^{-1} light intensity), with temperatures regimes of 22°C (day) and 20°C (night). After 10 days, seedlings were transferred into larger hydroponic boxes containing fresh 10x Heeg solution.

Following 3 weeks of growth under control conditions, polyethylene glycol (PEG)-induced osmotic stress treatments were initiated. The nutrient solution was replaced with fresh 10x Heeg supplemented with 3% (w/v) PEG. Nutrient solutions were replenished regularly to maintain both nutrient and osmotic consistency throughout the experiment. After 2 weeks of stress exposure, rosette leaves were collected for RNA isolation. Plants were harvested after 3 weeks in stress conditions for FW determination and anthocyanin content analysis.

3. Gene Expression Analysis

3.1. RNA isolation

Total RNA was isolated from rosette leaves of *A. thaliana* Col-0 WT, *nat1, nat2*, *doxc21-a* and *doxc21-b* mutant lines, using the GeneMATRIX Universal RNA Purification Kit (EURx, Poland) following the manufacturer's protocol. Leaf tissue was homogenized under liquid nitrogen with stainless-steel beads to ensure through cellular disruption. An on-column DNase digestion step was included to eliminate genomic DNA contamination. RNA concentration and purity were initially assessed using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA).

3.2. Quantitative real-time PCR

First-strand complementary DNA (cDNA) synthesis was performed using 1 µg of total RNA and the LunaScript® RT Master Mix Kit (NEB, USA), following the manufacturer's instructions, with oligo d(T) primers. Primer specificity was verified using Primer-BLAST (NCBI, https://www.ncbi.nlm.nih.gov/tools/primer-blast/) and validated experimentally by melt curve analysis. Amplification efficiency was calculated from five-point serial dilution series of pooled cDNA. A final 8-fold cDNA dilution was selected for subsequent analyses as it provided optimal signal linearity across targets. Reactions were conducted using the Luminaris Color HiGreen qPCR Master Mix (Thermo Scientific, USA), and cycling conditions were as follows:

- Uracil-DNA glycosylase (UDG) activation: 50°C for 2 min
- Initial denaturation: 95°C for 10 min
- Amplification (40 cycles):
    - Denaturation: 95°C for 10 s
    - Annealing/extension: 60°C for 5 s with fluorescence acquisition
- Final cooling: 40°C for 30 s
- Melt curve analysis: continuous fluorescence acquisition from 60°C to 97°C

All reactions were performed in triplicate (technical replicates) for each biological replicate (n=4). Primer sequences are listed in Supplementary Tab. S1.

3.2.1. *DOXC21-A* and *DOXC21-B* expression in Col-0 WT

To evaluate *DOXC21-A* and *DOXC21-B* transcript levels, qRT-PCR was conducted on Col-0 WT plants cultivated under two distinct abiotic stress conditions: (I) *in vitro* on ½ MS medium supplemented with 200 mM mannitol (osmotic stress) and (II) in soil under drought stress. Reactions were carried out on the CFX96 Real-Time PCR Detection System (Bio-Rad, USA). Expression levels were calculated using the $2^{-\Delta CT}$ method, with normalization to the *ACT2* (*ACTIN 2*) reference gene. Three biological replicates were analyzed for *in vitro* grown plants and five for soil grown plants. Each biological replicate was tested in three technical replicates. Statistical significance was assessed using two-way ANOVA followed by Bonferroni's multiple comparisons test.

3.3. Transcriptome profiling (RNA-Seq) and expression validation

3.3.1. RNA isolation, library construction and quality assessment

Transcriptomic profiling was performed for *A. thaliana* Col-0 WT and *nat1, nat2,* and *doxc21-a* mutant lines cultivated hydroponically under osmotic stress conditions (10× Heeg, pH 4.5, 3% [w/v] PEG). Rosette leaves were harvested one week after stress induction from five biological replicates of *doxc21-a* and six replicates of Col-0 WT, *nat1,* and *nat2* lines. Total RNA was isolated from rosette leaves and quality controlled was done as previously described. RNA integrity was evaluated *via* agarose gel electrophoresis and the absence of DNA contamination was confirmed by qPCR on non-reverse-transcribed RNA samples.

Samples passing quality control were shipped on dry ice to the external sequencing provider weSEQ.IT (Rybnik, Poland; https://weseq.it), where final RNA integrity verification, library preparation, sequencing, and bioinformatic analysis were performed.

RNA quality was re-assessed using the Agilent 2100 Bioanalyzer system (Agilent Technologies, CA, USA) with the RNA 6000 Nano Assay Kit. For each library, 1 µg of total RNA was used as input. Libraries were constructed using the NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (New England Biolabs, USA), according to the manufacturer's protocol. Unique index codes were incorporated into each sample for multiplexing. Briefly, mRNA was enriched using poly-T oligo-attached magnetic beads and fragmented at elevated temperature using divalent cations. First Strand Synthesis Reaction Buffer（5X). First strand cDNA synthesis was performed using random hexamer primers and M-MuLV Reverse Transcriptase, followed by second-strand synthesis using DNA Polymerase I and RNase H. End repair was conducted to generate blunt ends, followed by adenylation of 3' ends and ligation of NEBNext adaptors with hairpin structures. Fragment size selection (~240 bp) was performed using AMPure XP beads (Beckman Coulter, USA). USER enzyme treatment (37°C for 15 min; 95°C for 5 min) was used to remove uracil containing bases from adaptors. PCR enrichment of adaptor ligated DNA fragments was conducted using Phusion High-Fidelity DNA Polymerase and indexed primers. Final libraries were purified using AMPure XP beads and validated for quality and fragment size distribution using the Agilent Bioanalyzer 2100 system.

3.3.2. RNA sequencing and transcriptome analysis

Following library quality assessment, indexed RNA libraries were clustered on a cBot Cluster Generation System using the TruSeq PE Cluster Kit v4-cBot-HS (Illumina, USA) and sequenced on the Illumina NovaSeq 6000 platform to generate 150 bp paired-end reads (PE150). Raw sequencing data in FASTQ format were filtered using custom in-house Perl scripts to remove adapter sequences, poly-N regions, and low-quality reads. The quality of clean reads was assessed using FastQC and RSeQC, evaluating parameters such as GC content, base quality scores (Q20, Q30), and mapping statistics.

High-quality reads were aligned to the *A. thaliana* reference genome (TAIR10.45) using HISAT2. Transcript assembly and quantification were performed with StringTie. Expression values were normalized and reported as FPKM (fragments *per* kb of transcript *per* million mapped reads). Differential gene expression analysis was conducted using DESeq2 and edgeR, applying a significance threshold of $\log_2$ fold change (FC) ≥ 1 and an adjusted p-value<0.05.

Alternative splicing (AS) events were identified using ASprofile, which classifies splicing events such as exon skipping, intron retention, and usage of alternative splice donor/acceptor sites.

Functional annotation of differentially expressed genes (DEGs) was carried out through sequence alignment against multiple biological databases, including NCBI non-redundant protein (Nr) and nucleotide (Nt) databases, Pfam, KOG/COG, Swiss-Prot, Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG). These annotations supported the functional interpretation of transcriptomic changes and enabled pathway enrichment analysis.

DESeq2 implemented a model based on the negative binomial distribution, using shrinkage estimation to improve variance and $\log_2$ FC estimates. *P*-values were adjusted for multiple testing using the Benjamini–Hochberg method, with FDR (False Discovery Rate)-adjusted p-value<0.05 considered significantly significant.

GO enrichment analysis was conducted using the GOseq R package, which corrects for gene length bias *via* the Wallenius non-central hypergeometric distribution [30]. For pathway enrichment, KEGG analysis was performed using the KOBAS tool [31], against the KEGG database [32].

To investigate putative protein level interactions among DEGs, protein–protein interaction (PPI) networks were predicted by mapping DEG protein sequences *via* BLASTx to orthologs in the STRING database (https://string-db.org/). Visualization and topological network analyses were performed using Cytoscape [33].

3.3.3. RNA-seq validation

To validate transcriptomic data obtained from RNA-seq, qRT-PCR was conducted using cDNA from hydroponically grow plants subjected to 3% (w/v) PEG-induced osmotic stress (pH 4.5). Reactions were run on the Roche® LightCycler® 480 system (Roche, Switzerland), using the Luminaris Color HiGreen qPCR Master Mix (Thermo Scientific, USA). Expression values were normalized against two internal controls: *ACT2* and *EF1-α* (*Elongation Factor 1-alpha*),

and expression was calculated using the $2^{-\Delta CT}$ method, normalized to the geometric mean of both reference genes. Four biological replicates *per* genotype and treatment condition were analyzed, with three technical replicates each (n=12). Statistical analysis was performed *via* two-way ANOVA followed by Bonferroni's post hoc test for multiple comparisons.

4. Untargeted metabolomics analysis

4.1. Metabolite extraction

Leaf and root tissues were harvested from *A. thaliana* Col-0 WT, *nat1, nat2*, and *doxc21-a* plants cultivated hydroponically under control (pH 4.5) and PEG-induced osmotic stress conditions (3% [w/v] PEG, pH 4.5) following three weeks of treatment. Plant material was collected, immediately frozen in liquid nitrogen and homogenized using pre-chilled mortars and pestles. For each biological replicate, approximately 100 ± 2 mg of powdered leaf tissue or 50 ± 2 mg of root tissue was weighted. When tissue quantity was insufficient, samples from the same genotype and treatment group were either pooled or the final extraction volume was proportionally adjusted to maintain consistent solvent-to-tissue ratios across all samples.

Metabolites were extracted using 0.8 mL of 80% (v/v) methanol. The homogenates were sonicated for 10 min at 70% amplitude using an ultrasonic bath equipped with a sweep function, followed by overnight incubation at 4°C in the dark. The next day, samples were centrifuged at 16,000 × g for 15 min at 4°C, and the supernatants were transferred into fresh vials. To improve extraction efficiency, a second round of extraction was performed by adding an additional 0.75 mL of 80% methanol to the residual pellet, followed by vortexing, sonication, and centrifugation as described above. Supernatants from both extraction steps were combined and subsequently evaporated to dryness using a vacuum concentrator. Dried metabolite extracts were stored at -20°C and shipped to the analytical facility (Laboratoire Agronomie et Environnement, Université de Lorraine-INRAE, France) for untargeted metabolomic profiling using UHPLC Coupled to a high-resolution Orbitrap mass spectrometer.

4.2. Orbitrap Mass Spectrometry conditions

Liquid chromatographic analyses were performed on a quaternary pump Vanquish UHPLC system, with an autosampler and a temperature-controlled column oven. We used a core-shell Kinetex XB-C18 column (150 × 2.1 mm, 2.6 µm) (Phenomenex Inc., Torrance, CA, USA) to separate the metabolites contained in the extracts, using a mobile phase gradient composed of 0.1% formic acid in water (A) and 0.1% formic acid in methanol (B) at a flow rate of 200 µl min^{-1}, at a column temperature of 40°C. The elution program consisted of starting with 10% B for 2 min, then linearly increasing from 10 to 30% B in 10 min, then to 95% B in 10 min. The column was rinsed for 5 min with 95% B and re-equilibrated to the initial conditions for 4 min prior to the next run. The samples were analyzed randomly.

High-resolution mass spectrometry detection was performed on an Orbitrap ID-X mass spectrometer (ThermoFisher Scientific, Bremen, Germany) in positive and negative electrospray

ionization (ESI) modes. The capillary voltages were set at 3.5 and 2.5 kV for positive and negative modes. The source gases were set (in arbitrary unit min^{-1}) at 40, 8 and 1 (for, respectively, sheath, auxiliary and sweep gas) and the vaporizer temperature was 320°C. Full-scan mass spectra were acquired from 120 to 1200 m/z at a resolution of 60 000.

MS/MS analysis was performed for characterization on a QC sample using AcquireX workflow (ThermoFisher TM software) consisting of a succession of analyses in order to collect MS2 spectra for most of the compounds in the sample. MS/MS spectrum was acquired at a resolution of 15 000 after a higher-energy C-trap dissociation and a collision-induced dissociation of the MS1 signals.

4.3. Molecular network analysis

For molecular network analysis, the processed feature table containing m/z values, retention times, and peak areas was used to generate a molecular similarity network based on MS/MS spectral alignment. The workflow was performed in accordance with the procedure described at https://github.com/Florent-1/Molecular-Networks-from-Compound-Discoverer-to-Cytoscape. Network construction was based on spectral similarity scores, linking features with comparable fragmentation patterns. The resulting molecular network was visualized in Cytoscape v3.9.1, where nodes represent individual metabolite features and edges reflect spectral relatedness, thus allowing for the identification of compound clusters potentially related to DOXC21-dependent metabolic pathways.

5. Heterologous expression and functional enzymatic activity assays

5.1. Cloning and *Agrobacterium tumefaciens*-mediated expression in *Nicotiana benthamiana*

To assess the functional roles of *DOXC21-A* and *DOXC21-B,* transient heterologous expression assays were conducted in *N. benthamiana* leaves *via A. tumefaciens*-mediated infiltration. Full-length open reading frames (ORFs) of both genes were amplified from *A. thaliana* Col-0 WT cDNA using Phusion™ High-Fidelity DNA Polymerase (Thermo Scientific, USA). Primer sequences and amplification conditions are listed in Supplementary Tab. S1.

The resulting PCR products were cloned into the pCR8 entry vector using the pCR8®/GW/TOPO® TA Cloning Kit (Invitrogen, USA) as described by Vialart *et al*. [34]. Verified entry clones were subsequently recombined into the binary expression vector pBIN-GW *via* Gateway® LR Clonase® II (Invitrogen, USA), yielding pBIN-GW:*DOXC21-A* and pBIN-GW:*DOXC21-B* constructs.

Expression constructs were introduced into *A. tumefaciens* strain GV3101 *via* electroporation. Additional constructs used included pBIN:*GFP* (green fluorescent protein as expression control) and pBIN61-*P19* (a viral silencing suppressor). Transformed *A. tumefaciens* strains were cultured in LB medium containing kanamycin (50 mg/L), rifampicin (50 mg/L), and gentamycin (25 mg/L) at 28°C for 48 h with shaking (200 rpm). Bacterial cultures were harvested by centrifugation (4 000 g, 10 min), and pellets were washed three times in sterile distilled water to remove antibiotic residues.

Bacterial pellets were resuspended in infiltration buffer (5 mM $MgCl_2$, 5 mM MES, pH 5.7) to an optical density (OD_{600}) of 0.2 for the *DOXC21* constructs and pBIN:*GFP*, and 0.4 for pBIN61-*P19*. For co-infiltration, appropriate bacterial cultures were mixed and introduced into fully expanded leaves of 4–6-week-old *N. benthamiana* plants using 1 mL needleless syringe. Each construct was tested on a minimum of six plants (three leaves *per* plant). Negative controls included infiltration buffer only (mock), or pBIN-GW (empty vector) together with pBIN:*GFP* and pBIN61-*P19*.

Following infiltration, plants were incubated in the dark overnight in sealed boxes to enhance transformation efficiency, then transferred to standard growth conditions (16 h light / 8 h dark; 22°C / 20°C). Five days after infiltration, successful transformation was confirmed by GFP fluorescence. Infiltrated leaf tissues were harvested, frozen in liquid nitrogen, and processed for metabolite extraction using 80% methanol as previously described. Dried extracts were submitted for untargeted metabolomic profiling at the Laboratoire Agronomie et Environnement, Université de Lorraine-INRAE (France).

5.2. Protein overproduction and purification

To evaluate the enzymatic properties of DOXC21-A and DOXC21-B, heterologous protein expression was performed in *Escherichia coli.* Coding sequences previously cloned into the pCR8 entry vector were subcloned into the pET28a(+) expression vector (Invitrogen, Thermo Fisher Scientific) *via Eco*RI for *DOXC21-A*, and *Bam*HI and *Xhol*I for *DOXC21-B* restriction sites, followed by T4 DNA ligase mediated ligation. Resulting constructs (pET28a+:*DOXC21-A* and pET28a+:*DOXC21-B*) were introduced into *E. coli* Rosetta 2 cells *via* electroporation.

Transformed colonies were selected on LB agar containing 100 mg/L kanamycin and 33 mg/L chloramphenicol. Single colonies were used to inoculate into 10 mL of LB starter cultures supplemented with the same antibiotics and grown overnight at 37°C with shaking (200 rpm). The next day, 2 mL of overnight culture was transferred into 1 L of fresh LB medium and cultured under identical conditions until the OD_{600} reached 0.4–0.6. Protein expression was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and incubation continued overnight at 37 °C.

Cells were harvested by centrifugation at 4 000 g for 20 min at 4°C and resuspended in 1 mL of lysis buffer (0.1 M Tris-HCl pH 8.0; 70 µM lysozyme; 0.2 M NaCl). Following brief vortexing, cells were sonicated on ice (5 cycles of 30 sec) and lysates were centrifuged at 20 000 g for 20 min at 4 °C. Insoluble pellets were resuspended in 2 mL of denaturing buffer (0.1 M Tris-HCl pH 8.0; 0.2 M NaCl; 8 M urea; 1 mM DTT) using a Potter-Elvehjem homogenizer. The denaturated proteins were subjected to a two-step dialysis renaturation protocol using dialysis tubing (20 kDa MW cut off) [35].

Dialysis was performed against 1 L of renaturation buffer 1 (0.1 M Tris-HCl pH 8.0; 1M urea; 1 mM DTT) for 6 h, followed by renaturation buffer 2 (0.1 M Tris-HCl pH 8.0; 1 mM DTT)

overnight, both at 4°C with gentle stirring. After dialysis, samples were centrifuged again (20 000 g, 20 min, 4°C), and supernatant were concentrated using Amicon® Ultra-15 centrifugal filter units (10 kDa MWCO; Merck Millipore, Germany).

His-tag purification of recombinant proteins was carried out using Ni-NTA affinity chromatography. The efficiency of expression, solubilization and partial renaturation was assessed by sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) (Supplementary Fig. S2). Protein samples from each purification step were separated on 12% SDS-PAGE and stained with Coomassie Brilliant Blue. The presence of ~40–45 kDa bands in renatured and purified fractions confirmed successful expression and partial recovery of DOXC21-*A* and DOXC21-B.

5.3. *In vitro* enzymatic activity assays

To assess the catalytic potential of DOXC21-A and DOXC21-B, *in vitro* enzymatic assays were conducted using renatured, purified recombinant proteins. Reactions were carried out in a total volume of 100 µL, containing 0.1 M Tris-HCl buffer (pH 6.0) and supplemented with essential cofactors for 2OGD activity at saturated concentrations: (I) 0.5 mM $FeSO_4$ (ferrous sulfate), (II) 5 mM 2-oxoglutarate (2-OG), and (III) 5 mM ascorbic acid. Test substrates were added at a final concentration of 100 µM and included a chemically diverse panel of plant-derived compounds listed in **Tab. 1**, spanning phenylpropanoid intermediates (e.g., cinnamic acid derivatives, CoA esters), coumarins and flavonoids, monolignols, indole-molecules, and phytohormones.

Reactions were initiated by the addition of 20–40 µg of purified recombinant DOXC21-A or DOXC21-B protein and incubated at 28°C for 24 h in dark to avoid light-induced degradation. All assays were performed in triplicate. Negative controls were included for each substrate and consisted of parallel reactions lacking either the recombinant protein or 2-OG to assess spontaneous degradation or background reactivity. Reactions were terminated by the addition of 10 µL of 1M HCl, followed by vortexing and centrifugation at 10 000 g for 10 min at 4°C. Supernatants were filtered (0.2 µM) and transferred to fresh vials for analysis. Metabolite profiles were analyzed using UHPLC. Potential enzymatic activity was assessed by monitoring substrate depletion and/or product formation, using extracted ion chromatograms and comparison to authentic standards or database predictions.

Table 1. Classification of substrates used for DOXC21-A and DOXC21-B *in vitro* assays.
The table lists all chemical compounds either already tested or chosen for upcoming *in vitro* enzymatic activity assays with recombinant DOXC21-A and DOXC21-B proteins. Substrates are grouped into relevant biosynthetic or chemical classes, including phenylpropanoids, coumarins, flavonoids, phytohormones, and other specialized metabolites. The selection reflects key branches of secondary metabolism potentially modulated by DOXC21 enzymes. Underlined compounds were tested with DOXC21-A only, while those marked with an asterisk ('*') indicate substrates selected for future *in vitro* testing with DOXC21-A. This classification supports ongoing biochemical investigation of DOXC21 substrate specificity and functional relevance in pathways such as lignin biosynthesis, coumarin biosynthesis, and indole/hormonal metabolism.

Category	Compound Name
Monolignols & aldehydes	Coniferyl alcohol; Sinapyl alcohol; Coniferaldehyde; Sinapaldehyde
Cinnamic acid derivatives	trans-Ferulic acid; Isoferulic acid; 5-Hydroxyferulic acid; Sinapic acid; Cinnamic acid; Caffeic acid; 3,4-Dimethoxycinnamic acid; 3,5-Dimethoxycinnamic acid; 2,3,4-Trimethoxycinnamic acid; 3,4,5-Trimethoxycinnamic acid; o-Coumarate; p-Coumarate; o-Coumarate glucoside
Simple coumarins	Umbelliferone, Coumarin, Scopoletin, Isoscopoletin, Esculetin, Daphnetin; 6-Hydroxycoumarin; 5,7-Dihydroxycoumarin; 6,7,8-Trihydroxycoumarin; 4-Methylumbelliferone; 5-Hydroxy-7-methoxycoumarin; 5-Methoxy-7-hydroxycoumarin
Methoxylated/modified coumarins	Scoparone, Dimethylfraxetine, Daphnetin dimethylether; Daphnetin-7-methylether; Daphnetin-8-methylether; Fraxetin, Fraxidin, Isofraxidin, Osthol, Limettin, Auraptene; 7-Methylsuberosine, O-Prenylumbelliferone, Suberosine
Isoprenylated coumarins	6,7-Diisoprenyloxycoumarin; 6-Isoprenyloxy-7-methoxycoumarin; 5-Geranyloxy-7-methoxycoumarin
Glycosylated coumarins	Esculin, Fraxin, Scopolin, Skimmin
Other benzenoids	4-Hydroxyherniarine; Columbianetin, Colombianadine
Flavonoids	Luteolin, Apigenin, Eriodictyol, Naringenin, Taxifolin, Pinocembrin, Isoliquiritigenin, Sakuranetin, Flavanone; 5-Hydroxyflavanone; 6-Hydroxyflavanone; 7-Hydroxyflavanone
Phytohormones and precursors	Salicylic acid, Benzylaminopurine (cytokinin), Indoleacetamide (IAA precursor), Indole-3-butyric acid (IBA)
Amino acids	Tryptophan
Oxylipins	16-HPOT, 18-HPOT
Phenylpropanoid-CoA esters	Caffeoyl-CoA, *p*-Coumaroyl-CoA, Feruloyl-CoA
Indole derivatives*	Indole-3-carbaldehyde, Indole-3-carboxylic acid, Camalexin, Malonyltryptophan, Acetyltryptophan, Melatonin

RESULTS

1. *DOXC21-A* and *DOXC21-B* are conserved across land plants and linked to secondary metabolism evolution

Kawai *et al.* [15] proposed a classification of plant 2OGDs into three major clades: DOXA, DOXB and DOXC. The DOXC clade include genes encoding enzymes involved in specialized metabolism, particularly the oxidative transformation of secondary metabolites. Our previous study on the functional characterization of scopoletin 8-hydroxylase (S8H), a 2OGD enzyme involved in fraxetin biosynthesis [28], led us to investigate two closely related genes: *AT3G19000* and *AT3G19010*. We performed the phylogenetic analysis on available nucleic acid sequences of genes encoding 2OGDs from *A. thaliana* genome and revealed that these two enzymes form a distinct clade. It was in accordance with classification previously done by Kawai *et al.* [15], which also pointed that DOXC21 is one of the few clades conserved across all major land plant lineages – including bryophytes (*Physcomitrium*), lycophytes (*Selaginella*), gymnosperms (*Picea*),

and angiosperms (*Oryza, Arabidopsis*) – suggesting its evolutionary persistence since the colonization of terrestrial habitats.

We named the studied genes *DOXC21-A* (*AT3G19000*) and *DOXC21-B* (*AT3G19010*), following the classification system for the DOXC21 clade.

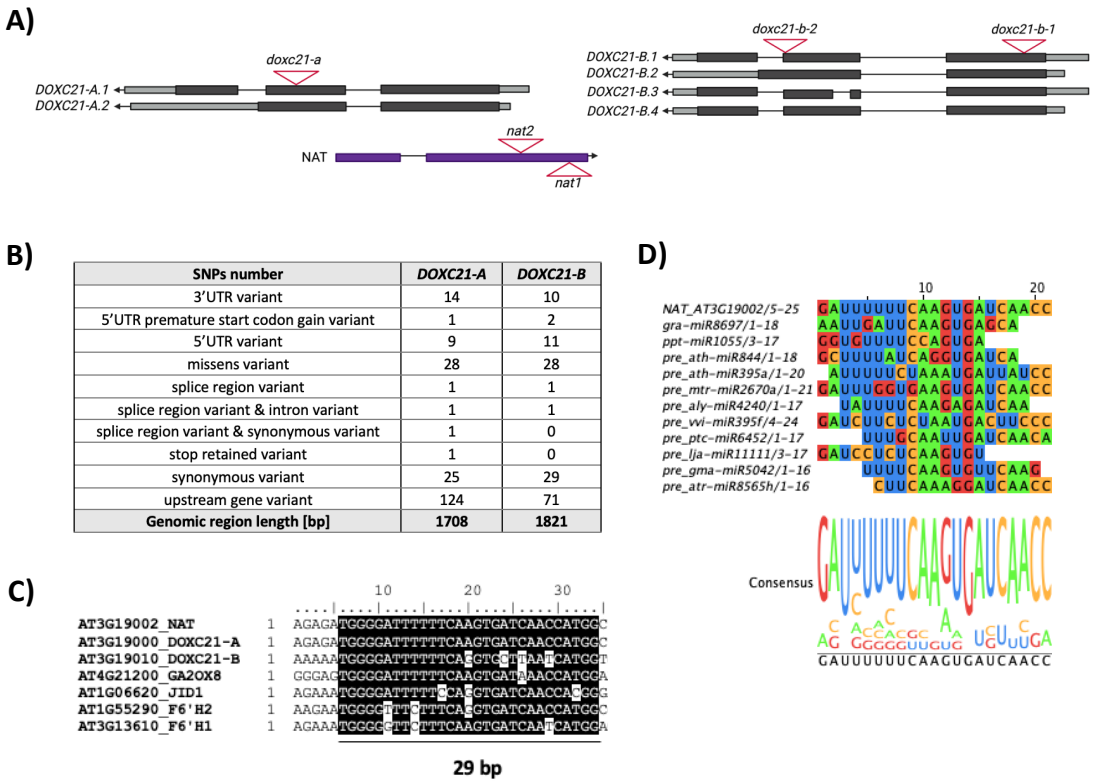
1.1. Conservation and antisense regulation of *DOXC21-A* in the *A. thaliana* genome

Gene structure analysis (**Fig. 1A**) revealed that *DOXC21-A* encodes two transcript isoforms (five exons, three introns) and is uniquely overlapped by a long non-coding NAT, *AT3G19002*. This NAT spans the 5′ region of *DOXC21-A*, includes seven exons, five annotated coding sequences (CDSs), and overlaps untranslated regions (UTRs). In contrast, *DOXC21-B* (*AT3G19010*) shows greater transcript complexity with four isoforms (13 exons, 9 introns) but lacks any antisense transcript overlap. Sequence query coverage for *DOXC21-A* and *DOXC21-B* was 79%, with 72.2% identity. Structural data were retrieved from the BAR ePlant and ThaleMine database (University of Toronto).

SNP analysis (**Fig. 1B**) showed that both genes display a comparable number of 28 missense variants, but *DOXC21-A* harbors more upstream regulatory SNPs (124 vs. 71), indicating higher potential for transcriptional variability, suggesting increased regulatory variation. Notably, *DOXC21-B* includes a unique stop-retained variant specific to the Rennes-11 accession.

To evaluate functional conservation across the 2OGD family, a 35-bp region from the *DOXC21-A* NAT was aligned to the *A. thaliana* genome. This revealed a highly conserved 29-bp core sequence highly similar to *GA2OX8* (*AT4G21200*) (28/29), *JID1* (*AT1G05260*), *F6'H2* (*AT1G55290*), *F6'H1* (*AT3G13610*) (all 26/29), and *DOXC21-B* (25/29) (**Fig. 1C**). This conservation suggests potential regulatory crosstalk among dioxygenases *via* shared sequence regulatory motifs or post-transcriptional interactions.

A further search in miRBase uncovered a conserved 21-nt region within the NAT (GAUUUUUUCAAGUGAUCAACC) that aligns with multiple known plant miRNAs (**Fig. 1D**), including gra-miR8697 (*G. raimondii*), ppt-miR1055 (*P. patens*), and ath-miR844 (*A. thaliana*), the latter being implicated in biotic stress signaling [36]. Partial alignments were also found in species such as *M. truncatula, V. vinifera, P. trichocarpa, L. japonicus, G. max,* and *A. trichopoda*, indicating deep evolutionary conservation and suggests a conserved functional core potentially linked to miRNA mimicry or regulatory crosstalk.



SNPs number	*DOXC21-A*	*DOXC21-B*
3'UTR variant	14	10
5'UTR premature start codon gain variant	1	2
5'UTR variant	9	11
missens variant	28	28
splice region variant	1	1
splice region variant & intron variant	1	1
splice region variant & synonymous variant	1	0
stop retained variant	1	0
synonymous variant	25	29
upstream gene variant	124	71
Genomic region length [bp]	**1708**	**1821**

Figure 1. Genetic and regulatory conservation of the *DOXC21* locus in *A. thaliana*. **(A)** Schematic representation of gene structures and transcript isoforms of *DOXC21-A* and *DOXC21-B* in *A. thaliana*. Exons are shown in dark gray, while the NAT (*AT3G19002*) overlapping the 5′ region of *DOXC21-A* is shown in purple. Red triangles indicate T-DNA insertion sites: *doxc21-a* disrupts an exon of *DOXC21-A*; *nat1* and *nat2* insertions localize within the NAT; *doxc21-b-1* and *doxc21-b-2* disrupt the first and second exons of *DOXC21-B*, respectively. **(B)** Summary table of SNPs identified in the *DOXC21-A* and *DOXC21-B* genomic regions across 2029 *A. thaliana* accessions. SNPs were extracted using a custom R script (GitHub: SNP_extractor). **(C)** Multiple sequence alignment of a 35-bp region from the NAT (*AT3G19002*) with related 2-oxoglutarate-dependent dioxygenase (2OGD) family genes in *A. thaliana*, including *DOXC21-A*, *DOXC21-B*, *GA2OX8*, *JID1*, *12*, and *F6'H1*. A 29-bp conserved core is highlighted, black boxes indicate conserved residues, and white letters denote mismatches. **(D)** Sequence alignment and sequence logo of positions 5–20 from the conserved NAT segment, compared with homologous regions in predicted or annotated plant microRNAs. Included are sequences from *Gossypium raimondii* (gra-miR8697), *Physcomitrium patens* (ppt-miR1055), *A. thaliana* (pre-miR844, pre-miR395a), *Medicago truncatula* (pre-miR2670a), *Arabidopsis lyrata* (pre-miR4240), *Vitis vinifera* (pre-miR395f), *Populus trichocarpa* (pre-miR6452), *Lotus japonicus* (pre-miR11111), *Glycine max* (pre-miR5042), and *Amborella trichopoda* (pre-miR8565h). The WebLogo illustrates base conservation across species.

1.2. Phylogenetic analysis reveals ancient divergence of DOXC21-A and DOXC21-B

To investigate the evolutionary history of DOXC21-A and DOXC21-B, we performed a phylogenetic analysis of homologous amino acid sequences retrieved from the *A. thaliana* Gramene protein database using a BLASTp identity cutoff of ≥28%. Sequences were aligned using MUSCLE, and a Neighbor-Joining (NJ) tree was constructed with 500 bootstrap replicates in MEGA11 using the JTT model with Gamma+I (**Fig. 2**). The resulting topology showed that DOXC21-A and DOXC21-B form a closely related sub-branch within distinct and well-supported clade with homologs from eudicots, monocots, and basal angiosperms, separate from canonical gibberellic

acid (GA) and flavonoid biosynthetic dioxygenases. Taxa span major plant lineages, including bryophytes, lycophytes, gymnosperms, monocots, and eudicots. The broad conservation of DOXC21 paralogs supports an ancient origin, while their phylogenetic separation from hormone-associated 2OGDs suggests functional divergence. Despite moderate and overall sequence identity, two *A. thaliana* DOXC21-A and DOXC21-B proteins clustered together in phylogenetic tree, suggesting that the corresponding genes likely originated from a gene duplication event – a common mechanism driving the expansion of the 2ODG family.

Notably, the DOXC21-specific clade is placed within a distinct branch of 2OGD dioxygenases, closely neighboring SAN1A-like proteins and DLO1/DLO2-related sequences from *G. max*. SAN1A-like proteins and DLO1/DLO2-related dioxygenases have been previously implicated in specialized metabolism and regulatory functions in response to stress, suggesting that the proximity of DOXC21s to these sequences may reflect a conserved role in adaptive phenylpropanoid pathway modulation. Although enzymes like F6'H1, F6'H2, and S8H – key players in phenylpropanoid metabolism – appear in a nearby but separate clade, their more distant placement suggests potential evolutionary divergence. The phylogeny was rooted with *M. polymorpha*, an early-diverging bryophyte, to provide a basal framework for interpreting the evolutionary context of the DOXC21 lineage.

Conservation analysis (Supplementary Fig. S3) highlighted discrete regions of high similarity, particularly in the N-terminal region and Fe^{2+}-binding domains. All analyzed proteins contained the conserved H–X–[DE]–X(2–20)–H motif, which is characteristic of catalytically active 2OGDs.

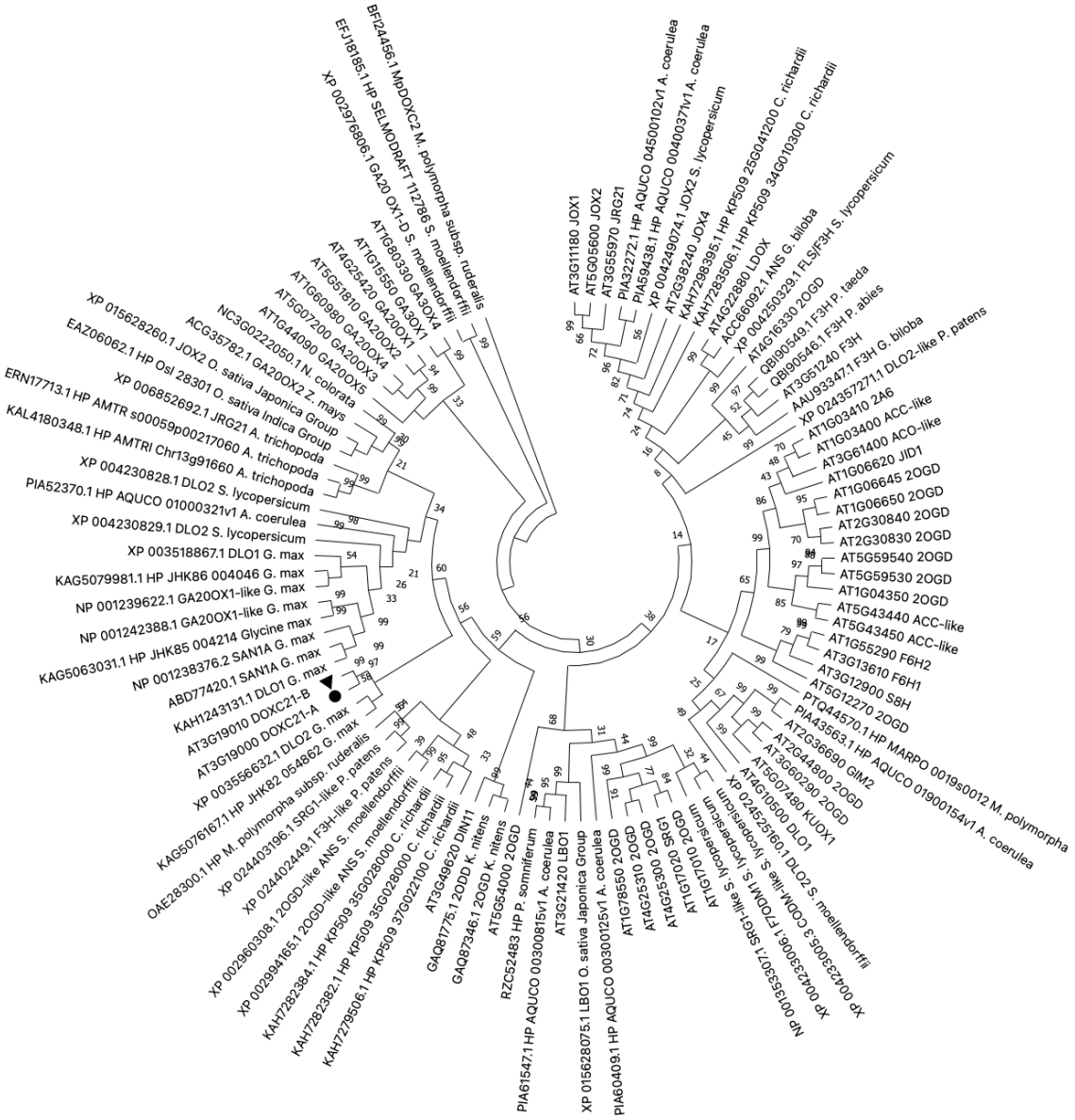


Figure 2. Phylogenetic reconstruction of DOXC21-A and DOXC21-B homologs across land plants.
A Neighbor-Joining (NJ) tree was generated from 100 protein sequences showing ≥28% identity to *A. thaliana* DOXC21-A (●) and DOXC21-B (▼), aligned using MUSCLE. Phylogenetic inference was performed in MEGA11 (JTT + G+I model, 500 bootstraps). The tree was rooted with a dioxygenase homolog from *Marchantia polymorpha* subsp. *ruderalis*, representing an early-diverging land plant.

1.3. Promoter regions of *DOXC21-A* and *DOXC21-B* contain stress-responsive *cis*-regulatory elements

To investigate regulatory features that might explain the expression dynamics of *DOXC21-A*, *DOXC21-B*, and NAT, we analyzed their promoter regions using PlantPromoterDB (https://ppdb.agr.gifu-u.ac.jp/ppdb/cgi-bin/index.cgi). Approximately 1 kb upstream sequences from the transcription start site were scanned for conserved TF binding motifs.

The *DOXC21-A* promoter contained several cis-elements associated with abiotic stress and hormone signaling, including ABRE-like (abscisic acid-responsive) motifs and DRE/CRT

elements linked to drought and cold responses. Binding sites for MYB and TF were also identified, suggesting responsiveness to both hormonal and biotic/abiotic cues. In contrast, the *DOXC21-B* promoter was enriched in elements associated with light signaling (e.g., G-box motifs), auxin response, and general growth regulation, indicating differential transcriptional control mechanisms. The NAT promoter also contained several stress-associated elements, including MYC and NAC recognition sequences, supporting its inducibility under abiotic stress.

1.4. Expression dynamics of *DOXC21-A* and *DOXC21-B* and predicted protein–protein interaction networks

To investigate the regulatory specialization of *DOXC21-A* and *DOXC21-B* under abiotic stress, we analyzed absolute transcript levels using the AtGenExpress microarray dataset. Temporal expression profiles in shoots and roots under various stress conditions (0–24 h) revealed contrasting patterns (**Fig. 3A**). At baseline, *DOXC21-A* was more highly expressed in shoots (234.1 units) than roots (98.9), while *DOXC21-B* expression was similar across tissues (~200 units). *DOXC21-A* was rapidly induced by cold stress, peaked early, and declined over time. It also showed dynamic, stress-specific regulation under drought, osmotic, and oxidative stress. UV-B strongly repressed its expression, while root expression declined progressively under oxidative and heat stress. In contrast, *DOXC21-B* was robustly upregulated by UV-B in both shoots and roots, with peak expression reaching 979.2 units in shoots at 3h. It was moderately induced by cold and wounding and downregulated by heat stress.

Predicted protein–protein interaction (PPI) networks (STRING v11.5) revealed distinct molecular context for each gene. The DOXC21-A network (**Fig. 3B**) was composed entirely of literature-derived associations and included partners involved in hormone biosynthesis (GA1), transcriptional regulation (WRKY22, WRKY27), redox balance (NPQ1, XRN3), and carotenoid metabolism (LUT1). Additional associations with PFK2 and STOP2 suggest links to energy metabolism and acid stress response.

The DOXC21-B network (**Fig. 3C**) displayed a broader and partially co-expression-supported profile, featuring oxidative signaling (OGOX2, NDA2), calcium-mediated responses (EDA39), and components of translational and nucleotide metabolism (OVA9, ZEU1, WRNEXO). Both paralogs were linked to GA1, indicating potential hormonal overlap.

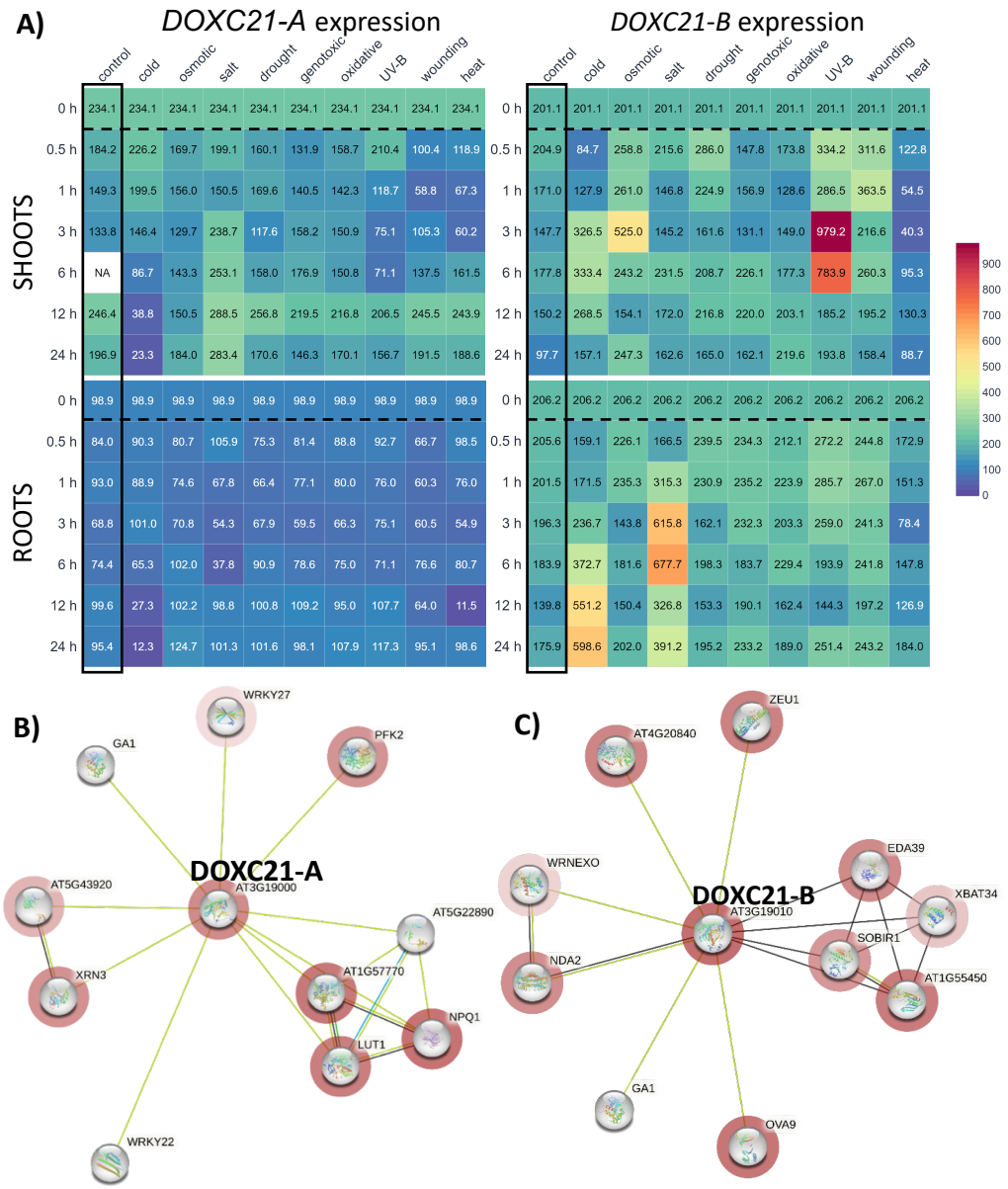


Figure 3. Differential expression *of DOXC21-A* and *DOXC21-B* and predicted protein interaction networks under abiotic stress. (A) Heatmaps of absolute transcript levels for *DOXC21-A* and *DOXC21-B* in shoots and roots under selected abiotic stresses (0–24 h), based on AtGenExpress ATH1 array data. Red and blue indicate high and low expression, respectively. *DOXC21-A* is responsive to cold, drought, and oxidative stress, particularly in shoots, whereas *DOXC21-B* shows strong UV-B-induced upregulation in both tissues. **(B, C)** STRING-derived protein–protein interaction (PPI) networks for DOXC21-A (B) and DOXC21-B (C), retrieved from STRING v11.5. Nodes represent proteins; edges indicate predicted associations. DOXC21-A interactors are primarily text-mined (yellow edges) and include partners linked to hormone biosynthesis, redox control, and photoprotection. DOXC21-B interacts with proteins involved in oxidative stress, calcium signaling, and core metabolism, with additional co-expression support (black edges).

2. Phenotypic characterization of Col-0 WT and all mutant lines grown in soil under control and drought stress conditions

2.1. Contrasting stress responses in *nat* and *doxc21-a* mutants grown in soil

To evaluate the potential involvement of *DOXC21-A*, NAT and *DOXC21-B* in drought stress responses, we performed phenotypic characterization of all tested *A. thaliana* genotypes in soil: Col-0 WT, two NAT T-DNA insertion lines (*nat1*, *nat2*), a knockout line for *DOXC21-A* (*doxc21-a*), and two lines for *DOXC21-B* (*doxc21-b-1, doxc21-b-2*). Stress treatments were selected based on prior expression analyses. We focused first on drought as an environmental factor previously shown to modulate *DOXC21-A* and *DOXC21-B* expression and that is evolutionarily relevant to plant adaptation from aquatic to terrestrial environments – specifically oxidative, cold, and drought stress.

To evaluate drought responses, plants were grown in soil and exposed to a 3-week water deficit followed by rehydration. Visual phenotypic differences were minimal under control conditions but became more pronounced under drought, with Col-0 WT and *doxc21-a* lines exhibiting the most drought-sensitive phenotypes. Notably, the *doxc21-a* mutants displayed severe drought sensitivity after rehydration, with wilting, chlorosis, and poor survival (**Fig. 4A**). In contrast, *nat* and *doxc21-b* mutants maintained turgid, green leaves and recovered more effectively than Col-0 WT and *doxc21-a*.

Survival rate assessed three days after rewatering (**Fig. 4C**) confirmed significantly higher recovery in *nat* and *doxc21-b* lines compared to *doxc21-a* and Col-0 WT. Shoot FW (fresh weight) measurements mirrored these results. While no statistically significant differences in shoot FW were observed among genotypes under control conditions, drought stress significantly reduced FW shoot mass across all lines, except *nat* and *doxc21-b* lines that retained significantly more biomass (P<0.05–0.01; **Fig. 4B**). Drought stress increases differences in water content (WC) across genotypes. Leaf WC measurement under drought conditions further supported this trend, with *nat* and *doxc21-b* lines maintaining higher WC compared to *doxc21-a* and Col-0 WT (P<0.05–0.0001; **Fig. 4D**). Electrolyte leakage assays revealed mild membrane destabilization in *nat1* but not in *doxc21-a* or *doxc21-b* (Supplementary Fig. S4). Phenotypic characterization of plants grown in soil under oxidative and cold stresses are presented in Supplementary Fig. S5-S8.

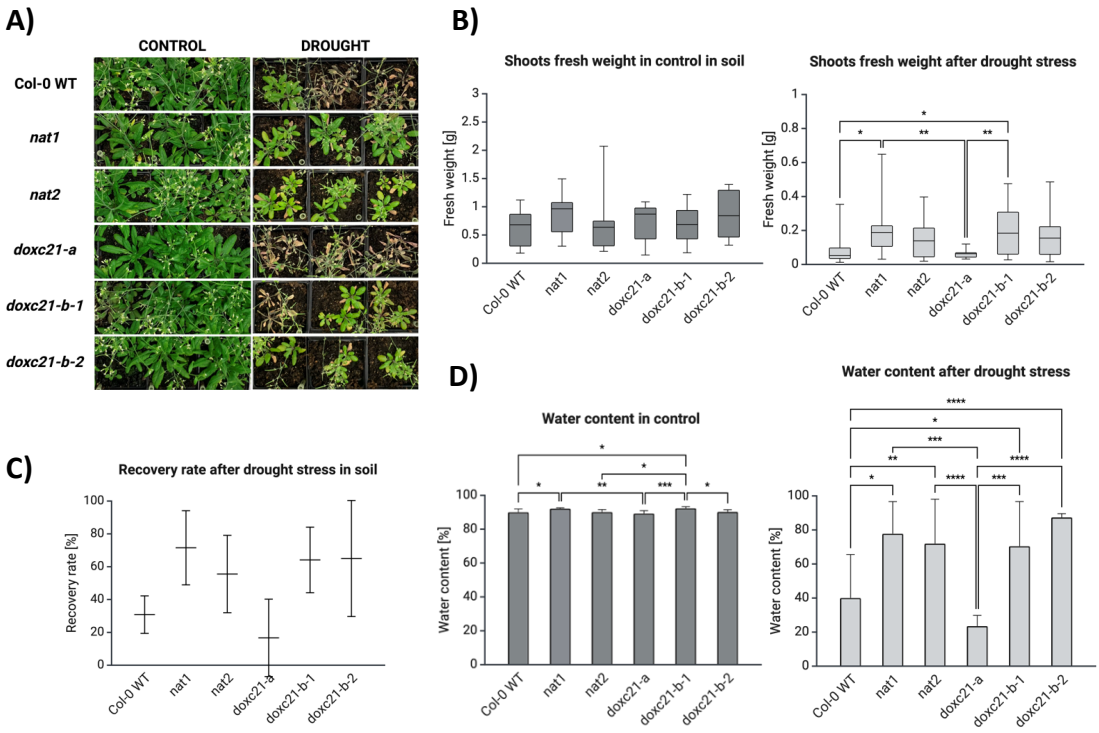


Figure 4. Phenotypic characterization of Col-0 WT, *nat1, nat2, doxc21-a, doxc21-b-1* and *doxc21-b-2* grown in soil under control and drought stress conditions. **(A)** Representative phenotypes of Col-0 WT and tested mutant lines grown in soil under control conditions (left) and following drought stress (right). Images show plants from one experimental replicate, captured after re-watering. **(B)** Shoot fresh weight (FW) was measured three days after re-watering. Data are from one (out of two) experimental replicate, with a minimum of eight plants per genotype per condition. **(C)** Recovery rate assessed in two independent experimental replicates. Recovery was visually scored one day after re-watering based on the presence of green, viable rosette leaves. Graphs show mean percentage values ± standard deviation (SD). **(D)** Water content (WC) of rosettes measured in Col-0 WT and mutant lines under control and drought conditions. WC was calculated based on FW and dry weight. All measurements were performed on a minimum of eight biological replicates across two independent experiments. Statistical significance was determined using one-way ANOVA with Tukey's multiple comparisons test for parametric data (control conditions), and with Kruskal-Wallis test with Dunn's multiple comparisons test for non-parametric data (drought stress). Asterisks indicate significant differences between treatments (**P*<0.05, ***P*<0.01, ****P*<0.001; *****P*<0.0001).

3. Expression analysis of *DOXC21-A* and *DOXC21-B* genes by qRT-PCR under drought (soil-grown conditions) and osmotic stress (*in vitro*) in Col-0 genetic background

To assess the impact of drought-related stress on *DOXC21-A* and *DOXC21-B* expression, we analyzed transcript levels in Col-0 WT plants exposed to physiological drought in soil (**Fig. 5A**) or in osmotic stress *in vitro* (**Fig. 5B**). Osmotic stress was induced with 200 mM mannitol.

Gene expression analysis under *in vitro* conditions revealed that *DOXC21-B* consistently showed higher expression than *DOXC21-A,* across both control and osmotic treatments, with only a slight, non-significant increase in *DOXC21-A* under stress. In contrast, drought-treated soil-grown plants

exhibited strong upregulation of *DOXC21-A* relative to well-watered controls ($P<0.01$), while *DOXC21-B* remained unchanged, suggesting stress specific transcriptional activation of *DOXC21-A* under physiological relevant drought. To sum up, drought stress significantly upregulated *DOXC21-A* expression in soil-grown plants ($P<0.01$), while *DOXC21-B* showed significantly higher expression *in vitro* compared to soil, irrespective of treatment.

To explore potential regulation by NAT and *DOXC21-A*, *DOXC21-B* expression was also measured in *nat* and *doxc21-b* mutant lines under both control and drought conditions in soil. Notably, *nat2* showed significantly elevated *DOXC21-A* transcript levels under both conditions, whereas *nat1* displayed consistently reduced expression compared to Col-0 WT, suggesting genotype-specific regulation of *DOXC21-A* (Supplementary Fig. S9).

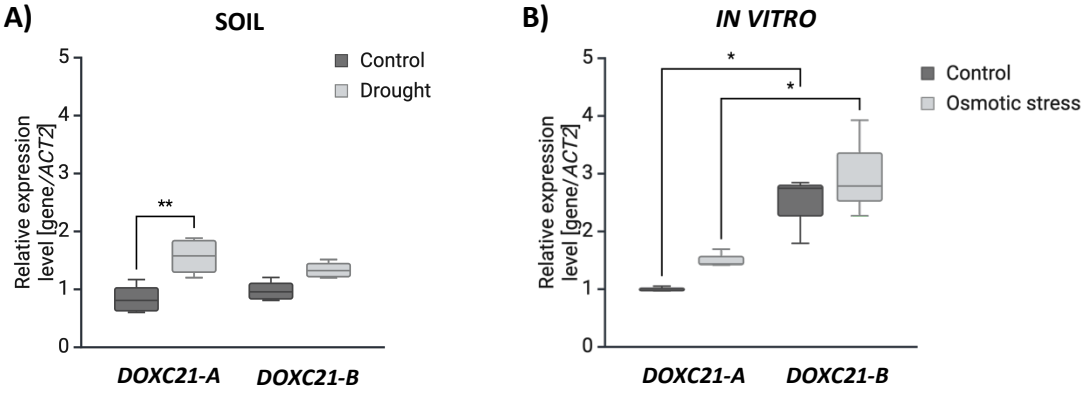


Figure 5. Relative expression levels of *DOXC21-A* and *DOXC21-B* under control and stress conditions in *A. thaliana* leaves grown in soil (A) or *in vitro* (B). Gene expression was quantified in Col-0 WT plants exposed to osmotic stress (*in vitro*, 200 mM mannitol on ½ MS medium) or drought stress (soil, water withheld for 3 weeks). Expression levels of *DOXC21-A* (left) and *DOXC21-B* (right) were normalized to the reference gene *ACT2* and are presented as box plots ($2^{-\Delta\Delta CT}$ values) for control (dark gray) and stress (light gray) conditions. Data represent three biological replicates for *in vitro* and four for soil-grown samples. Statistical significance was assessed using two-way ANOVA followed by Bonferroni's post hoc test (*$P<0.05$, **$P<0.01$).

4. Functional analysis of *DOXC21-A* in drought and osmotic stress responses: hydroponic growth and anthocyanin quantification followed by RNA-Seq transcriptome profiling

To investigate the role of DOXC21-A under abiotic stress conditions typical of terrestrial conditions, we decided to investigate its response to osmotic stress in hydroponic systems. While drought stress poses a serious challenge to plant water homeostasis, osmotic stress, induced by the application of osmotic agents such as mannitol or, more commonly used in hydroponics, PEG, mimics the physiological effects of water limitation, including reduced turgor pressure and altered cellular water potential. Hydroponics, although technically more difficult, is easier to control than soil cultivation in a controlled and repeatable manner in terms of the composition of the growing medium. In this part of the work, we focused on the *DOXC21-A* gene, not only because its expression is strongly dependent on osmotic stress, but also on the presence of overlapping NAT

at the *DOXC21-A* locus, suggesting a complex transcriptional architecture, potentially contributing to gene regulation under stress conditions.

4.1. Growth performance and anthocyanin accumulation differ among Col-0 WT, *nat1, nat2,* and *doxc21-a* in hydroponic cultures

To investigate the contribution of *DOXC21-A* and NAT to plant performance under selected stress conditions, we established a hydroponic system incorporating mild osmotic stress and low pH conditions. A pH of 4.5 was selected as the baseline condition, as it simulates the acidic environment of nutrient-poor or stressed soils, which is particularly relevant to root development and stress signaling in *A. thaliana*. Although pH 4.5 is mildly stressful, it does not severely impact plant growth on its own and thus serves as a suitable control. To further mimic natural stress combinations, the nutrient medium was based on a modified Heeg solution, which includes 10-fold concentrated micronutrients. This modification ensures sufficient availability of essential trace elements under low-pH conditions, where nutrient uptake and solubility may be compromised. Maintaining micronutrient balance is particularly important in hydroponics to avoid confounding deficiencies that could hidden stress-related phenotypes. To introduce mild osmotic stress, we supplemented the acidic medium with 3% (w/v) PEG, which reduces water potential without being taken up by the plant or interfering with ion homeostasis. This concentration is commonly used to simulate early-stage drought-like conditions in *A. thaliana* without causing severe or irreversible damage, allowing for observation of subtle differences in stress sensitivity among genotypes.

Under control conditions, all genotypes, including Col-0 WT, *nat1*, *nat2*, and *doxc21-a,* grew normally, displaying no obvious developmental or pigment-related phenotypes (**Fig. 6A**). Under stress conditions, *nat1* and *nat2* mutants displayed milder leaf senescence and chlorosis than Col-0 WT, indicating improved stress tolerance, while *doxc21-a* mutants showed the most severe phenotype with reduced fitness, premature leaf senescence and pronounced chlorosis (**Fig. 6B**), indicating hypersensitivity to combined osmotic and acidic stress. In contrast, *nat1* and *nat2* mutants exhibited milder stress responses, with *nat2* maintaining the greenest and most turgid rosettes. Col-0 WT showed an intermediate phenotype.

To assess further physiological responses to osmotic stress under low pH conditions, we quantified shoot FW (**Fig. 6C**) and anthocyanin content (**Fig. 6D**) in Col-0 WT, *nat1*, *nat2*, and *doxc21-a* lines grown hydroponically with or without 3% (w/v) PEG. Under control conditions, no significant differences in shoot biomass were observed among genotypes. However, PEG-induced osmotic stress led to a significant reduction in FW across all lines. The *doxc21-a* mutant was the most affected, displaying a marked decrease in shoot mass, while *nat1* and *nat2* were moderately impacted. Notably, *nat1* retained the highest biomass under stress, suggesting improved tolerance compared to other lines.

Anthocyanin accumulation, a well-established marker of stress responses and secondary metabolism, was also measured. Under control conditions, *doxc21-a* exhibited significantly elevated anthocyanin content relative to Col-0 WT and *nat* mutants. Under PEG-induced osmotic stress, anthocyanin levels remained relatively stable across all genotypes, suggesting that the differences observed were predominantly genotype-specific (**Fig. 6D**).

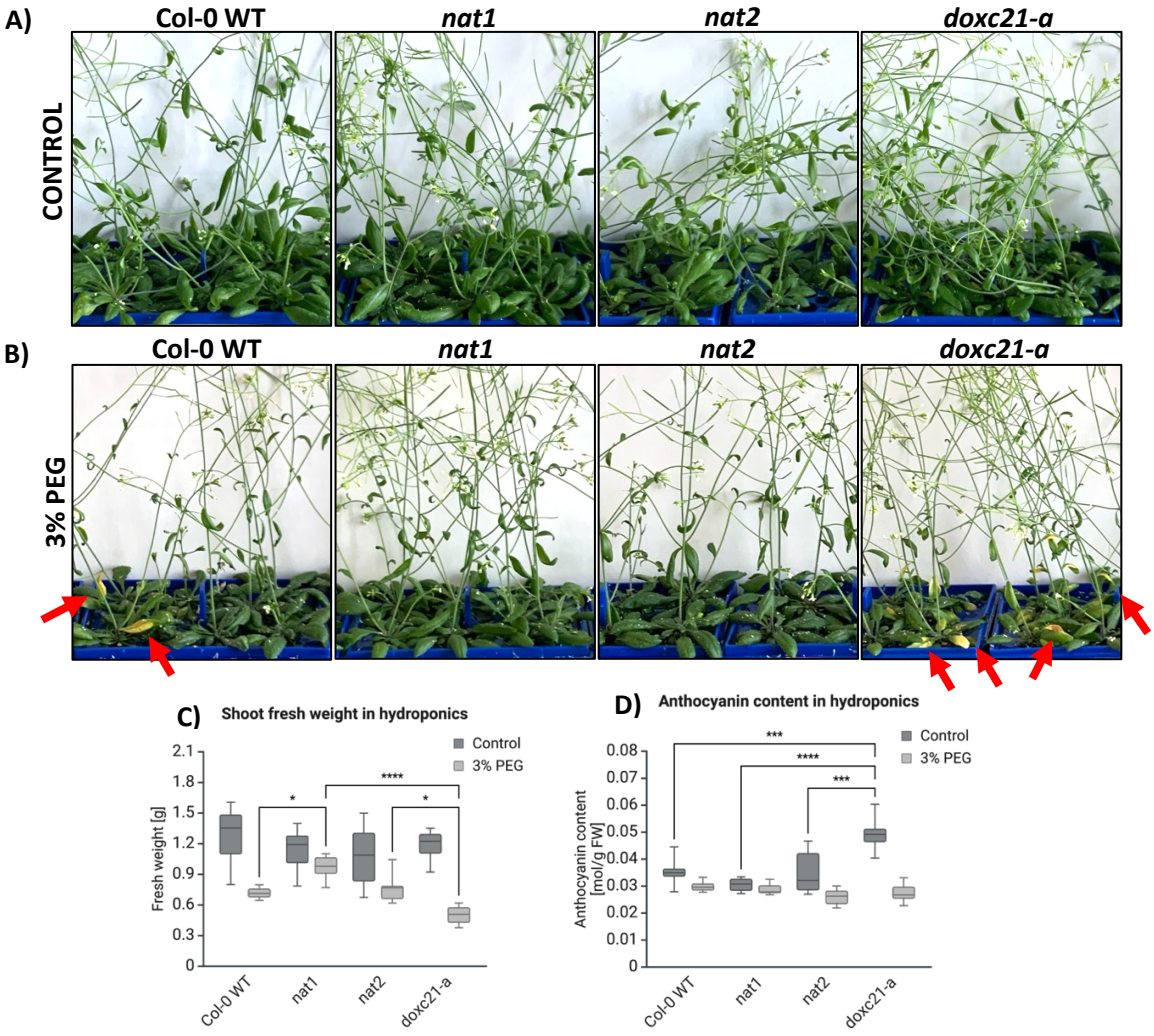


Figure 6. Phenotypic characterization of Col-0 WT, *nat1, nat2,* and *doxc21-a* mutant lines cultivated in hydroponic culture under control (A) and osmotic stress (B) conditions. Plants were cultivated under a 16 h light/ 8 h dark photoperiod at 22°C in modified Heeg nutrient solution (pH 4.5) with 10-fold concentrated micronutrients. Osmotic stress was induced by supplementing the medium with 3% (w/v) PEG 8000. Red arrows indicate chlorotic rosette leaves. **(C)** Shoot fresh weight (FW) was measured at the end of the cultivation period. While no significant differences were observed among genotypes under control (dark grey), PEG-induced osmotic stress (light grey) significantly reduced biomass across all lines. Data are presented as means from ten biological replicates, shown as box plots. **(D)** Total anthocyanin content was quantified in shoot tissue and expressed as mmol g^{-1} FW. Statistical significance was determined using two-way ANOVA followed by Bonferroni's multiple comparisons test (**P*<0.05, ****P*<0.001, *****P*<0.0001).

4.2. Transcriptomic profiling under osmotic stress reveals genotype-specific responses

4.2.1. RNA-Seq quality control and refinement of gene annotations

RNA-seq analysis was conducted to investigate transcriptional changes across *A. thaliana* mutant lines grown in hydroponics and subjected to PEG-induced osmotic stress. Transcriptome sequencing was performed on 23 samples, including six biological replicates for Col-0 WT, *nat1, nat2*, and five biological replicates for *doxc21-a* generating 369.50 Gb high-quality reads. A total of 2.47 billion raw pair-end reads were obtained across all samples. The number of reads per sample ranged between 96.2 and 120 million, with an average GC content of 46.67%. Quality assessment indicated that sequencing data was of high reliability, with base quality scores exceeding Q20 (96.53%) and Q30 (93.95%) in all samples.

After quality filtering, clean reads were aligned to the *A. thaliana* reference genome with high efficiency. The overall mapping rate ranged from 98.06% to 98.57%, and uniquely mapped reads constituted 93.02% to 96.17% of total reads, indicating minimal off-target alignment. Multiple mapped reads accounted for 2.25% to 5.38%. Strand specificity analysis showed an even distribution of reads across the positive and negative strands (50.58% to 52.26%). Detailed statistics of sequencing data are provided in Supplementary Tab. S3-S4.

To refine genome annotations accuracy, gene structures were refined based on RNA-Seq read coverage, allowing extension of 5′ and 3′ UTRs in 1,684 genes. Additionally, StringTie-based transcript assembly revealed 772 novel genes absent from the reference genome. Of the novel genes, 445 were functionally annotated using DIAMOND, InterProScan, and HMMER (NR, Swiss-Prot, Pfam, GO, KEGG databases), expanding the number of putative protein-coding loci for downstream gene expression analyses (Supplementary Tab. S5).

4.2.2. Global patterns of differential gene expression in Col-0 WT, *nat1, nat2,* and *doxc21-a* under osmotic stress

Differential gene expression analysis revealed distinct transcriptional changes across Col-0 WT and mutant lines (*nat1, nat2,* and *doxc21-a*) under the tested conditions. The number of DEGs varied among comparisons, with the highest number observed in *nat1 vs doxc21-a* (809 DEGs) and *nat2 vs. doxc21-a* (760 DEGs), suggesting substantial transcriptional reprogramming upon *doxc21-a* disruption (**Fig. 7**). The Col-0 WT *vs doxc21-a* comparison also exhibited a high number of DEGs (741), emphasizing the impact of the mutation on global gene regulation. In contrast, the lowest number of DEGs was detected in Col-0 WT vs *nat2* (204 DEGs), indicating a milder transcriptional shift.

Across all conditions, downregulated genes constituted a significant proportion of total DEGs, with Col-0 WT vs. *doxc21-a*, *nat1* vs. *doxc21-a* and *nat2* vs. *doxc21-a* showing the highest number of repressed genes (529, 614 and, 522 respectively). The *doxc21-a* genotype showed the highest number of repressed genes, suggesting a strong transcriptional response under stress.

To investigate the transcriptional impact of *DOXC21-A* and NAT disruption, subset of genes exhibiting a minimum $\log_2$ FC (>2.5) and statistical significance (P<0.05) across all comparisons was identified (**Tab. 2**). The results revealed significant alterations in key pathways, including stress responses, hormone signaling, secondary metabolism, and transcriptional regulation. Several genes involved in abiotic and biotic stress responses showed differential expression across mutant lines. Interestingly, in *nat2* mutant line expression level of *DOXC21-A* was significantly upregulated in comparison to all other lines including Col-0 WT, *nat1* and *nat2* ko.

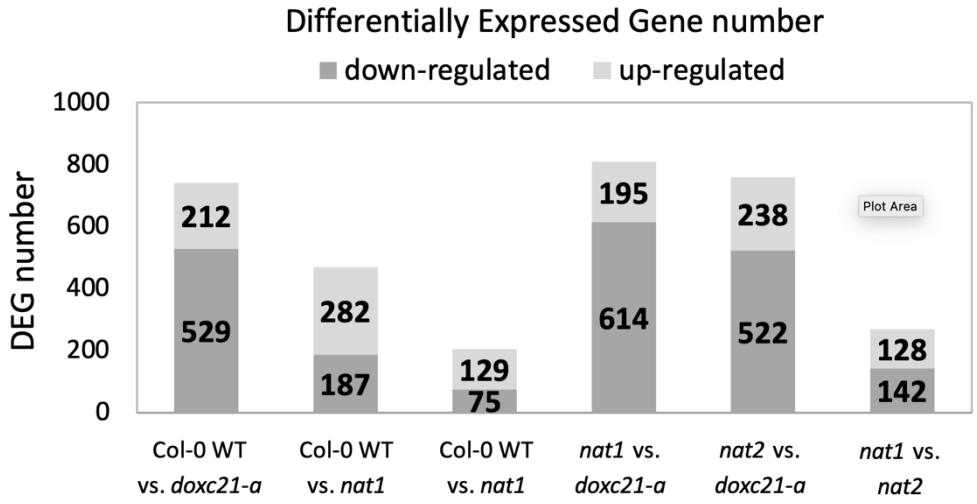


Figure 7. Number of differentially expressed genes (DEGs) identified in pairwise comparisons between *DOXC21-A* mutant lines and Col-0 WT. Dark gray bars represent downregulated genes, while light gray bars denote upregulated genes. The highest number of DEGs was observed in *nat1* vs. *doxc21-a* (809) and *nat2* vs. *doxc21-a* (760), while the lowest number was detected in Col-0 WT vs. *nat2* (204).

Several genes associated with camalexin biosynthesis and tryptophan metabolism exhibited differential expression patterns. *ABCG40* (*AT1G15520*), a pleiotropic drug resistance transporter involved in camalexin secretion [37], showed a strong downregulation in *nat2* and *doxc21-a* compared to Col-0 WT, suggesting altered phytoalexin accumulation. Similarly, *CYP71A13* (*AT2G30770*), encoding a key enzyme in camalexin biosynthesis [38], was significantly repressed in mutant lines, indicating compromised defense responses. *NIT2* (*AT3G44300*), involved in indole-3-acetic acid (IAA) biosynthesis from indole-3-acetonitrile [39], displayed significant downregulation in *doxc21-a* while in *nat2* - upregulation, which may be associated with auxin-related pathways. *MES9* (*AT4G37150*), a methyl-IAA esterase responsible for auxin turnover [40], was repressed in *doxc21-a*, further confirming disruptions in auxin metabolism. Additionally, *TSPO* (*AT2G47770*), an outer membrane tryptophan-rich sensory protein-related with roles in stress adaptation [41], was notably reduced in *doxc21-a* mutant.

Table 2. Expression profiles of selected genes that are differentially expressed (Differentially Expressed Genes, DEGs) under osmotic stress (pH 4.5) conditions between the studied genotypes, namely Col-0 WT, *nat1, nat2,* and *doxc21-a*. Genes were selected based on significant differential expression ($\log_2FC \geq \pm 2.5$, $P<0.05$) identified from RNA-seq analysis using [DESeq2/edgeR]. Values represent FPKM (fragments per kb of transcript per million mapped reads)-normalized transcript abundance to illustrate relative expression across genotypes. Color coding reflects expression level (yellow: moderate levels; green: upregulation; red: downregulation) was applied individually per gene, to illustrate relative expression shifts between genotypes for each gene separately. Directional arrows highlight expression trends. FPKM values are shown for descriptive comparison only. Statistical significance was determined independently *via* DE analysis (see Supplementary Tables S6).

#ID	gene name	Col-0 WT	*nat1*	*nat2*	*doxc21-a*
AT1G02310	MAN1	1.86	2.88	3.06	0.33
AT1G05340	YUP8H12.4	43.05	54.08	41.52	4.96
AT1G11785	AT1G11785	1.52	6.71	8.14	1.05
AT1G12570	AT1G12570	0.30	0.28	0.17	2.06
AT1G15520	ABCG40	6.23	4.52	1.73	0.34
AT1G17020	SRG1	5.50	5.72	4.72	0.77
AT1G18300	NUDT4	9.78	14.23	38.95	6.65
AT1G19250	FMO1	16.45	9.87	3.89	0.16
AT1G22990	HIPP22	67.66	82.70	110.68	11.99
AT1G29640	AT1G29640	5.50	7.17	7.09	1.00
AT1G32350	AOX3	6.92	4.85	3.77	0.43
AT1G33960	AIG1	12.19	10.95	14.51	2.33
AT1G47980	AT1G47980	3.66	1.49	2.44	0.13
AT1G51670	HTT5	0.39	1.84	0.31	0.16
AT1G53480	MRD1	23.95	25.74	0.74	0.22
AT1G53490	HEI10	2.27	2.40	0.51	0.41
AT1G60960	IRT3	0.09	3.04	3.07	2.87
AT1G62510	AT1G62510	12.27	38.44	17.86	5.84
AT1G64795	AT1G64795	0.05	0.06	0.07	80.24
AT1G67856	AT1G67856	2.16	2.50	7.77	1.02
AT1G73540	NUDT21	30.41	36.89	99.52	16.86
AT1G74590	GSTU10	14.27	9.86	13.42	1.97
AT1G74930	ORA47	5.68	13.39	31.27	4.76
AT1G75945	AT1G75945	0.30	15.06	15.46	14.31
AT1G76650	CML38	13.30	26.29	77.55	9.03
AT2G13810	ALD1	12.71	5.37	5.22	2.08
AT2G14610	PR1	455.96	412.27	973.33	99.86
AT2G19900	NADP-ME1	6.12	7.70	4.65	0.64
AT2G22470	AGP2	8.66	15.39	3.24	1.46
AT2G26400	ARD	17.21	20.76	22.37	2.00
AT2G29350	SAG13	63.93	30.08	45.98	11.11
AT2G29460	GSTU4	45.54	36.38	29.64	7.63
AT2G30770	CYP71A13	22.21	17.61	4.91	1.17
AT2G43570	CHI	250.70	147.68	243.85	35.77
AT2G44240	AT2G44240	1.66	0.74	7.09	0.43
AT2G45220	PME17	10.40	3.67	2.14	0.19
AT2G45570	CYP76C2	5.06	4.64	2.30	0.03
AT2G45760	BAP2	6.59	6.36	4.56	0.67
AT2G46820	PSI-P	864.10	906.00	840.22	130.30
AT2G47770	TSPO	21.37	36.39	15.09	3.00
AT3G02840	AT3G02840	2.01	3.42	5.38	0.58
AT3G08860	PYD4	18.34	17.86	24.14	4.21
AT3G10930	AT3G10930	5.78	5.26	19.62	1.49
AT3G11340	UGT76B1	3.09	3.44	0.77	0.27
AT3G13610	F6'H1	3.16	2.45	2.44	0.41
AT3G17520	AT3G17520	5.10	12.58	4.41	0.44
AT3G18250	AT3G18250	25.77	19.32	31.11	4.72
AT3G19000	**DOXC21-A**	**19.48**	**19.62**	**125.43**	**3.89**
AT3G22600	LTPG5	41.58	25.07	20.41	3.19
AT3G22840	ELIP1	13.89	28.77	13.26	4.80
AT3G27473	AT3G27473	0.35	0.25	1.64	0.20

#ID	gene name	Col-0 WT	*nat1*	*nat2*	*doxc21-a*
AT3G29000	CML45	25.54	18.61	47.01	6.77
AT3G44260	CAF1-9	28.74	27.79	87.91	10.39
AT3G44300	NIT2	48.62	20.31	59.01	0.60
AT3G46080	ZAT8	4.81	4.06	5.83	0.53
AT3G48360	BT2	4.29	30.26	9.51	17.12
AT3G48850	PHT3;2	3.34	2.09	1.60	0.51
AT3G50980	XERO1	5.48	5.67	5.18	0.67
AT3G53150	UGT73D1	2.63	2.13	2.36	0.29
AT3G56970	bHLH38	2.79	4.40	5.91	0.67
AT3G61190	BAP1	13.56	11.25	30.37	3.41
AT4G11211	AT4G11211	19.16	4.49	21.14	27.83
AT4G12580	AT4G12580	7.31	17.35	4.41	2.22
AT4G18422	AT4G18422	8.18	6.59	18.55	3.07
AT4G23680	AT4G23680	11.02	13.44	13.20	1.33
AT4G23700	CHX17	1.75	1.31	1.41	0.15
AT4G24570	PUMP4	22.67	44.59	76.15	9.39
AT4G25000	AMY1	24.43	3.98	13.85	7.34
AT4G27280	KRP1	17.54	17.70	33.70	5.94
AT4G27652	AT4G27652	7.37	4.93	24.61	4.26
AT4G27654	AT4G27654	1.84	2.33	11.93	1.13
AT4G28790	BHLH23	1.90	2.28	12.39	0.98
AT4G28800	AT4G28800	0.01	1.83	27.40	0.01
AT4G29780	AT4G29780	5.41	11.28	19.86	3.01
AT4G33790	FAR3	4.21	2.35	0.99	8.44
AT4G34410	ERF109	0.83	3.40	9.84	1.69
AT4G37150	MES9	4.45	3.70	6.09	0.52
AT4G39670	GLTP	10.53	8.52	11.75	1.48
AT5G01300	AT5G01300	9.14	14.92	9.05	1.73
AT5G07010	SOT15	5.24	1.70	1.78	0.43
AT5G10625	FLP2	4.11	3.23	3.40	0.64
AT5G13080	WRKY75	11.99	8.00	7.48	0.73
AT5G13170	SWEET15	6.53	7.82	4.13	0.76
AT5G17350	AT5G17350	0.97	2.02	5.14	0.49
AT5G28080	WNK9	1.46	4.45	2.16	0.70
AT5G39520	AT5G39520	18.11	12.23	17.55	2.29
AT5G42380	CML37	3.66	4.52	7.16	1.08
AT5G44420	PDF1.2A	12.13	1.92	6.01	20.58
AT5G45630	AT5G45630	2.75	5.07	12.16	1.55
AT5G57560	XTH22	14.93	26.90	29.93	5.24
AT5G59820	ZAT12	5.95	5.18	16.35	2.40
AT5G62150	AT5G62150	4.57	4.05	5.45	0.95
AT5G62165	AGL42	6.06	1.91	0.85	11.26
AT5G62480	GSTU9	5.25	5.22	4.88	0.65
AT5G62520	SRO5	0.83	3.55	5.05	1.30
NewGene_114	--	0.11	0.34	0.23	0.84
NewGene_142	--	3.09	0.76	0.48	0.57
NewGene_159	--	0.00	0.00	31.83	0.00
NewGene_280	--	0.06	0.36	0.54	0.13
NewGene_710	--	0.02	3.66	3.77	4.29
NewGene_1147	--	8.29	4.54	4.30	0.70
NewGene_1284	--	3.28	1.09	0.84	0.54

Key genes implicated in plant defence and systemic acquired resistance (SAR) were significantly misregulated. *FMO1* (*AT1G19250*), a flavin monooxygenase essential for SAR induction [42], was downregulated in *doxc21-a* mutants, indicating a suppression of immune priming. Similarly, *ALD1* (*AT2G13810*), a pipecolic acid biosynthesis gene critical for SAR signalling [43], exhibited decreased expression in *doxc21-a* mutants. *PR1* (*AT2G14610*), a canonical SAR marker [44], exhibited a dynamic expression pattern in response to *DOXC21-A* disruption. It was downregulated in *doxc21-a*, maintained at moderate levels in Col-0 WT and *nat1*, and strongly induced in *nat2*, suggesting complex transcriptional regulation linked to stress adaptation.

Multiple genes involved in micronutrient acquisition and transport displayed differential expression. *HIPP22* (*AT1G22990*), encoding a heavy metal-associated isoprenylated plant protein involved in Zn homeostasis [45], was strongly downregulated in *doxc21-a* and upregulated in *nat1* and *nat2* in comparison to Col-0 WT, suggesting altered micronutrient management. *IRT3* (*AT1G60960*), an iron-regulated transporter [46], was differentially expressed (upregulation *nat1, nat2* and *doxc21-a* mutant lines), indicating disruptions in Fe uptake pathways. Moreover, *F6'H1* (*AT3G13610*), an enzyme that hydroxylates feruloyl-CoA [20], exhibited altered expression levels, implicating its role in phenylpropanoid-related stress responses. Notably, *F6'H1* overexpression has been shown to redirect feruloyl-CoA flux toward scopoletin biosynthesis, thereby reducing lignin content and modifying lignin composition and structure [47]. *bHLH38* (*AT3G56970*), a TF regulating Fe homeostasis [48], displayed significant downregulation in *doxc21-a,* moderate expression in Col-0 WT and upregulation in in *nat1* and in *nat2*, suggesting an Fe-deficiency response. Structural gene such as *XTH22* (*AT5G57560*), involved in cell wall remodelling under boron deficiency [49], and *ZAT12* (*AT5G59820*), a zinc-finger TF interacting with FIT (FER-like iron deficiency-induced TF) linking Fe deficiency and oxidative stress response [50], were also misregulated.

Genes involved in RNA metabolism and transcriptional regulation were significantly altered. *NUDT21* (*AT1G73540*), a Nudix hydrolase involved in mRNA capping [51], exhibited significant upregulation in *nat2*, indicating post-transcriptional regulation under stress. *CAF1-9* (*AT3G44260*), a component of the CCR4-NOT complex responsible for mRNA deadenylation and degradation [52], exhibited significant upregulation in *nat2* mutant line, suggesting potential alterations in RNA turnover dynamics and post-transcriptional regulation.

Several genes linked to JA biosynthesis and fatty acid metabolism exhibited significant differential expression. *ORA47* (*AT1G74930*), a key JA-responsive TF [53], was significantly upregulated in *nat2*, suggesting enhanced jasmonate signaling pathway. *FAR3/CER4* (*AT4G33790*), encoding fatty acyl-CoA reductase involved in cuticular wax biosynthesis [54], showed repression in *nat1* and *nat2* mutant lines and upregulation in *doxc21-a*, indicating modifications in lipid metabolism. *SOT15* (*AT5G07010*), a sulfotransferase involved in JA metabolism [55], and *CML37* (*AT5G42380*), a calmodulin-like protein that modulates jasmonate responses [56], were significantly

affected, especially in *doxc21-a*. Additionally, *PDF1.2* (*AT5G44420*), a JA-responsive defense gene [57], was upregulated in *doxc21-a*, indicating possible involvement of JA-related regulation.

Acidic growth conditions led to differential regulation of genes linked to pH stress adaptation and macronutrient transport. *CML38* (*AT1G76650*), encoding a calmodulin-like protein responsive to Ca^{2+} fluctuations under low pH [58], exhibited increased expression in *nat2*. *PYD4* (*AT3G08860*), a pyrimidine biosynthesis gene linked to nucleotide metabolism [59], was significantly repressed in *doxc21-a. PHT3;2* (*AT3G48850*), encoding a phosphate (P) transporter [60], was differentially expressed, suggesting P homeostasis disruptions in *DOXC21-A* mutant lines. Furthermore, *WRKY75* (*AT5G13080*), a TF modulating root architecture under P starvation [61], displayed altered expression patterns with strong downregulation in *doxc21-a*. *CHX17* (*AT4G23700*), a Na+/H+ exchanger involved in pH regulation and K+ homeostasis [62], was significantly misregulated in *doxc21-a*.

Photosynthesis-associated genes were also affected by *DOXC21-A* mutations. *PSI-P* (*AT2G46820*), encoding a photosystem I subunit [63], exhibited reduced expression in *doxc21-a* mutants, a pattern consistent with stress-induced adjustments in photosynthesis-related gene expression. *ELIP1* (*AT3G22840*), an early light-inducible protein involved in chloroplast photoprotection [64], was also downregulated in *doxc21-a*, implying increased susceptibility to oxidative damage.

Three other functionally relevant genes exhibited notable expression shifts. *AGL42* (*AT5G62165*), a MADS-box TF involved in developmental processes [65], was differentially expressed, hinting at possible developmental plasticity. *SRO5* (*SIMILAR TO RCD-ONE*, *AT5G62520*), a redox-regulatory protein involved in oxidative stress responses [66], showed increased expression in *nat1* and *nat2*, which is commonly associated with redox regulation. *HEI10* (*AT1G53490*), a RING/U-box superfamily E3 ubiquitin ligase involved in meiotic recombination [67], was downregulated in *nat2* and *doxc21-a,* suggesting a potential link between ubiquitin-mediated chromatin regulation and stress adaptation.

Supplementary Tables 6A-F provide an overview of the top DEGs identified in RNA-seq analysis, including upregulated and downregulated genes across different comparisons. Each table includes gene annotations, FC values, *p*-values, GO classifications, KEGG pathway annotations, and protein function predictions.

4.3. KEGG pathway mapping highlights differential activity in phenylpropanoid biosynthesis

KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis was conducted to classify the DEGs and identify biological pathways that are significantly enriched with the list of genes of interest. The analysis revealed that DEGs were predominantly associated with key biological processes, including environmental signal transduction, metabolic pathways, and stress response mechanisms.

The **Col-0 WT vs. *doxc21-a*** comparison revealed that the plant hormone signal transduction pathway was among the most enriched categories, with 23 DEGs (8.21%), indicating significant disruptions in hormone-mediated regulatory networks. Similarly, MAPK signaling (16 DEGs, 5.71%) was significantly altered, suggesting changes in stress perception and defense signaling. Metabolic processes, including photosynthesis (15 DEGs, 5.36%), phenylpropanoid biosynthesis (13 DEGs, 4.64%), and oxidative phosphorylation (12 DEGs, 4.29%), were also affected. The plant-pathogen interaction pathway (25 DEGs, 8.93%) was significantly enriched, highlighting a substantial reprogramming of immune responses.

In the **Col-0 WT vs. *nat1*** comparison, plant hormone signaling (16 DEGs, 10.13%) and MAPK signaling (11 DEGs, 6.96%) were among the top enriched pathways, further confirming the role of *DOXC21-A* in stress response coordination. Metabolic processes such as starch and sucrose metabolism (11 DEGs, 6.96%), biosynthesis of amino acids (8 DEGs, 5.06%), phenylpropanoid biosynthesis (8 DEGs, 5.06%), and galactose metabolism (7 DEGs, 4.43%) were also differentially regulated. The plant-pathogen interaction pathway (19 DEGs, 12.03%) was significantly enriched, pointing to alterations in defense mechanisms.

The **Col-0 WT vs. *nat2*** comparison revealed strong enrichment in MAPK signaling (7 DEGs, 11.86%), phosphatidylinositol signaling (3 DEGs, 5.08%), and plant hormone signal transduction (3 DEGs, 5.08%), suggesting modifications in environmental response pathways. Additionally, cysteine and methionine metabolism (3 DEGs, 5.08%), starch and sucrose metabolism (3 DEGs, 5.08%), and circadian rhythm regulation (3 DEGs, 5.08%) were affected, indicating a shift in metabolic and regulatory processes. The plant-pathogen interaction pathway (10 DEGs, 16.95%) was among the most highly enriched categories, reinforcing the impact of *DOXC21-A* on pathogen defense.

The ***nat1* vs. *doxc21-a*** comparison showed an even stronger impact on hormone signaling (29 DEGs, 10.90%), MAPK signaling (17 DEGs, 6.39%), and phenylpropanoid biosynthesis (15 DEGs, 5.64%), than observed in the Col-0 WT vs. *nats* comparisons, reinforcing the role of *DOXC21-A* in regulating stress responses and metabolic pathways. Notably, starch and sucrose metabolism (13 DEGs, 4.89%) and glucosinolate biosynthesis (10 DEGs, 3.76%) were also affected, suggesting an interplay between secondary metabolite production and stress adaptation. The plant-pathogen interaction pathway (17 DEGs, 6.39%) was significantly enriched, emphasizing immune-related transcriptional shifts.

A similar pattern was observed in the ***nat2* vs. *doxc21-a*** comparison, where plant hormone signal transduction (24 DEGs, 9.52%), MAPK signaling (18 DEGs, 7.14%), and phenylpropanoid biosynthesis (18 DEGs, 7.14%) were among the most differentially enriched categories. Taking into account the centrally located T-DNA insertion within the NAT sequence in the ***nat2*** line, it was particularly interesting to conduct a more detailed comparison of DEGs between this line and the ***doxc21-a*** knockout mutant (**Fig. 8**).

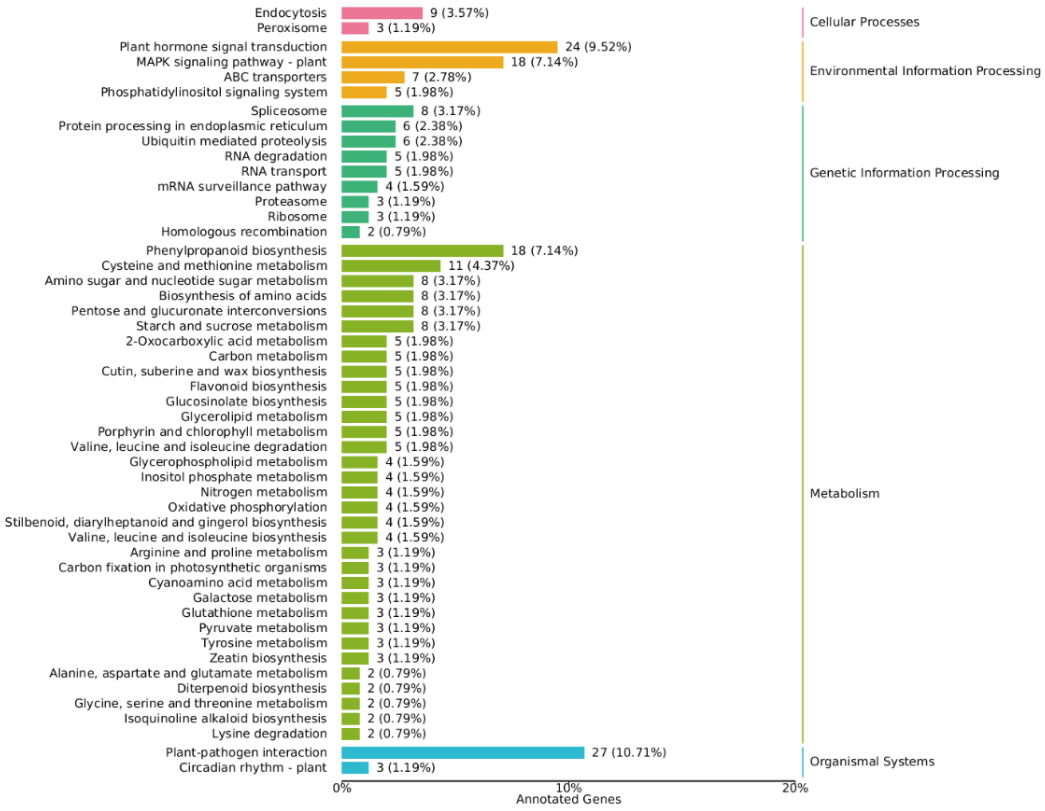


Figure 8. KEGG classification of Differentially Expressed Genes (DEGs) identified in comparisons of *nat2* and *doxc21-a*. The KEGG pathways are grouped into major functional categories, including Genetic Information Processing, Environmental Information Processing, Metabolism, and Organismal Systems. The most significantly enriched pathways include plant hormone signal transduction, MAPK signaling, phenylpropanoid biosynthesis, cysteine and methionine metabolism, and plant-pathogen interactions. The percentage of annotated DEGs in each pathway is indicated in parentheses.

A further comparison of DEGs among ***nat2*** vs. ***doxc21-a*** revealed particularly interesting insights into phenylpropanoid biosynthesis (**Fig. 9**). Several enzymatic steps in lignin and flavonoid biosynthesis were differentially regulated, suggesting altered pathway flux between these genotypes under osmotic stress conditions. The cysteine and methionine metabolism pathway (11 DEGs, 4.37%) was also significantly regulated, indicating changes in sulfur metabolism and polyamine biosynthesis. The plant-pathogen interaction pathway (27 DEGs, 10.71%) exhibited a notable increase in differentially regulated genes, suggesting that *DOXC21-A* is involved in modulating immune responses.

Lastly, in the ***nat1* vs. *nat2*** comparison, pathways related to metabolic and genetic processing were differentially regulated. The plant hormone signal transduction category accounted for 10 DEGs (12.99%), demonstrating alterations in hormonal regulation. Genetic information processing pathways, such as ribosome function (7 DEGs, 9.09%) and protein processing in the endoplasmic reticulum (5 DEGs, 6.49%), were also significantly enriched. Notably, ubiquitin-mediated proteolysis (4 DEGs, 5.19%), phenylpropanoid biosynthesis (4 DEGs, 5.19%), and starch

and sucrose metabolism (4 DEGs, 5.19%) indicated potential modifications in protein degradation and carbohydrate metabolism. Additionally, plant-pathogen interaction (8 DEGs, 10.39%) was significantly enriched.

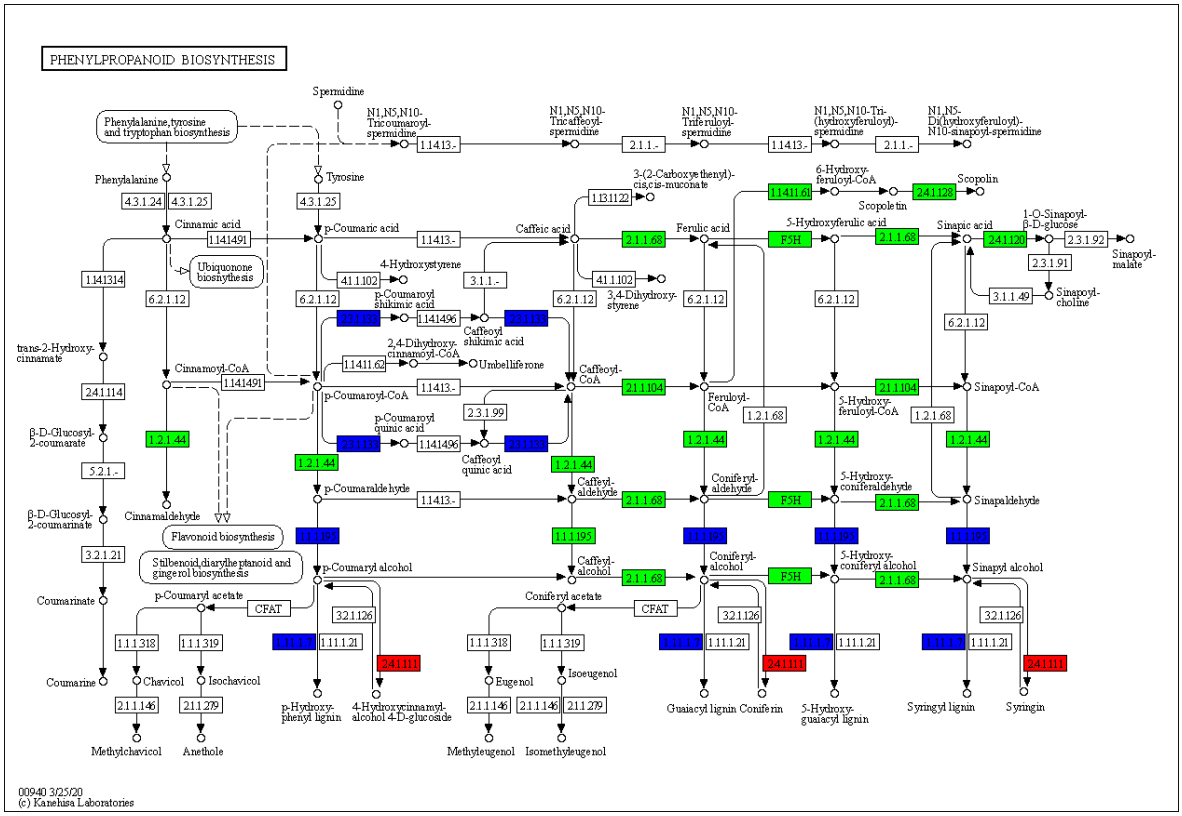


Figure 9. Differential regulation of the phenylpropanoid biosynthesis pathway in *nat2* and *doxc21-a* mutant lines under osmotic stress conditions. Transcriptomic data were mapped onto the KEGG phenylpropanoid biosynthesis pathway (KEGG Pathway: ath00940) to illustrate differential gene regulation in hydroponically grown *nat2* and *doxc21-a* plants exposed to 3% (w/v) PEG-induced osmotic stress. Genes upregulated in *nat2* are marked in green, those upregulated in *doxc21-a* in red, and genes showing mixed regulation between the two genotypes are highlighted in blue.

4.4. Differential Exon Usage (DEU) analysis reveals isoform-specific regulation

Alternative splicing (AS) serves as a crucial regulatory mechanism that expands transcriptome complexity and enables plants to fine-tune gene expression in response to environmental stressors. To investigate AS variation across *A. thaliana* Col-0 WT, *nat1, nat2*, and *doxc21-a*, a transcriptome-wide analysis was done, classifying a total of 1,067,430 AS events into 12 distinct AS types, based on variations in exon skipping, intron retention, alternative exon boundaries, and transcription start/termination site modifications (**Tab. 3**).

Table 3. Alternative Splicing (AS) event statistics. The average number of alternative splicing (AS) events detected in Col-0 WT, *nat1, nat2,* and *doxc21-a* mutant lines. XSKIP (approx. single exon skipping, fuzzy boundary); XMSKIP (approx. multi-exon skipping, fuzzy boundary); XMIR (approx. multi-intron retention, fuzzy boundary); XIR (approx. single intron retention, fuzzy boundary); XAE (approx. alternative exon variable 5' or 3' end, fuzzy boundary); TTS (alternative 3' last exon/transcription termination site, the last exon splicing); TSS (alternative 5' exon/transcription start site, the first exon splicing); SKIP (precise single exon skipping); MSKIP (precise multi-exon skipping); MIR (multi-intron retention); IR (single intron retention); AE (alternative exon ends).

AS	Alternative splicing	Col-0 WT	*nat1*	*nat2*	*doxc21-a*
XSKIP	approx. single exon skipping, fuzzy boundary	201	201	200	197
XMSKIP	approx. multi-exon skipping, fuzzy boundary	10.7	15.8	10.8	11.2
XMIR	approx. multi-intron retention, fuzzy boundary	37	35	38	33
XIR	approx. single intron retention, fuzzy boundary	586	593	607	577
XAE	approx. alternative exon variable 5' or 3' end, fuzzy boundary	814	827	805	805
TTS	alternative 3' last exon/transcription termination site, the last exon splicing	18557	18492	18554	18428
TSS	alternative 5' exon/transcription start site, the first exon splicing	19310	19273	19317	19208
SKIP	precise single exon skipping	535	546	558	540
MSKIP	precise multi-exon skipping	36.8	40.2	36.3	36
MIR	multi-intron retention	196	194	200	201
IR	single intron retention	2260	2294	2291	2285
AE	alternative exon ends	3872	3922	3965	3850

XSKIP (Approx. Exon Skipping) represents single exon skipping with a fuzzy boundary, while XMSKIP (Approx. Multi-Exon Skipping) involves the skipping of multiple exons. XMIR (Approx. Multi-Intron Retention) and XIR (Approx. Intron Retention) describe cases of multi-intron or single intron retention, respectively. XAE (Approx. Alternative Exon Ends) reflects alternative exon boundary variations. TTS (Alternative Transcription Termination Site) and TSS (Alternative Transcription Start Site) denote variability in the last 3' and first 5' exons, respectively. SKIP (Precise Exon Skipping) and MSKIP (Precise Multi-Exon Skipping) correspond to well-defined exon skipping events. MIR (Multi-Intron Retention) and IR (Single Intron Retention) indicate intron retention events, while AE (Alternative Exon Ends) represents variability in either or both exon termini. Among all genotypes, TTS and TSS were the most frequently detected AS events, both exceeding 18,400 and 19,200 occurrences, respectively, across samples. IR was also highly represented, followed by AE, reflecting extensive transcriptomic flexibility. IR and AE frequencies were slightly elevated in *nat1* and *nat2* mutant lines relative to Col-0 WT, suggesting altered splicing patterns upon *DOXC21-A* disruption. In contrast, *doxc21-a* exhibited a reduced occurrence of these events, which may indicate distinct regulatory consequences of this mutation. Additionally, SKIP although not statistically significant (p=0.079) was more abundant in *nat2* in comparison to Col-0 WT, while MSKIP events were more abundant in *nat1* in comparison to *doxc21-a*. XMSKIP was more abundant in *nat1* when compared to other lines. Differences in the statistics on alternative

splicing events indicate transcriptomic plasticity associated with *DOXC21-A* gene expression regulation.

Differential AS analysis identified 152 significantly differentially spliced genes (DSGs) (FDR<0.01). DEU analysis revealed multiple genes exhibiting significant exon-level changes, particularly those involved in stress response, transcriptional regulation, and secondary metabolism. DEU described in following paragraph are collected in **Tab. 4**.

DEU analysis between **Col-0 WT and *doxc21-a*** identified significant exon usage shifts in genes involved in transcriptional regulation, stress response, and metabolic processes. Among the most upregulated exons, *DOXC21-A* (*AT3G19000*) exhibited the highest exon usage variation. Additionally, *HDS* (*AT5G60600*), encoding 4-hydroxy-3-methylbut-2-enyl diphosphate synthase involved in isoprenoid biosynthesis [68], and *SMP2* (*AT4G37120*), encoding a pre-mRNA splicing Prp18-interacting factor [69], displayed elevated exon inclusion, suggesting a role in post-transcriptional regulation under osmotic stress conditions. Conversely, the most downregulated exons included *PDE327* (*AT4G30720*), encoding a putative oxidoreductase/electron carrier detected in the chloroplast stroma [70], and *AK-HSDH I* (*AT1G31230*), encoding aspartate kinase-homoserine dehydrogenase I involved in essential amino acid biosynthesis [71].

The most notable changes in DEU between **Col-0 WT and *nat1*** were observed in *LARP1* (*AT5G21160*), which exhibited a strong reduction in exon inclusion, suggesting a role in stress adaptation. Similarly, *RABE1c* (*AT3G46060*), encoding a small GTP-binding protein, displayed reduced exon retention [72], potentially affecting intracellular signaling pathways. *AT1G01210*, encoding a DNA-directed RNA polymerase subunit, also showed altered exon usage, indicating possible transcriptional regulatory changes.

Table 4. Differential Exon Usage (DEU) analysis reveals genotype-specific transcript isoform regulation in *doxc21-a, nat1,* and *nat2* mutant lines under osmotic stress conditions. The table presents genes exhibiting significant DEU across pairwise comparisons between Col-0 WT and the indicated mutant lines (*doxc21-a, nat1, nat2*), based on RNA-seq data from hydroponically grown rosette leaves treated with 3% (w/v) PEG. A plus sign ("+") indicates a significant change in exon usage (adjusted P<0.05), as detected by DEXSeq.

Gene ID	Gene name	Col-0 WT vs. *doxc21-a*	Col-0 WT vs. *nat1*	Col-0 WT vs. *nat2*	*nat1 vs. doxc21-a*	*nat2 vs. doxc21-a*	*nat1 vs. nat2*
AT1G01210			+				
AT1G19650				+			
AT1G25145	*LpxC4*						+
AT1G31230	*AK-HSDH I*	+					
AT1G44575	*NPQ4*				+		
AT1G53490	*HEI10*						+
AT1G56300							+
AT1G64790	*ILITHYIA*				+		
AT1G79530	*GAPCP-1*			+			
AT2G22720					+		
AT2G45570	*CYP76C2*					+	
AT3G15450						+	
AT3G19000	***DOXC21-A***	**+**			**+**	**+**	
AT3G46060	*RABE1c*		+				
AT4G28790	*bHLH23*					+	+
AT4G30720	*PDE327*	+					
AT4G35770	*SEN1*						+
AT4G37120	*SMP2*	+					
AT4G37280						+	
AT5G04590	*SIR*			+		+	
AT5G21160	*LARP1A*		+		+		+
AT5G37850	*SOS4*			+			
AT5G60600	*HDS*	+					
AT5G67030	*ABA1*			+			

DEU analysis between **Col-0 WT and *nat2*** mutants revealed significant exon usage alterations in genes associated with metabolic regulation, stress response, and secondary metabolism. The most upregulated exon usage was observed in *GAPCP-1* (*AT1G79530*), encoding glyceraldehyde-3-phosphate dehydrogenase of plastid 1 [73], followed by *SIR* (*AT5G04590*), encoding sulfite reductase [74], and *ABA*1 (*AT5G67030*), encoding zeaxanthin epoxidase, which functions in the first step of ABA biosynthesis [75]. Conversely, *AT1G19650*, Sec14p-like phosphatidylinositol transfer family protein, showed the strongest exon repression, highlighting its potential involvement in stress-related metabolic shifts. *SOS4* (*AT5G37850*), encoding a pyridoxal kinase required for root hair development [76], exhibited decreased exon inclusion.

Genes with significant splicing variations between ***nat1*** **and** ***doxc21-a*** included *DOXC21-A* (*AT3G19000*), *LARP1A* (*AT5G21160*) involved in mRNA degradation in response to stress [77], *AT2G22720* encoding *SPT2* chromatin protein, and *ILITHYIA* (*AT1G64790*), a HEAT repeat protein involved in plant immunity [78]. Furthermore, *NPQ4* (*AT1G44575*), encoding PSII-S (CP22), a pigment-binding protein associated with photosystem II (PSII) [79], exhibited distinct alternative splicing patterns.

Additionally, *DOXC21-A* (*AT3G19000*), *bHLH23* (*AT4G28790*) and *CYP76C2* (*AT2G45570*, an enzyme involved in monoterpenol metabolism [80]), displayed altered exon composition between ***nat2* and *doxc21-a***. Other affected genes included *AT3G15450*, *SIR* (*AT5G04590*), encoding sulfite reductase essential for sulfur metabolism [74], and *AT4G37280*, encoding an MRG family chromatin-binding protein.

DEU between the ***nat1* and *nat2*** mutant lines identified multiple genes exhibiting significant exon-level changes. Among them, TF *bHLH23* (*AT4G28790*), *LpxC4* (*AT1G25145*), encoding a UDP-3-O-acyl N-acetylglycosamine deacetylase family protein [81], *LARP1A* (*AT5G21160*), encoding an RNA-binding protein involved in mRNA degradation under stress [77], and *SEN1* (*AT4G35770*), a senescence-associated gene strongly induced by P starvation [82], displayed alternative exon usage. Additionally, *HEI10* (*AT1G53490*), encoding a RING/U-box superfamily protein that regulates the condensation of class I crossover factors during meiosis [67], exhibited substantial exon usage shifts, suggesting post-transcriptional regulation potentially impacting meiotic gene expression. Genes associated with protein modification, metabolic pathways, and signal transduction, such as *AT1G56300*, encoding a chaperone DnaJ- domain protein, were also differentially spliced.

These findings suggest that the disruption of *DOXC21-A* influences exon usage in key regulatory genes, potentially affecting transcriptional networks involved in stress responses and secondary metabolism. As expected, *DOXC21-A* (*AT3G19000*) displayed DEU specifically in comparisons involving the *doxc21-a* knockout line. Several genes with roles in abiotic stress signaling, development, and transcriptional regulation were affected, including *ABA1* (*AT5G67030*), *SMP2* (*AT4G37120*), and *NPQ4* (*AT1G44575*). Notably, isoform shifts were also observed in genes involved in hormone biosynthesis (*HDS*), translational control (*RABE1c, SIR*), and stress signaling (*SOS4, ILITHYIA*). These findings suggest that the loss *of DOXC21-A* and disruption of NAT transcription influence post-transcriptional regulatory mechanisms, likely contributing to the fine-tuning of stress adaptation in *A. thaliana*.

4.5. Gene family analysis reveals altered expression of selected transcription factors (TFs), ABC transporters, and cytochrome P450s (CYPs)

To elucidate the transcriptional regulatory responses associated with *DOXC21-A* disruption, we analyzed differentially expressed TFs from the MYB, WRKY, bHLH, and bZIP families. Heat maps were generated to illustrate their expression profiles across Col-0 WT, *nat* and *doxc21-a* mutant lines (**Fig. 10**). These TF families are known to mediate responses to environmental cues and developmental processes in plants. Specifically, MYB TFs regulate secondary metabolism, hormone signaling, and stress responses, often regulating phenylpropanoid biosynthesis and drought adaptation [83, 84]. WRKY TFs are central to plant immunity and systemic acquired resistance (SAR) [85], bZIP TFs control ABA signaling and energy homeostasis [86, 87], while bHLH TFs participate in light and hormone signaling, Fe homeostasis, and root development [88].

Among the MYB family (**Fig. 10A**), *MYBL2* (*AT1G71030*), a key repressor of proanthocyanidin biosynthesis [89], showed the highest expression in *nat1,* suggesting its involvement in modulating flavonoid pathway. *MYB59* (*AT5G59780*), associated with K^+/NO^{3-} translocation under potassium deficiency [90], was strongly downregulated in *doxc21-a*, indicating potential perturbation of ion homeostasis. *MYB112* (*AT1G48000*), which promotes anthocyanin accumulation and abiotic stress responses [91], was markedly upregulated in *nat1* and *nat2*. Conversely, *MYB7* (*AT2G16720*), a repressor of flavonol biosynthesis, was downregulated in *doxc21-a*, suggesting metabolic reprogramming in secondary metabolite pathways. *MYB31* (*AT1G74650*), involved in wax biosynthesis during reproductive development [92], was upregulated in *doxc21-a*. Notably, *MYB4* (*AT4G38620*), a suppressor of phenylpropanoid metabolism [93], showed reduced expression in *nat1* and *nat2*, pointing toward potential activation of phenylpropanoid-related pathways.

WRKY TFs (**Fig. 10B**) exhibited extensive differential expression, highlighting their role in stress-related transcriptional regulation. *WRKY18* (*AT4G31800*) displayed the most pronounced expression variation - strongly upregulated in *nat2* and Col-0 WT but repressed in *doxc21-a* suggesting involvement in biotic stress adaptation [94]. *WRKY40* (*AT1G80840*), a negative regulator of ABA signaling [95], exhibited a similar trend, with upregulation in *nat1* and *nat2* but strong downregulation in *doxc21-a*, indicating a potential shift in ABA homeostasis. *WRKY22* (*AT4G01250*), involved in ethylene signaling [96], was induced in *nat1* and *nat2* as it was shown in PPI *in silico* analysis (**Fig. 3B**). *WRKY46* (*AT2G46400*), regulator of hormone crosstalk [97], was downregulated in *doxc21-a*, but elevated in other genotypes. *WRKY75* (*AT5G13080*), which modulates P starvation and senescence [61], was repressed in *doxc21-a*, suggesting altered nutrient stress responses.

In the bZIP TF family (**Fig. 10C**), *bZIP1* (*AT5G49450*), known for its role in nitrogen (N)-responsive metabolic regulation [98], was markedly upregulated in *nat1*, suggesting an involvement in nutrient signaling reprogramming. *bZIP63* (*AT5G28770*), a key regulator of ABA-dependent energy signaling [99], also displayed elevated expression in *nat1*, indicating activation of stress-related metabolic adaptation mechanisms.

Within the bHLH TF (**Fig. 10D**), *bHLH* (*AT5G57150*), *bHLH23* (*AT4G28790*) and *bHLH56-like* (*AT4G28800*) showed the highest expression increases in *nat2*, although their functional roles remain uncharacterized. Notably, *bHLH150* (*AT3G05800*), associated with brassinosteroid signaling pathway [100], was significantly induced in both *nat2* and *doxc21-a* lines, potentially linking *DOXC21-A* disruption to hormone-regulated growth and stress response.

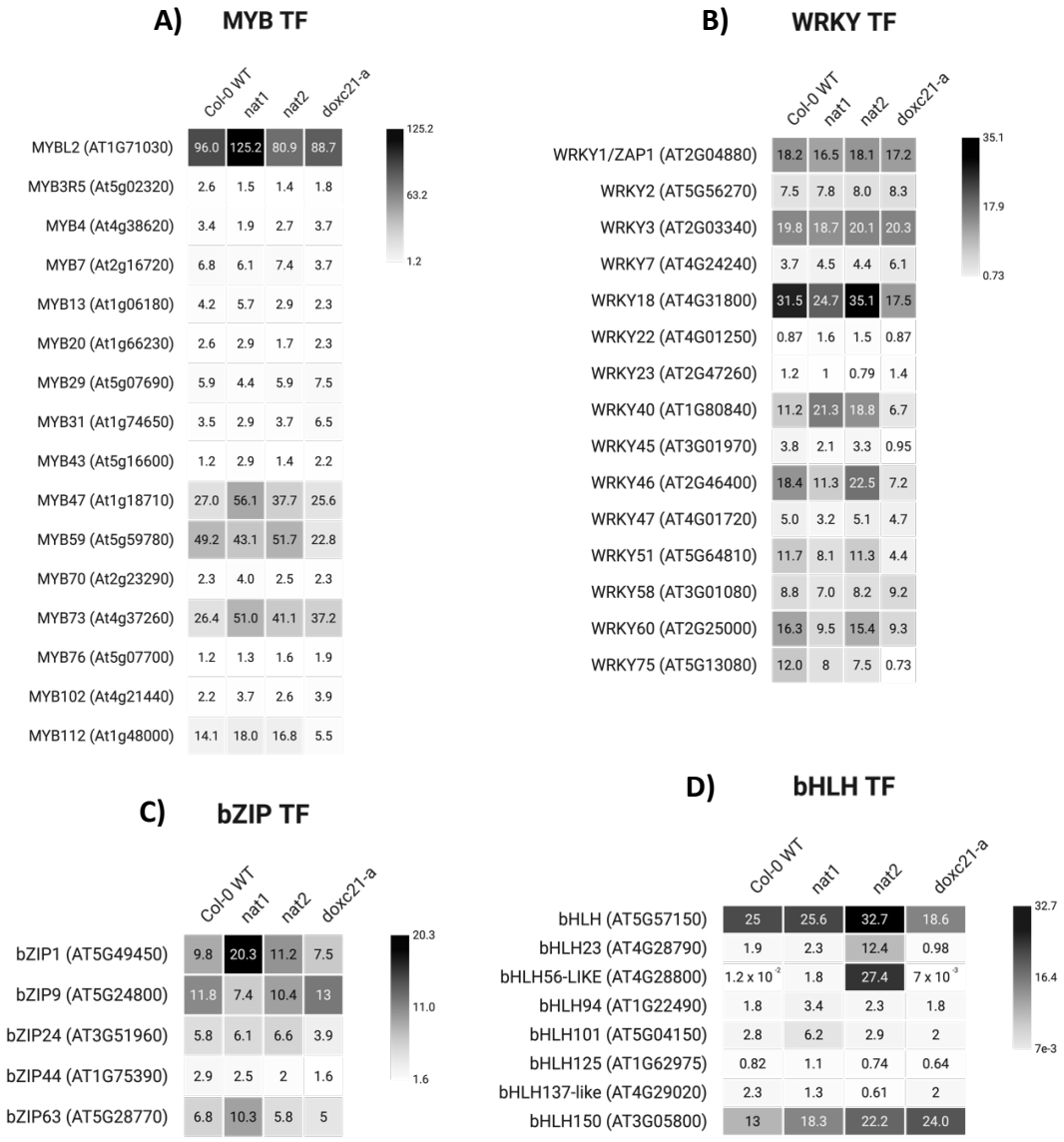


Figure 10. Expression profiles of selected TF families across Col-0 WT, *nat1*, *nat2*, and *doxc21-a* mutant lines under osmotic stress. Heatmaps display the average FPKM values derived from RNA-seq data for DEG belonging to the (A) MYB, (B) WRKY, (C) bZIP, and (D) bHLH TF families. Genes were selected based on statistically significant expression changes ($log_2FC \geq 1$, $p<0.05$) identified through differential expression analysis. FPKM values represent normalized transcript abundance and are shown here for comparative visualization of genotype-specific expression patterns. Color intensity reflects relative expression levels across samples.

Together, these findings highlight the distinct regulation of key TSs in *doxc21-a* and *nat* mutants, particularly those involved in hormonal signaling, nutrient regulation, secondary metabolism, and abiotic stress adaptation. These transcriptional shifts underscore the potential role of *DOXC21-A* as a central integrator of environmental signaling and stress-responsive transcriptional networks in *A. thaliana*.

ABC (ATP-binding cassette) transporters and CYP monooxygenases are two major families involved in plant metabolism, stress responses, and development [101] (**Fig. 11**). ABC transporters

actively shuttle phytohormones, secondary metabolites, and xenobiotics across cellular membranes, playing crucial roles in detoxification, nutrient transport, and defense [102]. CYP enzymes catalyze a wide range of oxidative reactions involved in the biosynthesis and modification of hormones (e.g., auxins, gibberellins, and brassinosteroids), and specialized metabolites such as glucosinolates and flavonoids [103]. Given their importance in stress adaptation, signaling, and metabolic regulation, both gene families were analyzed to determine the impact of *DOXC21-A* disruption on membrane transport and metabolic pathways across Col-0 WT, *nat* and *doxc21-a* mutant lines.

ABC transporter gene expression (**Fig. 11A**) showed clear genotype-dependent shifts. *ABCG36* (*PEN3/PDR8; AT1G59870*), which transports camalexin and indole-type metabolites and contributes to pathogen defense [37], was strongly upregulated in *doxc21-a* in comparison to all other genotypes, indicating that *DOXC21-A* may influence immunity by modulating antimicrobial compound secretion, ROS signaling, and auxin precursor efflux. In contrast, *ABCG40* (*PDR12; AT1G15520*), a multifunctional transporter implicated in camalexin export, ABA uptake, and heavy metal detoxification [37], was markedly downregulated in *doxc21-a*, suggesting suppression of ABA signaling. *ABCF5* (*GCN5; AT5G64840*), a histone acetyltransferase required for P starvation signaling, cuticular wax biosynthesis, and Fe homeostasis [104, 105], exhibited reduced expression in *doxc21-a*, reinforcing a potential role for *DOXC21-A* in transcriptional regulation *via* chromatin remodeling. Notably, *ABCC14* (*AT3G62700*), encoding a transporter of acylated anthocyanins [106], was significantly upregulated in *doxc21-a*, suggesting enhanced vacuolar sequestration of anthocyanins as a component of antioxidant defense. Additionally, *ABCG29* (*AT3G16340*), a monolignol exporter involved in lignin biosynthesis [107], was upregulated in *doxc21-a*, linking *DOXC21-A* to cell wall remodeling and structural adaptation during stress.

CYP expression patterns (**Fig. 11B**) revealed transcriptional alterations in genes central to hormone biosynthesis and defense-related metabolism. *CYP71A13* (*AT2G30770*), a key enzyme in camalexin biosynthesis [37], was strongly downregulated in *nat2* and *doxc21-a,* pointing to suppressed camalexin-mediated defense. Likewise, *CYP83B1* (*SUR2; AT4G31500*), a pivotal regulator of glucosinolate precursor formation, auxin/brassinosteroid balance, and phenylpropanoid metabolism [108, 109], showed a pronounced decrease in *nat1* and *nat2,* reflecting potential hormonal and metabolic dysregulation. Intriguingly, *CYP707A3* (*AT5G45340*), involved in ABA catabolism [110], was upregulated in *nat1* and *nat2*, while *CYP707A1* (*AT4G19230*), another key enzyme in ABA metabolism [111], remained relatively stable, indicating genotype-specific modulation of ABA turnover. *CYP84A1* which encodes *ferulate 5-hydroxylase* (*F5H, At4G36220*), a CYP-dependent monooxygenase that participates in the biosynthesis of the monolignol sinapyl alcohol, thereby influencing the accumulation of syringyl (S) lignin monomers in vascular plants [112], was significantly downregulated in *doxc21-a* in comparison to Col-0 WT and *nat* mutants. Finally, *CYP79F1* (*BUS1; AT1G16410*), a glucosinolate biosynthetic gene [113], was downregulated

in *nat1* and *nat2* but elevated in *doxc21-a*, highlighting potential complex regulation of sulfur-containing secondary metabolism by *DOXC21-A.*

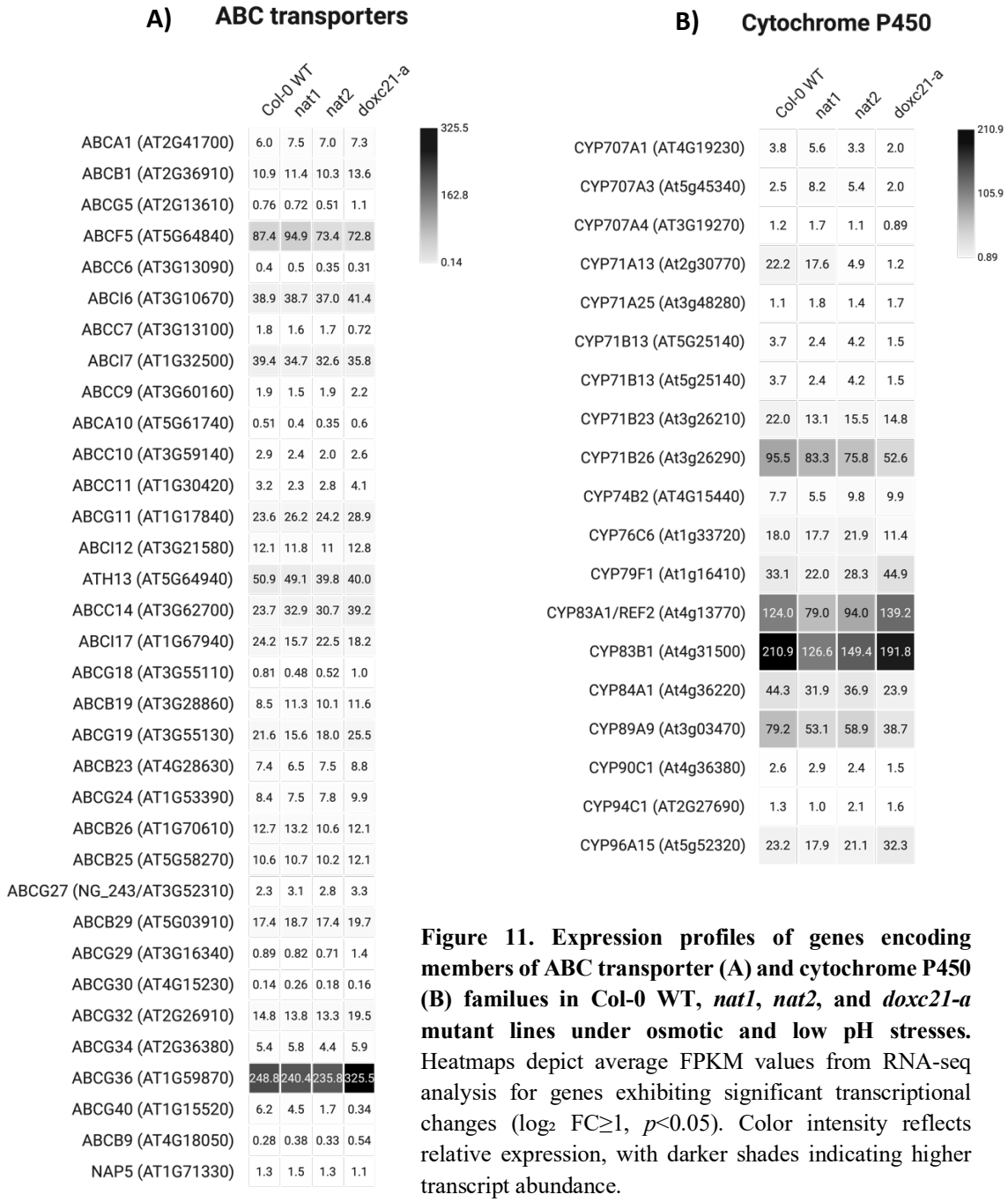


Figure 11. Expression profiles of genes encoding members of ABC transporter (A) and cytochrome P450 (B) familues in Col-0 WT, *nat1*, *nat2*, and *doxc21-a* mutant lines under osmotic and low pH stresses. Heatmaps depict average FPKM values from RNA-seq analysis for genes exhibiting significant transcriptional changes ($\log_2$ FC≥1, p<0.05). Color intensity reflects relative expression, with darker shades indicating higher transcript abundance.

Taken together, these results demonstrate that *DOXC21-A* disruption leads to extensive transcriptional reprogramming of genes in the ABC transporter and CYP families. Notable genotype-dependent differences were observed for *ABCG40, ABCB19, CYP71A13*, and *CYP707A3,* implicating shifts in hormone transport, defense response, and specialized metabolite biosynthesis. The altered expression of genes involved in hormone signaling, detoxification, and secondary

metabolite biosynthesis and transport suggests that *DOXC21-A* may play an integrative role in coordinating plant stress responses at the interface of metabolic and transcriptional regulation.

5. Targeted qRT-PCR validation confirms transcriptomic trends

5.1. *DOXC21-A*, *DOXC21-B*, *NAT*, and stress-responsive dioxygenases are differentially regulated

We performed qPCR analysis (**Fig. 12**) to gain deeper insights into the transcriptional regulation of *DOXC21-A* and *NAT*, and selected genes involved in stress response and secondary metabolism. Rosette leaves were collected from Col-0 WT, *nat1*, *nat2*, and *doxc21-a* plants grown hydroponically under control conditions (pH 4.5) and under mild osmotic stress (3% [w/v] PEG). This analysis aimed to validate transcriptomic trends and assess how disruption of *DOXC21-A* and NAT affects the expression of key regulatory and biosynthetic genes, particularly in the phenylpropanoid and lignin biosynthesis pathways.

As expected, expression analysis confirmed altered regulation of the targeted dioxygenase genes *DOXC21-A* and *DOXC21-B*, along with their putative regulatory lncRNA NAT (**Fig. 12A**). *DOXC21-A* expression was significantly upregulated in *nat1* ($P<0.0001$) and in *nat2* ($P<0.001$) under stress, while remaining unchanged in Col-0 WT. NAT transcript levels were significantly reduced in both *nat* mutants, confirming successful T-DNA disruption. In Col-0 WT, NAT expression showed a slight, non-significant increase under PEG treatment, whereas in *doxc21-a* it was significantly downregulated by this stress ($P<0.0001$ vs. Col-0 WT and control). *DOXC21-B* expression was modestly but significantly reduced in *nat2* under control conditions ($P<0.05$ *vs. nat1*).

We next examined three 2OGDs involved in specialized phenylpropanoid metabolism − F6'H1, F6'H2, and F3'H (**Fig. 12B**). *F6'H1*, a coumarin biosynthesis gene [20, 47], was strongly induced in Col-0 WT under osmotic stress ($P<0.0001$), a pattern not observed in mutant lines. *F6'H2* expression remained relatively stable across genotypes and treatments. While *F3'H* expression remained unchanged under stress in Col-0 WT, but in mutant lines showed a significant upregulation under stress compared to control ($P<0.001$ to $P<0.0001$), suggesting differential regulation of the flavonoid pathway.

5.2. Hormone- and phenylpropanoid-related genes show genotype-specific expression patterns

Genes involved in hormonal signaling and growth-stress integration, including *GA20OX1*, *GA20OX8* (*Gibberellin 20-oxidases*), and *JID1* (*Jasmonate-Induced Dioxygenase 1*), also displayed genotype- and stress-dependent expression patterns (**Fig. 12C**). Under PEG-induced stress conditions, *GA20OX1* was significantly upregulated under stress in *nat1*, *nat2,* and *doxc21-a* ($P<0.001$), with higher expression in *nat2* and *doxc21-a* compared to Col-0 WT. Notably, *GA20OX8* expression was significantly induced only in *nat2* under stress treatment. Interestingly, *JID1*

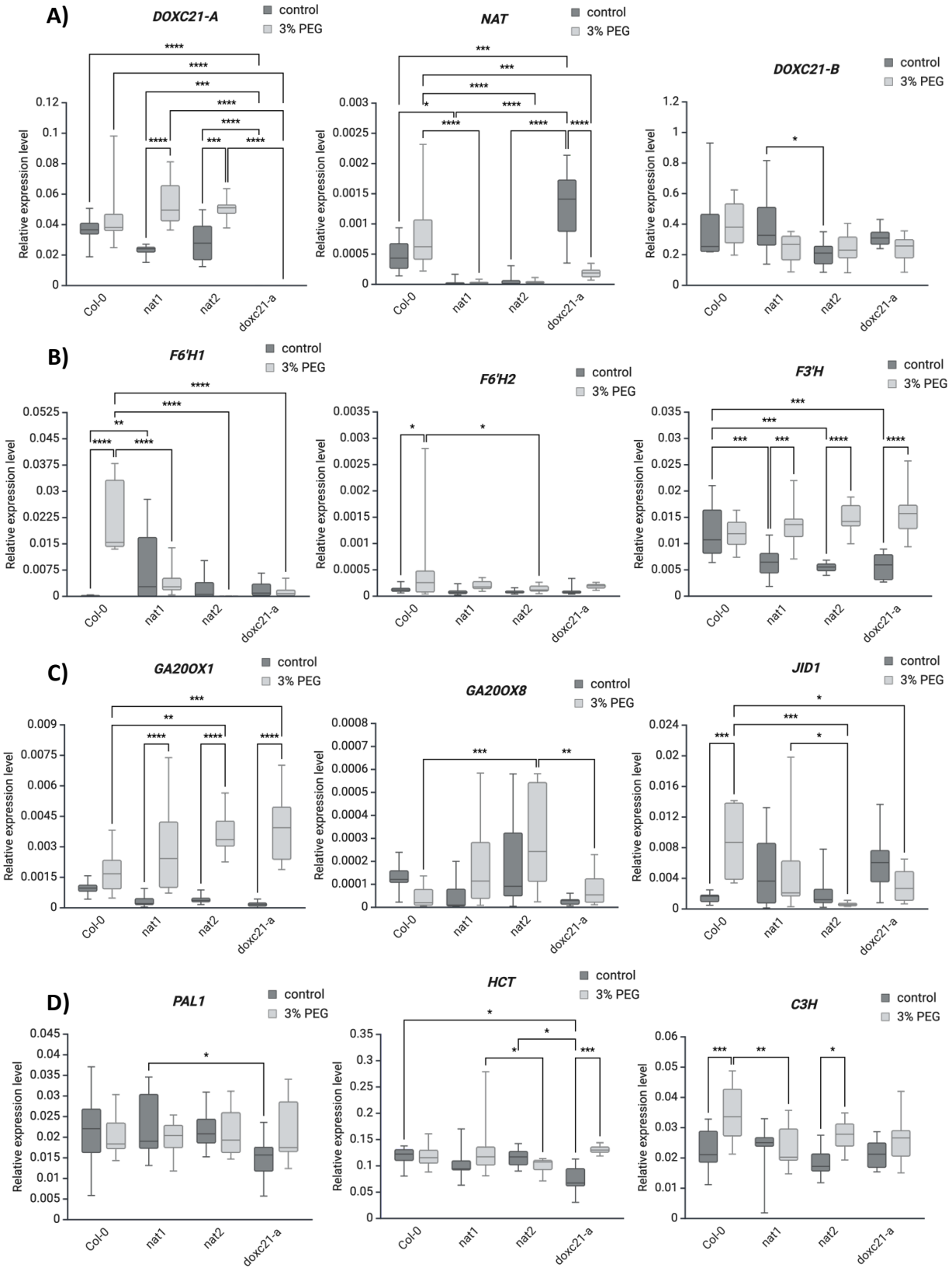
expression was significantly stress-induced in Col-0 WT (P<0.001), but remained unchanged in all mutant lines, indicating a possible disruption of JA-mediated signaling in the mutants.

We also examined early core phenylpropanoid pathway genes (**Fig. 12D**), including *PAL1 (Phenylalanine Ammonia-Lyase*), *C3H* (*p-Coumarate 3-Hydroxylase*) and HCT (*Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase*). *PAL1* was significantly downregulated in *doxc21-a* under control conditions (P<0.05). *C3H* expression increased under stress in Col-0 WT (P<0.001) and to a lesser extend in *nat2* (P<0.05). Under PEG-induced stress, *HCT* expression was specifically upregulated in *doxc21-a* under PEG treatment, but was not stress responsive in the other genotypes.

Analysis of lignin pathway genes (**Fig. 12E**), including *CCoAOMT1* (*Caffeoyl-CoA O-methyltransferase*), which converts caffeoyl-CoA to feruloyl-CoA and *CCR1* (*Cinnamoyl-CoA reductase*) responsible for the conversion of feruloyl-CoA to coniferaldehyde were also differentially regulated in tested lines. In response to stress, *CCoAOMT1* showed a strong induction in Col-0 WT under stress (P<0.0001), but this induction was either absent or reduced in all mutant lines. *CCR1* was significantly downregulated in *nat1* and *doxc21-a* under stress, suggesting impaired lignin biosynthetic potential. This downregulation pattern mirrored that of *GA20OX1* and *GA20OX8*, further indicating a broader repression of growth-related processes and secondary cell wall biosynthesis in these genotypes.

Altogether, these qRT-PCR results confirm the regulatory impact of *DOXC21-A* and NAT on phenylpropanoid metabolism and hormone-related gene expression. The observed patterns support the hypothesis that *DOXC21-A* and NAT function antagonistically to coordinate the transcriptional response to osmotic and low-pH stress by modulating growth and secondary metabolism pathways.

6.


A)
DOXC21-A
NAT
DOXC21-B
B)
F6'H1
F6'H2
F3'H
C)
GA20OX1
GA20OX8
JID1
D)
PAL1
HCT
C3H
control
3% PEG
Relative expression level
Col-0
nat1
nat2
doxc21-a

7.

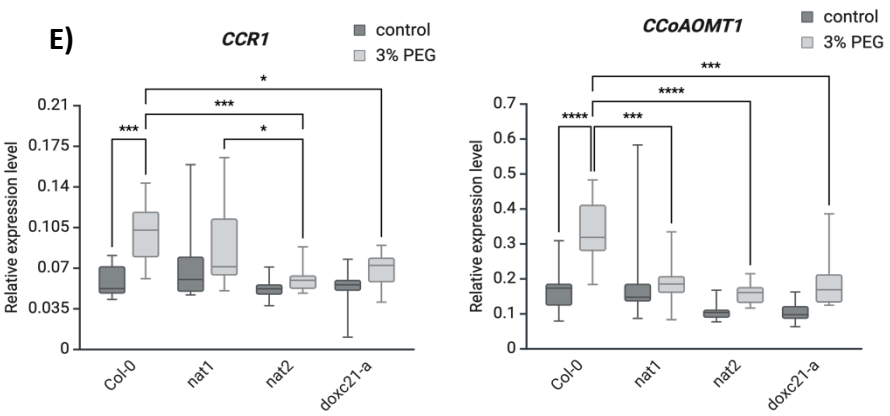


Figure 12. Expression analysis of *DOXC21-A*, *DOXC21-B*, NAT, and selected stress-responsive, phenylpropanoid, and hormone-related genes in rosette leaves of Col-0 WT and mutant lines under control and osmotic stress conditions. Rosette leaves were harvested from Col-0 WT, *nat1*, *nat2*, and *doxc21-a* plants cultivated hydroponically under control (pH 4.5) or 3% (w/v) PEG-induced osmotic stress. Gene expression was quantified *via* qRT-PCR, normalized to *ACT2* and *EF1-α*, and calculated using the $2^{-\Delta CT}$ method. Data are shown as box plots (n = 12; 4 biological × 3 technical replicates per condition). Statistical analysis was performed using two-way ANOVA with Bonferroni's multiple comparisons test ($P<0.05$, $P<0.01$, $P<0.001$, $P<0.0001$). **(A)** *DOXC21-A*, NAT and *DOXC21-B* expression level. **(B)** Expression level of 2OGDs involved in specialized phenylpropanoid metabolism − *F6'H1, F6'H2,* and *F3'H.* **(C)** Expression level of genes involved in hormonal signaling and growth-stress integration, including *GA20OX1*, *GA20OX8* (*Gibberellin 20-oxidases*), and *JID1*(*Jasmonate-Induced Dioxygenase 1*). **(D)** Expression level of early core phenylpropanoid pathway genes, including *PAL1 (Phenylalanine Ammonia-Lyase*), *C3H* (*p-Coumarate 3-Hydroxylase*) and HCT (*Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase*). **(E)** Expression level of downstream lignin biosynthesis genes *CCR1* (*Cinnamoyl-CoA reductase*) and *CCoAOMT1* (*Caffeoyl-CoA O-methyltransferase*).

7. Untargeted metabolomics reveals metabolomic reprogramming in mutants under osmotic stress

7.1. Orbitrap-MS-based metabolite profiling reveals distinct metabolite accumulation patterns in *nat2* and *doxc21-a*

To complement transcriptomic findings, untargeted metabolomics profiling was conducted on *A. thaliana* rosette leaves of Col-0 WT, *nat1, nat2*, and *doxc21-a* mutant lines cultivated hydroponically under control and 3% (w/v) PEG-induced osmotic stress conditions (pH 4.5). High-resolution Orbitrap-MS data were processed using the SIRIUS platform for compound prediction and pathway classification. This analysis revealed a diverse set of differentially accumulated metabolites ranging across primary and secondary metabolic pathways including phenylpropanoids, amino acids, and alkaloid derivatives (**Tab. 5**).

Consistent with transcriptomic patterns, *nat2* mutants exhibited elevated increased accumulation of several phenylpropanoid-related metabolites under both control and stress conditions. These included ferulic aldehyde (compound 356, $C_{11}H_{12}O_3$), a sulfur-containing coumarin derivative (compound 3015, $C_{16}H_{10}O_3S$), scoparone isomer (compound 95, $C_{11}H_{10}O_4$), and 4-methylumbelliferyl-β-D-glycoside (compound 1224, $C_{16}H_{18}O_8$). Conversely, the abundance

of a putative diglycosylated phenylpropanoid (compound 841, $C_{30}H_{30}O_{18}$) was significantly reduced in *nat2*.

Beyond phenylpropanoids, the *nat2* metabolome displayed enrichment in amino acid- and peptide-related compounds, suggesting broader metabolic reprogramming. Accumulation of N-(1-deoxy-1-fructosyl)phenylalanine (compound 295, $C_{15}H_{21}NO_7$) suggest activation of early Maillard-type reactions, likely reflecting enhanced sugar–amino acid conjugation under osmotic stress. Other upregulated N-containing compounds included compounds 521 ($C_{15}H_{19}NO_6$), compound 1347 ($C_{28}H_{30}N_8O_5$), and compound 2741 ($C_{14}H_{17}NO_8$) showing varying levels of annotation confidence.

In contrast to *nat2*, the *doxc21-a* line preferentially accumulated distinct N-rich metabolites. Notable examples include N-acetyl-L-tryptophan (compound 1087, $C_{13}H_{14}N_2O_3$, 100% match), the tripeptide H-Ile-Val-Pro-OH (compound 2454, $C_{16}H_{29}N_3O_4$), and additional amino acid derivatives such as compound 485 ($C_{14}H_{14}N_2O_5$) and 1991 ($C_{17}H_{24}N_2O_6$), indicative of altered peptide turnover or N recycling pathways.

Importantly, two alkaloid-like compounds were identified: compound 2913 ($C_{11}H_{18}N_7O_3P$), containing a phosphoramidate moiety, and compound 3245 ($C_{14}H_{12}N_8O_5$), consistent with a N-rich heterocyclic scaffold. Additionally, compound 2759 ($C_{16}H_{24}O_4$), classified as a putative terpenoid or modified benzenoid derivative, and two benzenoid-related compounds – compound 1754 ($C_{17}H_{23}NO_8$) and 2510 ($C_{18}H_{18}N_4O_{10}$) – were enriched specifically in the *nat2* mutant under PEG-induced stress, reinforcing genotype-specific shifts in specialized metabolism.

Finally, compound 2903 ($C_{18}H_{18}N_4O_{10}$) classified as an organic oxygen compound, was significantly upregulated in both *nat2* and *doxc21-a* mutants under stress, while compound 1523 ($C_{16}H_{20}N_6O$) assigned as indole derivative was markedly downregulated in *doxc21-a* during osmotic stress.

Altogether, the untargeted metabolomic data support the transcriptomic observations, confirming that *nat2* and *doxc21-a* lines activate distinct stress-related metabolic programs under low pH and mild osmotic stresses. These two mutant lines differ in *DOXC21-A* function, with *doxc21-a* representing a knockout and *nat2* carrying a centrally located T-DNA insertion within the NAT sequence that leads to *DOXC21-A* overexpression. These genotype-specific chemical signatures highlight a coordinated rewiring of transcriptional, post-transcriptional, and enzymatic processes that shape the plant's stress-responsive metabolic landscape. The observed patterns support genotype-specific reprogramming of specialized metabolism in response to osmotic stress.

Table 5. Differentially accumulated metabolites identified by untargeted metabolomics in *nat2* and *doxc21-a* mutants under osmotic stress. The table presents a selection of annotated metabolites that showed significant differential accumulation in *nat2* and *doxc21-a* mutant lines compared to Col-0 WT under 3% (w/v) PEG-induced osmotic stress in hydroponic culture (pH 4.5). Compound identification was based on high-resolution Orbitrap-MS analysis and processed using the SIRIUS platform for molecular formula prediction, chemical class assignment, and structural annotation confidence scoring. Only compounds with SIRIUS annotation confidence ≥45% and relevance to secondary or amino acid-related metabolism are included. For each entry, compound ID, molecular formula, assigned chemical class (e.g., phenylpropanoid, alkaloid, peptide), predicted compound name or closest match (when available), and annotation confidence score (%) are provided. Metabolites highlighted in *nat2* reflect upregulation of phenylpropanoid and nitrogen-containing conjugates, whereas *doxc21-a* is enriched in amino acid derivatives and stress-related peptides.

Genotype	Culture conditions	SIRIUS	UP/ DOWN	RT [min]	Compound assignment	SIRIUS assignment HCD	SIRIUS assignment CID	Compound. SIRIUS molecular formula assignment	m/z	Mass	SIRIUS Pathway/ SuperClass› assignment HCD/CID
nat2	3% PEG	95	UP	16.82	CNP0186764			C11 H10 O4	207.065	206.058	Shikimate/phenylpropanoid
nat2	Control	295	UP	4.35		CNP0288003 (95%)	135070649 (Pubchem. 48%)	C15 H21 N O7	328.139	327.132	Amino acids/small peptides
nat2	Control	356	UP	12.87	LTS0061649	CNP0207087 (74%)	CNP0207087 (74%)	C11 H12 O3	193.086	192.079	Shikimate/phenylpropanoid
nat2	3% PEG	373	DOWN	21.981		24805711 (Pubchem) (42%)	66875512 (Pubchem. 38%)	C16 H24 O4	303.157	280.168	Terpenoid (HCD). Benzenoid (CID)
doxc21-a	Control/3% PEG	485	UP	17.241		CNP0292306 (58%)		C14 H14 N2 O5	291.098	290.09	Amino acids/small peptides
nat2	Control	521	UP	4.35			73898384 (Pubchem. 46%)	C15 H19 N O6	310.129	309.121	Amino acids/small peptides
doxc21-a	3% PEG	615	UP	6.294				C27 H47 O6 P3	561.265	560.258	
nat2	Control	726	UP	17.15				C12 H22 N6 O6	347.168	346.16	
nat2	3% PEG	841	DOWN	16.718		131831345 (Pubchem. 78%)		C30 H30 O18	701.131	678.142	Shikimate/phenylpropanoid
doxc21-a	Control	918	UP	21.892					327.206	328.214	
nat2	Control/3% PEG	948	UP	8.53					223.017	224.025	
nat2	3% PEG	1011	DOWN	21.976				C14 H18 N6 O2	303.157	302.15	
doxc21-a	3% PEG	1042	UP	9.093				C15 H35 O8 P3	437.163	436.156	
doxc21-a	Control/3% PEG	1049	UP	14.869				C12 H13 N O5	252.087	251.079	
doxc21-a	Control/3% PEG	1087	UP	16.58	CNP0068645	CNP0293251 (100%)	CNP0293251 (100%)	C13 H14 N2 O3	247.108	246.101	Amino acids/small peptides
nat2	3% PEG	1154	DOWN	16.733				C28 H31 N4 O16 P	709.141	710.148	
nat2	Control	1160	UP	15.36	LTS0061649			C13 H28 O7	319.173	296.184	
nat2	Control	1215	UP	13.69				C11 H24 O6	275.147	252.157	
nat2	3% PEG	1224	UP	16.072	CNP0349962		131837581 (Pubchem. 65%)	C16 H18 O8	339.107	338.1	Shikimate/phenylpropanoid
nat2	3% PEG	1260	DOWN	17.21				C33 H30 N6 O10 P2	733.158	732.151	
nat2	3% PEG	1347	UP	13.993			68639828 (Pubchem. 58%)	C28 H30 N8 O5	559.24	558.233	Amino acids/small peptides
nat2	3% PEG	1420	UP	16.302				C54 H58 O23 P2	1137.292	1136.284	
doxc21-a	Control/3% PEG	1508	UP	14.871					250.064	251.071	
doxc21-a	3% PEG	1523	DOWN	20.425			46968347 (Pubchem. 60%)	C16 H20 N6 O	313.178	312.17	Indoles and derivatives

Genotype	Culture conditions	SIRIUS	UP/ DOWN	RT [min]	Compound assignment	SIRIUS assignment HCD	SIRIUS assignment CID	Compound. SIRIUS molecular formula assignment	m/z	Mass	SIRIUS Pathway/ SuperClass› assignment HCD/CID
doxc21-a	Control	1533	UP	4.734				C14 H19 N O6	298.129	297.121	
nat2	3% PEG	1666	DOWN	16.726				C31 H33 N4 O13 P S	733.158	732.151	
doxc21-a	Control	1689	UP	4.231				C10 H21 O6 P	267.1	268.107	
doxc21-a	3% PEG	1753	DOWN	20.877				C15 H20 O	217.159	216.152	
nat2/doxc21-a	3% PEG	1754	UP	14.86			133655698 (Pubchem. 50%)	C17 H23 N O8	370.15	369.142	Benzenoids
doxc21-a	Control	1761	UP	18.561				C22 H39 N O7	430.28	429.273	
doxc21-a	3% PEG	1765	DOWN	17.234					679.15	678.143	
nat2	3% PEG	1830	DOWN	16.722				C37 H20 N2 O14	717.099	716.091	
doxc21-a	3% PEG	1991	UP	15.922			88715888 (Pubchem. 45%)	C17 H24 N2 O6	375.153	352.164	Small peptides
nat2	3% PEG	2115	DOWN	16.735				C31 H30 N2 O13 P2	701.131	700.123	
nat2/doxc21-a	3% PEG	2122	UP	15.686				C16 H30 O8	351.201	350.194	
doxc21-a	3% PEG	2145	DOWN	20.872				C15 H18	199.148	198.141	
nat2/doxc21-a	3% PEG	2159	DOWN	17.197				C29 H34 O19 P2	749.125	748.118	
nat2/doxc21-a	3% PEG	2301	UP	19.846				C27 H40 O10 P2	585.202	586.209	
nat2	3% PEG	2306	UP	12.665				C16 H27 O6 P3	409.11	408.103	
nat2/doxc21-a	3% PEG	2363	DOWN	12.635				C7 H6 N6 O	191.068	190.061	
nat2	3% PEG	2401	DOWN	17.376				C22 H20 O15	525.088	524.08	
doxc21-a	Control	2454	UP	11.361		16035882 (Pubchem 89%)		C16 H29 N3 O4	328.223	327.216	Amino acids/small peptides
nat2/doxc21-a	3% PEG	2510	UP	14.624			88887463 (Pubchem. 49%)	C18 H18 N4 O10	449.094	450.102	Benzenoids
nat2/doxc21-a	3% PEG	2626	UP	17.618				C18 H18 N4 O9	433.1	434.107	
nat2	Control	2741	UP	10.8			148966 (Pubchem. 45%)	C16 H17 N O8	352.103	351.095	Amino acids
doxc21-a	3% PEG	2871	UP	9.35				C25 H36 O7 P2	511.2	510.193	
nat2/doxc21-a	3% PEG	2903	UP	12.948			54413794 (Pubchem. 45%)	C18 H18 N4 O10	449.094	450.102	Organic oxygen compounds
nat2	Control	2913	UP	4.37			135994930 (Pubchem. 48%)	C11 H18 N7 O3 P	326.113	327.121	Alkaloids
nat2	3% PEG	2932	DOWN	9.43				C22 H20 N8 O6 P2	555.107	554.1	
nat2	3% PEG	2955	UP	11.613				C26 H39 O22 P	733.159	734.167	
nat2	Control	3015	UP	7.57	CNP0283089	101906853 (Pubchem. 64%)		C16 H10 O3 S	283.043	282.035	Phenylpropanoid
nat2	3% PEG	3045	DOWN	15.664				C28 H33 N4 O17 P	727.151	728.158	
nat2/doxc21-a	Control	3100	DOWN	13.534	CNP0210534			C15 H16 O6	293.102	292.095	
nat2	3% PEG	3100	DOWN	13.534	CNP0210534			C15 H16 O6	293.102	292.095	
doxc21-a	Control	3231	UP	9.237				C17 H31 N3 O6	374.229	373.221	
nat2	Control	3245	UP	16.23		23478947 (Pubchem. 65%)		C14 H12 N8 O5	371.086	372.093	Alkaloid/pseudo alkaloid

6.2. Network-based clustering highlights genotype-specific shifts in metabolite classes in *nat2* and *doxc21-a*

To further explore the organization of differentially accumulated metabolites, Cytoscape-based network analysis was performed on untargeted Orbitrap-MS data using SIRIUS-derived fragmentation trees and chemical structural similarity (**Fig. 13**). This approach enabled the identification of metabolite subnetworks based on fragmentation pattern homology, revealing chemically related clusters altered under osmotic stress.

In the first set of subnetworks (**Fig. 13A**), which compares *nat2* mutant to Col-0 WT, several highly organized and structurally interconnected metabolite clusters were detected. These included nodes centered around compound 95 ($C_{11}H_{10}O_4$), and compound 356 ($C_{11}H_{12}O_3$, ferulic aldehyde), both of which were upregulated in *nat2*, indicating coordinated changes in phenylpropanoid metabolism. Compound 1224 ($C_{16}H_{18}O_8$, 4-methylumbelliferyl-β-D-glycoside) and compound 295 ($C_{15}H_{21}NO_7$, N-(1-deoxy-1-fructosyl)phenylalanine) were embedded within clusters linked to sugar–phenylpropanoid conjugation or Maillard-type reactivity.

The second group of subnetworks (**Fig. 13B**), reflecting the *doxc21-a* vs. Col-0 WT comparison, highlighted clusters enriched in N-containing small molecules. Compound 1087 ($C_{13}H_{14}N_2O_3$), confidently annotated (100%) as N-acetyl-L-tryptophan, formed a linear subnetwork together with additional tryptophan-related features, suggesting alterations in aromatic amino acid metabolism. A dense cluster around compound 2626 included structurally related nitrogenous molecules, possibly corresponding to flavonoid or alkaloid derivatives. Other subnetworks contained features such as compound 1042 ($C_{14}H_{14}N_2O_5$), 485 ($C_{14}H_{14}N_2O_5$), and peptide-like compound 2454 ($C_{16}H_{29}N_3O_4$), indicating shifts in N metabolism and peptide turnover in the *doxc21-a* line.

Together, these network-based analyses underscore distinct genotype-specific reprogramming of secondary metabolism under osmotic stress. Structural clustering provides high-resolution insights into chemical diversity modulated by *DOXC21-A* and its antisense regulator, revealing metabolite signatures consistent with transcriptomic alterations in phenylpropanoid, alkaloid, and N-related biosynthetic pathways.

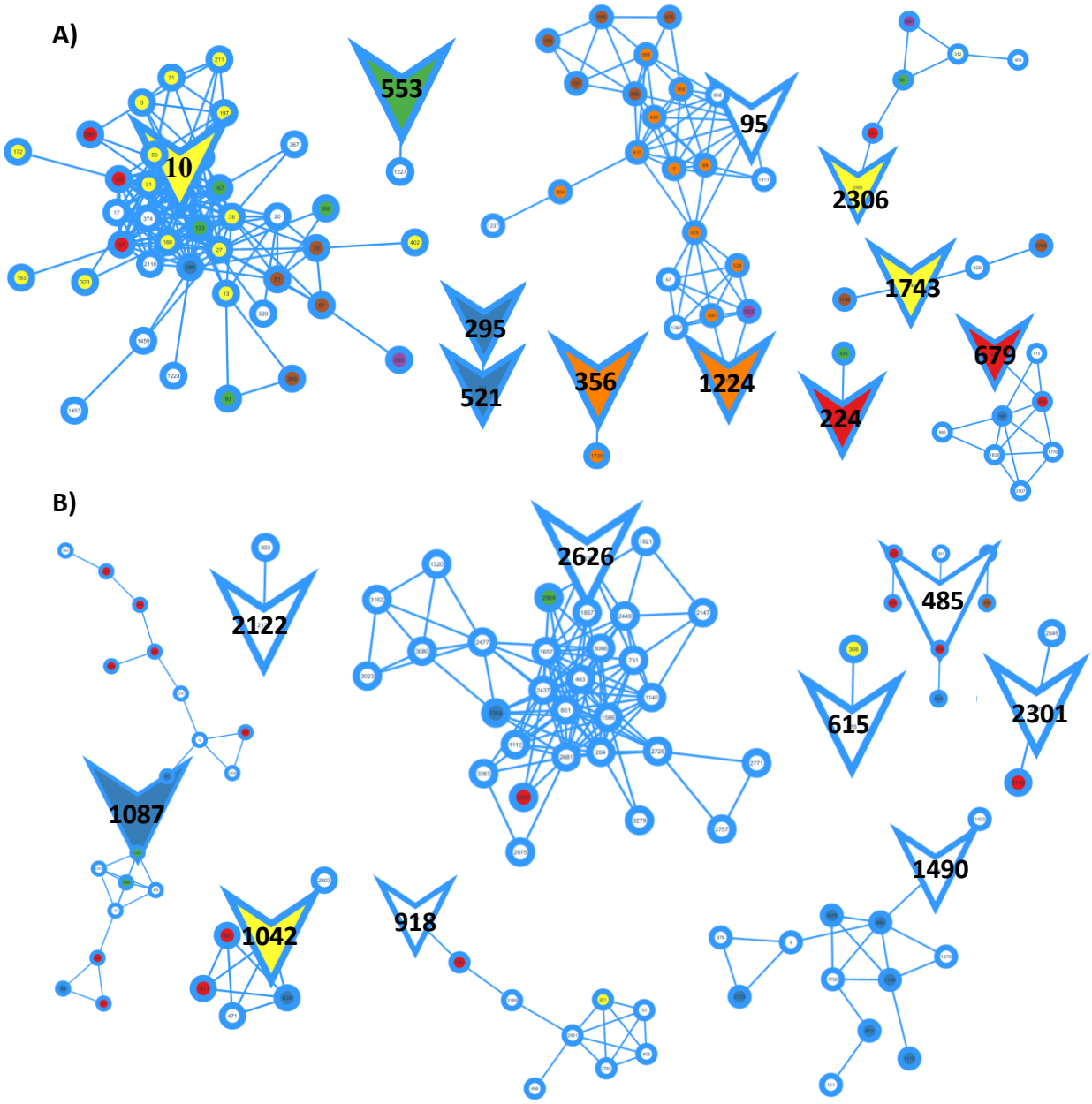


Figure 13. Network-based clustering of differentially accumulated metabolites in *nat2* (A) and *doxc21-a* (B) mutants under osmotic stress. Cytoscape networks represent structurally related metabolite clusters derived from untargeted Orbitrap-MS data, annotated using SIRIUS and grouped by fragmentation pattern similarity. Each node corresponds to a putative metabolite, and edges indicate structural relationships. Triangular markers highlight differentially abundant, annotated compounds in mutants compared to Col-0 WT under 3% (w/v) PEG-induced osmotic stress. **(A)** In *nat2*, selected clusters include phenylpropanoid-related glycosides (compounds 1224, 356) and sugar–amino acid conjugates (compounds 295, 521), consistent with enhanced secondary metabolism. **(B)** In *doxc21-a*, enriched subnetworks feature nitrogen (N)-containing metabolites such as compound 1087 (N-acetyl-L-tryptophan, 100%), compound 1042, and compound 485, as well as densely connected clusters centered on compound 2626 and compound 2301, indicative of altered amino acid turnover and N-rich metabolite pools.

8. Functional characterization of DOXC21 enzymes in heterologous expression systems

7.1. Transient expression of *DOXC21* homologs in *Nicotiana benthamiana* alters accumulation of non-nitrogenous and nitrogenous metabolites

To assess the metabolic impact of DOXC21 homologs *in planta*, we transiently expressed *DOXC21-A* and *DOXC21-B* in *N. benthamiana* leaves under a constitutive 35S promoter. Metabolites were extracted and profiled *via* UHPLC-MS across four independent experiments. Compound identification and quantification were performed using Compound Discoverer, and normalized peak areas were used to evaluate abundance.

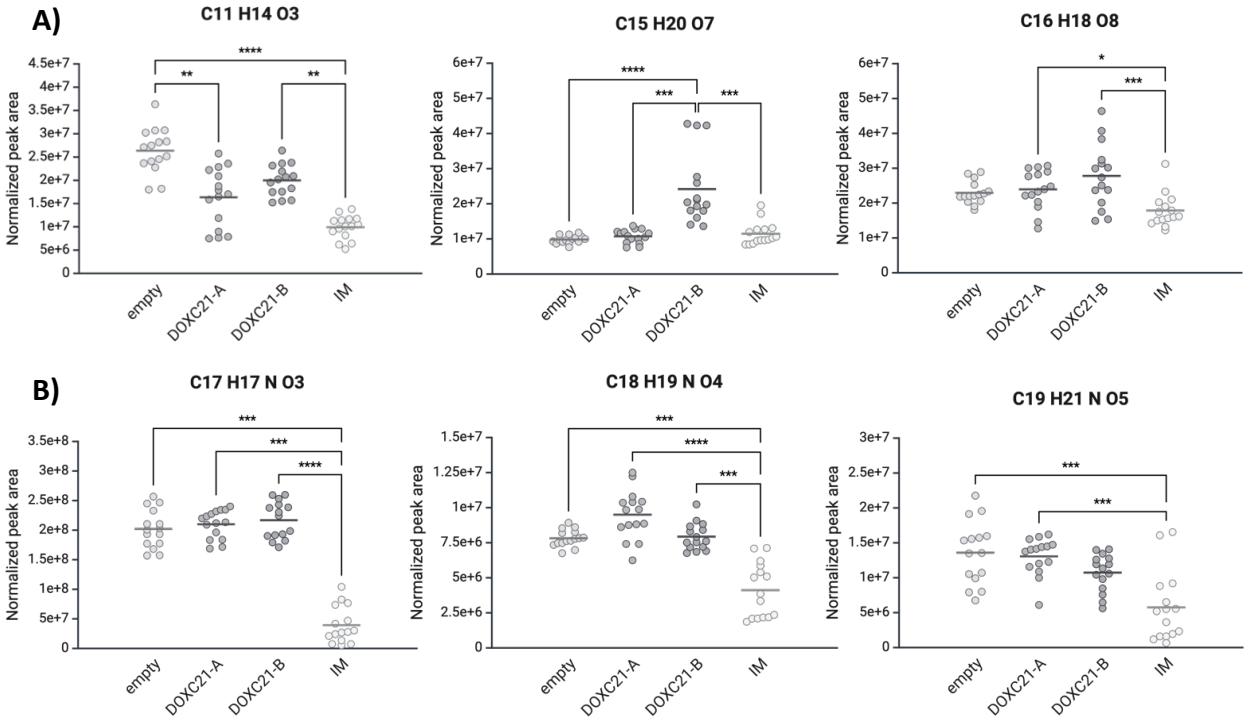


Figure 14. Metabolite profiling of *Nicotiana benthamiana* leaves transiently expressing *DOXC21-A* and *DOXC21-B*. Targeted metabolite analysis was performed on *N. benthamiana* leaves infiltrated with *Agrobacterium* strains carrying 35S:*DOXC21-A*, 35S:*DOXC21-B*, an empty vector control or infiltration medium (IM) as a mock. Leaves were harvested four days post-infiltration, and methanolic extracts were analyzed using high-resolution Orbitrap-MS across two independent biological experiments. Peak areas were normalized using Compound Discoverer software. **(A)** Non-nitrogenous metabolites showing significant differences across treatments, including compounds $C_{11}H_{14}O_3$, $C_{15}H_{20}O_7$, and $C_{16}H_{18}O_8$. Notably, *DOXC21-B* expression significantly elevated $C_{15}H_{20}O_7$ and $C_{16}H_{18}O_8$, while both *DOXC21-A* and *DOXC21-B* altered levels of $C_{11}H_{14}O_3$ compared to control. **(B)** Nitrogen-containing metabolites, including $C_{17}H_{17}N_3O_3$, $C_{18}H_{19}NO_4$, and $C_{19}H_{21}NO_5$, were also differentially accumulated, particularly in *DOXC21-B*-expressing samples. Each dot corresponds to an individual biological replicate. Statistical significance was determined using Kruskal-Wallis test with Dunn's multiple comparisons test: **P*<0.05*, ***P*<*0.01, ****P*<*0.001*, *****P*<0.0001.

As shown in **Fig. 14** (data from two experimental replicates), independent transient expression of *DOXC21-A* and *DOXC21-B* both altered the accumulation of several molecules when compared to the empty vector and mock controls. Panel A displays three non-nitrogenous compounds that showed significant genotype-specific differences. These included compounds annotated

as cinnamic acid derivatives $C_{11}H_{14}O_3$ (LTS0177974), significantly downregulated in DOXC21-A line (P<0.01) when compared to empty vector transformed leaves, and $C_{15}H_{20}O_7$ (LTS0221173), which was significantly upregulated in *DOXC21-B* transformed plants compared to control (P <0.001). Compound $C_{16}H_{18}O_8$ (CNP0349962, simple coumarin derivative) which accumulated also in *A. thaliana nat2* mutant (**Tab. 5**), was slightly upregulated in *DOXC-B* transformed leaves. Panel B focuses on N-containing compounds, revealing additional genotype-specific effects. For instance, $C_{18}H_{19}NO_4$ (CNP0280683) and $C_{17}H_{17}NO_3$ (CNP0416867), annotated as coumarin derivatives showed slight upregulation in *DOXC21-A* transformed leaves. On the other hand, $C_{19}H_{21}N_2O_5$ (CNP0238844) annotated as furanocumarin derivative, showed slightly decreased level in *DOXC21-B*.

Altogether, these data indicate that *DOXC21-A* and *DOXC21-B* modulate different branches of specialized metabolism when ectopically expressed in *N. benthamiana*. The observed shifts in N- and non-N-containing metabolites underscore the functional divergence between the two homologs and suggest that they may regulate complementary biosynthetic pathways relevant to stress response and secondary metabolism in *A. thaliana*.

7.2. *In vitro* enzymatic activity assays fail to confirm predicted dioxygenase activity for *DOXC21-A* and *DOXC21-B*

To biochemically validate the putative dioxygenase function of *DOXC21-A* and *DOXC21-B*, we performed *in vitro* enzymatic activity assays using recombinant proteins expressed and purified from *E. coli*. A curated panel of 70 structurally diverse substrates spanning monolignols, cinnamic acid derivatives, simple and methoxylated coumarins, flavonoids, benzenoids, phytohormones, indole-related metabolites, and phenylpropanoid-CoA esters was chosen. Substrates were selected to reflect key intermediates and branch points of phenylpropanoid metabolism and hormonal signaling, including caffeoyl-CoA, feruloyl-CoA, scopoletin, camalexin, sinapyl alcohol, and N-acetyltryptophan (**Tab. 1**).

DISCUSSION

The transition of plants from aquatic to terrestrial environments required extensive physiological, molecular, and regulatory adaptations. The expansion and diversification of 2OGDs was one of the central components to this evolutionary transition, as these enzymes played essential roles in several key adaptations, such as enhanced plant tolerance to drought, UV radiation, and temperature extremes, primarily by regulating the biosynthesis of secondary metabolites like flavonoids, lignins and other compounds. The 2OGDs are recognized for their pivotal roles not only in modulating stress responses and secondary metabolism, but they also contribute to the regulation of plant hormones, which are crucial for growth and development in the terrestrial environment. By fine-tuning metabolic pathways, plants could better survive in more challenging conditions on land.

In this study, we characterized DOXC21-A and DOXC21-B, two members of the DOXC21 clade, alongside the DOXC21-A associated NAT, proposing that together they form a multilayered regulatory unit integrating environmental stress signals with transcriptional and metabolic reprogramming in *A. thaliana*.

Our phylogenetic analysis rooted using *M. polymorpha* subsp. *ruderalis* − a representative of early-diverging liverworts – places DOXC21-A and DOXC21-B in a distinct clade conserved across monocots, eudicots, and basal angiosperms **(Fig. 2)**. This clade is clearly separated from canonical hormone biosynthesis dioxygenases such as GA20OXs or flavonoid-related enzymes (FLS, LDOX), suggesting an evolutionarily derived role for DOXC21-type proteins. The distinct topology supports a gene duplication followed by subfunctionalization model – a well-documented mechanism of stress-responsive gene evolution [114-116]. Such duplication events allow paralogs to specialize, providing selective advantages under fluctuating environmental conditions [117]. Dioxygenases frequently exhibit neo- or sub-functionalization, especially in stress-associated metabolic pathways [118, 119], with diversification in gene regulatory networks contributing to phenotype plasticity [120, 121]. The glucosinolate-modifying AOP dioxygenases, as well as enzymes involved in coumarin biosynthesis and lignin modification, provide illustrative examples of how duplication and regulatory divergence shape metabolic responses to environmental stimuli [118, 122, 123].

Despite high sequence similarity between DOXC21-A and DOXC21-B, our results highlight possible functional divergence between these paralogs. DOXC21-A, associated with an overlapping NAT, demonstrates regulatory flexibility through post-transcriptional modulation, while DOXC21-B lacks this antisense feature, potentially limiting its responsiveness to environmental stimuli [124]. The presence or absence of NATs is increasingly recognized as a key factor driving functional divergence among gene duplicates, often shaping their developmental timing and stress responsiveness [125, 126]. This supports emerging models where NATs influence paralog divergence by modifying mRNA stability, chromatin accessibility, or even local subnuclear architecture [127]. Protein interaction data further support their functional separation, with DOXC21-A integrated within stress and secondary metabolism networks, whereas DOXC21-B associates with components of light signaling, photosynthetic, and redox homeostasis **(Fig. 3BC)** [128-130]. These distinctions likely reflect partitioning of ancestral roles, with DOXC21-B specializing in light-dependent or photoprotective, consistent with established models of duplicate gene retention and functional specialization [117, 131]. Notably, *DOXC21-B* is upregulated in light-exposed *in vitro* root conditions **(Fig. 5B)** suggesting its involvement in light-responsive metabolic adjustment.

The identification of a conserved 29-nucleotide segment within the *AT3G19002* NAT transcript, overlapping the 5′ region of *DOXC21-A*, points to a potential regulatory mechanism involving sequence-based mimicry of plant miRNAs. The strong sequence similarity between this NAT region

and annotated or predicted miRNAs from diverse land plant species, including gra-miR8697 (*Gossypium*), ppt-miR1055 (*Physcomitrium*), and ath-miR844 (*Arabidopsis*) – suggests an evolutionarily conserved motif that may engage in post-transcriptional regulatory interactions. Such mimicry could enable the NAT to act either as a miRNA decoy or as a miRNA precursor itself, thereby influencing the stability or translation of stress-related transcripts. The presence of homologous sequences in both eudicots and basal land plants like *Amborella* and *Physcomitrium* underscores the deep conservation of this motif and reinforces the hypothesis that NATs contribute to ancient regulatory circuits in plant stress responses. Given that ath-miR844 has been previously implicated in biotic stress signaling [36], its partial sequence homology with the *DOXC21-A* NAT may reflect co-option of miRNA-like elements for abiotic stress adaptation. The alignment-derived sequence logo further highlights base preferences in the 5–20 position window, a region critical for target recognition in canonical miRNA interactions. Together, these findings suggest that *AT3G19002* may function beyond conventional antisense repression by modulating small RNA pathways or serving as a molecular decoy, adding a layer of regulatory complexity to the *DOXC21-A* locus [127]. Further experimental validation – such as small RNA sequencing and AGO1-binding assays – will be crucial to determine whether this NAT indeed participates in miRNA-mediated regulatory networks.

Beyond evolutionary insights, our work contributes to a broader understanding of the role of NATs in stress regulation. The reciprocal expression between DOXC21-A and its NAT mirrors similar inverse transcriptional relationships reported for other NAT-regulated gene pairs [132-134]. This regulatory arrangement likely provides a mechanism for fine-tuning gene expression dynamics, as exemplified by *SVALKA* NAT modulating cold-responsive *CBF* genes in *A. thaliana* [135]. NATs may operate through multiple mechanisms, including transcriptional interference, RNA duplex formation, siRNA-mediated pathways, or recruitment of chromatin-modifying complexes [124, 136-138], as documented in stress-responsive loci in *A. thaliana*, where histone modifications play a role in antisense-mediated gene silencing [133].

Under *in vitro* osmotic stress, *DOXC21-B* was the dominant transcript, but under soil-grown physiological drought, *DOXC21-A* expression rose sharply. These findings align with *DOXC21-B*'s potential light-responsiveness *in vitro* (roots exposed to ambient light, **Fig. 5B**), while *DOXC21-A* may respond to root-specific water deficits, mediated *via* ABA or auxin pathways. The TF binding profiles of their promoters reinforce this divergence: DOXC21-A contains ABRE, DRE/CRT, and WRKY sites, while *DOXC21-B* promoter is enriched in light-regulatory and redox-responsive motifs, suggesting photoperception sensitivity (see Results section 1.3).

The transcriptomic profiles of the *doxc21-a* and *nat* mutants reinforce the hypothesis that these two components function within distinct, yet overlapping, regulatory hierarchies. Although both mutants showed substantial transcriptomic reprogramming, the limited overlap between their DEGs suggest that DOXC21-A and its NAT target distinct aspects of stress adaptation [139, 140]. Pathways

consistently affected across both mutants included phenylpropanoid biosynthesis, ABC transporter activity, RNA splicing, and TF networks – core components of stress mitigation and developmental plasticity [141, 142]. However, regulatory directionality varied between genotypes: downregulation of lignin biosynthesis and upregulation of flavonoid metabolism were particularly prominent in the *nat* lines, aligning with an adaptive shift toward metabolic flexibility under osmotic stress [143-145].

This divergence in transcriptomic responses was reflected at the metabolomic level. In *nat2* mutant, we observed an enrichment of flavonoids and shikimate pathway derivatives – metabolites associated with antioxidant functions and hormone crosstalk, particularly with auxin and jasmonate signaling [146-148]. Conversely, the *doxc21-a* mutant exhibited broader metabolic remodeling, characterized by the accumulation of non-canonical stress-associated metabolites, including benzenoids, stress-related peptides, and atypical N-containing compounds, many of which are associated with generalized stress responses such as detoxification, pathogen defense, or signaling reprogramming [149]. Such distinct metabolic signatures suggest that the NAT directs flux toward specific defensive outputs, whereas DOXC21-Ainfluences broader metabolic adjustments, including potential cross-pathway regulation, consistent with patterns reported for F6′H dioxygenases and related 2OGDs [150, 151].

Transient expression of *DOXC21-A* and *DOXC21-B* in *N. benthamiana* further support their involvement in specialized metabolism, as indicated by reproducible alterations in metabolite profiles, particularly within phenylpropanoid and amino acid-derived pathways. Nevertheless, despite these metabolomic signatures, our attempts to detect direct enzymatic activity *in vitro* using a broad panel of candidate substrates were inconclusive, reflecting a common challenge within the 2OGD superfamily where context-specific activity, cofactor dependency, or protein-protein interactions are often required for full functionality [14, 15]. Enzymes of this class may rely on metabolite channeling through multi-protein assemblies or be regulated by local subcellular factors such as redox status or metal ion availability [152-155]. There is also possibility that DOXC21 proteins participate in regulatory roles beyond classical enzymatic activity – a phenomenon described as “moonlighting” behavior in other plant metabolic enzymes [156].

Despite extensive screening, no substrate conversion was observed for any tested compound with either *DOXC21-A* or *DOXC21-B* under standard assay conditions. In particular, no oxidative cleavage, hydroxylation, or rearrangement products were detected that would suggest classical 2OGD activity. These results are consistent across multiple independent replicates and enzyme batches. The lack of detectable activity suggests either: (I) the tested metabolites are not physiological substrates of *DOXC21* enzymes, (II) additional cofactors or interacting partners are required *in planta* for catalytic function, or (III) *DOXC21-A* and *DOXC21-B* may perform non-canonical or regulatory roles beyond classical enzymatic activity.

2OGDs are well known for their enzymatic roles in plant metabolism, however, a growing body of research now supports the idea that they may also have regulatory, and potentially non-enzymatic, roles in plant signaling, development, and stress responses [120, 157-161]. The lack of conclusive *in vitro* activity, combined with the transcriptomic and metabolomic evidence from mutant lines, supports the idea that *DOXC21* genes may function within a larger metabolic or signaling complex whose activity cannot be recapitulated in simplified biochemical assays.

CONCLUSIONS

Beyond their classical enzymatic roles, some 2OGDs emerge as regulatory hubs in plant signaling, development, and adaptation to stress, indicating a broader functional versatility of this gene family. Our integrative approach combining phylogenetic analysis, transcriptomics, metabolomics, and transient expression highlights the role of DOXC21-A as a regulatory hub linking abiotic stress signals with transcriptional and metabolomic responses. The dual-layer control *via* DOXC21-A and its NAT enables both structural and functional plasticity, allowing the plant to fine-tune resource allocation between growth and defense trade-offs models, where pathways such as jasmonate, auxin, and salicylic acid signaling dynamically adjust plant response to stress **(Fig. 15)** [162, 163].

These insights position DOXC21-A and its NAT as promising targets for regulatory engineering strategies aimed at enhancing crop resilience to climate variability [164, 165]. Unlike classical overexpression approaches, manipulation of endogenous regulatory modules offers the advantage of preserving developmental balance while promoting resilience. To fully realize the biotechnological potential of this system, future research should focus on: (I) Substrate discovery and protein interaction mapping; (II) miRNA and NAT mechanistic studies; (III) Characterizing the epigenetic landscape and chromatin-binding properties of the NAT and the *DOXC21-A* promoter region; (IV) Carbon assimilation studies under stress to connect DOXC21 function to growth–defense trade-offs and CO_2 homeostasis.

These approaches will be instrumental in uncovering the precise functional architecture of this regulatory hub and may provide a framework for similar strategies targeting other members of the 2OGD family. Advances in this area could contribute significantly to our understanding of how plants integrate environmental information at the molecular level and may inform novel strategies for enhancing crop resilience through regulatory reprogramming.

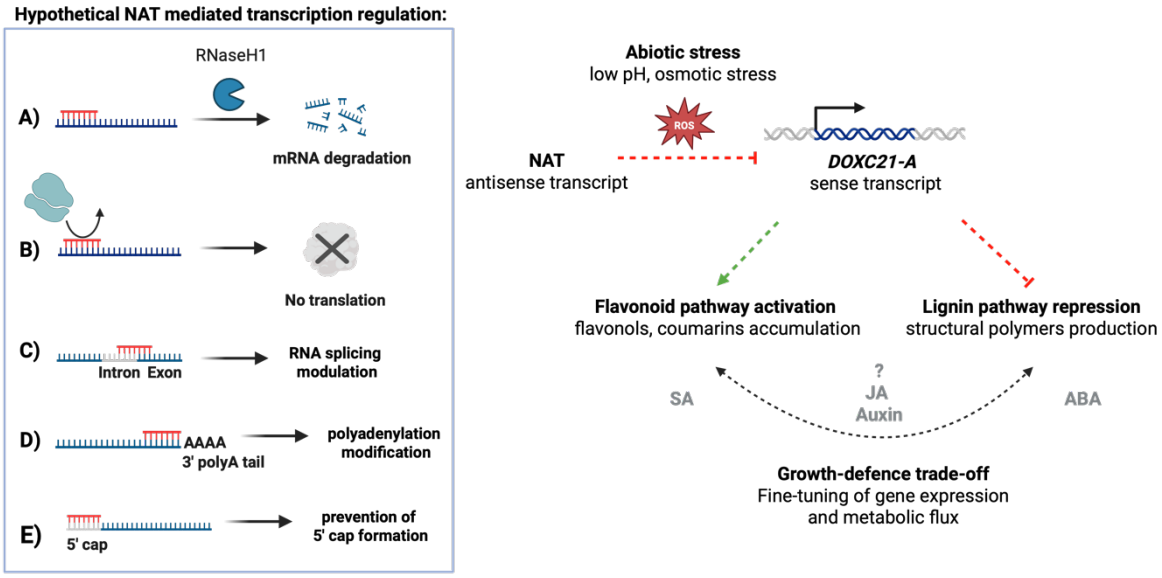


Figure 15. Proposed regulatory model of DOXC21-A/NAT-mediated transcriptional and metabolic regulation under environmental stress. The schematic illustrates the dual-layer regulatory mechanism involving the sense transcript *DOXC21-A* and its associated natural antisense transcript (NAT) in *A. thaliana* stress responses. Abiotic stress induces ROS production and influence NAT expression, which applies post-transcriptional repression on *DOXC21-A* through potential mechanisms including RNaseH1-dependent mRNA degradation (A), translation inhibition (B), modulation of RNA splicing (C), interference with polyadenylation (D), or prevention of 5' cap formation (E). *DOXC21-A* may act as a key regulatory hub, integrating stress cues with transcriptional control of the phenylpropanoid pathway. Activation of *DOXC21-A* promotes the flavonoid biosynthesis branch (e.g., flavonols, coumarins), enhancing antioxidant and signaling functions, while repressing lignin biosynthesis, an energetically demanding pathway involved in structural reinforcement. This metabolic partitioning contributes to the growth-defense trade-off, fine-tuning gene expression and metabolic flux in coordination with hormonal signals, including jasmonic acid (JA), salicylic acid (SA), abscisic acid (ABA), and auxin. The model highlights how NAT-mediated regulation adds flexibility and dynamic responsiveness to DOXC21-A function, allowing adaptive modulation of metabolic and developmental programs under abiotic stress.

DATA AVAILABILITY

All RNA sequencing data generated in this study will be deposited in the NCBI Sequence Read Archive (SRA) under the BioProject accession. Additional supporting datasets are provided in the Supplementary Materials. MSA used for phylogenetic analysis will be provided as separate Supplementary File.

ACKNOWLEDGEMENTS

We would like to thank Dr. Arthur Korte for his support with SNP analysis. We are also grateful to MSc Emilia Witwicka and MSc Alicja Januchowska for their valuable experimental input. Figures were generated with BioRender (https://www.biorender.com/). An artificial intelligence (AI) tool (ChatGPT) was employed exclusively to enhance the style and grammar of text segments originally written by the non-native English-speaking authors themselves. Its use was limited to language refinement, analogous to conventional word-processing or grammar-checking software.

FUNDING

This work was supported by the Polish National Science Centre grants: 2019/35/O/NZ1/02751 (EL) and UMO-2022/47/B/NZ2/01835 (AI); and the Polish National Agency for Academic Exchange grant PPN/BFR/2019/1/00050 (AI).

ATHORS CONTRIBUTION

I.P.: project conceptualization, methodology, investigation, data analysis, validation, writing – original draft preparation, writing – review and editing. A.O.: investigation, methodology optimization, supervision and manuscript review. F.M.: data analysis. C.C., T.B., A.D., and J.S.: investigation. E.L.: project conceptualization, supervision, funding acquisition, writing – review and editing; A.I.: project conceptualization, methodology, supervision, data analysis, validation, writing – review and editing, funding acquisition.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interests.

REFERENCES

[1] De Vries, J., & Archibald, J. M. (2018). Plant evolution: landmarks on the path to terrestrial life. *New Phytologist*, 217(4), 1428-1434.

[2] Foyer, C. H., & Kranner, I. (2023). Plant adaptation to climate change. *Biochemical Journal*, 480(22), 1865-1869.

[3] Kumar, P., Kumar, P., Verma, V., Irfan, M., Sharma, R., & Bhargava, B. (2023). How plants conquered land: evolution of terrestrial adaptation. *Journal of Evolutionary Biology*, 35(5), 5-14.

[4] Ménard, D., Blaschek, L., Kriechbaum, K., Lee, C. C., Serk, H., Zhu, C., ... & Pesquet, E. (2022). Plant biomechanics and resilience to environmental changes are controlled by specific lignin chemistries in each vascular cell type and morphotype. *The Plant Cell*, 34(12), 4877-4896.

[5] Schreiber, M., Rensing, S. A., & Gould, S. B. (2022). The greening ashore. *Trends in Plant Science*, 27(9), 847-857.

[6] Graham, L., Lewis, L. A., Taylor, W., Wellman, C., & Cook, M. (2014). Early terrestrialization: transition from algal to bryophyte grade. *Photosynthesis in bryophytes and early land plants*, 9-28.

[7] Takshak, S., & Agrawal, S. B. (2019). Defense potential of secondary metabolites in medicinal plants under UV-B stress. *Journal of Photochemistry and Photobiology B: Biology*, 193, 51-88.

[8] Wahab, A., Abdi, G., Saleem, M. H., Ali, B., Ullah, S., Shah, W., ... & Marc, R. A. (2022). Plants' physio-biochemical and phyto-hormonal responses to alleviate the adverse effects of drought stress: A comprehensive review. *Plants,* 11(13), 1620.

[9] Alseekh, S., & Fernie, A. R. (2023). Expanding our coverage: strategies to detect a greater range of metabolites. *Current Opinion in Plant Biology, 73*, 102335.

[10] Wu, M., Northen, T. R., & Ding, Y. (2023). Stressing the importance of plant specialized metabolites: omics-based approaches for discovering specialized metabolism in plant stress responses. *Frontiers in Plant Science*, 14, 1272363.

[11] Maeda, H. A. (2019). Evolutionary diversification of primary metabolism and its contribution to plant chemical diversity. *Frontiers in Plant Science*, 10, 881.

[12] Jan, R., Asaf, S., Numan, M., Lubna, & Kim, K. M. (2021). Plant secondary metabolite biosynthesis and transcriptional regulation in response to biotic and abiotic stress conditions. *Agronomy*, 11(5), 968

[13] Ha, C. M., Rao, X., Saxena, G., & Dixon, R. A. (2021). Growth–defense trade-offs and yield loss in plants with engineered cell walls. *New Phytologist,* 231(1), 60-74.

[14] Farrow, S. C., & Facchini, P. J. (2014). Functional diversity of 2-oxoglutarate/Fe (II)-dependent dioxygenases in plant metabolism. *Frontiers in plant science*, 5, 524.

[15] Kawai, Y., Ono, E., & Mizutani, M. (2014). Evolution and diversity of the 2–oxoglutarate-dependent dioxygenase superfamily in plants. *The Plant Journal*, 78(2), 328-343.

[16] Pyne, M. E., Narcross, L., & Martin, V. J. (2019). Engineering plant secondary metabolism in microbial systems. *Plant physiology*, 179(3), 844-861.

[17] Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., ... & Harter, K. (2007). The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *The Plant Journal*, 50(2), 347-363.

[18] Deforges, J., Reis, R. S., Jacquet, P., Sheppard, S., Gadekar, V. P., Hart-Smith, G., ... & Poirier, Y. (2019). Control of cognate sense mRNA translation by cis-natural antisense RNAs. *Plant physiology*, 180(1), 305-322.

[19] Owens, D. K., Alerding, A. B., Crosby, K. C., Bandara, A. B., Westwood, J. H., & Winkel, B. S. (2008). Functional analysis of a predicted flavonol synthase gene family in *Arabidopsis*. *Plant physiology*, 147(3), 1046-1061.

[20] Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., ... & Shimizu, B. I. (2008). Scopoletin is biosynthesized *via* ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in *Arabidopsis thaliana*. *The Plant Journal*, 55(6), 989-999

[21] Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant methods*, 9, 1-9.

[22] Tello-Ruiz MK, Jaiswal P, Ware D. (2022). Gramene: A Resource for Comparative Analysis of Plants Genomes and Pathways. *Methods Mol Biol*. 2443:101-131

[23] Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797.

[24] Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., & Barton, G. J. (2009). Jalview Version 2−a multiple sequence alignment editor and analysis workbench. *Bioinformatics,* 25(9), 1189–1191.

[25] Capra, J. A., & Singh, M. (2007). Predicting functionally important residues from sequence conservation. *Bioinformatics,* 23(15), 1875–1882

[26] Shenkin, P. S., Erman, B., & Mastrandrea, L. D. (1991). Information-theoretical entropy as a measure of sequence variability. *Proteins: Structure, Function, and Bioinformatics*, 11(4), 297-313.

[27] Heeg, C., Kruse, C., Jost, R., Gutensohn, M., Ruppert, T., Wirtz, M., & Hell, R. (2008). Analysis of the Arabidopsis O-acetylserine (thiol) lyase gene family demonstrates compartment-specific differences in the regulation of cysteine synthesis. *The Plant Cell*, 20(1), 168-185.

[28] Siwinska, J., Siatkowska, K., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., ... & Ihnatowicz, A. (2018). Scopoletin 8-hydroxylase: a novel enzyme involved in coumarin biosynthesis and iron-deficiency responses in Arabidopsis. *Journal of Experimental Botany*, 69(7), 1735-1748.

[29] Gillespie, D. P., Papio, G., & Kubota, C. (2021). High nutrient concentrations of hydroponic solution can improve growth and nutrient uptake of spinach (Spinacia oleracea L.) grown in acidic nutrient solution. *HortScience*, 56(6), 687-694.

[30] Young, M. D., Wakefield, M. J., Smyth, G. K., & Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome biology*, 11, 1-12.

[31] Mao, X., Cai, T., Olyarchuk, J. G., & Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics*, 21(19), 3787-3793.

[32] Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., ... & Yamanishi, Y. (2007). KEGG for linking genomes to life and the environment. *Nucleic acids research*, 36(suppl_1), D480-D484.

[33] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 13(11), 2498-2504.

[34] Vialart, G., Hehn, A., Olry, A., Ito, K., Krieger, C., Larbat, R., ... & Bourgaud, F. (2012). A 2-oxoglutarate-dependent dioxygenase from Ruta graveolens L. exhibits p-coumaroyl CoA 2′-hydroxylase activity (C2′ H): a missing step in the synthesis of umbelliferone in plants. *The Plant Journal*, 70(3), 460-470.

[35] Olry, A., Boschi-Muller, S., Yu, H., Burnel, D., & Branlant, G. (2005). Insights into the role of the metal binding site in methionine-R-sulfoxide reductases B. *Protein science*, 14(11), 2828-2837.

[36] Niu, D., Xia, J., Jiang, C., Qi, B., Ling, X., Lin, S., ... & Zhao, H. (2016). Bacillus cereus AR156 primes induced systemic resistance by suppressing miR825/825* and activating defense-related genes in Arabidopsis. *Journal of integrative plant biology*, 58(4), 426-439.

[37] He, Y., Xu, J., Wang, X., He, X., Wang, Y., Zhou, J., ... & Meng, X. (2019). The Arabidopsis pleiotropic drug resistance transporters PEN3 and PDR12 mediate camalexin secretion for resistance to *Botrytis cinerea*. *The Plant Cell*, 31(9), 2206-2222.

[38] Mucha, S., Heinzlmeir, S., Kriechbaumer, V., Strickland, B., Kirchhelle, C., Choudhary, M., ... & Glawischnig, E. (2019). The formation of a camalexin biosynthetic metabolon. *The Plant Cell*, 31(11), 2697-2710.

[39] Vorwerk, S., Biernacki, S., Hillebrand, H., Janzik, I., Müller, A., Weiler, E. W., & Piotrowski, M. (2001). Enzymatic characterization of the recombinant *Arabidopsis thaliana* nitrilase subfamily encoded by the NIT 2/NIT 1/NIT 3-gene cluster. *Planta,* 212, 508-516.

[40] Yang, Y., Xu, R., Ma, C. J., Vlot, A. C., Klessig, D. F., & Pichersky, E. (2008). Inactive methyl indole-3-acetic acid ester can be hydrolyzed and activated by several esterases belonging to the At MES esterase family of Arabidopsis. *Plant physiology*, 147(3), 1034-1045.

[41] Balsemão-Pires, E., Jaillais, Y., Olson, B. J., Andrade, L. R., Umen, J. G., Chory, J., & Sachetto-Martins, G. (2011). The Arabidopsis translocator protein (At TSPO) is regulated at multiple levels in response to salt stress and perturbations in tetrapyrrole metabolism. *BMC plant biology*, 11, 1-17.

[42] Hartmann, M., Zeier, T., Bernsdorff, F., Reichel-Deland, V., Kim, D., Hohmann, M., ... & Zeier, J. (2018). Flavin monooxygenase-generated N-hydroxypipecolic acid is a critical element of plant systemic immunity. *Cell*, 173(2), 456-469.

[43] Song, J. T., Lu, H., McDowell, J. M., & Greenberg, J. T. (2004). A key role for ALD1 in activation of local and systemic defenses in Arabidopsis. *The Plant Journal*, 40(2), 200-212.

[44] Chen, Y. L., Lin, F. W., Cheng, K. T., Chang, C. H., Hung, S. C., Efferth, T., & Chen, Y. R. (2023). XCP1 cleaves Pathogenesis-related protein 1 into CAPE9 for systemic immunity in Arabidopsis. *Nature Communications*, 14(1), 4697.

[45] Tehseen, M., Cairns, N., Sherson, S., & Cobbett, C. S. (2010). Metallochaperone-like genes in Arabidopsis thaliana. *Metallomics* 2, 556–564.

[46] Lin, Y. F., Liang, H. M., Yang, S. Y., Boch, A., Clemens, S., Chen, C. C., ... & Yeh, K. C. (2009). Arabidopsis IRT3 is a zinc-regulated and plasma membrane localized zinc/iron transporter. *New Phytologist,* 182(2), 392-404.

[47] Wang, N., Takada, M., Sakamoto, S., Vanholme, R., Goeminne, G., Kim, H., ... & Kajita, S. (2025). Effects of feruloyl-CoA 6′-hydroxylase 1 overexpression on lignin and cell wall characteristics in transgenic hybrid aspen. *Frontiers in Plant Science*, 16, 1543168.

[48] Cui, Y., Chen, C. L., Cui, M., Zhou, W. J., Wu, H. L., & Ling, H. Q. (2018). Four IVa bHLH transcription factors are novel interactors of FIT and mediate JA inhibition of iron uptake in Arabidopsis. *Molecular Plant*, 11(9), 1166-1183.

[49] Zhang, C., He, M., Jiang, Z., Liu, L., Pu, J., Zhang, W., ... & Xu, F. (2022). The xyloglucan endotransglucosylase/hydrolase gene XTH22/TCH4 regulates plant growth by disrupting the cell wall homeostasis in Arabidopsis under boron deficiency. *International Journal of Molecular Sciences*, 23(3), 1250.

[50] Le, C. T. T., Brumbarova, T., Ivanov, R., Stoof, C., Weber, E., Mohrbacher, J., ... & Bauer, P. (2016). Zinc finger of Arabidopsis thaliana12 (ZAT12) interacts with FER-like iron deficiency-induced transcription factor (FIT) linking iron deficiency and oxidative stress responses. *Plant Physiology*, 170(1), 540-557.

[51] Song, M. G., Bail, S., & Kiledjian, M. (2013). Multiple Nudix family proteins possess mRNA decapping activity. *Rna*, 19(3), 390-399.

[52] Liang, W., Li, C., Liu, F., Jiang, H., Li, S., Sun, J., ... & Li, C. (2009). The Arabidopsis homologs of CCR4-associated factor 1 show mRNA deadenylation activity and play a role in plant defence responses. *Cell research*, 19(3), 307-316.

[53] Chen, H. Y., Hsieh, E. J., Cheng, M. C., Chen, C. Y., Hwang, S. Y., & Lin, T. P. (2016). ORA47 (octadecanoid-responsive AP2/ERF-domain transcription factor 47) regulates jasmonic acid and abscisic acid biosynthesis and signaling through binding to a novel cis-element. *New Phytologist*, 211(2), 599-613.

[54] Rowland, O., Zheng, H., Hepworth, S. R., Lam, P., Jetter, R., & Kunst, L. (2006). CER4 encodes an alcohol-forming fatty acyl-coenzyme A reductase involved in cuticular wax production in Arabidopsis. *Plant physiology*, 142(3), 866-877.

[55] Gidda, S. K., Miersch, O., Levitin, A., Schmidt, J., Wasternack, C., & Varin, L. (2003). Biochemical and molecular characterization of a hydroxyjasmonate sulfotransferase from *Arabidopsis thaliana*. *Journal of Biological Chemistry*, 278(20), 17895-17900.

[56] Scholz, S. S., Vadassery, J., Heyer, M., Reichelt, M., Bender, K. W., Snedden, W. A., ... & Mithöfer, A. (2014). Mutation of the Arabidopsis calmodulin-like protein CML37 deregulates the jasmonate pathway and enhances susceptibility to herbivory. *Molecular plant*, 7(12), 1712-1726.

[57] Pré, M., Atallah, M., Champion, A., De Vos, M., Pieterse, C. M., & Memelink, J. (2008). The AP2/ERF domain transcription factor ORA59 integrates jasmonic acid and ethylene signals in plant defense. *Plant physiology*, 147(3), 1347-1357.

[58] Lokdarshi, A., Conner, W. C., McClintock, C., Li, T., & Roberts, D. M. (2016). Arabidopsis CML38, a calcium sensor that localizes to ribonucleoprotein complexes under hypoxia stress. *Plant physiology*, 170(2), 1046-1059.

[59] Zrenner, R., Riegler, H., Marquard, C. R., Lange, P. R., Geserick, C., Bartosz, C. E., ... & Slocum, R. D. (2009). A functional analysis of the pyrimidine catabolic pathway in Arabidopsis. *New phytologist*, 183(1), 117-132.

[60] Guo, B., Irigoyen, S., Fowler, T. B., & Versaw, W. K. (2008). Differential expression and phylogenetic analysis suggest specialization of plastid-localized members of the PHT4 phosphate transporter family for photosynthetic and heterotrophic tissues. *Plant signaling & behavior*, 3(10), 784-790.

[61] Guo, P., Li, Z., Huang, P., Li, B., Fang, S., Chu, J., & Guo, H. (2017). A tripartite amplification loop involving the transcription factor WRKY75, salicylic acid, and reactive oxygen species accelerates leaf senescence. *The Plant Cell*, 29(11), 2854-2870.

[62] Cellier, F., Conéjéro, G., Ricaud, L., Luu, D. T., Lepetit, M., Gosti, F., & Casse, F. (2004). Characterization of AtCHX17, a member of the cation/H+ exchangers, CHX family, from Arabidopsis thaliana suggests a role in K+ homeostasis. *The Plant Journal*, 39(6), 834-846.

[63] Khrouchtchova, A., Hansson, M., Paakkarinen, V., Vainonen, J. P., Zhang, S., Jensen, P. E., ... & Haldrup, A. (2005). A previously found thylakoid membrane protein of 14 kDa (TMP14) is a novel subunit of plant photosystem I and is designated PSI-P. *FEBS letters*, 579(21), 4808-4812.

[64] Rossini, S., Casazza, A. P., Engelmann, E. C., Havaux, M., Jennings, R. C., & Soave, C. (2006). Suppression of both ELIP1 and ELIP2 in Arabidopsis does not affect tolerance to photoinhibition and photooxidative stress. *Plant Physiology*, 141(4), 1264-1273.

[65] Dorca-Fornell, C., Gregis, V., Grandi, V., Coupland, G., Colombo, L., & Kater, M. M. (2011). The Arabidopsis SOC1-like genes AGL42, AGL71 and AGL72 promote flowering in the shoot apical and axillary meristems. *The Plant Journal*, 67(6), 1006-1017.

[66] Zhao, X., Gao, L., Ren, J., & Pan, F. (2019). Arabidopsis SIMILAR TO RCD-ONE genes are ubiquitous and respond to multiple abiotic stresses through diverse signaling pathways. *Journal of Biosciences*, 44(6), 129.

[67] Wang, T., Wang, H., Lian, Q., Jia, Q., You, C., Copenhaver, G. P., ... & Wang, Y. (2023). HEI10 is subject to phase separation and mediates RPA1a degradation during meiotic interference-sensitive crossover formation. *Proceedings of the National Academy of Sciences*, 120(52), e2310542120.

[68] Seemann, M., Wegner, P., Schünemann, V., Bui, B. T. S., Wolff, M., Marquet, A., ... & Rohmer, M. (2005). Isoprenoid biosynthesis in chloroplasts via the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GcpE) from Arabidopsis thaliana is a [4Fe–4S] protein. *JBIC Journal of Biological Inorganic Chemistry*, 10, 131-137.

[69] Yan, T., Heng, Y., Wang, W., Li, J., & Deng, X. W. (2022). SWELLMAP 2, a phyB-interacting splicing factor, negatively regulates seedling photomorphogenesis in Arabidopsis. *Frontiers in Plant Science*, 13, 836519.

[70] Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., & van Wijk, K. J. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PloS one*, 3(4), e1994.

[71] Clark, T. J., & Lu, Y. (2015). Analysis of loss-of-function mutants in aspartate kinase and homoserine dehydrogenase genes points to complexity in the regulation of aspartate-derived amino acid contents. *Plant Physiology*, 168(4), 1512-1526.

[72] Chen, D., He, L., Lin, M., Jing, Y., Liang, C., Liu, H., ... & Wang, M. (2021). A ras-related small GTP-binding protein, RabE1c, regulates stomatal movements and drought stress responses by mediating the interaction with ABA receptors. *Plant Science*, 306, 110858.

[73] Wang, J., Wang, Y., Lv, Q., Wang, L., Du, J., Bao, F., & He, Y. K. (2017). Nitric oxide modifies root growth by S-nitrosylation of plastidial glyceraldehyde-3-phosphate dehydrogenase. *Biochemical and Biophysical Research Communications,* 488(1), 88-94.

[74] Khan, M. S., Haas, F. H., Allboje Samami, A., Moghaddas Gholami, A., Bauer, A., Fellenberg, K., ... & Hell, R. (2010). Sulfite reductase defines a newly discovered bottleneck for assimilatory sulfate reduction and is essential for growth and development in *Arabidopsis thaliana*. *The Plant Cell*, 22(4), 1216-1231.

[75] Sánchez-Corrionero, Á., Sánchez-Vicente, I., González-Pérez, S., Corrales, A., Krieger-Liszkay, A., Lorenzo, Ó., & Arellano, J. B. (2017). Singlet oxygen triggers chloroplast rupture and cell death in the zeaxanthin epoxidase defective mutant aba1 of Arabidopsis thaliana under high light stress. *Journal of Plant Physiology,* 216, 188-196.

[76] Shi, H., & Zhu, J. K. (2002). SOS4, a pyridoxal kinase gene, is required for root hair development in Arabidopsis. *Plant Physiology*, 129(2), 585-593.

[77] Merret, R., Descombin, J., Juan, Y. T., Favory, J. J., Carpentier, M. C., Chaparro, C., ... & Bousquet-Antonelli, C. (2013). XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. *Cell reports*, 5(5), 1279-1293.

[78] Monaghan, J., & Li, X. (2010). The HEAT repeat protein ILITYHIA is required for plant immunity. *Plant and cell physiology*, 51(5), 742-753.

[79] Correa-Galvis, V., Poschmann, G., Melzer, M., Stühler, K., & Jahns, P. (2016). PsbS interactions involved in the activation of energy dissipation in Arabidopsis. *Nature plants*, 2(2), 1-8.

[80] Swaminathan, S., Morrone, D., Wang, Q., Fulton, D. B., & Peters, R. J. (2009). CYP76M7 is an ent-cassadiene C11α-hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. *The Plant Cell*, 21(10), 3315-3325.

[81] Li, C., Guan, Z., Liu, D., & Raetz, C. R. (2011). Pathway for lipid A biosynthesis in Arabidopsis thaliana resembling that of Escherichia coli. Proceedings of the National Academy of Sciences, 108(28), 11387-11392.

[82] Schenk, P. M., Kazan, K., Rusu, A. G., Manners, J. M., & Maclean, D. J. (2005). The SEN1 gene of Arabidopsis is regulated by signals that link plant defence responses and senescence. *Plant Physiology and Biochemistry*, 43(10-11), 997-1005.

[83] Ambawat, S., Sharma, P., Yadav, N. R., & Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. *Physiology and Molecular Biology of Plants*, 19, 307-321.

[84] Liu, J., Osbourn, A., & Ma, P. (2015). MYB transcription factors as regulators of phenylpropanoid metabolism in plants. *Molecular plant*, 8(5), 689-708.

[85] Saha, B., Nayak, J., Srivastava, R., Samal, S., Kumar, D., Chanwala, J., ... & Giri, M. K. (2024). Unraveling the involvement of WRKY TFs in regulating plant disease defense signaling. *Planta*, 259(1), 7.

[86] Dröge-Laser, W., & Weiste, C. (2018). The C/S1 bZIP network: a regulatory hub orchestrating plant energy homeostasis. *Trends in plant science*, 23(5), 422-433.

[87] Liu, H., Tang, X., Zhang, N., Li, S., & Si, H. (2023). Role of bZIP transcription factors in plant salt stress. *International Journal of Molecular Sciences*, 24(9), 7893.

[88] Gao, F., & Dubos, C. (2024). The arabidopsis bHLH transcription factor family. *Trends in plant science*, 29(6), 668-680.

[89] Jiang, W., Yin, Q., Liu, J., Su, X., Han, X., Li, Q., ... & Pang, Y. (2024). The APETALA2–MYBL2 module represses proanthocyanidin biosynthesis by affecting formation of the MBW complex in seeds of Arabidopsis thaliana. *Plant Communications*, *5*(3).

[90] Du, X. Q., Wang, F. L., Li, H., Jing, S., Yu, M., Li, J., ... & Wang, Y. (2019). The transcription factor MYB59 regulates K+/NO3− translocation in the Arabidopsis response to low K+ stress. *The Plant Cell*, 31(3), 699-714.

[91] Lotkowska, M. E., Tohge, T., Fernie, A. R., Xue, G. P., Balazadeh, S., & Mueller-Roeber, B. (2015). The Arabidopsis transcription factor MYB112 promotes anthocyanin formation during salinity and under high light stress. *Plant physiology*, *169*(3), 1862-1880.

[92] Shi, L., Chen, Y., Hong, J., Shen, G., Schreiber, L., Cohen, H., ... & Shi, J. (2022). AtMYB31 is a wax regulator associated with reproductive development in Arabidopsis. *Planta*, 256(2), 28.

[93] Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., ... & Martin, C. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *The EMBO journal.*

[94] Birkenbihl, R. P., Kracher, B., Roccaro, M., & Somssich, I. E. (2017). Induced genome-wide binding of three Arabidopsis WRKY transcription factors during early MAMP-triggered immunity. *The Plant Cell*, 29(1), 20-38.

[95] Geilen, K., & Böhmer, M. (2015). Dynamic subnuclear relocalization of WRKY40, a potential new mechanism of ABA-dependent transcription factor regulation. *Plant signaling & behavior,* 10(11), e1106659.

[96] Wang, Z., Wei, X., Cui, X., Wang, J., Wang, Y., Sun, M., ... & Jiang, Y. Q. (2024). The transcription factor WRKY22 modulates ethylene biosynthesis and root development through transactivating the transcription of ACS5 and ACO5 in Arabidopsis. *Physiologia Plantarum*, 176(3), e14371.

[97] Ding, Z. J., Yan, J. Y., Li, C. X., Li, G. X., Wu, Y. R., & Zheng, S. J. (2015). Transcription factor WRKY 46 modulates the development of Arabidopsis lateral roots in osmotic/salt stress conditions *via* regulation of ABA signaling and auxin homeostasis. *The Plant Journal*, 84(1), 56-69.

[98] Dietrich, K., Weltmeier, F., Ehlert, A., Weiste, C., Stahl, M., Harter, K., & Dröge-Laser, W. (2011). Heterodimers of the Arabidopsis transcription factors bZIP1 and bZIP53 reprogram amino acid metabolism during low energy stress. *The Plant Cell*, 23(1), 381-395.

[99] Viana, A. J., Matiolli, C. C., Newman, D. W., Vieira, J. G., Duarte, G. T., Martins, M. C., ... & Vincentz, M. (2021). The sugar-responsive circadian clock regulator bZIP63 modulates plant growth. *New Phytologist*, 231(5), 1875-1889.

[100] Wang, H., Zhu, Y., Fujioka, S., Asami, T., Li, J., & Li, J. (2009). Regulation of Arabidopsis brassinosteroid signaling by atypical basic helix-loop-helix proteins. *The Plant Cell*, 21(12), 3781-3791.

[101] Glombitza, S., Dubuis, P. H., Thulke, O., Welzl, G., Bovet, L., Götz, M., ... & Schäffner, A. (2004). Crosstalk and differential response to abiotic and biotic stressors reflected at the transcriptional level of effector genes from secondary metabolism. *Plant molecular biology*, 54, 817-835.

[102] Kretzschmar, T., Burla, B., Lee, Y., Martinoia, E., & Nagy, R. (2011). Functions of ABC transporters in plants. *Essays in Biochemistry,* 50, 145-160.

[103] Singh, A., Panwar, R., Mittal, P., Hassan, M. I., & Singh, I. K. (2021). Plant cytochrome P450s: Role in stress tolerance and potential applications for human welfare. *International Journal of Biological Macromolecules*, 184, 874-886.

[104] Wang, T., Xing, J., Liu, X., Yao, Y., Hu, Z., Peng, H., ... & Ni, Z. (2018). GCN5 contributes to stem cuticular wax biosynthesis by histone acetylation of CER3 in Arabidopsis. *Journal of experimental botany*, 69(12), 2911-2922.

[105] Wang, T., Xing, J., Liu, Z., Zheng, M., Yao, Y., Hu, Z., ... & Ni, Z. (2019). Histone acetyltransferase GCN5-mediated regulation of long non-coding RNA At4 contributes to phosphate starvation response in Arabidopsis. *Journal of experimental botany*, 70(21), 6337-6348.

[106] Dean, J. V., Willis, M., & Shaban, L. (2022). Transport of acylated anthocyanins by the Arabidopsis ATP-binding cassette transporters AtABCC1, AtABCC2, and AtABCC14. *Physiologia plantarum*, 174(5), e13780.

[107] Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Park, J., ... & Martinoia, E. (2012). AtABCG29 is a monolignol transporter involved in lignin biosynthesis. *Current Biology*, 22(13), 1207-1212.

[108] Goto, C., Ikegami, A., Goh, T., Maruyama, K., Kasahara, H., Takebayashi, Y., ... & Fukaki, H. (2023). Genetic interaction between Arabidopsis SUR2/CYP83B1 and GNOM indicates the importance of stabilizing local auxin accumulation in lateral root initiation. *Plant and Cell Physiology*, 64(10), 1178-1188.

[109] Kim, J. I., Dolan, W. L., Anderson, N. A., & Chapple, C. (2015). Indole glucosinolate biosynthesis limits phenylpropanoid accumulation in *Arabidopsis thaliana*. *The Plant Cell,* 27(5), 1529-1546.

[110] Umezawa, T., Okamoto, M., Kushiro, T., Nambara, E., Oono, Y., Seki, M., ... & Shinozaki, K. (2006). CYP707A3, a major ABA 8′-hydroxylase involved in dehydration and rehydration response in *Arabidopsis thaliana*. *The Plant Journal*, 46(2), 171-182.

[111] Saito, S., Hirai, N., Matsumoto, C., Ohigashi, H., Ohta, D., Sakata, K., & Mizutani, M. (2004). Arabidopsis CYP707A s encode (+)-abscisic acid 8′-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. *Plant physiology*, 134(4), 1439-1449.

[112] Fan, D., Li, C., Fan, C., Hu, J., Li, J., Yao, S., ... & Luo, K. (2020). MicroRNA6443-mediated regulation of FERULATE 5-HYDROXYLASE gene alters lignin composition and enhances saccharification in *Populus tomentosa*. *New Phytologist*, 226(2), 410-425.

[113] Chen, S., Glawischnig, E., Jørgensen, K., Naur, P., Jørgensen, B., Olsen, C. E., ... & Halkier, B. A. (2003). CYP79F1 and CYP79F2 have distinct functions in the biosynthesis of aliphatic glucosinolates in *Arabidopsis*. *The Plant Journal*, 33(5), 923-937.

[114] Kliebenstein, D. J. (2008). A role for gene duplication and natural variation of gene expression in the evolution of metabolism. PloS one, 3(3), e1838.

[115] Hanada, K., Kuromori, T., Myouga, F., Toyoda, T., Li, W. H., & Shinozaki, K. (2009). Evolutionary persistence of functional compensation by duplicate genes in *Arabidopsis. Genome biology and evolution*, 1, 409-414.

[116] Panchy, N., Lehti-Shiu, M., & Shiu, S. H. (2016). Evolution of gene duplication in plants. *Plant physiology*, 171(4), 2294-2316.

[117] Qian, W., & Zhang, J. (2014). Genomic evidence for adaptation by gene duplication. *Genome research*, 24(8), 1356-1362.

[118] Barco, B., & Clay, N. K. (2019). Evolution of glucosinolate diversity via whole-genome duplications, gene rearrangements, and substrate promiscuity. Annual Review of Plant Biology, 70(1), 585-604.

[119] Liu, C., Wu, Y., Liu, Y., Yang, L., Dong, R., Jiang, L., ... & Luo, L. (2021). Genome-wide analysis of tandem duplicated genes and their contribution to stress resistance in pigeonpea (Cajanus cajan). Genomics, 113(1), 728-735.

[120] Hagel, J. M., & Facchini, P. J. (2018). Expanding the roles for 2-oxoglutarate-dependent oxygenases in plant metabolism. *Natural Product Reports*, 35(8), 721-734.

[121] Jiang, D., Li, G., Chen, G., Lei, J., Cao, B., & Chen, C. (2021). Genome-wide identification and expression profiling of 2OGD superfamily genes from three brassica plants. *Genes*, 12(9), 1399.

[122] Kliebenstein, D. J., & Osbourn, A. (2012). Making new molecules–evolution of pathways for novel metabolites in plants. *Current opinion in plant biology*, 15(4), 415-423.

[123] Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151(4), 1531-1545.

[124] Pelechano, V., & Steinmetz, L. M. (2013). Gene regulation by antisense transcription. *Nature Reviews Genetics,* 14(12), 880-893.

[125] Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., ... & Chua, N. H. (2012). Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. *The Plant Cell*, 24(11), 4333-4345.

[126] Wang, H. L. V., & Chekanova, J. A. (2017). Long noncoding RNAs in plants. *Long non coding RNA biology*, 133-154.

[127] Magistri, M., Faghihi, M. A., St Laurent, G., & Wahlestedt, C. (2012). Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. *Trends in genetics*, 28(8), 389-396.

[128] Noctor, G., & Foyer, C. H. (1998). Ascorbate and glutathione: keeping active oxygen under control. *Annual review of plant biology*, 49(1), 249-279.

[129] Dietz, K. J. (2015). Efficient high light acclimation involves rapid processes at multiple mechanistic levels. Journal of experimental botany, 66(9), 2401-2414.

[130] Jing, Y., & Lin, R. (2020). Transcriptional regulatory network of the light signaling pathways. New Phytologist, 227(3), 683-697.
[131] Moghe, G. D., & Last, R. L. (2015). Something old, something new: conserved enzymes and the evolution of novelty in plant specialized metabolism. *Plant Physiology*, 169(3), 1512-1523.
[132] Lapidot, M., & Pilpel, Y. (2006). Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. EMBO reports, 7(12), 1216-1222.
[133] Swiezewski, S., Liu, F., Magusin, A., & Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature, 462(7274), 799-802.
[134] Yuan, C., Wang, J., Harrison, A. P., Meng, X., Chen, D., & Chen, M. (2015). Genome-wide view of natural antisense transcripts in *Arabidopsis thaliana. DNA research*, 22(3), 233-243.
[135] Kiger, N. M., & Schroeder, S. J. (2024). SVALKA: A Long Noncoding Cis-Natural Antisense RNA That Plays a Role in the Regulation of the Cold Response of *Arabidopsis thaliana. Non-Coding RNA*, 10(6), 59.
[136] Faghihi, M. A., & Wahlestedt, C. (2009). Regulatory roles of natural antisense transcripts. *Nature reviews Molecular cell biology*, 10(9), 637-643.
[137] Heo, J. B., & Sung, S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, 331(6013), 76-79.
[138] Rinn, J. L., & Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. *Annual review of biochemistry*, 81(1), 145-166.
[139] Wang, H., Chung, P. J., Liu, J., Jang, I. C., Kean, M. J., Xu, J., & Chua, N. H. (2014). Genome-wide identification of long noncoding natural antisense transcripts and their responses to light in Arabidopsis. *Genome research*, 24(3), 444-453.
[140] Chekanova, J. A. (2015). Long non-coding RNAs and their functions in plants. *Current opinion in plant biology*, 27, 207-216.
[141] Badri, D. V., Loyola-Vargas, V. M., Du, J., Stermitz, F. R., Broeckling, C. D., Iglesias-Andreu, L., & Vivanco, J. M. (2008). Transcriptome analysis of Arabidopsis roots treated with signaling compounds: a focus on signal transduction, metabolic regulation and secretion. *New Phytologist*, 179(1), 209-223.
[142] Chezem, W. R., & Clay, N. K. (2016). Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs. *Phytochemistry*, 131, 26-43.
[143] Falcone Ferreyra, M. L., Rius, S. P., & Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Frontiers in plant science*, 3, 222.
[144] Wu, P., Liu, A., & Li, L. (2021). Metabolomics and transcriptome analysis of the biosynthesis mechanism of flavonoids in the seeds of Euryale ferox Salisb at different developmental stages. *Molecular Genetics and Genomics*, 296(4), 953-970.
[145] Zhu, S., Cui, M., & Zhao, Q. (2024). Characterization of the 2ODD genes of DOXC subfamily and its members involved in flavonoids biosynthesis in S cutellaria baicalensis. *BMC Plant Biology*, 24(1), 804.
[146] Tohge, T., Watanabe, M., Hoefgen, R., & Fernie, A. R. (2013). Shikimate and phenylalanine biosynthesis in the green lineage. *Frontiers in plant science*, 4, 62.
[147] Brunetti, C., Fini, A., Sebastiani, F., Gori, A., & Tattini, M. (2018). Modulation of phytohormone signaling: A primary function of flavonoids in plant–environment interactions. *Frontiers in Plant science*, 9, 1042.

[148] Kuhn, B. M., Errafi, S., Bucher, R., Dobrev, P., Geisler, M., Bigler, L., ... & Ringli, C. (2016). 7-Rhamnosylated flavonols modulate homeostasis of the plant hormone auxin and affect plant development. *Journal of Biological Chemistry*, 291(10), 5385-5395.

[149] Obata, T., & Fernie, A. R. (2012). The use of metabolomics to dissect plant responses to abiotic stresses. *Cellular and Molecular Life Sciences,* 69, 3225-3243.

[150] Miedes, E., Vanholme, R., Boerjan, W., & Molina, A. (2014). The role of the secondary cell wall in plant resistance to pathogens. *Frontiers in plant science*, 5, 358.

[151] Wang, N., Takada, M., Sakamoto, S., Vanholme, R., Goeminne, G., Kim, H., ... & Kajita, S. (2025). Effects of feruloyl-CoA 6′-hydroxylase 1 overexpression on lignin and cell wall characteristics in transgenic hybrid aspen. *Frontiers in Plant Science*, 16, 1543168.

[152] Winkel, B. S. (2004). Metabolic channeling in plants. *Annu. Rev. Plant Biol*., 55(1), 85-107.

[153] Laursen, T., Borch, J., Knudsen, C., Bavishi, K., Torta, F., Martens, H. J., ... & Bassard, J. E. (2016). Characterization of a dynamic metabolon producing the defense compound dhurrin in sorghum. *Science*, 354(6314), 890-893.

[154] Obata, T. (2019). Metabolons in plant primary and secondary metabolism. *Phytochemistry Reviews*, 18(6), 1483-1507.

[155] Zhang, Y., & Fernie, A. R. (2021). Metabolons, enzyme–enzyme assemblies that mediate substrate channeling, and their roles in plant metabolism. *Plant Communications*, 2(1).

[156] d Moore, B. (2004). Bifunctional and moonlighting enzymes: lighting the way to regulatory control. *Trends in plant science*, 9(5), 221-228.

[157] Iacopino, S., & Licausi, F. (2020). The contribution of plant dioxygenases to hypoxia signaling. *Frontiers in Plant Science*, 11, 1008.

[158] Islam, M. S., Leissing, T. M., Chowdhury, R., Hopkinson, R. J., & Schofield, C. J. (2018). 2-Oxoglutarate-dependent oxygenases. *Annual review of biochemistry*, 87(1), 585-620.

[159] Frost, J., Frost, M., Batie, M., Jiang, H., & Rocha, S. (2021). Roles of HIF and 2-oxoglutarate-dependent dioxygenases in controlling gene expression in hypoxia. *Cancers*, 13(2), 350.

[160] Mahmood, A. M., & Dunwell, J. M. (2020). 2-oxoglutarate-dependent dioxygenases: A renaissance in attention for ascorbic acid in plants. *PLoS One*, 15(12), e0242833.

[161] Lange, T., & Pimenta Lange, M. J. (2020). The multifunctional dioxygenases of gibberellin synthesis. *Plant and Cell Physiology*, 61(11), 1869-1879.

[162] Verma, V., Ravindran, P., & Kumar, P. P. (2016). Plant hormone-mediated regulation of stress responses. *BMC plant biology*, 16, 1-10.

[163] Berens, M. L., Berry, H. M., Mine, A., Argueso, C. T., & Tsuda, K. (2017). Evolution of hormone signaling networks in plant defense. *Annual review of phytopathology*, 55(1), 401-425.

[164] Baillo, E. H., Kimotho, R. N., Zhang, Z., & Xu, P. (2019). Transcription factors associated with abiotic and biotic stress tolerance and their potential for crops improvement. *Genes*, 10(10), 771.

[165] Shankar, N., & Nath, U. (2024). Advantage looping: Gene regulatory circuits between microRNAs and their target transcription factors in plants. *Plant Physiology*, 196(4), 2304-2319.

[166] Wellburn, A. R. (1994). The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *Journal of plant physiology*, 144(3), 307-313.

[167] Roychoudhury, A., Basu, S., & Sengupta, D. N. (2012). Antioxidants and stress-related metabolites in the seedlings of two indica rice varieties exposed to cadmium chloride toxicity. *Acta Physiologiae Plantarum*, 34, 835-847.

[168] Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., & Scheible, W. R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant physiology*, 139(1), 5-17.

[169] Unterholzner, S. J., Rozhon, W., Papacek, M., Ciomas, J., Lange, T., Kugler, K. G., ... & Poppenberger, B. (2015). Brassinosteroids are master regulators of gibberellin biosynthesis in Arabidopsis. *The Plant Cell*, 27(8), 2261-2272.

[170] Wang, L., Ma, C., Wang, S., Yang, F., Sun, Y., Tang, J., ... & Wu, J. (2024). Ethylene and jasmonate signaling converge on gibberellin catabolism during thigmomorphogenesis in Arabidopsis. *Plant Physiology*, 194(2), 758-773.

SUPPLEMENTARY FIGURES

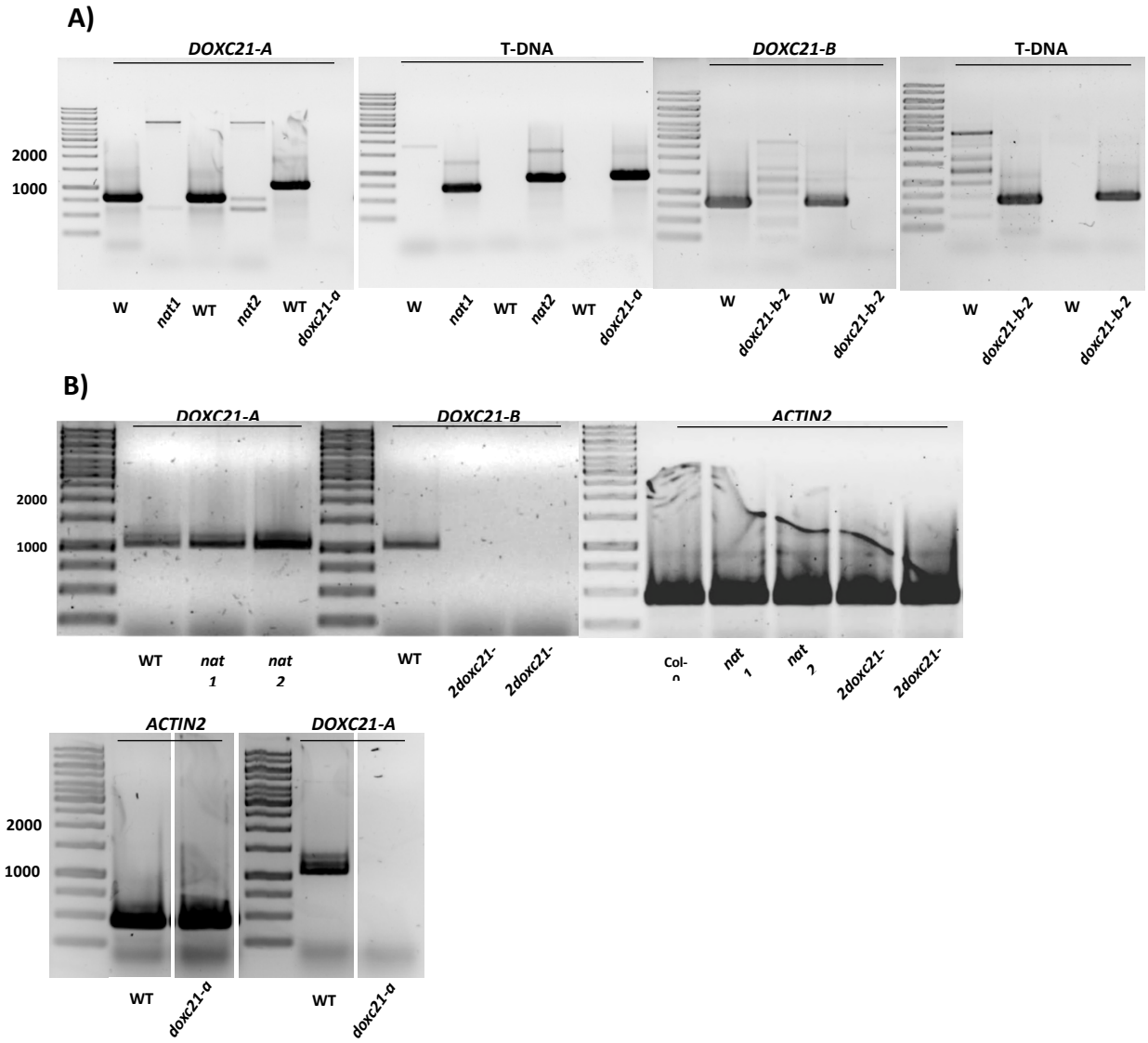
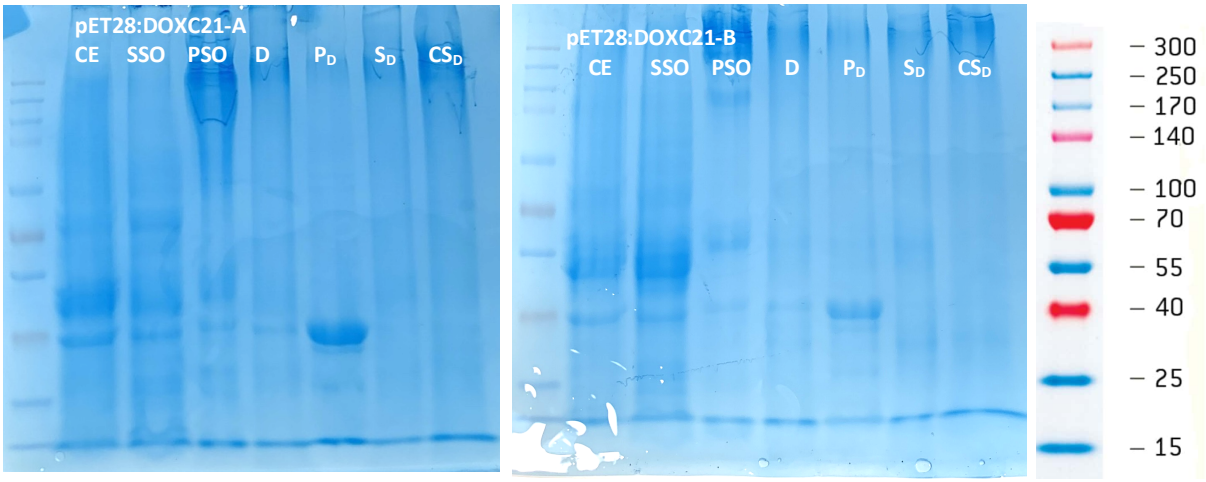
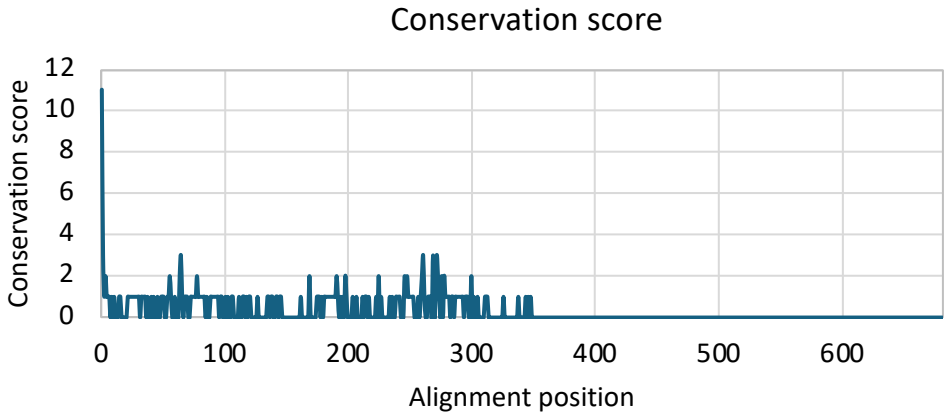


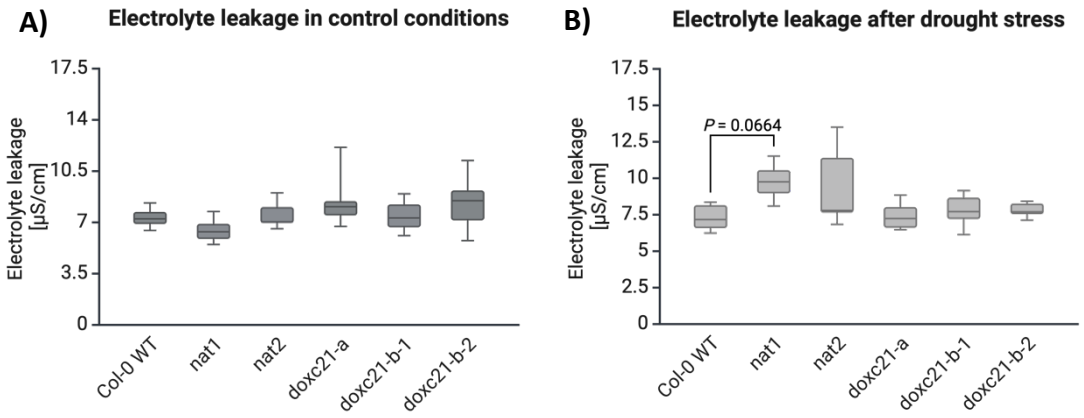
Figure S1. Validation of T-DNA insertional mutants for *DOXC21-A*, *DOXC21-B*, and *NAT* by genomic PCR and RT-PCR. (**A)** Genomic PCR analysis of Col-0 WT and mutant lines using primers flanking the predicted T-DNA insertion sites. T-DNA-specific amplicons were detected in *nat1*, *nat2, doxc21-a*, and *doxc21-b* mutants, confirming the insertional disruption of *DOXC21-A, DOXC21-B* and *NAT* genomic loci. **(B)** Reverse transcription PCR (RT-PCR) on cDNA from rosette leaves assessed *DOXC21-A* and *DOXC21-B* expression. *DOXC21-A* transcripts were detected in Col-0 WT and *nat* mutants, whereas *DOXC21-A* and *DOXC21-B* transcripts were absent in *doxc21-a* and *doxc21-b* lines, respectively, confirming effective gene knock-out. *ACTIN2* was used as a constitutive reference control.



Supplementary Figure S2. SDS-PAGE analysis of DOXC21-A and DOXC21-B recombinant protein purification. Recombinant proteins were expressed in *Escherichia coli* Rosetta 2 cells carrying pET28:DOXC21-A or pET28:DOXC21-B constructs. Samples were collected from different stages of the purification process and analyzed on 12% SDS-PAGE gels stained with Coomassie Brilliant Blue. The expected molecular weight of the His-tagged DOXC21 proteins is approximately ~40–45 kDa region. Lanes represent the following fractions: CE – crude extract of sonication; SSO – soluble supernatant after sonication; PSO – pellet after sonication; D – dialysate; P_D – pellet of dialysate; S_D – supernatant of dialysate; CS_D – concentrated supernatant of dialysate. A pre-stained protein ladder (PageRuler™ Plus, Thermo Scientific™, USA) was used for molecular weight reference (right panel). The prominent bands near 40 kDa in renatured fractions indicate partial recovery of soluble DOXC21-A and DOXC21-B proteins.



Supplementary Figure S3. Conservation score profile across aligned DOXC21-related protein sequences. This line plot illustrates conservation scores across alignment positions derived from 100 homologous DOXC21 protein sequences spanning major land plant lineages. The analysis was performed using Jalview with the Shenkin entropy-based conservation metric. Elevated conservation scores, particularly in the N-terminal region (positions ~1–300), indicate evolutionary constraint and suggest the presence of functionally important residues. The decline in conservation across the C-terminal half reflects greater sequence divergence, potentially corresponding to species-specific functional adaptations.

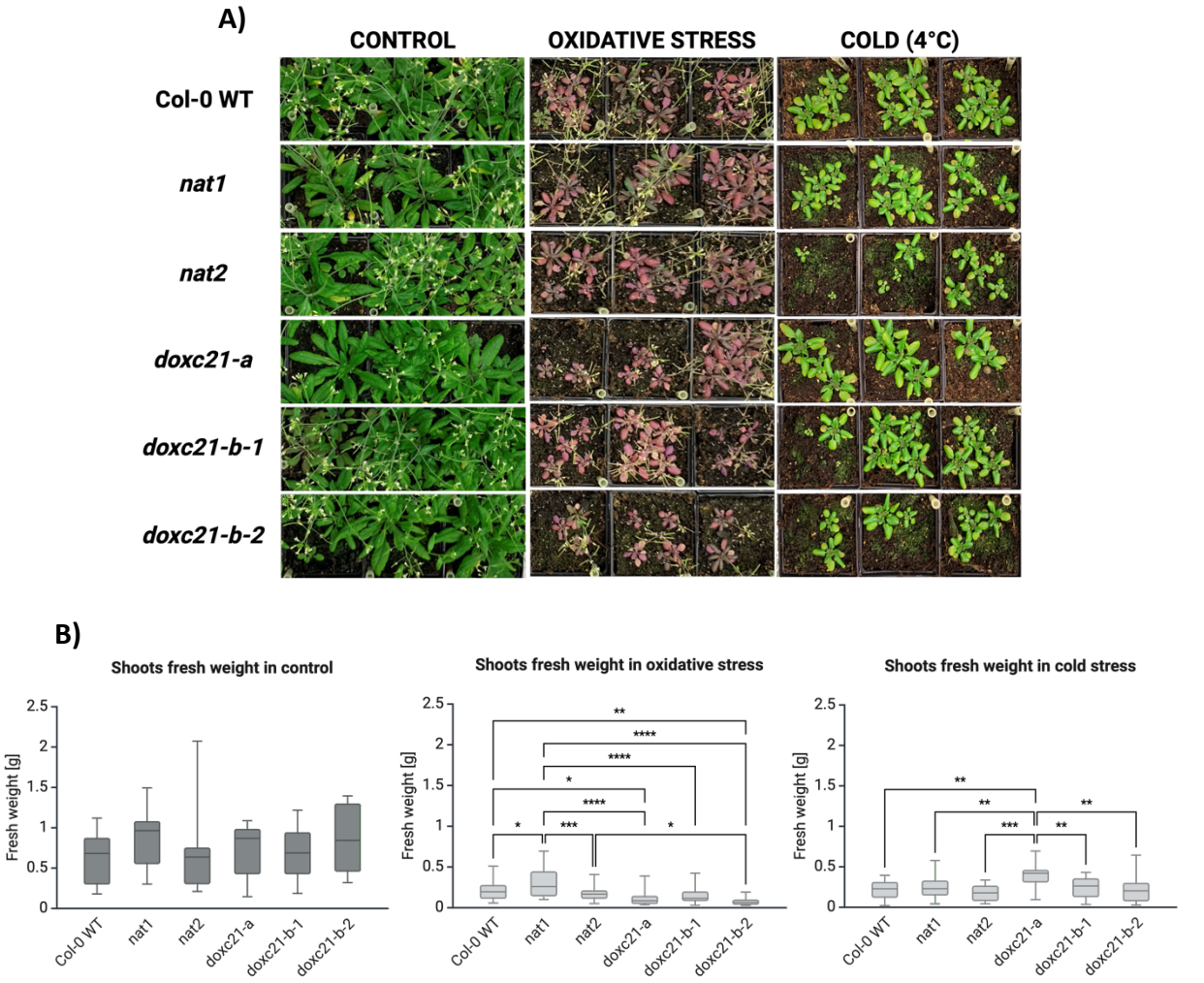


Supplementary Figure S4. Electrolyte leakage in Col-0 WT and mutant lines under control and drought stress conditions. Electrolyte leakage (EL) was measured from rosette leaves of soil-grown Arabidopsis plants following 21 days of drought stress. Leaves were incubated in deionized water, and conductivity was measured before and after sonication (30 minutes at 80 °C) to determine total electrolyte release. Conductivity values (µS/cm) reflect plasma membrane integrity and cellular damage. **(A)** Under control conditions, all genotypes exhibited similarly low baseline electrolyte leakage. **(B)** Following drought stress, *nat1* mutant displayed increased (P=0.0664) electrolyte leakage, suggesting higher membrane damage. In contrast, *doxc21-a* and *doxc21-b* mutants maintained similar electrolyte leakage to Col-0 WT, suggesting preserved membrane stability under drought. Boxplots represent biological replicates per genotype (n = 6). Statistical significance was tested with one-way ANOVA with Tukey multiple comparisons test.

Oxidative stress in soil cultures was induced by growing plants at higher light intensity of 300 µmol m^{-2} s^{-1} using LED lightning panels, starting from the third week of growth. Plants were grown under these high-light conditions for an additional three weeks. At the end of the stress treatment, shoot FW was measured. Leaf material was harvested, frozen in liquid nitrogen, and stored at -80°C until further biochemical analyses (chlorophyll, carotenoid, and anthocyanin quantification) (**Supplementary Fig. S5**).

Cold stress was applied from the third week of development by transferring plants to a dedicated chamber maintained at 6°C during the day and 4°C during the night. After 4 weeks of growth under cold conditions, shoots were harvested for FW measurements and biochemical analyses. Leaf samples were immediately frozen in liquid nitrogen and stored in -80°C for subsequent pigment quantification (**Supplementary Fig. S5**).

Under oxidative stress, *nat1* retained significantly higher shoot biomass compared do Col-0 WT and other lines, while *doxc21-a* showed enhanced growth under cold conditions (**Supplementary Fig. S5**). No significant differences were observed under control conditions. Chlorophyll and carotenoid content under non-stress conditions were altered in *nat2* and *doxc21-a*, with all genotypes showing reduced pigment levels under stress (**Supplementary Figs. S6–S7**). Anthocyanin accumulation increased across all lines in response to oxidative and cold stress, with *doxc21-a* showing a variable pattern and *nat1/nat2* exhibiting moderate elevation (**Supplementary Fig. S8**).



Supplementary Figure S5. Phenotypic characterization of Col-0 WT and mutant lines under control, oxidative, and cold stress conditions in soil-grown Arabidopsis plants.

(A) Representative phenotypes of Col-0 WT, *nat1, nat2, doxc21-a, doxc21-b-1*, and *doxc21-b-2* grown in soil under control conditions (left), oxidative stress (middle), and cold stress (right). Images were taken after 3 weeks of growth under the indicated stress conditions. Mutant lines exhibited variable degrees of stress sensitivity. **(B)** Shoot fresh weight (FW) of Col-0 WT and mutant genotypes grown under control (left), oxidative (middle), and cold stress (right) conditions. No significant differences in FW were detected under control conditions. In oxidative stress *nat1* exhibited significantly higher biomass compared to other lines, while in cold - *doxc21-a* showed significantly higher biomass than other genotypes. Box plots represent data from one experimental replicate (min. six plants per genotype per treatment).

Statistical significance was determined using one-way ANOVA with Tukey's multiple comparisons test. Asterisks denote statistically significant differences between groups (*P<0.05, **P<0.01, ***P<0.001, ****P < 0.0001).

Biochemical analyses

Chlorophyll and carotenoid contents were quantified according to Wellburn [166], using leaf tissue collected from plants grown under oxidative and cold stress conditions. Frozen rosette leaves were homogenized using liquid nitrogen in pre-chilled mortars, and approximately 100±2 mg of powdered tissue was weighed per biological replicate. Each sample was extracted with 1 mL of freshly prepared, ice-cold 80% (v/v) acetone saturated with Na_2CO_3 followed by a 10 min incubation in darkness at 4 °C. Homogenates were centrifuged at 10,000 rpm for 10 min at 4°C, and the supernatants were transferred to fresh tubes. The remaining pellets were re-extracted with an additional 1 mL of 80% (v/v) acetone, shaken, and centrifuged again under the same conditions. Both supernatants were pooled and diluted 1:4 prior to spectrophotometric measurements. Absorbance was measured at 663.2 nm, 646.8 nm, and 470 nm using a UV-Vis spectrophotometer (Specord 200 Plus, Analytik Jena, Germany). Chlorophylls concentrations were calculated using the following equations:

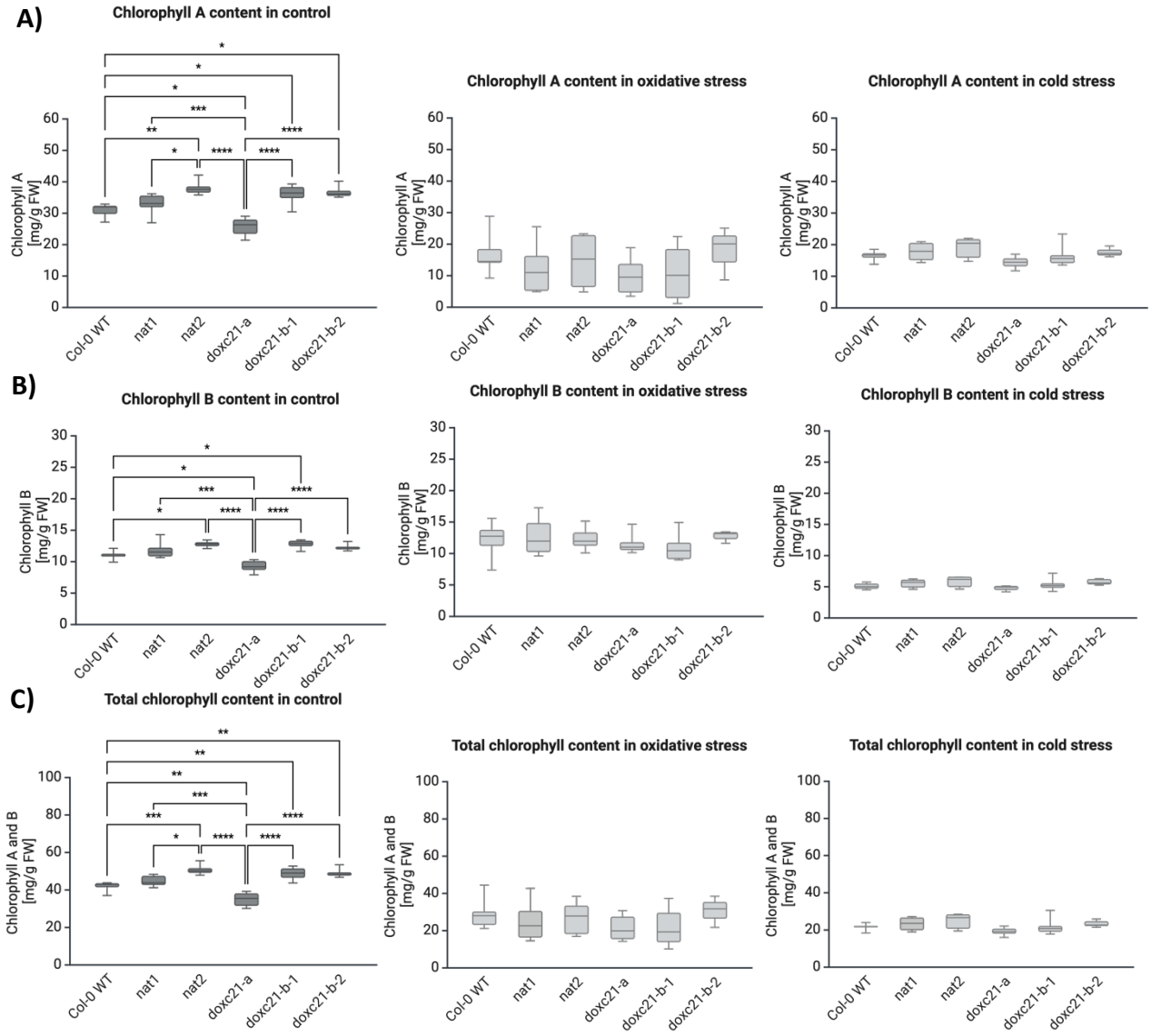
- Chlorophyll A (Chl A) = $12.25 \times A_{663.2} - 2.79 \times A_{646.8}$
- Chlorophyll B (Chl B) = $21.50 \times A_{646.8} - 5.10 \times A_{663.2}$
- Total chlorophyll = Chl A + Chl B

Carotenoid content was calculated using the formula:

- Carotenoid = $(1000 \times A_{470} - 1.82 \times \text{Chl A} - 85.02 \times \text{Chl B}) / 198$

Anthocyanin content was determined following the method of Roychoudhury *et al.* [167]. Leaf samples frozen at -80°C were homogenized in mortars using liquid nitrogen and 100±2 mg of tissue was extracted with 1 mL of methanol containing 1% (v/v) HCl. Samples were incubated overnight at 4°C in darkness, then centrifuged at 10.000 rpm in 4°C for 10 min. The absorbance of the resulting supernatant was measured at 525 nm using quartz cuvettes. Anthocyanin content was calculated using millimolar extinction coefficient of 31.6.

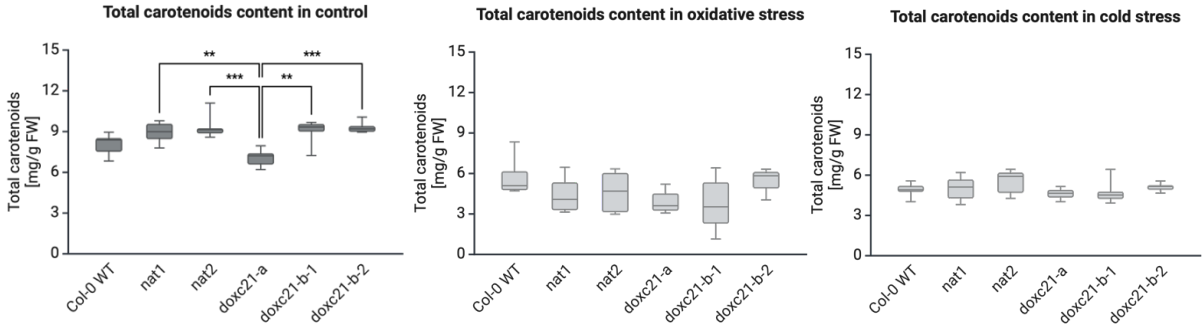
All measurements were conducted using a minimum of six biological replicates, except for the *doxc21-b-2* line, for which n=3 (oxidative stress) or n=5 (cold stress). Statistical analyses were performed using one-way ANOVA followed by Tukey's multiple comparisons test.



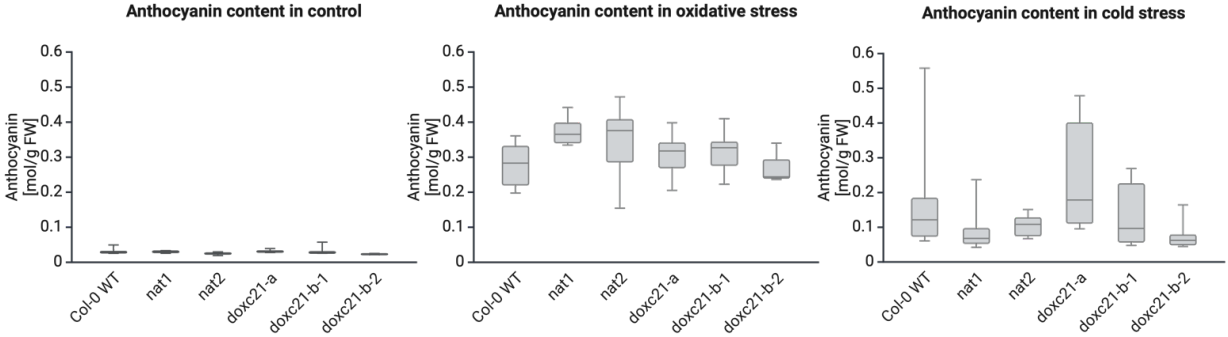
Supplementary Figure S6. Quantification of chlorophyll content in Col-0 WT and *nat1, nat2, doxc21-a, doxc21-b-1,* and *doxc21-b-2* mutant lines under control and stress conditions.

(A) Chlorophyll A (Chl A) content measured in rosette leaves under control, oxidative, and cold stress conditions. Under control conditions, mutants exhibited significant differences in Chl A levels compared to Col-0 WT, particularly in the nat2 and doxc21-a lines. Stress conditions, however, led to an overall reduction in Chl A across all genotypes, with no statistically significant variation observed between lines. **(B)** Chlorophyll B (Chl B) content showed a similar pattern, with significant genotype-dependent differences detected in the control group. Under oxidative and cold stress, Chl B levels were reduced across all genotypes, with no major outliers. **(C)** Total chlorophyll content (Chl A + Chl B) showed the combined effect of Chl A and Chl B measurements. As with individual chlorophyll types, significant reductions were detected in *doxc21-a* mutant line under control conditions. Stress conditions suppressed total chlorophyll content in all genotypes to a comparable level.

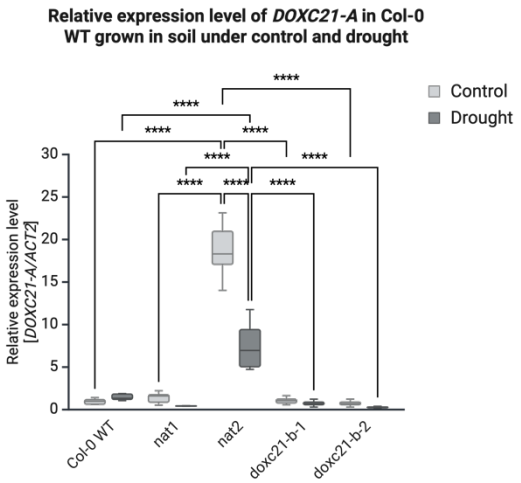
Each boxplot represents the distribution of chlorophyll content from at least six biological replicates. Statistical significance was assessed using one-way ANOVA followed by Tukey's post hoc test. Asterisks denote levels of significance (**P*<0.05, ***P*<0.01, ****P*<0.001; *****P*<0.0001).



Supplementary Figure S7. Total carotenoid content in Col-0 WT and *nat1, nat2, doxc21-a, doxc21-b-1*, and *doxc21-b-2* mutant lines under control, oxidative, and cold stress conditions. Total carotenoids were extracted from rosette leaves and quantified spectrophotometrically in Arabidopsis plants subjected to control, oxidative, and cold stress treatments. Under control conditions (left), statistically significant lower carotenoid contents were observed between *doxc21-a* and other mutant lines. Exposure to oxidative stress (middle) reduced total carotenoid levels across all genotypes, however no statistically significant differences were detected between lines. Under cold stress (right), carotenoid levels were generally uniform across genotypes, with only minor variations and no statistically significant changes. Carotenoid concentrations are expressed as mg per g of fresh weight (FW). Each boxplot is based on at least six biological replicates per genotype and condition (except *doxc21-b-2*, where in oxidative stress n=5, and in control n=3). Asterisks denote statistically significant differences determined by one-way ANOVA with Tukey's post-hoc test (**P*<0.05, ***P*<0.01, ****P*<0.001).



Supplementary Figure S8. Anthocyanin accumulation in Col-0 WT *and nat1, nat2, doxc21-a, doxc21-b-1*, and *doxc21-b-2* mutant lines under control and stress conditions. Anthocyanin content was measured in rosette leaves of Col-0 WT and mutant genotypes exposed to control, oxidative, and cold stress conditions. Under control conditions (left), anthocyanin levels remained uniformly low across all genotypes, with minimal variation observed. Oxidative stress (middle) led to a substantial increase in anthocyanin accumulation across all tested lines, with *nat1* and *nat2* displaying higher levels compared to Col-0 WT, however not significantly. Under cold stress (right), anthocyanin content increased moderately in all genotypes. The *doxc21-a* line showed the most variable response, suggesting possible genotype-specific regulation under low temperature. Measurements were performed using at least six biological replicates per genotype and condition (except *doxc21-b-2*, where in oxidative stress n=5, and in control n=3). Boxplots show median, interquartile range, and minimum/maximum values. For the statistical significance one-way ANOVA followed by Tukey's post hoc test was applied.



Supplementary Figure S9. Relative expression of *DOXC21-A* in Col-0 WT and mutant genotypes under control and drought conditions in soil-grown *A. thaliana*. Expression of *DOXC21-A* was quantified by RT-qPCR in rosette leaves of Col-0 WT, *nat1*, *nat2*, *doxc21-b-1*, and *doxc21-b-2* plants following 3 weeks of drought stress or well-watered control conditions. Expression levels were normalized to the reference gene *ACT2* and are shown as $2^{-\Delta\Delta CT}$ values. Only *nat2* exhibited a strong induction of *DOXC21-A* under both control and drought conditions, significantly exceeding expression levels in Col-0 WT and all other lines (****P<0.0001). No significant induction was observed in other genotypes. Data represent five biological replicates per condition. Statistical significance was determined using one-way ANOVA with Tukey's post hoc test.

Table S1. List of primers used in this study. This table provides sequences and details of all primers used for cloning, qRT-PCR, genotyping, and sequencing. For each primer pair, the corresponding gene target, application, and expected amplicon size are included. Primer specificity was validated *in silico* using NCBI Primer-BLAST and experimentally confirmed by melt-curve analysis. All sequences are presented in the 5′ to 3′ orientation.

Target gene	Locus ID	Primer name	Sequence (5′ – 3′)	Application	Reference
DOXC21-A	AT3G19000	2	AAATCGTCTTGTGGTCCTCCT	genotyping (*nat1, nat2*)	In-house design
DOXC21-A	AT3G19000	3	CTAATAGCTCGAAGGGCTGCT	genotyping (*nat1, nat2*)	In-house design
DOXC21-A	AT3G19000	341	TTCGGTAATCCACCTACACTATCT	genotyping (*doxc21-a*)	In-house design
DOXC21-A	AT3G19000	320	GCTGAACTACGATAAAGACCAAAGG	genotyping (*doxc21-a*)	In-house design
DOXC21-B	AT3G19010	163	ATGGAGGATCTTGATCCAAC	genotyping (*doxc21-b-1*)	In-house design
DOXC21-B	AT3G19000	183	ATATACAGTGGTGGGTGCAGC	genotyping (*doxc21-b-1*)	In-house design
DOXC21-B	AT3G19010	164	TTAAGTTTTAACTTTGAAGTCATCG	genotyping (*doxc21-b-2*)	In-house design
DOXC21-B	AT3G19000	185	CTGCACCCACCACTGTATATG	genotyping (*doxc21-b-2*)	In-house design
T-DNA	LB-1.3	1	ATTTTGCCGATTTCGGAAC	genotyping (*nat1, nat2, doxc21-b*)	SALK mutant lines
T-DNA	WiscDs_LB	384	TCCTCGAGTTTCTCCATAATAATGT	genotyping (*doxc21-a*)	WiscDsLox
DOXC21-A	AT3G19000	159	ATGGGAGAACTCGACGAAGCT	RT-PCR, pBIN cloning	In-house design
DOXC21-A	AT3G19000	160	CTAAGCCTTGAAGTGATCAATCTG	RT-PCR, pBIN cloning	In-house design
DOXC21-B	AT3G19010	163	ATGGAGGATCTTGATCCAAC	RT-PCR, pBIN cloning	In-house design
DOXC21-B	AT3G19010	164	TTAAGTTTTAACTTTGAAGTCATCG	RT-PCR, pBIN cloning	In-house design
DOXC21-A	AT3G19000	161	GAATTCGGAGAACTCGACGAAGCTT	pET cloning (*EcoR*I)	In-house design
DOXC21-A	AT3G19000	162	GAATTCAGCCTTGAAGTGATCAATCTGG	pET cloning (*EcoR*I)	In-house design
DOXC21-B	AT3G19010	165	GGATCCGAGGATCTTGATCCAACCTAC	pET cloning (*BamH*I)	In-house design
DOXC21-B	AT3G19010	166	CTCGAGAGTTTTAACTTTGAAGTCATCG	pET cloning (*XhoI*)	In-house design
ACT2	AT3G18780	50	TCCCAGTGTTGTTGGTAGGC	RT-PCR	[168]
ACT2	AT3G18780	51	CAAGACGGAGGATGGCATGA	RT-PCR	[168]
DOXC21-A	AT3G19000	201	TCCGAAAGCTCACGAACCA	qPCR (*in vitro*, soil)	In-house design
DOXC21-A	AT3G19000	202	TGGCATACCTCTCTGAAATGAGAA	qPCR (*in vitro,* soil)	In-house design
DOXC21-B	AT3G19010	203	TACATCCAAGCACCAGAGCATAG	qPCR (*in vitro,* soil, hydroponics)	In-house design
DOXC21-A	AT3G19000	204	GGGTATTTCTTCTGGTTGGTTAGG	qPCR (*in vitro*, soil, hydroponics)	In-house design
F6′H1	AT3G13610	206	GATGAGGACAGAGTCGCTGAA	qPCR (hydroponics)	In-house design
F6′H1	AT3G13610	207	CACTTGAAAGAACCCCCATTTC	qPCR (hydroponics)	In-house design
CCoAOMT1	AT4G34050	81	AGAAAGCCGGCGTTGCTCACA	qPCR (hydroponics)	In-house design
CCoAOMT1	AT4G34050	82	GCATCAGGAGGAGCCACGACA	qPCR (hydroponics)	In-house design
PAL1	AT2G37040	281	CCTACATCGCCGGACTTCTC	qPCR (hydroponics)	In-house design
PAL1	AT2G37040	282	CCGTTGGGACCAGTAGCTTT	qPCR (hydroponics)	In-house design
DOXC21-A	AT3G19000	273	TGACGTCCAACACCTAATGC	qPCR (hydroponics)	In-house design
DOXC21-A	AT3G19000	274	ATCGGCTAACCGGTTTCTTC	qPCR (hydroponics)	In-house design
F6′H2	AT1G55290	208	TCGGACGTCACTCTGATGTTTC	qPCR (hydroponics)	In-house design
F6′H2	AT1G55290	209	GAGACCACCGATCTCGTCTTG	qPCR (hydroponics)	In-house design
C3H	AT2G40890	79	CGTCGTGACCGCCTCACTCG	qPCR (hydroponics)	In-house design
C3H	AT2G40890	80	GCCATCGCCCATTCCGCTGT	qPCR (hydroponics)	In-house design
HCT	AT5G48930	59	AGCTTATTCCCGAAGTTGATCACT	qPCR (hydroponics)	In-house design
HCT	AT5G48930	59	CAAAACGAGAAGCGGGAAAG	qPCR (hydroponics)	In-house design
CCR1	AT1G15950	96	GCTCTTAAGGCGGCGATTG	qPCR (hydroponics)	In-house design
CCR1	AT1G15950	97	CAGGAGAAGCCGTGTGAAAG	qPCR (hydroponics)	In-house design
GA20OX1	AT4G25420	425	TGATGGCTCATGGATCTCTGTCC	qPCR (hydroponics)	[169]
GA20OX1	AT4G25420	426	CCTCCCATTTGTCATCACCTGGAG	qPCR (hydroponics)	[169]
GA20OX8	AT4G21200	427	TATCCGGCAGCTTTCTTGGT	qPCR (hydroponics)	[170]
GA20OX8	AT4G21200	428	TGCAAGAACCTCTGCCAACA	qPCR (hydroponics)	[170]
NAT	AT3G19002	435	TTCAGTGTTCACGTGTGGGA	qPCR (hydroponics)	In-house design
NAT	AT3G19002	436	ATCACAATCGTCGGATCTCCA	qPCR (hydroponics)	In-house design
F3′H	AT3G51240	433	GAGGAGCGTGACCACAAAGA	qPCR (hydroponics)	In-house design
F3′H	AT3G51240	434	AAAACACACCGAGCCTAGCA	qPCR (hydroponics)	In-house design
JID1	AT1G06620	429	CGCCACTTGTGGGGAGATG	qPCR (hydroponics)	In-house design
JID1	AT1G06620	430	GCAACAGCAAGGAGTTAGTGC	qPCR (hydroponics)	In-house design
ACT2	AT3G18780	83	CTTGCACCAAGCAGCATGAA	qPCR (in vitro, soil, hydroponics)	[168]
ACT2	AT3G18780	84	CCGATC CAGACACTGTACTTCCTT	qPCR (in vitro, soil, hydroponics)	[168]
EF-1	AT5G60390	291	TGAGCACGCTCTTCTTGCTTTCA	qPCR (hydroponics)	[168]
EF-1	AT5G60390	292	GGTGGTGGCATCCATCTTGTTACA	qPCR (hydroponics)	[168]

Supplementary Table S2. Composition of the Heeg nutrient solution used for hydroponic cultivation. The table presents the final concentrations of macronutrients and micronutrients in both standard 1× and modified 10× Heeg media used in this study. The 10× formulation was applied to enhance nutrient availability and stress responsiveness during hydroponic experiments. All concentrations are provided in millimolar (mM) for macronutrients and micromolar (µM) for micronutrients. For full experimental details and stress treatments, refer to the Materials and Methods section (Section 2.4).

	10x Heeg	1x Heeg
Macronutrients	**Final concentration (mM)**	**Final concentration (mM)**
KNO_3	2	2
$Ca(NO_3)_2$ x $4H_20$	2	2
$MgSO_4$	0.5	0.5
NH_4NO_3	0.4	0.4
KH_2PO_4	0.25	0.25
Micronutrients	**Final concentration (µM)**	**Final concentration (µM)**
FeEDTA	40	40
H3BO3	25	2.5
MnCl2•4H20	2	0.2
ZnSO4•7H20	2	0.2
CuSO4•5H20	0.5	0.05
KCl	4.5	0.45
NH4)6Mo7O24	0.08	0.008
CoCl2	0.26	0.026

Supplementary Table S3. Summary of sequencing data quality and mapping metrics for the analyzed samples. Sample refers to the unique identifier for each analyzed dataset. Total Reads represents the total number of raw reads generated per sample. Mapped Reads indicates the total number of reads that successfully aligned to the reference genome, with % Mapped Reads showing this as a percentage of total reads. Uniq Mapped Reads refers to the number of reads that mapped to a single genomic location, and % Uniq Mapped Reads represents the proportion of uniquely mapped reads relative to total reads. Multiple Map Reads captures the number of reads mapping to more than one genomic location, with % Multiple Map Reads indicating this as a percentage of total reads. Reads Map to '+' and Reads Map to '-' represent the proportion of reads mapped to the positive and negative genomic strands, respectively.

Sample name	Total Reads	Mapped Reads	% Mapped Reads	Uniq Mapped Reads	% Uniq Mapped Reads	Multiple Map Reads	% Multiple Map Reads	Reads Map to '+'	% Reads Map to '+'	Reads Map to '-'	% Reads Map to '-'	Clean reads	Clean bases	GC content	%≥Q20	%≥Q30
Col-0 WT 1	111,229,008	109,372,088	98.33%	106,867,797	96.08%	2,504,291	2.25%	56,258,980	50.58%	56,297,831	50.61%	55,614,504	16,625,066,747	46.49%	97.02%	94.69%
Col-0 WT 2	107,442,108	105,868,090	98.54%	101,331,344	94.31%	4,536,746	4.22%	55,526,645	51.68%	55,568,316	51.72%	53,721,054	16,045,930,356	46.62%	97.52%	95.59%
Col-0 WT 3	109,416,166	107,772,802	98.50%	105,222,857	96.17%	2,549,945	2.33%	55,546,554	50.77%	55,585,116	50.80%	54,708,083	16,329,673,388	46.48%	97.22%	95.05%
Col-0 WT 4	118,958,138	116,984,168	98.34%	113,840,212	95.70%	3,143,956	2.64%	60,483,784	50.84%	60,527,971	50.88%	59,479,069	17,780,431,360	46.41%	97.20%	95.01%
Col-0 WT 5	96,256,012	94,542,230	98.22%	92,268,245	95.86%	2,273,985	2.36%	48,699,902	50.59%	48,749,345	50.65%	48,128,006	14,394,504,940	46.72%	96.87%	94.49%
Col-0 WT 6	100,072,926	98,462,318	98.39%	95,993,214	95.92%	2,469,104	2.47%	50,784,983	50.75%	50,821,332	50.78%	50,036,463	14,962,503,424	46.73%	97.01%	94.72%
2ogd-h5-1 1	110,247,140	108,478,210	98.40%	102,546,820	93.02%	5,931,390	5.38%	57,551,185	52.20%	57,615,702	52.26%	55,123,570	16,465,549,834	46.86%	97.15%	94.96%
2ogd-h5-1 2	96,126,114	94,306,260	98.11%	91,763,970	95.46%	2,542,290	2.64%	48,764,961	50.73%	48,812,778	50.78%	48,063,057	14,366,402,408	46.40%	96.53%	93.95%
2ogd-h5-1 3	108,472,226	106,446,371	98.13%	102,811,375	94.78%	3,634,996	3.35%	55,366,726	51.04%	55,429,259	51.10%	54,236,113	16,212,919,508	46.66%	96.56%	94.02%
2ogd-h5-1 4	100,856,148	99,123,190	98.28%	96,472,979	95.65%	2,650,211	2.63%	51,268,717	50.83%	51,299,161	50.86%	50,428,074	15,071,045,783	46.48%	97.57%	95.65%
2ogd-h5-1 5	112,376,624	110,560,093	98.38%	107,475,319	95.64%	3,084,774	2.75%	57,205,225	50.90%	57,249,391	50.94%	56,188,312	16,800,685,346	46.77%	97.16%	94.97%
2ogd-h5-1 6	113,234,158	111,555,334	98.52%	108,612,264	95.92%	2,943,070	2.60%	57,622,606	50.89%	57,672,220	50.93%	56,617,079	16,913,029,745	47.20%	96.97%	94.69%
2ogd-h5-2 1	102,528,332	100,589,474	98.11%	97,889,852	95.48%	2,699,622	2.63%	51,988,820	50.71%	52,038,097	50.75%	51,264,166	15,320,695,595	46.50%	96.69%	94.22%
2ogd-h5-2 2	107,771,192	105,871,385	98.24%	102,732,393	95.32%	3,138,992	2.91%	54,919,084	50.96%	54,953,930	50.99%	53,885,596	16,090,151,548	46.45%	97.36%	95.28%
2ogd-h5-2 3	113,485,814	111,488,152	98.24%	108,135,898	95.29%	3,352,254	2.95%	57,817,598	50.95%	57,866,141	50.99%	56,742,907	16,965,982,732	46.27%	96.85%	94.64%
2ogd-h5-2 4	98,157,264	96,628,633	98.44%	93,392,366	95.15%	3,236,267	3.30%	50,254,608	51.20%	50,293,989	51.24%	49,078,632	14,652,177,551	46.99%	97.67%	95.78%
2ogd-h5-2 5	113,264,430	111,640,134	98.57%	108,568,637	95.85%	3,071,497	2.71%	57,731,921	50.97%	57,771,989	51.01%	56,632,215	16,924,125,464	46.92%	97.68%	95.83%
2ogd-h5-2 6	118,650,070	116,498,256	98.19%	113,177,503	95.39%	3,320,753	2.80%	60,295,335	50.82%	60,341,869	50.86%	59,325,035	17,696,174,982	46.62%	97.03%	94.76%
2ogd-h5-3 1	97,142,362	95,261,821	98.06%	92,634,129	95.36%	2,627,692	2.70%	49,260,169	50.71%	49,313,490	50.76%	48,571,181	14,508,282,610	46.93%	96.90%	94.51%
2ogd-h5-3 2	120,381,484	118,383,714	98.34%	115,133,124	95.64%	3,250,590	2.70%	61,207,110	50.84%	61,254,071	50.88%	60,190,742	17,996,042,623	46.77%	97.26%	95.13%
2ogd-h5-3 3	106,884,608	105,067,068	98.30%	102,060,909	95.49%	3,006,159	2.81%	54,417,674	50.91%	54,466,741	50.96%	53,442,304	15,946,676,367	46.70%	96.92%	94.60%
2ogd-h5-3 4	106,140,368	104,586,627	98.54%	101,766,108	95.88%	2,820,519	2.66%	54,079,582	50.95%	54,115,669	50.99%	53,070,184	15,857,995,453	46.74%	97.51%	95.56%
2ogd-h5-3 5	104,075,474	104,586,627	98.26%	98,438,273	94.58%	3,827,410	3.68%	53,404,788	51.31%	53,438,405	50.99%	52,037,737	15,572,197,834	46.73%	97.56%	95.59%
Sum/Avr	2,473,168,166	2,434,073,045	98.32%	2,359,135,588	95.39%	72,616,513	2.93%	1,260,456,957	50.96%	1,261,482,813	50.99%	1,236,584,083	369,498,245,598	46.67%	97.14%	94.94%

Supplementary Table S4. Strand-specific RNA-Seq library quality assessment. Table summarizes strand-specific mapping statistics for read pairs across all sequenced libraries. read1_vs_gene: The proportion of read 1 sequences mapped to the same strand as their corresponding annotated gene in the reference genome. read2_vs_gene: The proportion of read 2 sequences mapped to the opposite strand of the corresponding gene, consistent with strand-specific library preparation protocols. High proportions (close to 1.0) indicate strong strand specificity and overall high-quality library construction. Values are reported for individual biological replicates of Col-0 WT, *nat1, nat2* and *doxc21-a* mutant lines. Cases of ambiguous strand assignment (e.g., for genes with bidirectional transcription) are not shown in this table but are classified as "undefined" in the pipeline.

Sample name	read1_vs_gene	read2_vs_gene
Col-0 WT 1	0.0088	0.9912
Col-0 WT 2	0.0076	0.9924
Col-0 WT 3	0.0104	0.9896
Col-0 WT 4	0.0092	0.9908
Col-0 WT 5	0.0106	0.9894
Col-0 WT 6	0.0071	0.9929
***nat1* 1**	0.0096	0.9904
***nat1* 2**	0.0071	0.9929
***nat1* 3**	0.0120	0.9880
***nat1* 4**	0.0143	0.9858
***nat1* 5**	0.0102	0.9898
***nat1* 6**	0.0082	0.9918
***nat2* 1**	0.0082	0.9918
***nat2* 2**	0.0137	0.9863
***nat2* 3**	0.0095	0.9905
***nat2* 4**	0.0092	0.9908
***nat2* 5**	0.0130	0.9870
***nat2* 6**	0.0114	0.9886
***doxc21-a* 1**	0.0088	0.9912
***doxc21-a* 2**	0.0089	0.9911
***doxc21-a* 3**	0.0103	0.9897
doxc21-a 4	0.0114	0.9886
doxc21-a 5	0.0102	0.9898

Supplementary Table S5. Summary of functional annotation results for 772 novel genes identified by RNA-Seq-based transcript assembly. Of these, 445 genes were successfully annotated based on homology searches against major biological databases. The table lists the number and percentage of annotated genes for each database, including COG, GO, KEGG, KOG, Pfam, Swiss-Prot, TrEMBL, eggNOG, and NR. The annotations provide insights into putative functions, evolutionary relationships, and biochemical pathways, contributing to the expanded annotation of the Arabidopsis genome.

Database	Annotated number	Annotated percent [%]
COG	19	2.5
GO	238	30.8
KEGG	169	21.9
KOG	170	22
Pfam	197	25.5
Swissprot	144	18.7
TrEMBL	437	56.6
eggNOG	214	27.7
nr	440	57
All	445	57.6

Supplementary Tables S6. Overview of differentially expressed genes (DEGs) identified by RNA-seq analysis under osmotic stress. Following tables summarize the top DEGs identified across six pairwise comparisons: (A) Col-0 WT *vs*. *doxc21-a*, (B) Col-0 WT *vs*. *nat1*, (C) Col-0 WT v*s. nat2*, (D) *nat1 vs. doxc21-a*, (E) *nat2 vs*. *doxc21-a*, and (F) *nat1 vs. nat2*. Each table presents genes with an absolute $\log_2$ fold change ($\log_2$FC) of ≥ 2.5 and an adjusted p-value < 0.05. For each DEG, the following information is provided where available: gene identifier, fold change value, direction of regulation (up/down), p-value, and functional annotations. Gene Ontology (GO) classifications are included for molecular function, biological process, and cellular component. Additional annotations encompass KEGG orthology and pathway terms, KOG functional class predictions, and protein function descriptions based on Swiss-Prot and NCBI NR databases. Functional annotation was performed using DIAMOND, InterProScan, and HMMER against major biological databases including GO, KEGG, Pfam, COG, eggNOG, Swiss-Prot, and TrEMBL. Together, these tables provide insight into stress-responsive transcriptional changes, highlighting genes and pathways potentially regulated by *DOXC21-A* and *NAT* expression divergence in response to osmotic stress.

A) DEG between Col-0 WT *vs. doxc21-a*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
1	AT1G53480	1.26E-83	-6.691	down	--	--	--	--	--	mto 1 responding down 1 [Arabidopsis thaliana]
2	AT3G44300	2.07E-06	-6.307	down	MF: Hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds (GO:0016810) BP: Nitrogen compound metabolic process (GO:0006807)	K01501|3.0e-198|ath:AT3G44300|K01501 nitrilase [EC:3.5.5.1] | (RefSeq) NIT2; nitrilase 2	Tryptophan metabolism (ko00380); Cyanoamino acid metabolism (ko00460); Nitrogen metabolism (ko00910)	Amino acid transport and metabolism	Nitrilase 2 OS=Arabidopsis thaliana OX=3702 GN=NIT2 PE=1 SV=1	nitrilase 2 [Arabidopsis thaliana]
3	AT2G30770	8.42E-03	-4.240	down	MF: Iron ion binding (GO:0005506); Oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen (GO:0016705); Heme binding (GO:0020037)	K11868|1.0e-282|ath:AT2G30770|K11868 indoleacetaldoxime dehydratase [EC:4.99.1.6] | (RefSeq) CYP71A13; cytochrome P450 family 71 polypeptide	Tryptophan metabolism (ko00380)	Secondary metabolites biosynthesis, transport and catabolism	Indole acetaldoxime dehydratase OS=Arabidopsis thaliana OX=3702 GN=CYP71A13 PE=1 SV=1	cytochrome P450 family 71 polypeptide [Arabidopsis thaliana]
4	AT1G15520	6.61E-03	-4.189	down	MF: ATP binding (GO:0005524); ATPase activity (GO:0016887) CC: Integral component of membrane (GO:0016021)	K08712|1.8e-129|gsl:Gasu_29800|K08712 ATP-binding cassette, subfamily G (WHITE), member 2, SNQ2 | (RefSeq) ABC transporter, ATP-binding protein	ABC transporters (ko02010)	Secondary metabolites biosynthesis, transport and catabolism	ABC transporter G family member 40 OS=Arabidopsis thaliana OX=3702 GN=ABCG40 PE=1 SV=1	pleiotropic drug resistance 12 [Arabidopsis thaliana]
5	AT5G13080	3.11E-03	-4.025	down	MF: DNA-binding transcription factor activity (GO:0003700); Sequence-specific DNA binding (GO:0043565) CC: Nucleus (GO:0005634)	K12835|2.3e-25|cmax:111493042|K12835 ATP-dependent RNA helicase DDX42 [EC:3.6.4.13] | (RefSeq) DEAD-box ATP-dependent RNA helicase 24 isoform X1	Spliceosome (ko03040)	--	Probable WRKY transcription factor 75 OS=Arabidopsis thaliana OX=3702 GN=WRKY75 PE=1 SV=1	WRKY DNA-binding protein 75 [Arabidopsis thaliana]
6	AT3G22600	2.48E-03	-3.693	down	MF: Lipid binding (GO:0008289) BP: Lipid transport (GO:0006869) CC: Cytosol (GO:0005829); Anchored component of membrane (GO:0031225)	--	--	--	Xylogen-like protein 11 OS=Arabidopsis thaliana OX=3702 GN=XYP11 PE=2 SV=2	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein [Arabidopsis thaliana]
7	AT5G07010	6.92E-04	-3.561	down	MF: sulfotransferase activity (GO:0008146)	K22312|2.8e-210|ath:AT5G07010|K22312 hydroxyjasmonate sulfotransferase [EC:2.8.2.39] | (RefSeq) ST2A; sulfotransferase 2A	--	General function prediction only	Cytosolic sulfotransferase 15 OS=Arabidopsis thaliana OX=3702 GN=SOT15 PE=1 SV=1	sulfotransferase 2A [Arabidopsis thaliana]
8	NewGene_1147	7.69E-04	-3.525	down	--	K13065|7.64e-15|oeu:111404589|K13065 shikimate O-hydroxycinnamoyltransferase [EC:2.3.1.133] | (RefSeq) spermidine hydroxycinnamoyl transferase-like	Phenylpropanoid biosynthesis (ko00940); Flavonoid biosynthesis (ko00941);; Stilbenoid, diarylheptanoid and gingerol biosynthesis (ko00945)	--	--	At2g04460 [Arabidopsis thaliana]

A) DEG between Col-0 WT vs. *doxc21-a*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
9	AT3G11340	4.77E-03	-3.502	down	MF: UDP-glycosyltransferase activity (GO:0008194)	K13493|8.0e-130|brp:103846928|K13493 cytokinin-N-glucosyltransferase [EC:2.4.1.-] | (RefSeq) UDP-glycosyltransferase 76C2	Zeatin biosynthesis (ko00908)	Carbohydrate transport and metabolism;; Energy production and conversion	UDP-glycosyltransferase 76B1 OS=Arabidopsis thaliana OX=3702 GN=UGT76B1 PE=2 SV=1	UDP-Glycosyltransferase superfamily protein [Arabidopsis thaliana]
10	AT2G45760	5.62E-04	-3.277	down	MF: Phospholipid binding (GO:0005543); Lipid binding (GO:0008289) BP: Response to acid chemical (GO:0001101); Response to salicylic acid (GO:0009751); Regulation of defense response (GO:0031347) CC: Membrane (GO:0016020)	K03110|7.4e-14|cmax:111479767|K03110 fused signal recognition particle receptor | (RefSeq) cell division protein FtsY homolog, chloroplastic-like isoform X1	Protein export (ko03060)	--	BON1-associated protein 2 OS=Arabidopsis thaliana OX=3702 GN=BAP2 PE=1 SV=1	BON association protein 2 [Arabidopsis thaliana]
11	AT3G20810	1.20E-06	1.603	up	--	K10277|9.7e-189|ath:AT3G20810|K10277 lysine-specific demethylase 8 [EC:1.14.11.27] | (RefSeq) JMJD5; 2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	--	Chromatin structure and dynamics;; Signal transduction mechanisms	Lysine-specific demethylase JMJ30 OS=Arabidopsis thaliana OX=3702 GN=JMJ30 PE=1 SV=1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein [Arabidopsis thaliana]
12	AT1G08940	5.24E-06	1.642	up	MF: catalytic activity (GO:0003824)	--	--	--	Phosphoglycerate mutase-like protein AT74H OS=Arabidopsis thaliana OX=3702 GN=At1g08940 PE=3 SV=2	Phosphoglycerate mutase family protein [Arabidopsis thaliana]
13	AT1G65450	2.66E-02	1.844	up	MF: transferase activity, transferring acyl groups other than amino-acyl groups (GO:0016747)	K15400|5.6e-161|aly:9324472|K15400 omega-hydroxypalmitate O-feruloyl transferase [EC:2.3.1.188] | (RefSeq) omega-hydroxypalmitate O-feruloyl transferase	Cutin, suberine and wax biosynthesis (ko00073)	--	Hydroxycinnamoyltransferase OS=Narcissus pseudonarcissus OX=39639 GN=HCT PE=2 SV=1	HXXXD-type acyl-transferase family protein [Arabidopsis thaliana]
14	AT4G28755	1.46E-04	1.964	up	--	--	--	--	--	hypothetical protein AT4G28755, partial [Arabidopsis thaliana]
15	AT1G12570	1.22E-02	2.768	up	MF: oxidoreductase activity, acting on CH-OH group of donors (GO:0016614), flavin adenine dinucleotide binding (GO:0050660)	K00108|9.3e-208|vra:106768885|K00108 choline dehydrogenase [EC:1.1.99.1] | (RefSeq) protein HOTHEAD	Glycine, serine and threonine metabolism (ko00260)	General function prediction only	Protein HOTHEAD OS=Arabidopsis thaliana OX=3702 GN=HTH PE=1 SV=1	Glucose-methanol-choline (GMC) oxidoreductase family protein [Arabidopsis thaliana]
16	NewGene_114	5.72E-09	2.949	up	MF: nucleic acid binding (GO:0003676), RNA-DNA hybrid ribonuclease activity (GO:0004523)	--	--	General function prediction only	--	hypothetical protein AXX17_AT4G21170 [Arabidopsis thaliana]

A) DEG between Col-0 WT *vs. doxc21-a*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
17	AT1G60960	2.60E-36	4.957	up	MF: Metal ion transmembrane transporter activity (GO:0046873) CC: Integral component of membrane (GO:0016021)	K14709|5.9e-145|ath:AT1G60960|K14709 solute carrier family 39 (zinc transporter), member 1/2/3 | (RefSeq) IRT3; iron regulated transporter 3	--	Inorganic ion transport and metabolism	Fe(2+) transport protein 3, chloroplastic OS=Arabidopsis thaliana OX=3702 GN=IRT3 PE=2 SV=3	iron regulated transporter 3 [Arabidopsis thaliana]
18	AT1G75945	6.04E-39	5.520	up	--	--	--	--	--	hypothetical protein AT1G75945 [Arabidopsis thaliana]
19	NewGene_710	1.76E-84	7.934	up	BP: DNA integration (GO:0015074)	K12447|5.11e-52|thj:104808371|K12447 UDP-sugar pyrophosphorylase [EC:2.7.7.64] | (RefSeq) uncharacterized protein LOC104808371	Pentose and glucuronate interconversions (ko00040);; Galactose metabolism (ko00052);; Ascorbate and aldarate metabolism (ko00053);; Amino sugar and nucleotide sugar metabolism (ko00520)	General function prediction only	Retrovirus-related Pol polyprotein from transposon RE2 OS=Arabidopsis thaliana OX=3702 GN=RE2 PE=4 SV=1	uncharacterized protein LOC111829291 [Capsella rubella]
20	AT1G64795	2.38E-77	10.309	up	--	--	--	--	--	hypothetical protein AT1G64795 [Arabidopsis thaliana]

B) DEG between Col-0 WT *vs. nat1*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
1	AT5G44420	1.69E-10	-5.071	down	CC: extracellular region (GO:0005576) BP: defense response (GO:0006952)	K20727|2.3e-40|ath:AT5G44420|K20727 defensin-like protein 16 | (RefSeq) PDF1.2; plant defensin 1.2	MAPK signaling pathway - plant (ko04016)	--	Defensin-like protein 16 OS=Arabidopsis thaliana OX=3702 GN=PDF1.2A PE=1 SV=1	plant defensin 1.2 [Arabidopsis thaliana]
2	AT4G25000	4.57E-03	-2.715	down	MF: alpha-amylase activity (GO:0004556), calcium ion binding (GO:0005509), alpha-amylase activity (releasing maltohexaose) (GO:0103025) BP: carbohydrate metabolic process (GO:0005975)	K01176|2.2e-254|ath:AT4G25000|K01176 alpha-amylase [EC:3.2.1.1] | (RefSeq) AMY1; alpha-amylase-like protein	Starch and sucrose metabolism (ko00500)	Carbohydrate transport and metabolism	Alpha-amylase 1 OS=Arabidopsis thaliana OX=3702 GN=AMY1 PE=1 SV=1	alpha-amylase-like protein [Arabidopsis thaliana]
3	AT3G62290	2.69E-15	-2.229	down	MF: GTP binding (GO:0005525)	K07937|2.3e-99|aly:9318178|K07937 ADP-ribosylation factor 1/2 | (RefSeq) ADP-ribosylation factor 1	Endocytosis (ko04144)	Intracellular trafficking, secretion, and vesicular transport	ADP-ribosylation factor 1 OS=Arabidopsis thaliana OX=3702 GN=ARF1 PE=1 SV=2	PREDICTED: ADP-ribosylation factor 1-like isoform X1 [Brassica oleracea var. oleracea]
4	AT3G51860	1.06E-03	-1.807	down	MF: calcium:proton antiporter activity (GO:0015369) CC: vacuolar membrane (GO:0005774), integral component of membrane (GO:0016021)	K07300|3.0e-249|ath:AT3G51860|K07300 Ca2+:H+ antiporter | (RefSeq) CAX3; cation exchanger 3	--	Inorganic ion transport and metabolism	Vacuolar cation/proton exchanger 3 OS=Arabidopsis thaliana OX=3702 GN=CAX3 PE=1 SV=1	cation exchanger 3 [Arabidopsis thaliana]
5	NewGene 902	1.01E-03	-1.729	down	MF: protein kinase activity (GO:0004672), ATP binding (GO:0005524) CC: integral component of membrane (GO:0016021)	K19891|8.51e-11|gmx:100776826|K19891 glucan endo-1,3-beta-glucosidase 1/2/3 [EC:3.2.1.39] | (RefSeq) glucan endo-1,3-beta-glucosidase 1	Starch and sucrose metabolism (ko00500)	Signal transduction mechanisms	Probable LRR receptor-like serine/threonine-protein kinase At1g29720 OS=Arabidopsis thaliana OX=3702 GN=RFK1 PE=2 SV=3	hypothetical protein [Arabidopsis thaliana]
6	AT1G75580	6.87E-03	-1.664	down	BP: response to auxin (GO:0009733)	K14488|2.0e-55|ath:AT1G75580|K14488 SAUR family protein | (RefSeq) SAUR-like auxin-responsive protein family	Plant hormone signal transduction (ko04075)	--	Auxin-responsive protein SAUR50 OS=Arabidopsis thaliana OX=3702 GN=SAUR50 PE=1 SV=1	SAUR-like auxin-responsive protein family [Arabidopsis thaliana]
7	AT3G16640	8.73E-03	2.459	down	--	K00128|1.0e-45|ccav:112512809|K00128 aldehyde dehydrogenase (NAD+) [EC:1.2.1.3] | (RefSeq) aldehyde dehydrogenase 5, mitochondrial-like	Glycolysis / Gluconeogenesis (ko00010);; Ascorbate and aldarate metabolism (ko00053);; Fatty acid degradation (ko00071);; Valine, leucine and isoleucine degradation (ko00280);; Lysine degradation (ko00310);; Arginine and proline metabolism (ko00330);; Histidine metabolism (ko00340);; Tryptophan metabolism (ko00380);; beta-Alanine metabolism (ko00410);; Glycerolipid metabolism (ko00561);;	Cell cycle control, cell division, chromosome partitioning; Cytoskeleton	Translationally controlled tumor protein 1 OS=Arabidopsis thaliana OX=3702 GN=TCTP1 PE=1 SV=2	translationally controlled tumor protein [Arabidopsis thaliana]

							Pyruvate metabolism (ko00620);; Pantothenate and CoA biosynthesis (ko00770);; Limonene and pinene degradation (ko00903)			
B) DEG between Col-0 WT *vs. nat1*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
8	AT2G32487	1.67E-02	2.479	down	--	--	--	General function prediction only	--	hypothetical protein AT2G32487 [Arabidopsis thaliana]
9	AT5G64000	2.51E-03	2.521	down	MF: 3'(2'),5'-bisphosphate nucleotidase activity (GO:0008441) BP: sulfur compound metabolic process (GO:0006790), inositol phosphate dephosphorylation (GO:0046855)	K15422|2.1e-194|ath:AT5G64000|K15422 3'(2'), 5'-bisphosphate nucleotidase / inositol polyphosphate 1-phosphatase [EC:3.1.3.7 3.1.3.57] | (RefSeq) SAL2; Inositol monophosphatase family protein	Inositol phosphate metabolism (ko00562);; Sulfur metabolism (ko00920);; Phosphatidylinositol signaling system (ko04070)	Nucleotide transport and metabolism; Inorganic ion transport and metabolism	SAL2 phosphatase OS=Arabidopsis thaliana OX=3702 GN=SAL2 PE=2 SV=1	Inositol monophosphatase family protein [Arabidopsis thaliana]
10	AT1G63980	3.44E-03	2.548	down	MF: nucleic acid binding (GO:0003676)	K11135|4.0e-194|ath:AT1G63980|K11135 Pin2-interacting protein X1 | (RefSeq) D111/G-patch domain-containing protein	--	RNA processing and modification; Cell cycle control, cell division, chromosome partitioning	G-patch domain-containing protein 1 OS=Arabidopsis thaliana OX=3702 GN=GDP1 PE=2 SV=1	D111/G-patch domain-containing protein [Arabidopsis thaliana]
11	AT1G11785	1.07E-02	2.616	up	CC: integral component of membrane (GO:0016021)	--	--	--	--	transmembrane protein [Arabidopsis thaliana]
12	AT1G51670	4.42E-43	2.703	up	--	--	--	--	Protein HEAT-INDUCED TAS1 TARGET 5 OS=Arabidopsis thaliana OX=3702 GN=HTT5 PE=2 SV=1	hypothetical protein AT1G51670 [Arabidopsis thaliana]
13	AT2G20670	2.67E-13	2.721	up	--	--	--	--	--	sugar phosphate exchanger, putative (DUF506) [Arabidopsis thaliana]
14	AT4G28755	7.30E-03	2.766	up	--	--	--	--	--	hypothetical protein AT4G28755, partial [Arabidopsis thaliana]

B) DEG between Col-0 WT vs. *nat1*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
15	AT3G48360	8.84E-03	3.614	up	MF: transcription coregulator activity (GO:0003712), histone acetyltransferase activity (GO:0004402), zinc ion binding (GO:0008270) BP: regulation of transcription, DNA-templated (GO:0006355) CC: host cell nucleus (GO:0042025)	K10523|2.2e-07|ppp:112282598|K10523 speckle-type POZ protein | (RefSeq) BTB/POZ domain-containing protein At3g56230-like	--	Transcription	BTB/POZ and TAZ domain-containing protein 2 OS=Arabidopsis thaliana OX=3702 GN=BT2 PE=1 SV=1	BTB and TAZ domain protein 2 [Arabidopsis thaliana]
16	NewGene 280	4.84E-44	5.044	up	MF: nucleic acid binding (GO:0003676), zinc ion binding (GO:0008270) BP: DNA integration (GO:0015074)	K03006|3.16e-236|adu:107466525|K03006 DNA-directed RNA polymerase II subunit RPB1 [EC:2.7.7.6] | (RefSeq) LOW QUALITY PROTEIN: DNA-directed RNA polymerase II subunit 1-like	RNA polymerase (ko03020)	General function prediction only	Retrovirus-related Pol polyprotein from transposon TNT 1-94 OS=Nicotiana tabacum OX=4097 PE=2 SV=1	putative protein [Arabidopsis thaliana]
17	AT1G60960	1.09E-45	5.626	up	MF: metal ion transmembrane transporter activity (GO:0046873) CC: integral component of membrane (GO:0016021)	K14709|5.9e-145|ath:AT1G60960|K14709 solute carrier family 39 (zinc transporter), member 1/2/3 | (RefSeq) IRT3; iron regulated transporter 3	--	Inorganic ion transport and metabolism	Fe(2+) transport protein 3, chloroplastic OS=Arabidopsis thaliana OX=3702 GN=IRT3 PE=2 SV=3	iron regulated transporter 3 [Arabidopsis thaliana]
18	AT1G75945	2.75E-62	7.763	up	--	--	--	--	--	hypothetical protein AT1G75945 [Arabidopsis thaliana]
19	AT4G28800	1.24E-17	11.005	up	MF: DNA binding (GO:0003677), protein dimerization activity (GO:0046983)	K16189|1.1e-20|egr:104442415|K16189 phytochrome-interacting factor 4 | (RefSeq) transcription factor PIF5	Plant hormone signal transduction (ko04075)	--	Putative transcription factor bHLH056 OS=Arabidopsis thaliana OX=3702 GN=BHLH56 PE=3 SV=2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]
20	NewGene 710	2.61E-79	14.344	up	BP: DNA integration (GO:0015074)	K12447|5.11e-52|thj:104808371|K12447 UDP-sugar pyrophosphorylase [EC:2.7.7.64] | (RefSeq) uncharacterized protein LOC104808371	Pentose and glucuronate interconversions (ko00040);; Galactose metabolism (ko00052);; Ascorbate and aldarate metabolism (ko00053);; Amino sugar and nucleotide sugar metabolism (ko00520)	General function prediction only	Retrovirus-related Pol polyprotein from transposon RE2 OS=Arabidopsis thaliana OX=3702 GN=RE2 PE=4 SV=1	uncharacterized protein LOC111829291 [Capsella rubella]

C) DEG between Col-0 WT *vs. nat2*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
1	AT1G53480	1.7E-10	-5.071	down	--	--	--	--	--	mto 1 responding down 1 [Arabidopsis thaliana]
2	AT5G62165	4.6E-03	-2.715	down	MF: RNA polymerase II transcription regulatory region sequence-specific DNA binding (GO:0000977); DNA-binding transcription factor activity (GO:0003700); Protein dimerization activity (GO:0046983) BP: Positive regulation of transcription by RNA polymerase II (GO:0045944) CC: Nucleus (GO:0005634)	K09264|1.5e-35|gmx:100785089|K09264 MADS-box transcription factor, plant | (RefSeq) agamous-like MADS-box protein AGL11 isoform X1	--	Transcription	MADS-box protein AGL42 OS=Arabidopsis thaliana OX=3702 GN=AGL42 PE=1 SV=1	AGAMOUS-like 42 [Arabidopsis thaliana]
3	AT1G53490	2.7E-15	-2.229	down	MF: Nucleic acid metabolic process (GO:0090304) BP: Meiosis I (GO:0007127); Reciprocal meiotic recombination (GO:0007131); Chromosome segregation (GO:0007059) CC: Nucleus (GO:0005634); Chromosome (GO:0005694); Chiasma (GO:0005712)	K10639|1.7e-168|ath:AT1G53490|K10639 E3 ubiquitin-protein ligase CCNP1IP1 [EC:2.3.2.27] | (RefSeq) HEI10; RING/U-box superfamily protein	--	--	E3 ubiquitin-protein ligase CCNB1IP1 homolog OS=Arabidopsis thaliana OX=3702 GN=HEI10 PE=2 SV=1	RING/U-box superfamily protein [Arabidopsis thaliana]
4	AT3G23550	1.1E-03	-1.807	down	MF: Antiporter activity (GO:0015297); Xenobiotic transmembrane transporter activity (GO:0042910) CC: Integral component of membrane (GO:0016021)	K03327|3.3e-259|ath:AT3G23550|K03327 multidrug resistance protein, MATE family | (RefSeq) MATE efflux family protein	--	General function prediction only	Protein DETOXIFICATION 18 OS=Arabidopsis thaliana OX=3702 GN=DTX18 PE=2 SV=1	MATE efflux family protein [Arabidopsis thaliana]
5	AT5G50915	1.0E-03	-1.729	down	MF: DNA binding (GO:0003677); Protein dimerization activity (GO:0046983) BP: Regulation of transcription, DNA-templated (GO:0006355)	K18878|2.5e-38|cann:107866053|K18878 BHLH transcription factor Upa20 | (RefSeq) Upa20; transcription factor bHLH137-like	Plant-pathogen interaction (ko04626)	--	Transcription factor bHLH137 OS=Arabidopsis thaliana OX=3702 GN=BHLH137 PE=1 SV=1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]
6	AT5G17330	6.9E-03	-1.664	down	MF: Glutamate decarboxylase activity (GO:0004351); Pyridoxal phosphate binding (GO:0030170) BP: Glutamate metabolic process (GO:0006536)	K01580|1.1e-294|ath:AT5G17330|K01580 glutamate decarboxylase [EC:4.1.1.15] | (RefSeq) GAD; glutamate decarboxylase	Alanine, aspartate and glutamate metabolism (ko00250);; beta-Alanine metabolism (ko00410);; Taurine and hypotaurine metabolism (ko00430);; Butanoate metabolism (ko00650)	Amino acid transport and metabolism	Glutamate decarboxylase 1 OS=Arabidopsis thaliana OX=3702 GN=GAD1 PE=1 SV=2	glutamate decarboxylase [Arabidopsis thaliana]

C) DEG between Col-0 WT *vs. nat2*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
7	AT2G27080	8.7E-03	2.459	up	CC: integral component of membrane (GO:0016021)	K04077|1.6e-13|cpep:111806851|K04077 chaperonin GroEL | (RefSeq) chaperonin CPN60-like 2, mitochondrial	RNA degradation (ko03018)	--	NDR1/HIN1-like protein 13 OS=Arabidopsis thaliana OX=3702 GN=NHL13 PE=2 SV=1	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family [Arabidopsis thaliana]
8	AT5G17350	1.7E-02	2.479	up	--	K03327|3.6e-12|nta:107781954|K03327 multidrug resistance protein, MATE family | (RefSeq) protein DETOXIFICATION 14-like	--	--	--	hypothetical protein AT5G17350 [Arabidopsis thaliana]
9	AT1G74930	2.5E-03	2.521	up	MF: DNA binding (GO:0003677); DNA-binding transcription factor activity (GO:0003700) CC: Nucleus (GO:0005634)	K09286|8.2e-23|ccav:112509229|K09286 EREBP-like factor | (RefSeq) dehydration-responsive element-binding protein 3-like	--	--	Ethylene-responsive transcription factor ERF018 OS=Arabidopsis thaliana OX=3702 GN=ERF018 PE=2 SV=1	Integrase-type DNA-binding superfamily protein [Arabidopsis thaliana]
10	AT1G76650	3.4E-03	2.548	up	MF: calcium ion binding (GO:0005509)	K13448|1.3e-91|ath:AT1G76650|K13448 calcium-binding protein CML | (RefSeq) CML38; calmodulin-like 38	Plant-pathogen interaction (ko04626)	Signal transduction mechanisms	Calcium-binding protein CML38 OS=Arabidopsis thaliana OX=3702 GN=CML38 PE=1 SV=1	calmodulin-like 38 [Arabidopsis thaliana]
11	AT5G62520	1.1E-02	2.616	up	BP: Response to osmotic stress (GO:0006970); Response to salt stress (GO:0009651); Reactive oxygen species metabolic process (GO:0072593) CC: Cytoplasm (GO:0005737), Mitochondrion (GO:0005739)	--	--	--	Probable inactive poly [ADP-ribose] polymerase SRO5 OS=Arabidopsis thaliana OX=3702 GN=SRO5 PE=1 SV=1	similar to RCD one 5 [Arabidopsis thaliana]
12	AT3G19000	4.4E-43	2.703	up	MF: Oxidoreductase activity (GO:0016491); Metal ion binding (GO:0046872)	K24028|1.4e-39|cam:101505284|K24028 salicylic acid 3-hydroxylase [EC:1.14.11.-] | (RefSeq) 1-aminocyclopropane-1-carboxylate oxidase-like	--	Secondary metabolites biosynthesis, transport and catabolism;; General function prediction only	Protein DOWNY MILDEW RESISTANCE 6 OS=Arabidopsis thaliana OX=3702 GN=DMR6 PE=1 SV=1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein [Arabidopsis thaliana]
13	AT4G28790	2.7E-13	2.721	up	MF: DNA binding (GO:0003677)	--	--	--	Transcription factor bHLH23 OS=Arabidopsis thaliana OX=3702 GN=BHLH23 PE=2 SV=1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]

C) DEG between Col-0 WT *vs. nat2*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
14	AT4G27654	7.3E-03	2.766	up	CC: integral component of membrane (GO:0016021)	--	--	--	--	transmembrane protein [Arabidopsis thaliana]
15	AT4G34410	8.8E-03	3.614	up	MF: DNA binding (GO:0003677); DNA-binding transcription factor activity (GO:0003700) CC: Nucleus (GO:0005634)	K09286|1.6e-21|csv:101203357|K09286 EREBP-like factor | (RefSeq) ethylene-responsive transcription factor ERF073	--	--	Ethylene-responsive transcription factor ERF109 OS=Arabidopsis thaliana OX=3702 GN=ERF109 PE=1 SV=1	redox responsive transcription factor 1 [Arabidopsis thaliana]
16	AT1G60960	4.8E-44	5.044	up	MF: Metal ion transmembrane transporter activity (GO:0046873) CC: Integral component of membrane (GO:0016021)	K14709|5.9e-145|ath:AT1G60960|K14709 solute carrier family 39 (zinc transporter), member 1/2/3 | (RefSeq) IRT3; iron regulated transporter 3	--	Inorganic ion transport and metabolism	Fe(2+) transport protein 3, chloroplastic OS=Arabidopsis thaliana OX=3702 GN=IRT3 PE=2 SV=3	iron regulated transporter 3 [Arabidopsis thaliana]
17	AT1G75945	1.1E-45	5.626	up	--	--	--	--	--	hypothetical protein AT1G75945 [Arabidopsis thaliana]
18	NewGene_710	2.8E-62	7.763	up	BP: DNA integration (GO:0015074)	K12447|5.11e-52|thj:104808371|K12447 UDP-sugar pyrophosphorylase [EC:2.7.7.64] | (RefSeq) uncharacterized protein LOC104808371	Pentose and glucuronate interconversions (ko00040);; Galactose metabolism (ko00052);; Ascorbate and aldarate metabolism (ko00053);; Amino sugar and nucleotide sugar metabolism (ko00520)	General function prediction only	Retrovirus-related Pol polyprotein from transposon RE2 OS=Arabidopsis thaliana OX=3702 GN=RE2 PE=4 SV=1	uncharacterized protein LOC111829291 [Capsella rubella]
19	AT4G28800	1.2E-17	11.005	up	MF: DNA binding (GO:0003677); Protein dimerization activity (GO:0046983)	K16189|1.1e-20|egr:104442415|K16189 phytochrome-interacting factor 4 | (RefSeq) transcription factor PIF5	Plant hormone signal transduction (ko04075)	--	Putative transcription factor bHLH056 OS=Arabidopsis thaliana OX=3702 GN=BHLH56 PE=3 SV=2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]
20	NewGene_159	2.6E-79	14.344	up	--	K00558|9.43e-07|aof:109841697|K00558 DNA (cytosine-5)-methyltransferase 1 [EC:2.1.1.37] | (RefSeq) uncharacterized protein LOC109841697	Cysteine and methionine metabolism (ko00270)	General function prediction only	--	PREDICTED: uncharacterized protein LOC109131263 [Camelina sativa]

D) DEG between *nat1 vs. doxc21-a*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
1	AT2G16340	6.0E-16	-8.110	down	--	--	--	--	--	hypothetical protein AT2G16340 [Arabidopsis thaliana]
2	AT4G28800	7.0E-27	-8.000	down	MF: DNA binding (GO:0003677), protein dimerization activity (GO:0046983)	K16189|1.1e-20|egr:104442415|K16189 phytochrome-interacting factor 4 | (RefSeq) transcription factor PIF5	Plant hormone signal transduction (ko04075)	--	Putative transcription factor bHLH056 OS=Arabidopsis thaliana OX=3702 GN=BHLH56 PE=3 SV=2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]
3	AT1G53480	3.0E-97	-6.843	down	--	--	--	--	--	mto 1 responding down 1 [Arabidopsis thaliana]
4	AT3G44300	2.0E-05	-5.116	down	BP: nitrogen compound metabolic process (GO:0006807) MF: hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds (GO:0016810)	K01501|3.0e-198|ath:AT3G44300|K01501 nitrilase [EC:3.5.5.1] | (RefSeq) NIT2; nitrilase 2	Tryptophan metabolism (ko00380);; Cyanoamino acid metabolism (ko00460);; Nitrogen metabolism (ko00910)	Amino acid transport and metabolism	Nitrilase 2 OS=Arabidopsis thaliana OX=3702 GN=NIT2 PE=1 SV=1	nitrilase 2 [Arabidopsis thaliana]
5	AT2G30770	1.2E-02	-3.935	down	MF: iron ion binding (GO:0005506), oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen (GO:0016705), heme binding (GO:0020037)	K11868|1.0e-282|ath:AT2G30770|K11868 indoleacetaldoxime dehydratase [EC:4.99.1.6] | (RefSeq) CYP71A13; cytochrome P450 family 71 polypeptide	Tryptophan metabolism (ko00380)	Secondary metabolites biosynthesis, transport and catabolism	Indoleacetaldoxime dehydratase OS=Arabidopsis thaliana OX=3702 GN=CYP71A13 PE=1 SV=1	cytochrome P450 family 71 polypeptide [Arabidopsis thaliana]
6	AT1G15520	1.1E-02	-3.756	down	MF: ATP binding (GO:0005524), ATPase activity (GO:0016887) CC: integral component of membrane (GO:0016021)	K08712|1.8e-129|gsl:Gasu_29800|K08712 ATP-binding cassette, subfamily G (WHITE), member 2, SNQ2 | (RefSeq) ABC transporter, ATP-binding protein	ABC transporters (ko02010)	Secondary metabolites biosynthesis, transport and catabolism	ABC transporter G family member 40 OS=Arabidopsis thaliana OX=3702 GN=ABCG40 PE=1 SV=1	pleiotropic drug resistance 12 [Arabidopsis thaliana]
7	AT2G47770	2.0E-03	-3.603	down	CC: integral component of membrane (GO:0016021)	K05770|7.1e-91|bna:106436135|K05770 translocator protein | (RefSeq) translocator protein homolog	--	--	Translocator protein homolog OS=Arabidopsis thaliana OX=3702 GN=TSPO PE=1 SV=1	TSPO (outer membrane tryptophan-rich sensory protein)-like protein [Arabidopsis thaliana]
8	AT1G32350	2.0E-03	-3.527	down	MF: alternative oxidase activity (GO:0009916), ubiquinol:oxygen oxidoreductase activity (GO:0102721), metal ion binding (GO:0046872) CC: integral component of membrane (GO:0016021), respirasome (GO:0070469)	K17893|2.5e-175|ath:AT1G32350|K17893 ubiquinol oxidase [EC:1.10.3.11] | (RefSeq) AOX1D; alternative oxidase 1D	--	--	Ubiquinol oxidase 3, mitochondrial OS=Arabidopsis thaliana OX=3702 GN=AOX3 PE=1 SV=2	alternative oxidase 1D [Arabidopsis thaliana]
9	AT1G51670	1.0E-09	-3.515	down	--	--	--	--	Protein HEAT-INDUCED TAS1 TARGET 5 OS=Arabidopsis thaliana OX=3702 GN=HTT5 PE=2 SV=1	hypothetical protein AT1G51670 [Arabidopsis thaliana]
10	AT1G05340	6.0E-04	-3.456	down	CC: nucleus (GO:0005634), cytoplasm (GO:0005737), plasma membrane (GO:0005886)	--	--	--	--	cysteine-rich TM module stress tolerance protein [Arabidopsis thaliana]
11	AT2G22470	6.0E-03	-3.411	down	CC: cytoplasm (GO:0005737)	--	--	--	Classical arabinogalactan protein 2 OS=Arabidopsis thaliana OX=3702 GN=AGP2 PE=1 SV=1	arabinogalactan protein 2 [Arabidopsis thaliana]

D) DEG between *nat1 vs. doxc21-a*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
12	AT2G26400	1.5E-02	-3.406	down	MF: iron ion binding (GO:0005506), acireductone dioxygenase [iron(II)-requiring] activity (GO:0010309) BP: L-methionine salvage from methylthioadenosine (GO:0019509) CC: nucleus (GO:0005634), cytoplasm (GO:0005737)	K08967|3.2e-106|ath:AT2G26400|K08967 1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase [EC:1.13.11.53 1.13.11.54] | (RefSeq) ARD3; acireductone dioxygenase 3	Cysteine and methionine metabolism (ko00270)	ARD/ARD' family	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase 1 OS=Arabidopsis thaliana OX=3702 GN=ARD1 PE=2 SV=1	acireductone dioxygenase 3 [Arabidopsis thaliana]
13	AT4G23680	1.0E-10	-3.373	down	BP: defense response (GO:0006952)	K01742|1.7e-15|psom:113322169|K01742 thebaine synthase [EC:4.2.99.24] | (RefSeq) major latex protein 149-like	Isoquinoline alkaloid biosynthesis (ko00950)	Pathogenesis-related protein Bet v 1 family	MLP-like protein 328 OS=Arabidopsis thaliana OX=3702 GN=MLP328 PE=2 SV=1	Polyketide cyclase/dehydrase and lipid transport superfamily protein [Arabidopsis thaliana]
14	AT5G13170	3.7E-02	-3.352	down	BP: carbohydrate transport (GO:0008643) CC: plasma membrane (GO:0005886), integral component of membrane (GO:0016021)	K15382|1.0e-154|ath:AT5G13170|K15382 solute carrier family 50 (sugar transporter) | (RefSeq) SAG29; senescence-associated gene 29	--	Sugar efflux transporter for intercellular exchange	Bidirectional sugar transporter SWEET15 OS=Arabidopsis thaliana OX=3702 GN=SWEET15 PE=1 SV=1	senescence-associated gene 29 [Arabidopsis thaliana]
15	AT1G63980	4.0E-38	1.670	up	MF: nucleic acid binding (GO:0003676)	K11135|4.0e-194|ath:AT1G63980|K11135 Pin2-interacting protein X1 | (RefSeq) D111/G-patch domain-containing protein	--	RNA processing and modification; Cell cycle control, cell division, chromosome partitioning	G-patch domain-containing protein 1 OS=Arabidopsis thaliana OX=3702 GN=GDP1 PE=2 SV=1	D111/G-patch domain-containing protein [Arabidopsis thaliana]
16	AT3G16640	8.0E-21	1.778	up	--	K00128|1.0e-45|ccav:112512809|K00128 aldehyde dehydrogenase (NAD+) [EC:1.2.1.3] | (RefSeq) aldehyde dehydrogenase 5, mitochondrial-like	Glycolysis / Gluconeogenesis (ko00010);; Ascorbate and aldarate metabolism (ko00053);; Fatty acid degradation (ko00071);; Valine, leucine and isoleucine degradation (ko00280);; Lysine degradation (ko00310);; Arginine and proline metabolism (ko00330);; Histidine metabolism (ko00340);; Tryptophan metabolism (ko00380);; beta-Alanine metabolism (ko00410);; Glycerolipid metabolism (ko00561);; Pyruvate metabolism (ko00620);; Pantothenate and CoA biosynthesis (ko00770);; Limonene and pinene degradation (ko00903)	Cell cycle control, cell division, chromosome partitioning; Cytoskeleton	Translationally controlled tumor protein 1 OS=Arabidopsis thaliana OX=3702 GN=TCTP1 PE=1 SV=2	translationally controlled tumor protein [Arabidopsis thaliana]
17	AT5G43460	2.0E-33	1.783	up	CC: endoplasmic reticulum (GO:0005783), integral component of membrane (GO:0016021)	--	--	--	--	HR-like lesion-inducing protein-like protein [Arabidopsis thaliana]
18	AT3G62290	2.0E-39	1.993	up	MF: GTP binding (GO:0005525)	K07937|2.3e-99|aly:9318178|K07937 ADP-ribosylation factor 1/2 | (RefSeq) ADP-ribosylation factor 1	Endocytosis (ko04144)	Intracellular trafficking, secretion, and vesicular transport	ADP-ribosylation factor 1 OS=Arabidopsis thaliana OX=3702 GN=ARF1 PE=1 SV=2	PREDICTED: ADP-ribosylation factor 1-like isoform X1 [Brassica oleracea var. oleracea]
19	AT5G62165	5.0E-05	2.405	up	MF: RNA polymerase II transcription regulatory region sequence-specific DNA binding (GO:0000977), DNA-binding transcription factor activity (GO:0003700), protein dimerization activity (GO:0046983) BP: positive regulation of transcription by RNA polymerase II (GO:0045944) CC: nucleus (GO:0005634)	K09264|1.5e-35|gmx:100785089|K09264 MADS-box transcription factor, plant | (RefSeq) agamous-like MADS-box protein AGL11 isoform X1	--	Transcription	MADS-box protein AGL42 OS=Arabidopsis thaliana OX=3702 GN=AGL42 PE=1 SV=1	AGAMOUS-like 42 [Arabidopsis thaliana]

D) DEG between *nat1 vs. doxc21-a*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
20	AT1G64795	5.0E-77	10.235	up	--	--	--	--	--	hypothetical protein AT1G64795 [Arabidopsis thaliana]

E) DEG between *nat2 vs. doxc21-a*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
1	NewGene 159	2.7E-155	-12.365	down	--	K00558|9.43e-07|aof:109841697|K00558 DNA (cytosine-5)-methyltransferase 1 [EC:2.1.1.37] | (RefSeq) uncharacterized protein LOC109841697	Cysteine and methionine metabolism (ko00270)	General function prediction only	--	PREDICTED: uncharacterized protein LOC109131263 [Camelina sativa]
2	AT4G28800	1.7E-42	-11.830	down	MF: DNA binding (GO:0003677), protein dimerization activity (GO:0046983)	K16189|1.1e-20|egr:104442415|K16189 phytochrome-interacting factor 4 | (RefSeq) transcription factor PIF5	Plant hormone signal transduction (ko04075)	--	Putative transcription factor bHLH056 OS=Arabidopsis thaliana OX=3702 GN=BHLH56 PE=3 SV=2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]
3	AT3G44300	1.6E-07	-6.643	down	MF: hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds (GO:0016810) BP: nitrogen compound metabolic process (GO:0006807)	K01501|3.0e-198|ath:AT3G44300|K01501 nitrilase [EC:3.5.5.1] | (RefSeq) NIT2; nitrilase 2	Tryptophan metabolism (ko00380);; Cyanoamino acid metabolism (ko00460);; Nitrogen metabolism (ko00910)	Amino acid transport and metabolism	Nitrilase 2 OS=Arabidopsis thaliana OX=3702 GN=NIT2 PE=1 SV=1	nitrilase 2 [Arabidopsis thaliana]
4	AT2G45570	4.0E-07	-6.001	down	MF: iron ion binding (GO:0005506), oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen (GO:0016705), heme binding (GO:0020037) CC: integral component of membrane (GO:0016021)	K20556|2.2e-293|ath:AT2G45570|K20556 cytochrome P450 family 76 subfamily C | (RefSeq) CYP76C2; cytochrome P450, family 76, subfamily C, polypeptide 2	--	Secondary metabolites biosynthesis, transport and catabolism	Cytochrome P450 76C2 OS=Arabidopsis thaliana OX=3702 GN=CYP76C2 PE=2 SV=1	cytochrome P450, family 76, subfamily C, polypeptide 2 [Arabidopsis thaliana]
5	AT3G19000	7.7E-167	-5.008	down	MF: oxidoreductase activity (GO:0016491), metal ion binding (GO:0046872)	K24028|1.4e-39|cam:101505284|K24028 salicylic acid 3-hydroxylase [EC:1.14.11.-] | (RefSeq) 1-aminocyclopropane-1-carboxylate oxidase-like	--	Secondary metabolites biosynthesis, transport and catabolism;; General function prediction only	Protein DOWNY MILDEW RESISTANCE 6 OS=Arabidopsis thaliana OX=3702 GN=DMR6 PE=1 SV=1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein [Arabidopsis thaliana]
6	AT1G19250	2.9E-03	-4.625	down	MF: N,N-dimethylaniline monooxygenase activity (GO:0004499), flavin adenine dinucleotide binding (GO:0050660), NADP binding (GO:0050661)	K22324|7.0e-16|csat:104714592|K22324 aliphatic glucosinolate S-oxygenase [EC:1.14.13.237] | (RefSeq) flavin-containing monooxygenase FMO GS-OX2 isoform X1	--	Secondary metabolites biosynthesis, transport and catabolism	Probable flavin-containing monooxygenase 1 OS=Arabidopsis thaliana OX=3702 GN=FMO1 PE=2 SV=1	flavin-dependent monooxygenase 1 [Arabidopsis thaliana]
7	AT1G47980	3.1E-06	-4.172	down	--	K15255|9.2e-24|nsy:104233171|K15255 ATP-dependent DNA helicase PIF1 [EC:3.6.4.12] | (RefSeq) ATP-dependent DNA helicase PIF1-like	--	--	Desiccation-related protein PCC13-62 OS=Craterostigma plantagineum OX=4153 PE=2 SV=1	desiccation-like protein [Arabidopsis thaliana]
8	AT2G44240	2.4E-02	-4.052	down	BP: response to oxidative stress (GO:0006979)	K12403|6.5e-06|psom:113315245|K12403 AP-4 complex subunit sigma-1 | (RefSeq) AP-4 complex subunit sigma-like isoform X1	--	--	--	NEP-interacting protein (DUF239) [Arabidopsis thaliana]
9	AT3G10930	7.2E-05	-3.763	down	--	--	--	--	--	hypothetical protein AT3G10930 [Arabidopsis thaliana]
10	AT4G28790	1.1E-29	-3.740	down	MF: DNA binding (GO:0003677)	--	--	--	Transcription factor bHLH23 OS=Arabidopsis thaliana OX=3702 GN=BHLH23 PE=2 SV=1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]
11	AT4G29020	1.0E-02	1.705	up	--	--	--	--	--	glycine-rich protein [Arabidopsis thaliana]

E) DEG between *nat2 vs. doxc21-a*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
12	AT4G25850	1.5E-02	1.708	up	--	K22285|5.6e-155|ath:AT4G25850|K22285 oxysterol-binding protein-related protein 8 | (RefSeq) ORP4B; OSBP (oxysterol binding protein)-related protein 4B	--	Signal transduction mechanisms	Oxysterol-binding protein-related protein 4B OS=Arabidopsis thaliana OX=3702 GN=ORP4B PE=2 SV=2	OSBP (oxysterol binding protein)-related protein 4B [Arabidopsis thaliana]
13	AT4G19430	1.6E-02	1.771	up	--	--	--	--	--	hypothetical protein AT4G19430 [Arabidopsis thaliana]
14	NewGene_114	6.4E-03	1.871	up	MF: nucleic acid binding (GO:0003676), RNA-DNA hybrid ribonuclease activity (GO:0004523)	--	--	General function prediction only	--	hypothetical protein AXX17_AT4G21170 [Arabidopsis thaliana]
15	AT5G54190	2.6E-03	1.906	up	MF: protochlorophyllide reductase activity (GO:0016630) BP: photosynthesis (GO:0015979), chlorophyll biosynthetic process (GO:0015995) CC: chloroplast (GO:0009507)	K00218|5.5e-161|ath:AT5G54190|K00218 protochlorophyllide reductase [EC:1.3.1.33] | (RefSeq) PORA; protochlorophyllide oxidoreductase A	Porphyrin and chlorophyll metabolism (ko00860)	Secondary metabolites biosynthesis, transport and catabolism	Protochlorophyllide reductase A, chloroplastic OS=Arabidopsis thaliana OX=3702 GN=PORA PE=1 SV=2	protochlorophyllide oxidoreductase A [Arabidopsis thaliana]
16	AT1G65450	2.2E-02	2.023	up	MF: transferase activity, transferring acyl groups other than amino-acyl groups (GO:0016747)	K15400|5.6e-161|aly:9324472|K15400 omega-hydroxypalmitate O-feruloyl transferase [EC:2.3.1.188] | (RefSeq) omega-hydroxypalmitate O-feruloyl transferase	Cutin, suberine and wax biosynthesis (ko00073)	--	Hydroxycinnamoyl transferase OS=Narcissus pseudonarcissus OX=39639 GN=HCT PE=2 SV=1	HXXXD-type acyl-transferase family protein [Arabidopsis thaliana]
17	AT3G23550	2.3E-02	2.156	up	MF: antiporter activity (GO:0015297), xenobiotic transmembrane transporter activity (GO:0042910) CC: integral component of membrane (GO:0016021)	K03327|3.3e-259|ath:AT3G23550|K03327 multidrug resistance protein, MATE family | (RefSeq) MATE efflux family protein	--	General function prediction only	Protein DETOXIFICATION 18 OS=Arabidopsis thaliana OX=3702 GN=DTX18 PE=2 SV=1	MATE efflux family protein [Arabidopsis thaliana]
18	AT4G33790	1.0E-02	3.016	up	MF: fatty-acyl-CoA reductase (alcohol-forming) activity (GO:0080019), alcohol-forming fatty acyl-CoA reductase activity (GO:0102965) BP: lipid metabolic process (GO:0006629)	K13356|1.1e-217|ath:AT4G33790|K13356 alcohol-forming fatty acyl-CoA reductase [EC:1.2.1.84] | (RefSeq) CER4; Jojoba acyl CoA reductase-related male sterility protein	Cutin, suberine and wax biosynthesis (ko00073);; Peroxisome (ko04146)	Lipid transport and metabolism	Fatty acyl-CoA reductase 3 OS=Arabidopsis thaliana OX=3702 GN=FAR3 PE=1 SV=1	G7 [Arabidopsis thaliana]
19	AT5G62165	7.7E-12	3.636	up	MF: RNA polymerase II transcription regulatory region sequence-specific DNA binding (GO:0000977), DNA-binding transcription factor activity (GO:0003700), protein dimerization activity (GO:0046983) BP: positive regulation of transcription by RNA polymerase II (GO:0045944) CC: nucleus (GO:0005634)	K09264|1.5e-35|gmx:100785089|K09264 MADS-box transcription factor, plant | (RefSeq) agamous-like MADS-box protein AGL11 isoform X1	--	Transcription	MADS-box protein AGL42 OS=Arabidopsis thaliana OX=3702 GN=AGL42 PE=1 SV=1	AGAMOUS-like 42 [Arabidopsis thaliana]
20	AT1G64795	1.9E-93	9.859	up	--	--	--	--	--	hypothetical protein AT1G64795 [Arabidopsis thaliana]

F) DEG between *nat1 vs.* nat2										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
1	AT1G53480	3.5E-11	-5.221	down	--	--	--	--	--	mto 1 responding down 1 [Arabidopsis thaliana]
2	AT1G53490	3.4E-15	-2.373	down	MF: Nucleic acid metabolic process (GO:0090304) BP: Meiosis I (GO:0007127); Reciprocal meiotic recombination (GO:0007131); Chromosome segregation (GO:0007059) CC: Nucleus (GO:0005634); Chromosome (GO:0005694); Chiasma (GO:0005712)	K10639| 1.7e168| ath:AT1G53490| K10639 E3 ubiquitin-protein ligase CCNP1IP1 [EC:2.3.2.27] | (RefSeq) HEI10; RING/U-box superfamily protein	--	--	E3 ubiquitin-protein ligase CCNB1IP1 homolog OS=Arabidopsis thaliana OX=3702 GN=HEI10 PE=2 SV=1	RING/U-box superfamily protein [Arabidopsis thaliana]
3	AT2G20670	3.5E-03	-2.041	down	--	--	--	--	--	sugar phosphate exchanger, putative (DUF506) [Arabidopsis thaliana]
4	AT5G54190	4.5E-02	-1.944	down	MF: Protochlorophyllide reductase activity (GO:0016630) BP: Photosynthesis (GO:0015979); Chlorophyll biosynthetic process (GO:0015995) CC: Chloroplast (GO:0009507)	K00218|5.5e-161| ath:AT5G54190| K00218 protochlorophyllide reductase [EC:1.3.1.33] | (RefSeq) PORA; protochlorophyllide oxidoreductase A	Porphyrin and chlorophyll metabolism (ko00860)	Secondary metabolites biosynthesis, transport and catabolism	Protochlorophyllide reductase A, chloroplastic OS=Arabidopsis thaliana OX=3702 GN=PORA PE=1 SV=2	protochlorophyllide oxidoreductase A [Arabidopsis thaliana]
5	AT2G44130	9.0E-04	-1.829	down	MF: Protein binding (GO:0005515) BP: Negative regulation of cytokinin-activated signaling pathway (GO:0080037); Regulation of phenylpropanoid metabolic process (GO:2000762); Negative regulation of signal transduction (GO:0009968) CC: Cytoplasm (GO:0005737); Cytosol (GO:0005829)	--	--	General function prediction only	F-box/kelch-repeat protein At2g44130 OS=Arabidopsis thaliana OX=3702 GN=At2g44130 PE=2 SV=2	Galactose oxidase/kelch repeat superfamily protein [Arabidopsis thaliana]
6	AT3G15500	1.3E-03	-1.774	down	MF: DNA binding (GO:0003677) BP: Regulation of transcription, DNA-templated (GO:0006355) CC: Nucleus (GO:0005634)	K00521|5.6e-45|cmos:111447866|K00521 ferric-chelate reductase [EC:1.16.1.7] | (RefSeq) ferric reduction oxidase 7, chloroplastic-like	--	--	NAC domain-containing protein 55 OS=Arabidopsis thaliana OX=3702 GN=NAC055 PE=2 SV=1	NAC domain containing protein 3 [Arabidopsis thaliana]
7	AT3G48360	3.4E-02	-1.745	down	MF: Histone acetyltransferase activity (GO:0004402); Transcription coregulator activity (GO:0003712); Zinc ion binding (GO:0008270) BP: Regulation of transcription, DNA-templated (GO:0006355) CC: Host cell nucleus (GO:0042025)	K10523|2.2e-07|ppp:112282598|K10523 speckle-type POZ protein | (RefSeq) BTB/POZ domain-containing protein At3g56230-like	--	Transcription	BTB/POZ and TAZ domain-containing protein 2 OS=Arabidopsis thaliana OX=3702 GN=BT2 PE=1 SV=1	BTB and TAZ domain protein 2 [Arabidopsis thaliana]
8	AT5G19120	1.7E-05	-1.745	down	MF: Aspartic-type endopeptidase activity (GO:0004190) BP: Cellular response to hypoxia (GO:0071456); Protein catabolic process (GO:0030163); Proteolysis (GO:0006508)	K22683|2.8e-14|gmx:100812865|K22683 aspartyl protease family protein [EC:3.4.23.-] | (RefSeq) LOW QUALITY PROTEIN: aspartyl protease family protein 2	--	Posttranslational modification, protein turnover, chaperones	Probable aspartic proteinase GIP2 OS=Nicotiana benthamiana OX=4100 GN=GIP2 PE=1 SV=1	Eukaryotic aspartyl protease family protein [Arabidopsis thaliana]
9	AT4G28755	7.3E-03	-1.668	down	--	--	--	--	--	hypothetical protein AT4G28755, partial [Arabidopsis thaliana]
10	AT1G21326	1.9E-02	-1.639	down	BP: Defense response signaling pathway, resistance gene-dependent (GO:0009870); Signal transduction (GO:0007165) CC: Nucleus (GO:0005634)	K20725|8.2e-33|dzi:111301972|K20725 MAP kinase substrate 1 | (RefSeq) nuclear speckle RNA-binding protein B-like	MAPK signaling pathway - plant (ko04016)	--	Nuclear speckle RNA-binding protein B OS=Arabidopsis thaliana OX=3702 GN=NSRB PE=2 SV=1	VQ motif-containing protein [Arabidopsis thaliana]
11	AT4G13395	1.1E-02	1.998	up	--	--	--	--	--	ROTUNDIFOLIA like 12 [Arabidopsis thaliana]

F) DEG between *nat1 vs. nat2*										
No.	**Gene ID**	**P value**	**Log2FC**	**regulated**	**GO annotation**	**KEGG annotation**	**KEGG pathway annotation**	**KOG class annotation**	**Swiss-Prot annotation**	**NR annotation**
12	AT4G27657	2.3E-04	2.030	up	--	--	--	--	--	hypothetical protein AT4G27657 [Arabidopsis thaliana]
13	AT3G62290	1.2E-33	2.255	up	MF: GTP binding (GO:0005525)	K07937|2.3e-99|aly:9318178|K07937 ADP-ribosylation factor 1/2 | (RefSeq) ADP-ribosylation factor 1	Endocytosis (ko04144)	Intracellular trafficking, secretion, and vesicular transport	ADP-ribosylation factor 1 OS=Arabidopsis thaliana OX=3702 GN=ARF1 PE=1 SV=2	PREDICTED: ADP-ribosylation factor 1-like isoform X1 [Brassica oleracea var. oleracea]
14	AT4G27652	1.8E-03	2.332	up	--	--	--	--	--	hypothetical protein AT4G27652 [Arabidopsis thaliana]
15	AT4G27654	2.1E-02	2.368	up	CC: integral component of membrane (GO:0016021)	--	--	--	--	transmembrane protein [Arabidopsis thaliana]
16	AT4G28790	5.8E-13	2.430	up	MF: DNA binding (GO:0003677)	--	--	--	Transcription factor bHLH23 OS=Arabidopsis thaliana OX=3702 GN=BHLH23 PE=2 SV=1	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]
17	AT3G27473	7.0E-10	2.630	up	BP: intracellular signal transduction (GO:0035556) MF: metal ion binding (GO:0046872)	K15042|1.5e-31|gra:105761158|K15042 importin subunit alpha-6/7 | (RefSeq) uncharacterized LOC105761158	--	--	--	Cysteine/Histidine-rich C1 domain family protein [Arabidopsis thaliana]
18	AT3G19000	2.9E-34	2.649	up	MF: oxidoreductase activity (GO:0016491), metal ion binding (GO:0046872)	K24028|1.4e-39|cam:101505284|K24028 salicylic acid 3-hydroxylase [EC:1.14.11.-] | (RefSeq) 1-aminocyclopropane-1-carboxylate oxidase-like	--	Secondary metabolites biosynthesis, transport and catabolism; General function prediction only	Protein DOWNY MILDEW RESISTANCE 6 OS=Arabidopsis thaliana OX=3702 GN=DMR6 PE=1 SV=1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein [Arabidopsis thaliana]
19	AT4G28800	1.9E-32	3.842	up	MF: DNA binding (GO:0003677), protein dimerization activity (GO:0046983)	K16189|1.1e-20|egr:104442415|K16189 phytochrome-interacting factor 4 | (RefSeq) transcription factor PIF5	Plant hormone signal transduction (ko04075)	--	Putative transcription factor bHLH056 OS=Arabidopsis thaliana OX=3702 GN=BHLH56 PE=3 SV=2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein [Arabidopsis thaliana]
20	NewGene 159	2.6E-126	13.118	up	--	K00558|9.43e-07|aof:109841697|K00558 DNA (cytosine-5)-methyltransferase 1 [EC:2.1.1.37] | (RefSeq) uncharacterized protein LOC109841697	Cysteine and methionine metabolism (ko00270)	General function prediction only	--	PREDICTED: uncharacterized protein LOC109131263 [Camelina sativa]

Publication No. 5

AIM

4) **Characterization of *A. thaliana* UDP-glucosyltransferase 79B9 involved in Fe deficiency and osmotic stress response.**

Perkowska, I., Siwinska, J., Dobek, A, Magot, F., Grosjean, J., Hehn, A., Olry, A., Lojkowska, E., & Ihnatowicz, A. **(2025b)**

Arabidopsis UGT79B9-mediated crosstalk between secondary metabolism, root exudation and plant responses to osmotic stress and iron deficiency

(manuscript in preparation)

Short title: UGT79B9 in plant metabolism and stress response

Arabidopsis UGT79B9-mediated cross-talk between secondary metabolism, root exudation and plant responses to osmotic stress and iron deficiency

Izabela Perkowska[1], Joanna Siwinska[1], Alicja Dobek[1], Florent Magot[2], Jeremy Grosjean[3], Alain Hehn[3], Alexandre Olry[3], Ewa Łojkowska[1], Anna Ihnatowicz[1,*]

[1]Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk, Abrahama 58, 80-807 Gdansk, Poland;
[2]Université de Tours, Biomolécules et Biotechnologies Végétales, EA2106 F-37200 Tours, France;
[3]Université de Lorraine, INRAE, LAE, 54000 Nancy, France;

***Corresponding author details:** Anna Ihnatowicz (Intercollegiate Faculty of Biotechnology of University of Gdansk and Medical University of Gdansk), anna.ihnatowicz@biotech.ug.edu.pl

One sentence summary: Arabidopsis UDP-glucosyltransferase (AtUGT79B9) with a previously unknown function, mediates the interaction between plant secondary metabolites, accumulated in roots and root exudates, which influence plant responses to abiotic stress factors like osmotic stress and iron deficiency.

Keywords: abiotic stresses, coumarins, plant–environment interactions, rhizosphere, scopolin, secondary metabolism, UDP-glucosyltransferase

ABSTRACT

In recent years, our team has established *Arabidopsis thaliana* as a valuable model for studying the genetic and molecular basis of variation in coumarin biosynthesis and accumulation. We have not only demonstrated the existence of large natural variation in the accumulation of various simple coumarin compounds among natural Arabidopsis populations, but we have also mapped the associated genomic regions and identified new candidate loci involved in this biosynthetic pathway. Among the previously identified candidate genes, *AtUGT79B9*, a member of the UDP-glucosyltransferase (UGT) family, emerged as a particularly compelling candidate due to its elevated expression in roots, a tissue essential for the accumulation and secretion of various important secondary metabolites, including coumarins.

In this study, we further investigated the function of the previously uncharacterized AtUGT79B9 enzyme and its role in the accumulation of secondary metabolites in roots and root exudates in response to abiotic stress factors. We transiently overexpressed *UGT79B9* in *Nicotiana*

benthamiana and performed targeted metabolite profiling using UHPLC-MS, complemented by *in silico* analyses and functional characterization of the Arabidopsis *ugt79b9* knockout, as well as complementation and overexpression lines. These analyses revealed that UGT79B9 plays a role not only in iron (Fe) deficiency responses but also in osmotic stress tolerance. Importantly, in soil tests the *ugt79b9* knockout mutant was found to be more resistant to drought stress than both tested wild-type lines, Col-0 and Est-1 Arabidopsis accessions. To comprehensively assess global metabolic changes associated with the loss of UGT79B9 function, and to uncover potential metabolic changes beyond coumarin pathways, we employed a targeted and untargeted metabolomics approach to identify specifically differentially accumulated coumarins and other metabolites across genotypes and stress conditions. This analysis revealed distinct metabolic phenotypes across various compounds, which were associated with plant genotypes (*ugt79b9* knockout, wild-type, and transgenic lines) and influenced by Fe availability in the liquid medium or PEG-induced osmotic stress in hydroponics.

Ongoing metabolic network analysis and clustering are focused on identifying specific metabolite interactions and regulatory patterns. Under varying Fe conditions, certain compounds clustered with potential glycosylated molecules derived from flavonols, phenylpropanoids, and coumarins. Remarkably, untargeted metabolomic analyses of root extracts and exudates from plants grown under Fe-deficient *in vitro* conditions, revealed common metabolic features across the roots and exudates of studied genotypes. Our findings indicate that UGT79B9 may function as a metabolic switch in response to osmotic stress and Fe deficiency, potentially influencing the accumulation of secondary metabolites, such as coumarins, involved in stress tolerance.

INTRODUCTION

Secondary metabolites are a large and diverse group of chemical compounds produced by plants in response to various environmental stress factors. Coumarins are a group of metabolites that are accumulated and secreted by roots into the rhizosphere, where they perform many different functions, not all of which are yet fully understood. In recent years, many studies have been conducted on the model plant *Arabidopsis thaliana* (hereinafter referred to as Arabidopsis), where the main coumarins are scopoletin and its glycosylated form – scopolin (Rohde *et al.*, 2004; Bednarek *et al.*, 2005; Kai *et al*., 2006 and 2008; Schmid *et al.*, 2014; Siwinska *et al*., 2014; Sisó-Terraza *et al*., 2016; Stringlis *et al.*, 2019; Perkowska *et al.*, 2021a; Ihnatowicz *et al.*, 2024). Coumarins have been shown to play an important role in iron (Fe) acquisition and homeostasis, especially under Fe deficiency conditions (Fourcroy *et al.*, 2014, Schmidt *et al.*, 2014, Rajniak *et al.*, 2028, Siwinska *et al.*, 2018, Tsai *et al.*, 2018, Perkowska *et al.*, 2021b). Moreover, numerous reports highlight their antibacterial properties and their crucial role as key components of root exudates in shaping the rhizosphere microbiome (Stringlis *et al.*, 2018, Voges *et al.*, 2019, Harbort *et al.*, 2020, Stassen *et al.*, 2021).

Numerous enzymes belonging to distinct chemical classes contribute significantly to plant secondary metabolism. In Arabidopsis, the largest and most diverse enzyme family is cytochrome P450 (CYP) (Werck-Reichhart & Feyereisen 2000, Bourgaud *et al.*, 2006, Bak *et al.*, 2011), whose members are involved in the oxidation, hydroxylation, and biosynthesis of terpenoids, phenols, alkaloids, and other specialized metabolites (Mizutani & Ohta, 2010). Two other relatively large and functionally broad families are: [1] the 2-oxoglutarate-dependent Fe(II) dioxygenases (DOX) (Farrow & Facchini 2014, Kawai *et al.*, 2014, Shimizu *et al.*, 2014), which mediate oxidative cleavage, hydroxylation, and ring modifications, [2] and glycosyltransferases (GTs) that catalyze the transfer of sugar moieties to a wide range of acceptor molecules, thereby modulating their solubility, stability, and biological activity (Gachon *et al.*, 2005). These enzymes play a key role in shaping metabolite diversity and enable plant adaptation to environmental stress by modulating detoxification, transport, and storage of secondary metabolites.

UGT enzymes belong to the large family of GTs, which play a crucial role in the modification of secondary metabolites (Rehman *et al.*, 2018, Wang *et al.*, 2023). They catalyze glycosylation, which involves the attachment of a sugar group (from sugar molecules such as glucose, galactose, or other sugars) to various acceptor molecules, including flavonoids, glucosinolates, terpenoids, phytohormones, and other specialized metabolites. As a result, glycosylation affects the properties of the molecule by influencing its solubility, stability, as well as the transport and bioactivity of compounds. A lot of glycosylated molecules are stored in the vacuole and excreted in the cytosol under stress conditions (Gachon *et al.*, 2005). Although UGTs represent the largest subfamily within the GTs, the biological functions and significance of many UGTs, especially in relation to stress adaptation in terrestrial plants, remain poorly understood.

UGTs in Arabidopsis constitute a large family of enzymes that play critical roles in regulating plant growth, stress responses, and immunity. Several AtUGTs are strongly induced by abiotic stresses such as cold, salinity, and drought. For example, UGT79B2 and UGT79B3 have been shown to contribute to stress tolerance by modifying anthocyanins and enhancing antioxidant activity (Pan *et al.*, 2017), while UGT86A1 promotes stress adaptation by increasing soluble sugar levels and facilitating reactive oxygen species (ROS) scavenging (Ma *et al.*, 2025). Other UGTs are implicated in plant immunity. UGT73C7, for instance, is involved in plant immune responses by redirecting phenylpropanoid metabolism (Hang *et al.*, 2021), while UGT76B1 contributes to disease resistance in Arabidopsis by generating defense-related metabolites (Lennart *et al.*, 2021). Some UGTs were shown to be essential for the biosynthesis coumarin derivatives and their regulation in Arabidopsis. The members of UGT72E subfamily were shown to be able to use coumarins as substrates leading to formation of glycosylated coumarin derivatives (Wu *et al.*, 2022). UGT72E members have been shown to be able to utilize coumarins as substrates, and genetic evidence suggests that enzymes that belong to UGT72E subfamily provides plants with increased tolerance to low Fe availability under alkaline conditions. Although several enzymes from other families have

been characterized in the coumarin biosynthesis pathway, many aspects of its regulation remain unclear. Notably, this regulation may involve glycosylation steps catalyzed by yet uncharacterized UGTs.

In previous genetic mapping studies, we identified a number of candidate genes potentially involved in scopolin and scopoletin accumulation in Arabidopsis plants (Siwinska *et al.*, 2014). Among the candidate genes selected on the basis of *in silico* analysis of tissue distribution and expression level, we selected one particularly interesting gene, AT5G53990, encoding a UGT named UGT79B9 that is relatively highly expressed in Arabidopsis roots. According to the 1001 Genomes Project (https://1001genomes.org/) and resequencing performed in our laboratory, several single nucleotide polymorphisms (SNPs) are present in the coding sequences of the AT5G53990 gene in Arabidopsis natural populations differing in accumulation of coumarin (Siwinska *et al.*, 2015; Perkowska *et al.*, 2021).

Here, we performed a comparative biochemical analysis of UGT79B9 from two Arabidopsis accessions Col-0 and Est-1, which differ in their scopoletin and scopolin levels. Transient overexpression of *UGT79B9* from Col-0 in *N. benthamiana* leaves led to elevated scopolin levels. In Arabidopsis, metabolic profiling of the *ugt79b9* knockout revealed an impaired scopolin to scopoletin ratio. *In silico* analysis suggested a role for UGT79B9 in osmotic stress response, which was supported by stress assays showing that *ugt79b9* knockout mutants developed larger roots and rosettes that was associated with improved stress tolerance. We also conducted untargeted metabolic profiling of methanol extracts from roots and root exudates of *ugt79b9* knockout, wild-type, and UGT79B9-overexpressing lines grown under Fe deficiency and PEG-induced stress conditions. These results suggest that UGT79B9 may function as a metabolic switch in responses to osmotic stress and Fe deficiency, likely *via* its role in coumarin biosynthesis. Mutation in *UGT79B9* may alter the levels of various secondary metabolites, including coumarins linked to drought tolerance and Fe deficiency responses. This is consistent with recent studies on coumarins in root exudates under Fe stress and the interplay between Fe and osmotic stress signaling. Remarkably, this is the first evidence of a biological function of UGT79B9 in Arabidopsis, highlighting its significance in stress adaptation through changes in secondary metabolism.

RESULTS

In Vivo Activity of UGT79B9 in *Nicotiana benthamiana*

One of the first steps of UGT79B9 functional characterization was to test its enzyme activity and substrate specificity using transient expression in *N. benthamiana*, followed by targeted metabolite profiling by UHPLC-MS. Transient overexpression of *AtUGT79B9* from the Arabidopsis Col-0 reference line in tobacco leaves resulted in a significantly elevated level of scopolin compared to the control infiltration using *Agrobacterium tumefaciens* carrying an empty vector (**Figure 1**). The accumulation of the glycosylated compound, namely scopolin, suggests that the UGTT79B9 enzyme is active *in planta* and is metabolically connected to scopolin biosynthesis.

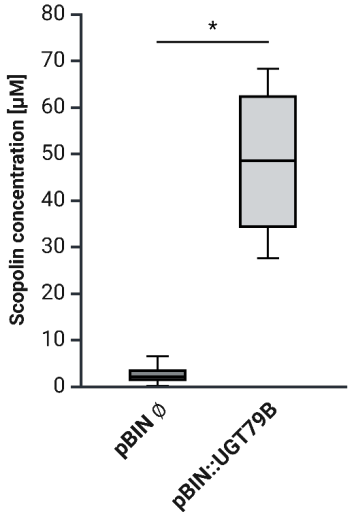


Figure 1. Scopolin content in *Nicotiana benthamiana* leaves quantified using UHPLC-MS. Transient expression was carried out using *Agrobacterium tumefaciens* transformed with the pBIN-GW empty vector (control) and pBIN-GW:*UGT79B9*. Whiskers indicate min and max values. * $P<0.05$ (Welch's t-test, 4 biological replicates).

In Vitro Enzyme Assays of UGT79B9

To further characterize the enzymatic properties of the UGT79B9, *in vitro* enzyme assays were conducted using the purified recombinant protein to determine its substrate specificity and kinetic parameters. We overexpressed the His-tagged UGT79B9 protein in a heterologous *Escherichia coli* system to obtain a purified protein fraction according to Siwinska et al. (2018). Although we were able to overproduce UGT79B9 using pET28a and pCold vectors (Supplementary Fig. S2), we were unable to determine the enzymatic activity of the renatured UGT79B9 protein (data not shown).

Characterization of *ugt79b9* Knockout Mutant Under Liquid Culture Conditions Promoting Root Growth

Next step was to perform biochemical characterization using Arabidopsis mutant line as a powerful approach to validate the *in vivo* function of the UGT79B9 enzyme identified through *in planta* assays in tobacco leaves. To gain insight into the biochemical and physiological role of UGT79B9 enzyme, we identified a mutant line with a non-functional At5g53990 gene (*ugt79b9*) and cultured it with two wild type lines - Col-0 and Est-1 accessions. The plants were grown *in vitro* in liquid cultures – conditions that not only support root growth (the tissues where coumarins and other important chemicals accumulate under stress) but also facilitate the harvesting of root material for further applications, including metabolic profiling. After 3 weeks of growth, quantification of scopoletin and its glycosylated form, scopolin, in methanol root extracts was conducted using UHPLC-MS. This analysis revealed different levels of the tested coumarin compounds between the three genotypes (**Figure 2**).

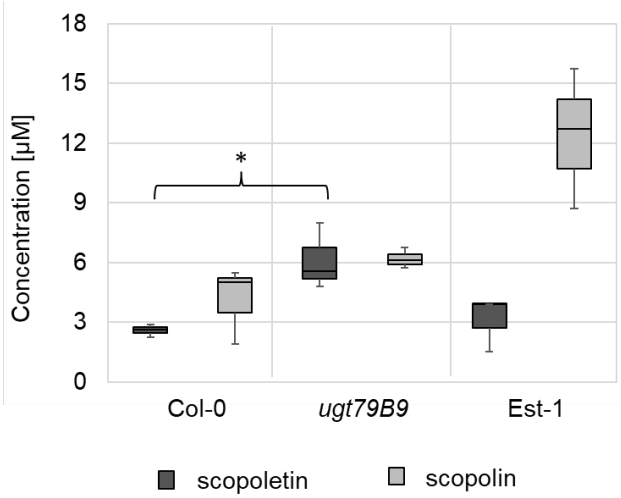


Figure 2. Relative levels of scopoletin and scopolin in methanol root extracts of Arabidopsis accessions (Col-0, Est-1) and *ugt79B9* mutant plants (Col-0 genetic background) grown *in vitro* in control ½ MS liquid cultures. Data are presented as mean ±SD. **P*<0.05.

Remarkably, the ratio of scopoletin to scopolin was significantly altered in the *ugt79b9* mutant compared to both wild-type lines, in which scopolin levels were clearly higher compared to scopoletin (**Figure 2**). Equal levels of both of these compounds, which are potential substrates and products of the UGT-catalyzed reaction, respectively, were observed in the *ugt79b9* mutant. This suggests a disruption in glycosylation activity attributed to the loss of UGT79B9 function.

***In Planta* Analysis of UGT79B9 Activity *via* Scopoletin Feeding**

To further investigate the enzymatic activity catalyzed *in planta* by the UGT79B9 enzyme, we conducted a similar experiment with the addition of exogenous scopoletin to assess its impact on glycosylation dynamics. Interestingly, we noticed that the addition of scopoletin (0.25 µM),

which is a potential substrate of the UGT79B9-catalyzed enzymatic reaction, significantly increased its possible product, namely scopoletin, but only in two wild-type lines, Col-0 and Est-1. In contrast, in *ugt79B9* knockout line, addition of exogenous scopoletin led to reduced levels of both scopoletin and scopolin compared to control conditions, indicating a possible disruption of coumarin metabolism or altered feedback regulation in the absence of functional UGT activity (**Figure 3**).

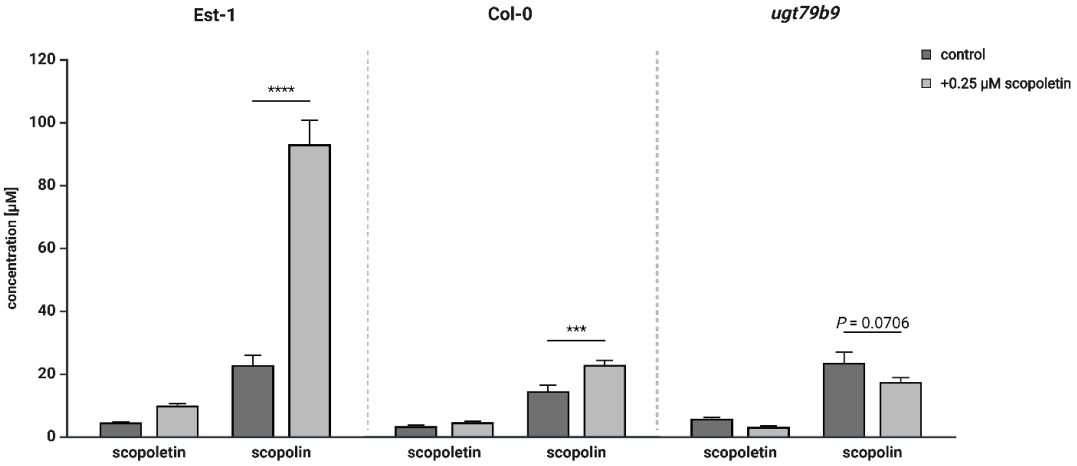


Figure 3. Relative levels of scopoletin and scopolin in methanol root extracts of *Arabidopsis thaliana* accessions (Col-0, Est-1) and *ugt79B9* mutant plants grown *in vitro* in ½ MS liquid cultures, in control conditions and with exogenous scopoletin (+0.25 µM). Data is presented as mean ±SEM. ***P<0.001, ****P<0.0001 (two-way ANOVA with Bonferroni multiple comparisons test, 4-6 replicates)

In Silico Analysis of Arabidopsis UGT79B9 UDP-Glucosyltransferase

To explore further the potential role of the previously uncharacterized UGT79B9 enzyme in coumarin biosynthesis we compared its protein sequence with other UGT proteins whose activity in this pathway has been experimentally confirmed (**Figure 4**). A comparison of multiple sequence alignment from Arabidopsis and tobacco, shows that AtUGT79B9 shares a high level of similarity with other coumarin-related UGTs within the 44 amino acid PSPG-box (plant secondary product glucosyltransferase-box) located in the C-terminal part of the protein, which is responsible for the majority of interactions with the sugar donor (Osmani *et al.*, 2008, Hansen *et al.*, 2009, Wang

et al., 2009). It should be noted that some key N-terminal residues with which acceptors might interact (Osmani *et al.*, 2009) are identical among all sequences tested.

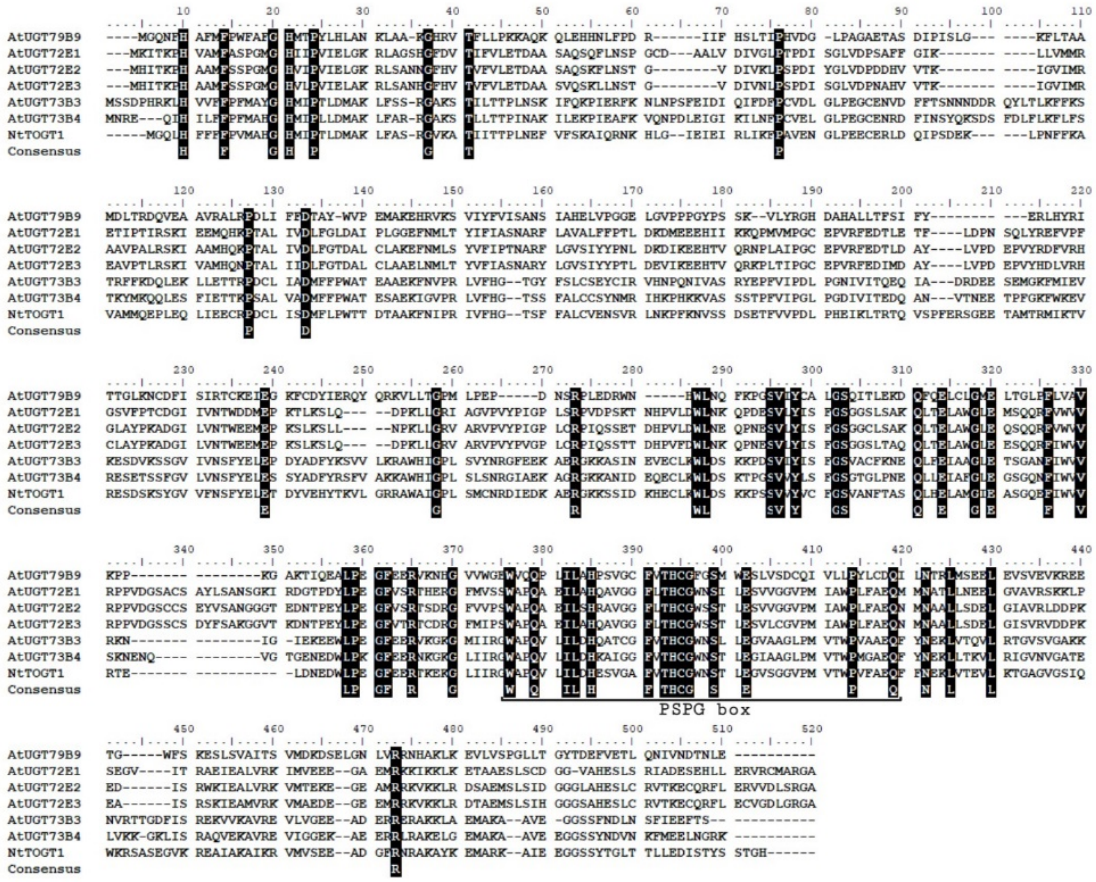


Figure 4. Amino acid sequence alignment of AtUGT79B9 and coumarin-related UGTs from Arabidopsis and Tobacco plants, namely AtUGT72E1-E3, AtUGT73B3-B4 and NtTOGT1.

Then, using STRING databases (https://string-db.org/), we checked the predicted functional association for UGT79B9 protein taking into account available gene co-expression and textmining data. Among the closest predicted functional partners of UGT79B9 are two, MYB40 (At5g14340) and WRKY72 (AT5G15130), which not only show the highest expression in vegetative rosette roots but were also among the candidate genes selected in our genetic mapping, likely underlying observed variability in coumarin content (Siwinska *et al.,* 2015). It should also be mentioned that previous microarray-based transcript abundance analysis showed that MYB40 is a potential THO1-dependent transcript. The Arabidopsis THO/TREX complex, whose component is the THO1 protein, mediates the accumulation of scopolin under abiotic stress conditions (Doll *et al.*, 2018). Even though, Doll *et al.* (2018), did not establish a consistent effect of mutations in MYB40 loci on scopolin accumulation, the potential link between the candidate gene *UGT79B7* and coumarin biosynthesis in Arabidopsis required further investigation. In particular, further studies are needed on the role of UGT79B7 under various environmental stress conditions that affect the accumulation of secondary metabolites – such as Fe deficiency, which is known to affect the synthesis and accumulation |

of coumarin, and osmotic stress, which, according to the cited studies, significantly affects the accumulation of scopolin in the roots. Therefore, the next step was an *in silico* gene expression analysis using Arabidopsis Affymetrix microarray data (http://www.bar.utoronto.ca/eplant/), focusing on the absolute expression level of *AT5G53990* under different abiotic stresses (**Figure 5**) to identify the environmental conditions regulating the expression of the gene encoding UGT79B7.

As shown in **Figure 5**, drought and osmotic stress are the two most important factors influencing *UGT79B7* gene expression levels, inducing it by approximately 3-4-fold. Therefore, subsequent analyses focused on these stress conditions, along with coumarin-related Fe deficiency stress.

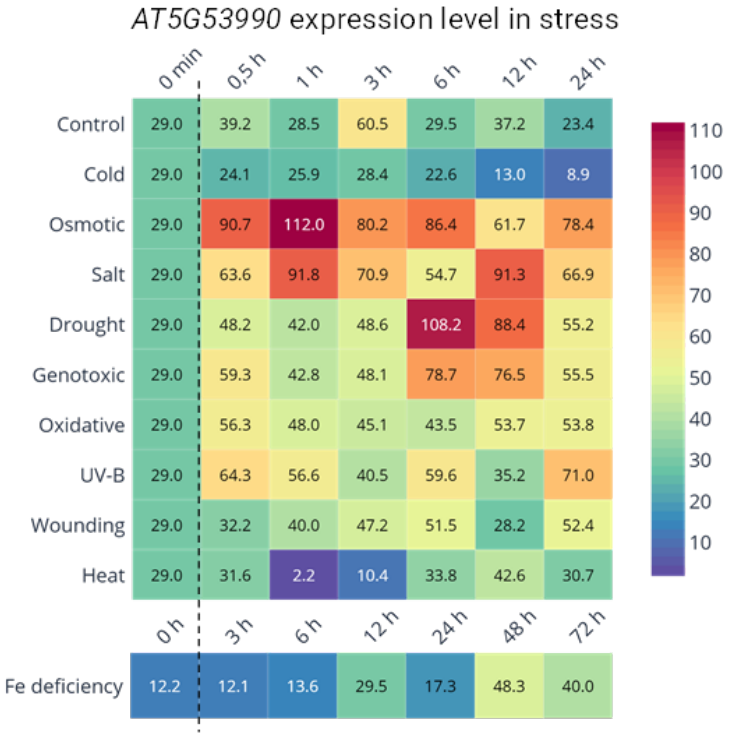


Figure 5. *In silico* gene expression analysis using Arabidopsis Affymetrix microarray data (www.bar.utoronto.ca/eplant/). Absolute expression level of *AT5G53990* under various abiotic stresses. (± standard deviation). Blue color indicates lover expression level in comparison to the expression level in plants from control conditions and red – higher expression level. All data were obtained from Kilian *et al.* (2007, Plant Journal 50:347-63) except iron deficiency results from Dinneny *et al.* (2008) Science 320:942-945.

Functional Analysis of *ugt79b9* Mutant Under Osmotic Stress in a Hydroponic System

The next step was to examine the growth and phenotypic characteristics of *ugt79b9* mutant plants under selected stress conditions, alongside control lines. These included Col-0, the genetic background of the mutant, and the Est-1 accession, which served as the second parental line in a previous genetic mapping study (Siwinska *et al*., 2015). Plants were grown in hydroponic cultures under control conditions and under osmotic stress induced by 3% (w/v) PEG 8000.

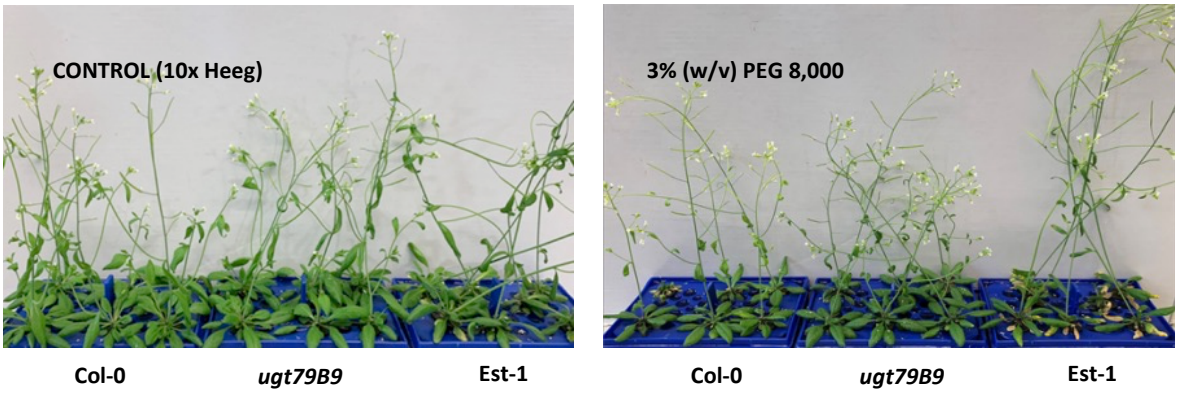


Figure 6. Phenotypic characterization of 5-week-old wild-type (Col-0, Est-1 accessions) and *ugt79B9* T-DNA mutant plants (in Col-0 genetic background) grown under osmotic stress induced by 3% (w/v) PEG 8,000. Col-0, Est-1 and *ugt79B9* mutant line were cultivated in 10x Heeg solution with (stress condition) or without (control) 3% PEG.

Under PEG-induced osmotic stress, *ugt79b9* knockout mutants exhibited slightly improved growth compared to Col-0 plants and were noticeably larger, with less symptoms of leaf senescence, than the second wild-type line, accession Est-1 (**Figure 6**). Importantly, under osmotic stress conditions, both the rosette mass and root length of the *ugt79B9* mutant were significantly increased compared to its genetics background (Col-0) (**Figure 7**), indicating that loss of UGT79B9 function can enhance growth under osmotic stress conditions. For both growth-related phenotypic traits, no statistically significant differences were observed between Col-0 and *ugt79B9* mutant under control conditions (**Figure 7**).

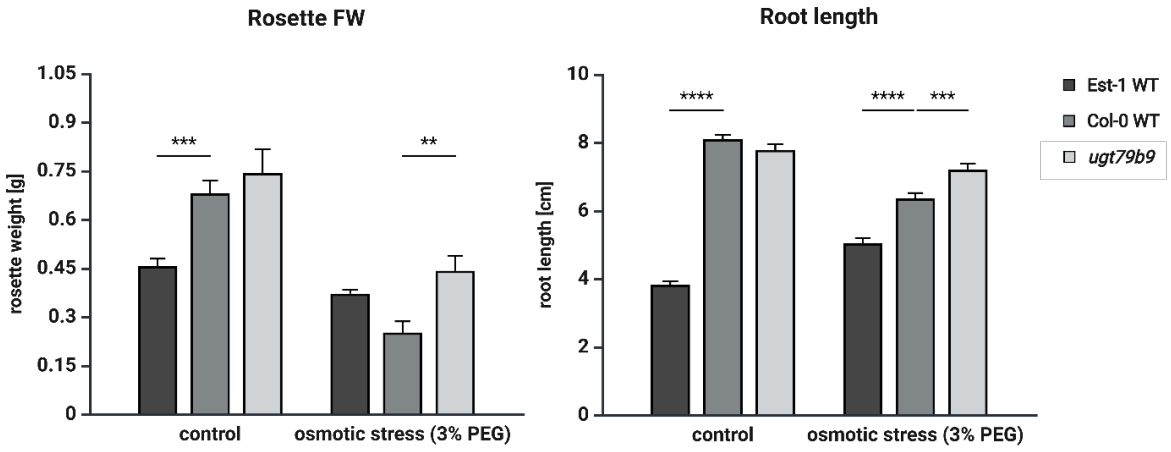


Figure 7. Rosette weight and root length of Col-0 (WT), *ugt79B9* mutants in Col-0 genetic background and Est-1 accession after cultivation in hydroponic cultures under control and osmotic stress conditions. Data is presented as mean ±SEM. ***P*<0.01, ****P*<0.001, ****P<0.0001 (two-way ANOVA with Bonferroni multiple comparisons test, 12-24 replicates)

AtUGT79B9 Expression Is Significantly Enhanced by PEG-Induced Osmotic Stress in an Accession-Dependent Manner in Hydroponic Culture

Considering the differential response to osmotic stress observed between two Arabidopsis wild-type populations (Col-0 and Est-1, **Figure 6 and 7**), the parental lines of the mapping population described by Siwińska *et al.* (2014), we investigated the relative expression of *UGT79B9* in the roots of 5-week-old plants grown in hydroponic culture under both control and osmotic stress conditions.

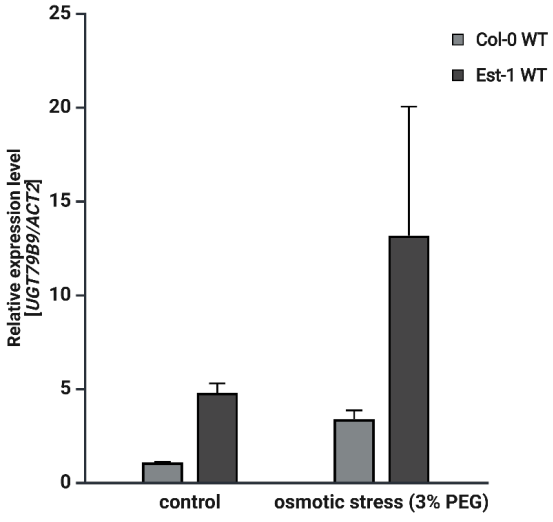


Figure 8. Relative expression level of *UGT79B9* in the roots of 5-week-old wild-type (Col-0, Est-1 accessions) grown under osmotic stress induced by 3% (w/v) PEG 8,000. Col-0 and Est-1 plants were cultivated in 10x Heeg solution with (stress condition) or without (control) 3% PEG. Data is presented as mean ±SEM. Two-way ANOVA with Bonferroni multiple comparisons test, 3 replicates

The expression level of the *UGT79B9* gene, which varied between Arabidopsis accessions under control conditions, was significantly upregulated by osmotic stress, regardless of genetic background (**Figure 8**). Notably, under control conditions, *UGT79B9* expression was higher in the Est-1 accession. Osmotic stress induced a significant increase in expression in both genotypes, with a highly significant induction in Est-1, which appears to be less tolerant to osmotic stress than Col-0 in PEG-containing hydroponic cultures (**Figures 6 and 7**).

Untargeted Metabolite Profiling of Methanolic Root Extracts from Col-0 and *ugt79b9* Plants Grown Hydroponically Under PEG-Induced Osmotic Stress

Given the lack of a clear binary relation between the *ugt79B9* knockout mutation and accumulation of coumarins, the inconclusive results from our *in vitro* enzyme assays, and the broad substrate specificity that is found in UGT enzyme family, we proceeded with untargeted metabolic profiling. We aimed to comprehensively assess global changes in the metabolite landscape associated with the loss of UGT79B9 function, and to uncover potential metabolic changes beyond coumarin pathways. To achieve that, we conducted untargeted metabolite profiling of methanolic root extracts from hydroponically grown Col-0 and *ugt79B9* knockout plants using a Linear Trap Quadrupole (LTQ) mass spectrometer, which has the advantage of high sensitivity. The resulting data were processed and quantified using the XCMS software.

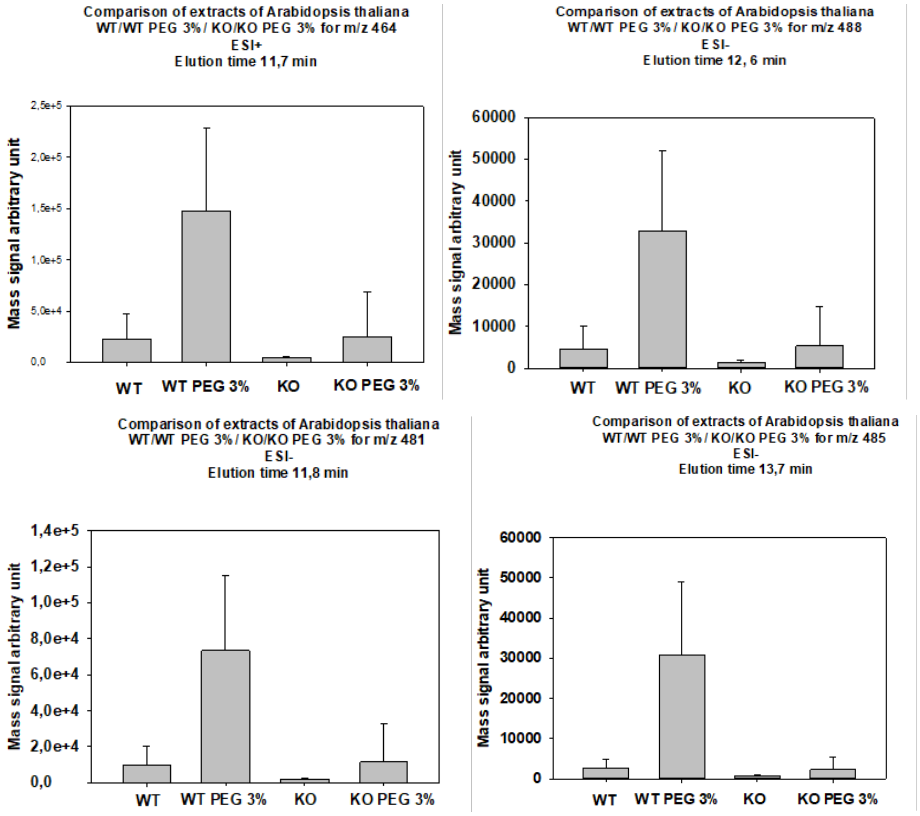


Figure 9. Metabolic phenotype of the *ugt79B9* knockout mutant (KO). Untargeted metabolomic profiling was conducted on methanolic root extracts from hydroponically grown Col-0 (wild-type, WT) and *ugt79B9* knockout (KO) plants under both control and 3% PEG-induced osmotic stress. A consistent disappearance of four molecules of four metabolites was observed in the *ugt79B9* mutant: three (m/z 488, 485, and 481) were detected in negative ionization mode (ESI−), and one (m/z 464) in positive ionization mode (ESI+).

Our results revealed a significant impact of the *ugt79b9* mutation on the accumulation of certain features (**Figure 9**). By comparing ion signals between Col-0 *vs. ugt79b9* samples under both control and PEG-induced osmotic stress conditions, four m/z features were identified that were consistently downregulated in the *ugt79b9* mutant. Notably, these molecules showed strong induction in response to PEG treatment in wild-type plants (**Figure 9**). Three of the affected molecules (m/z 488, 485, and 481) were detected in negative ionization mode (ESI−), while one (m/z 464) was detected in positive mode (ESI+). We did not detect any significant differences in the content of coumarins between Col-0 and the *ugt79b9* mutant line.

Generation of 35S::*UGT79B9*/Col-0 overexpression line and 35S::*UGT79B9*/*ugt79b9* complementation line

To further elucidate the functional role of UGT79B9 in Arabidopsis and its potential involvement in plant adaptation to stress *via* secondary metabolism biosynthesis, we generated two types

of overexpressing (OX) lines: (1) one to study the effects of increased *UGT79B9* gene expression on the plant's phenotype, metabolism, or stress response (35S::*UGT79B9*/Col-0), (2) and the other one in the *ugt79b9* mutant background (35S::*UGT79B9*/*ugt79b9*) to complement the mutant phenotype and further investigate the role of UGT79B9. All experiments and results described below were performed using these four plant genotypes: wild type (Col-0), a knockout mutant (*ugt79b9*), and the two overexpressing lines described above.

Untargeted metabolite profiling of roots and exudates from Col-0, *ugt79B9*, and transgenic lines under Fe-deficient conditions

Untargeted metabolite profiling was conducted on root and exudate extracts from Col-0, *ugt79B9* knockout plants, as well as 35S::*UGT79B9*/Col-0 overexpression and 35S::*UGT79B9*/*ugt79B9* complementation lines, all grown *in vitro* in liquid culture under control and Fe-deficient conditions. Using an LTQ mass spectrometer, we analyzed Arabidopsis root and root exudate extracts from *in vitro* liquid cultures to identify m/z signal intensities. The data were processed and compared using XCMS software. Samples were analyzed in both positive and negative ionization modes, comparing cultures grown with or without Fe. Genotypic comparisons included Col-0 wild type (WT) *vs. ugt79b9* knockout (KO), Col-0 *vs.* 35S::*UGT79B9*/Col-0 (Col-0::UGT, overexpression line), and Col-0 vs. 35S::*UGT79B9*/*ugt79b9* (ugt::UGT; complementation line). As a result, we detected the m/z 413.0 when we filtered the chromatogram between 412.5-413.5 (**Figure 10B**) with a loss of m/z 45.9 (**Figure 10A**) clearly identified in all the samples except for the KO mutants (**Figure 11**).

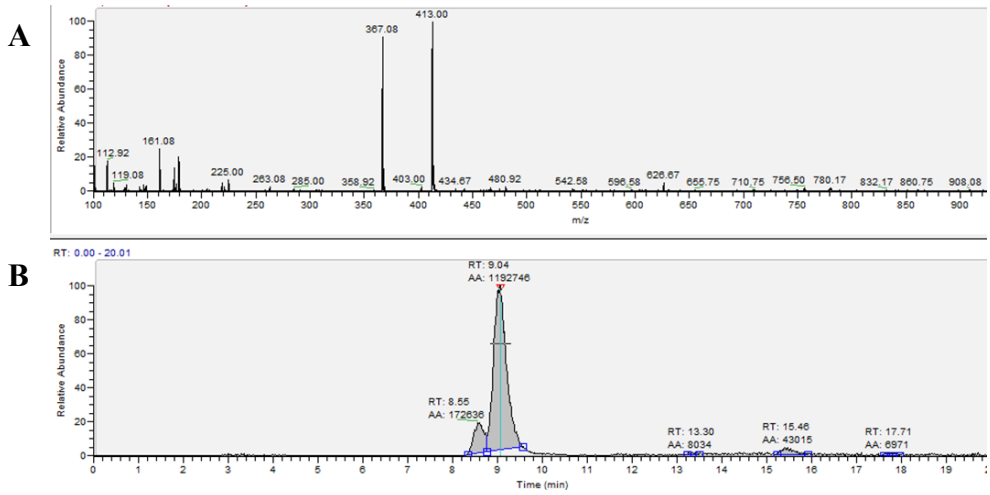


Figure 10. Mass spectrometry analysis of detected m/z 413.0. A. fragmentation pattern and B. the full filtered chromatogram between 412.5-413.5.

As shown in **Figure 11**, the m/z 413.0 signal is completely absent in the root extracts of the *ugt79b9* knockout (KO) line under both Fe-sufficient and Fe-deficient conditions in liquid cultures. In contrast, this signal accumulates significantly, about 2-3-fold, in both the overexpressing lines 35S::*UGT79B9*/Col-0 and 35S::*UGT79B9*/*ugt79B9* (Col-0::UGT and ugt::UGT, respectively).

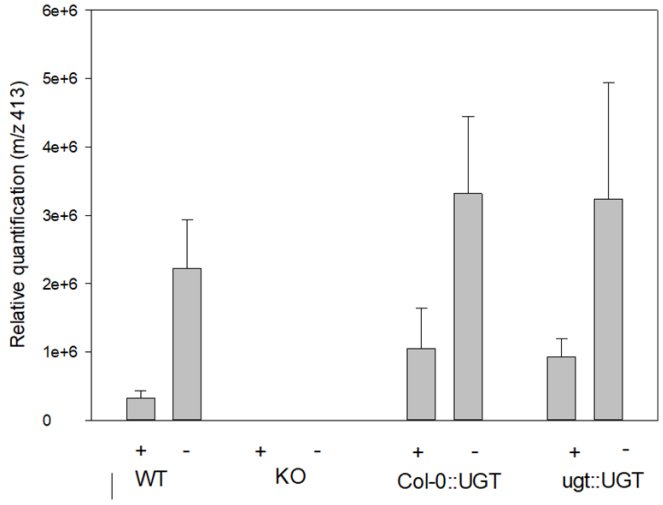


Figure 11. Root metabolic phenotype and relative quantification of m/z 413.0 of the WT, *ugt79B9* knockout mutant (KO) and *ugt79B9 over-expressing* (Col-0::UGT and ugt::UGT) *lines.* Targeted metabolomic profiling was conducted on methanolic root extracts from hydroponically grown Arabidopsis plants under both control and minus Fe stress.

Remarkably, under Fe-deficient conditions, the m/z 413.0 signal increases almost 9-fold, suggesting that the gene(s) involved in the biosynthesis of this compound are inducible under Fe deficiency conditions. Even under Fe-sufficient conditions, overexpression of *UGT79B9* in the Col-0 and *ugt79b9* mutant backgrounds leads to a 2- to 3-fold increase in accumulation compared to WT. Under Fe-deficient conditions, the compound corresponding to the m/z 413.0 signal accumulates strongly in overexpressing lines compared to WT, highlighting the contribution of the gene to metabolite production. Interestingly, the magnitude of induction under both Fe-sufficient and Fe-deficient conditions is similar in WT and overexpressing lines, suggesting that the upstream induction mechanism remains functional in the absence of UGT79B9, and m/z 413.16 likely represents a downstream product of UGT79B9 activity.

Notably, the compound corresponding to m/z 413.0 was also detected in the exudates of Col-0 plants, at similar levels under both Fe-sufficient and Fe-deficient conditions (**Figure 12**). In contrast, it is completely absent from the exudates of the *ugt79b9* knockout line, confirming the requirement of *UGT79B9* for its synthesis or secretion. Under Fe-sufficient conditions, m/z 413.0 accumulates approximately 10-fold in 35S::*UGT79B9*/Col-0 and 13-fold in 35S::*UGT79B9*/*ugt79B9* overexpression lines (Col-0::UGT and ugt::UGT, respectively) compared to wild type. Interestingly, Fe deficiency results in a reduced accumulation of m/z 413.0 in both overexpression lines.

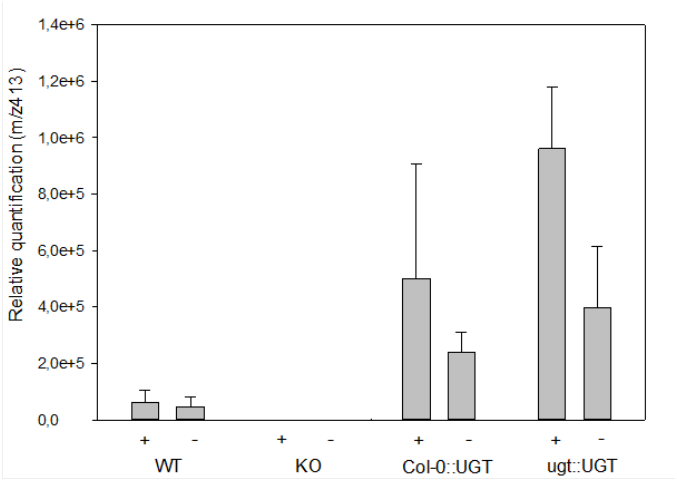


Figure 12. Root exudates metabolic phenotype and relative quantification of m/z 413.0 of the WT, *ugt79B9* knockout mutant (KO) and *ugt79B9 over-expressing* (Col-0::UGT and ugt::UGT) *lines.* Targeted metabolomic profiling was conducted on methanolic root extracts from hydroponically grown Arabidopsis plants under both control and minus Fe stress.

The observed pattern suggests that the secretion of the molecule corresponding to m/z 413.0 may be regulated by the cellular Fe concentration and its availability. This may indicate a possible feedback mechanism or resource allocation in response to plant nutritional status.

Metabolomic clustering of m/z 413.0

The initial metabolomic analyses using the LTQ mass spectrometer confirmed the impact of the mutation on the composition of certain compounds, in particular a m/z value of 413.0, which is completely absent in the *ugt79b9* mutant line. The acquisition of an Orbitrap mass spectrometer enabled the generation of high-resolution data, which were subsequently analyzed using advanced software (Compound Discoverer). Samples were injected and we focused on the 26 features that were showed to be significantly sensitive to the presence of *ugt79b9* mutated allele. Through clustering analysis, we identified about 20 molecules with closely related structures within this cluster (**Figure 13**). This result indicates a close structural similarity between almost all these features.

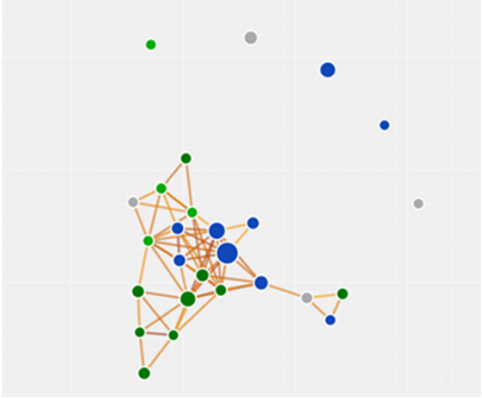


Figure 13. Metabolomic clustering of m/z 413.0.

To further investigate the physiological role of UGT79B9, we performed a comparative untargeted metabolomics analysis of Arabidopsis roots and exudates from Col-0 wild-type plants, *ugt79b9* knockout, as well as overexpression (35S::*UGT79B9*/Col-0) and complementation (35S::*UGT79B9*/*ugt79b9*) lines grown under Fe-deficient conditions.

Comparative Metabolomic Analyses of Extracts from Col-0, *ugt79B9* knockout and overexpressing lines grown *in vitro* under Fe-deficiency

By comparing Col-0 to *ugt79B9* compounds, of the approximately 350 m/z features detected in ESI- mode identified in the extracts, the relative abundance of 15 and 10 compounds in the roots and exudates (blue features in the green square), respectively was clearly reduced in the *ugt79B9* knockout lines - by at least a factor of 8, and in some cases to the limit of detection (approximately 100-fold decrease) (**Figure 14 A and B**). The m/z features in the red square represent those that disappear in the *ugt79B9* line. Remarkably, among the features that decrease significantly in the *ugt79B9* knockout lines, most of the compounds are postulated to be polyhydroxylated molecules in roots extracts and exudates (**Table 1 and 2**). When *UGT79B9* was overexpressed in the *ugt79b9* mutant background (**Figure 15 A and B**), all the features except one (in green) that disappeared in the *ugt79b9* line reappear in the roots extract and all the features reappear in the exudates (highlighted in the green square). This indicates a partial recovery of the metabolite profile associated with the UGT79B9 activity (**Figure 15 A and B**).

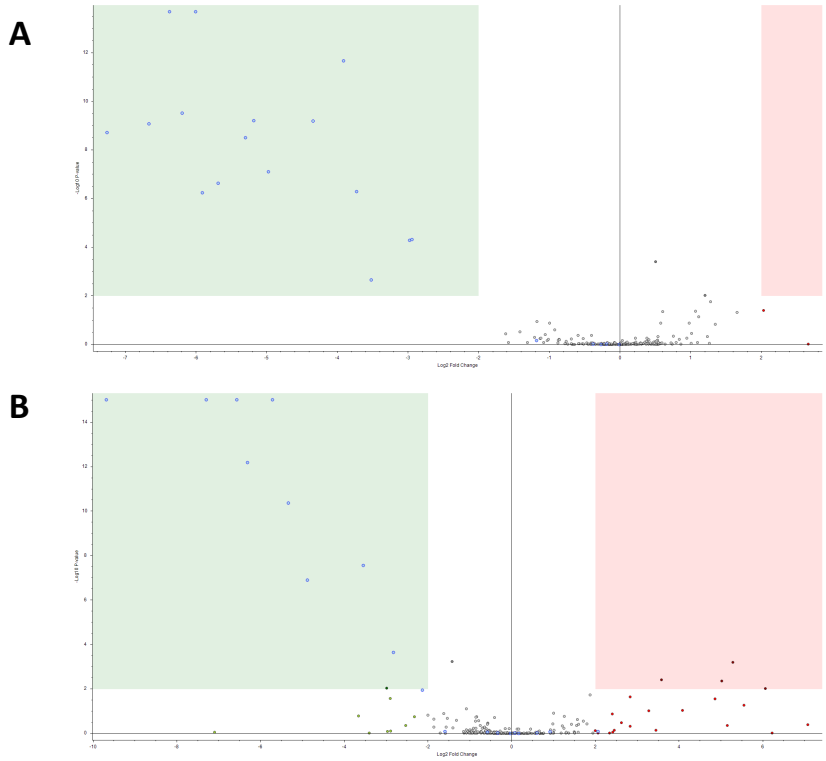


Figure 14. *ugt79b9 vs*. Col-0. Volcano plot for roots (A) and exudates (B). The blue circle represents m/z features. And those that disappear are included in the green square represent those that disappear in the *ugt79b9* mutant line.

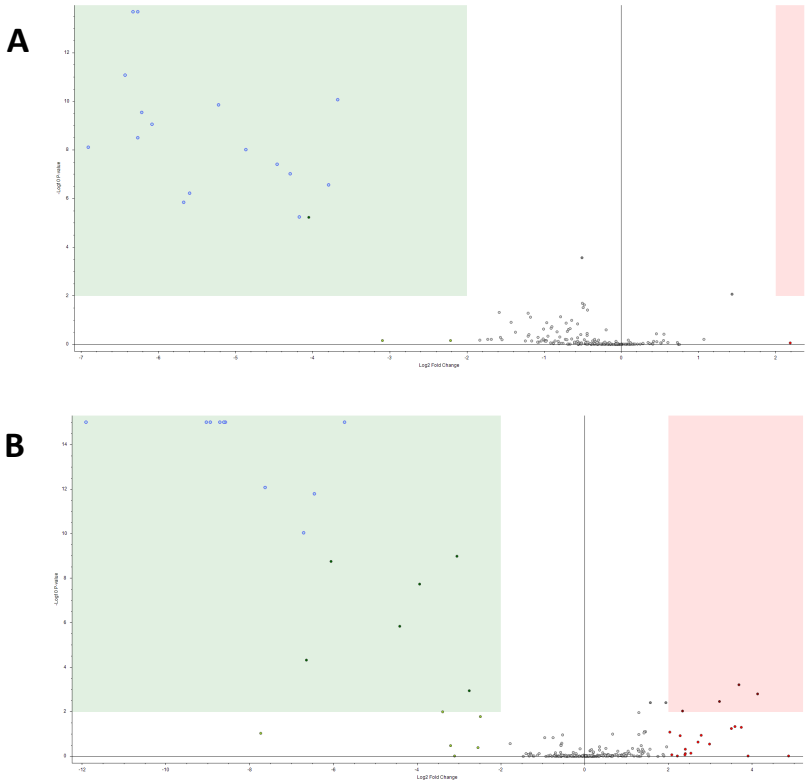


Figure 15. *ugt79b9* vs. 35S::*UGT79B9/ugt79b9* Volcano plot for roots (A) and exudates (B). The blue circle represents m/z features. And those that disappear are included in the green square represent those that disappear in the *ugt79b9* mutant line.

When comparing the feature content for the Col-0 line with the 35S::*UGT79B9*/Col-0 overexpression line in roots (**Figure 16A**), we did not observe a significant difference; however, we observed a trend towards accumulation of the same 15 peaks that disappear in the *ugt79b9* line (in blue) identified earlier, suggesting a dosage effect of *UGT79B9* overexpression. On the contrary, in exudates 8 features that disappear in the *ugt79b9* line reappear in the 35S::*UGT79B9* line (**Figure 16B**).

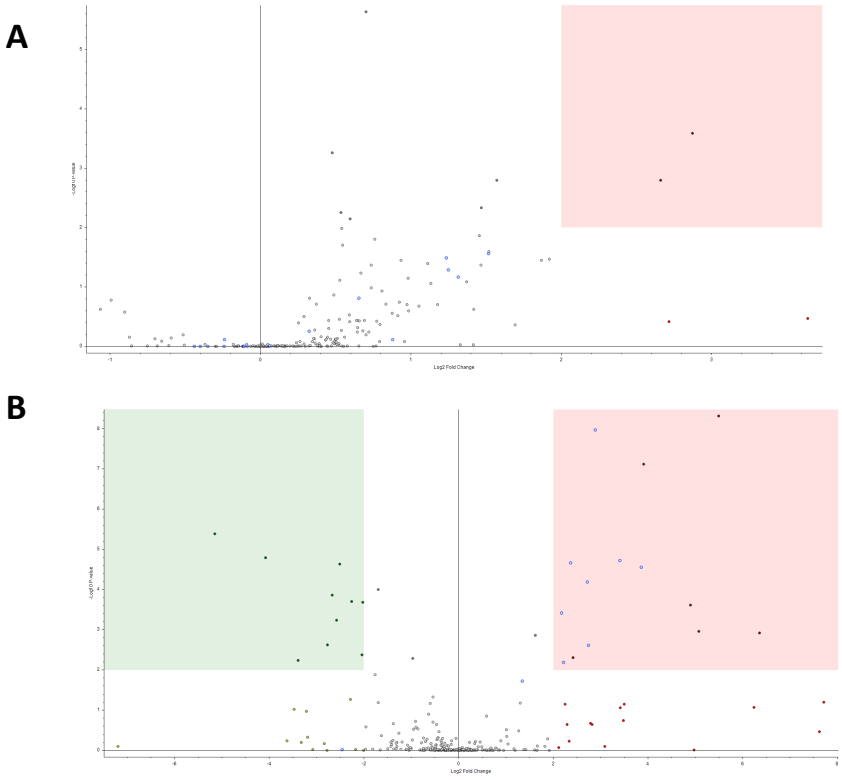


Figure 16. 35S::*UGT79B9*/Col-0 *vs*. Col-0. Volcano plot for roots (A) and exudates (B). The blue circle represents m/z features. And those that disappear are included in the green square represent those that disappear in the *ugt79b9* mutant line.

All the features described above that moves significantly between the lines, are summarized in the two tables attached below. Only 2 structures were proposed (LTS0110312 and CNP0383438) that correspond to an oxylipin derivative and a polycyclic aromatic compound respectively in root extracts. In exudates one structure was proposed (LTS0151534) that correspond to an oxylipin. These structural identifications were based on the Compound Discoverer analysis.

Table 1. List of roots compounds identified from plants cultivated *in vitro* that decrease significantly in the *ugt79B9* knockout line in comparison to Col-0 WT. Compound and molecular formula assignment was determined by Compound Discoverer software. Class compound assignment was determined from a SIRIUS analysis. RT: retention time. Peak: compound that move significantly (p value<0.01, log fold change>3). COCONUT (CNP) and LOTUS (LTS) databases were used for compound assignment. Red color indicates m/z features shared between root extracts and exudates.

Peak	RT (min)	Compound assignment	Molecular formula assignment	m/z [M-H]-
1	9.32		$C_{16}H_{30}O_{12}$	413.1668
2	9.32		$C_{15}H_{28}O_{10}$	367.1612
3	12.02	CNP0383438	$C_{25}H_{18}O_6$	413.1025
4	13.48		$C_{18}H_{32}O_{13}$	455.1773
5	13.71		$C_{17}H_{30}O_{11}$	409.1718
6	14.71		$C_{17}H_{30}O_{11}$	409.1718
7	14.71		$C_{18}H_{32}O_{13}$	455.1771
8	14.71		$C_{18}H_{30}O_{13}$	453.1618
9	15.27		$C_{18}H_{34}O_{12}S$	473.1700
10	17.28	LTS0110312	$C_9H_{16}O_5S$	235.0645
11	17.31		$C_{18}H_{34}O_{12}$	441.1981
12	17.46		$C_{22}H_{32}O_{13}$	503.1774
13	18.2		$C_{19}H_{34}O_{11}$	437.2032
14	18.61		$C_{19}H_{34}O_{11}$	437.2032
15	18.93		$C_{20}H_{36}O_{12}$	467.2137
16	19.00		$C_{19}H_{36}O_{12}$	455.2139
17	19.03		$C_{21}H_{38}O_{13}$	497.2245
18	19.57		$C_{22}H_{36}O_{13}$	507.2086
29	19.57		$C_{21}H_{36}O_{11}$	463.2189
20	19.70		$C_{20}H_{36}O_{11}$	451.2188
21	19.70		$C_{21}H_{36}O_{13}$	495.2088

Table 2. List of root exudates compounds identified from plants cultivated *in vitro* that decrease significantly in the *ugt79B9* knockout line in comparison to Col-0 WT. Compound and molecular formula assignment was determined by compound discoverer software. Class compound assignment was determined from a SIRIUS analysis. RT: retention time. Peak: compound that move significantly (p value<0.01, log fold change>3). COCONUT (CNP) and LOTUS (LTS) databases were used for compound assignment. Red color indicates m/z features shared between root extracts and exudates.

Peak	RT (min)	Compound assignment	Molecular formula assignment	m/z [M-H]-
1	9.21		$C_{15}H_{28}O_{10}$	367.1612
2	9.22		$C_{16}H_{30}O_{12}$	413.1667
3	10.61		$C_{16}H_{28}O_{11}$	395.1563
4	14.73		$C_{17}H_{30}O_{11}$	409.1719
5	14.73		$C_{18}H_{30}O_{13}$	453.1618
6	15.07		$C_{18}H_{32}O_{12}$	439.182
7	15.6		$C_{18}H_{32}O_{13}$	455.1776
8	17.26		$C_{21}H_{32}O_{12}$	475.1826
9	17.30		$C_{18}H_{34}O_{12}$	441.1982
10	18.92		$C_{20}H_{36}O_{12}$	467.2139
11	18.99		$C_{19}H_{36}O_{12}$	455.2140
12	19.55		$C_{16}H_{28}O_{9}$	363.1665
13	19.59		$C_{21}H_{36}O_{11}$	463.2192
14	19.62		$C_{15}H_{28}O_{9}$	351.1663
15	19.62		$C_{14}H_{26}O_{7}$	305.1607
16	19.72		$C_{20}H_{36}O_{11}$	451.2189
17	19.72		$C_{21}H_{36}O_{13}$	495.2089
18	21.02		$C_{21}H_{40}O_{12}$	483.2451
19	22.46	LTS0151534	$C_{18}H_{32}O_{5}$	327.2179
20	22.52		$C_{16}H_{28}O_{4}$	283.1916
21	22.66		$C_{10}H_{13}NO_{3}$	194.0822
22	22.73		$C_{14}H_{24}O_{3}$	239.1652

Interestingly, we observed that a large number of compounds were shared between root extracts and exudates (10 compounds, i.e., 50%). This result strongly suggests that UGT79B9 is involved in the biosynthesis of compounds produced in the roots and secreted into the external environment.

Metabolic network map using the available metabolomics data to visualize relationships among detected compounds

The Compound Discoverer software did not yield detailed structural or identification data for the previously detected molecules. Only preliminary structural information is currently available. This observation underscores the potential novelty of these compounds, which appear to be absent from existing chemical databases. We then constructed a metabolic network map using the available metabolomics data, with a particular focus on the MS^2 fragmentation patterns of each feature, which serve as molecular fingerprints. This allowed us to visualize potential structural relationships among the detected compounds. As a first step, a correlation heatmap was generated to identify compounds in root extracts and exudates that exhibited significant differences in accumulation between *in vitro* grown Arabidopsis genotypes under Fe-sufficient (+Fe) and Fe-deficient (–Fe) conditions (**Figure 17**).

The heatmaps of root extracts and exudates (**Figure 17A and B**) were organized along the horizontal axis according to genotype and Fe availability. Distinct sample groups emerged clearly. In both root extracts and exudates, mutation in *UGT79B9* had a marked effect on the selected metabolites, with a substantial reduction in their accumulation for most features, leading to a well-defined cluster linked with knockout mutation. For WT (Col-0), overexpressing (35S::*UGT79B9*/Col-0) and complementation (35S::*UGT79B9*/*ugt79B9*) lines, the patterns differed between roots and exudates. In root extracts, Fe deficiency was strongly associated with an increased accumulation of numerous features. In contrast, the exudate profiles were more complex: Col-0 samples clustered together with minimal influence from Fe status, while complementation and overexpressing samples grouped together, with Fe availability further driving the formation of sub-clusters.

Analysis of the 3 groups of features in the root extracts

For the root extracts heatmap, we could distinguish 3 groups of features. In the first group 2 metabolites were under-expressed in the presence of Fe and in the *ugt79b9* line samples (4, 7) most of which show little or no induction under Fe-deficient conditions. Within this group, a first subgroup (10) displays slight overproduction under Fe deficiency, with similar patterns in WT and *ugt79b9* lines, indicating that overexpression has a limited effect. A second group of 3 features (1, 5 and 8) showed a low intensity in *ugt79b9* lines samples and an over-accumulation in all the other samples growing under Fe-deficiency. This subgroup is characterized by a strong decrease in metabolite levels in the *ugt79b9* mutant, while both overexpression and Fe deficiency

lead to high accumulation—also observed in the Col-0—suggesting a synergistic effect. In the last group we also observed the under-expression of 9 features in the *ugt79b9* mutant in comparison with the other genotypes. Once again, in the Fe-deficient samples the accumulation of the features is the strongest.

In a third subgroup, metabolites show minimal accumulation under Fe-sufficient conditions. In the three groups, overexpression of *UGT79B9* in both Col-0 WT and *ugt79b9* mutant backgrounds leads to a restoration of metabolite accumulation.

Analysis of the 3 groups of features in the root exudates

From a general perspective, metabolite accumulation was higher in the overexpression and complementation lines. In more detail, the first group composed of 6 features (2, 5, 6, 9, 10, 11) exhibited low levels in *ugt79b9* mutant samples, with higher level in both overexpression and complementation lines. The second group, consisting of four metabolites (3, 7, 17, 18) showed very low level in both *ugt79b9* mutant and WT lines and no elevated levels in overexpression and complementation lines under Fe-deficient conditions. Finally, a third group of eight metabolites was present in the heatmap but displayed no clear pattern.

The next step involved metabolic network analysis, during which approximately 300 m/z features detected in ESI mode *via* untargeted metabolomics were used to construct the metabolic clustering presented below for root extracts and exudates (**Figure 18 A and B**).

In exudates, compounds No. 2 and 5 (**Figure 18B, Table 2**) were clustered with a known coumarin derivative, scopolin, and a glucosylated phenylpropane derivative, suggesting structural or biosynthetic relatedness. Compounds No. 6, 8, and 10 were identified as glycosylated molecules derived from flavonols, phenylpropanoids, and coumarins, respectively. In root extracts, a clustering of features 2, 5, 7, 11 (Table 1) was observed but no structural information could be attributed. All these features disappear in the *ugt79b9* mutant line.

Similar experiments (untargeted metabolic profiling using Thermo Scientific ORBITRAP followed by correlation heatmap and metabolic network analysis) were conducted with knockout (KO), wild-type (WT), and overexpression lines grown under PEG-induced osmotic stress. The data are not shown and are currently being processed.

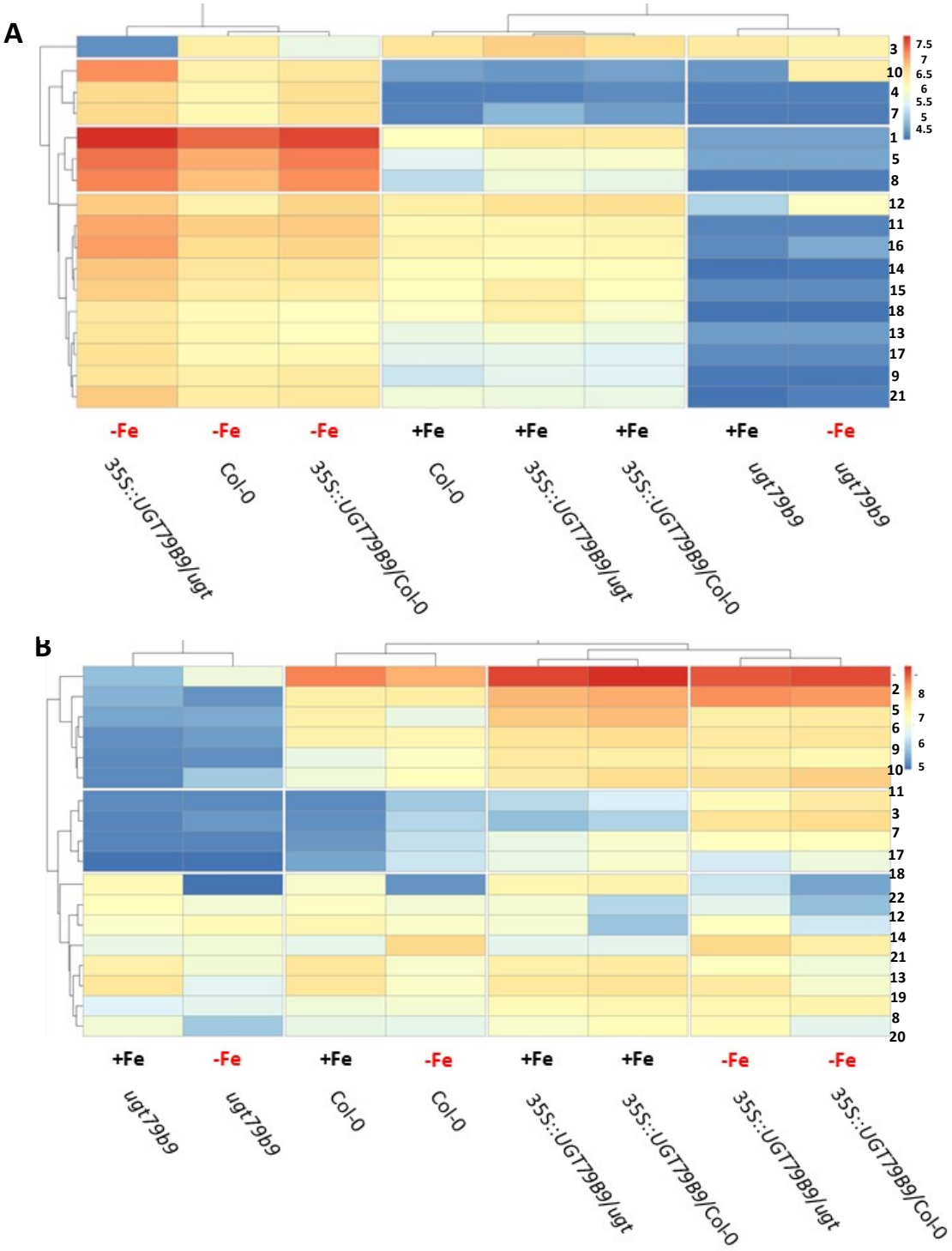


Figure 17. Hierarchically clustered heat map of selected features of Arabidopis roots **(A)** and exudates **(B)** differentially accumulated in response to Fe stress and plant genotype. To select differential levels, m/z area levels were examined using Student's *t*-test, and only those that were significantly different at the log2 fold change and $P<0.01$ level were selected. The cells color represents the changes of each metabolite relative to the mean control level. Red color represents up-accumulated and blue color down-accumulated m/z. The samples were analysed with Orbitrap and Compound Discoverer.

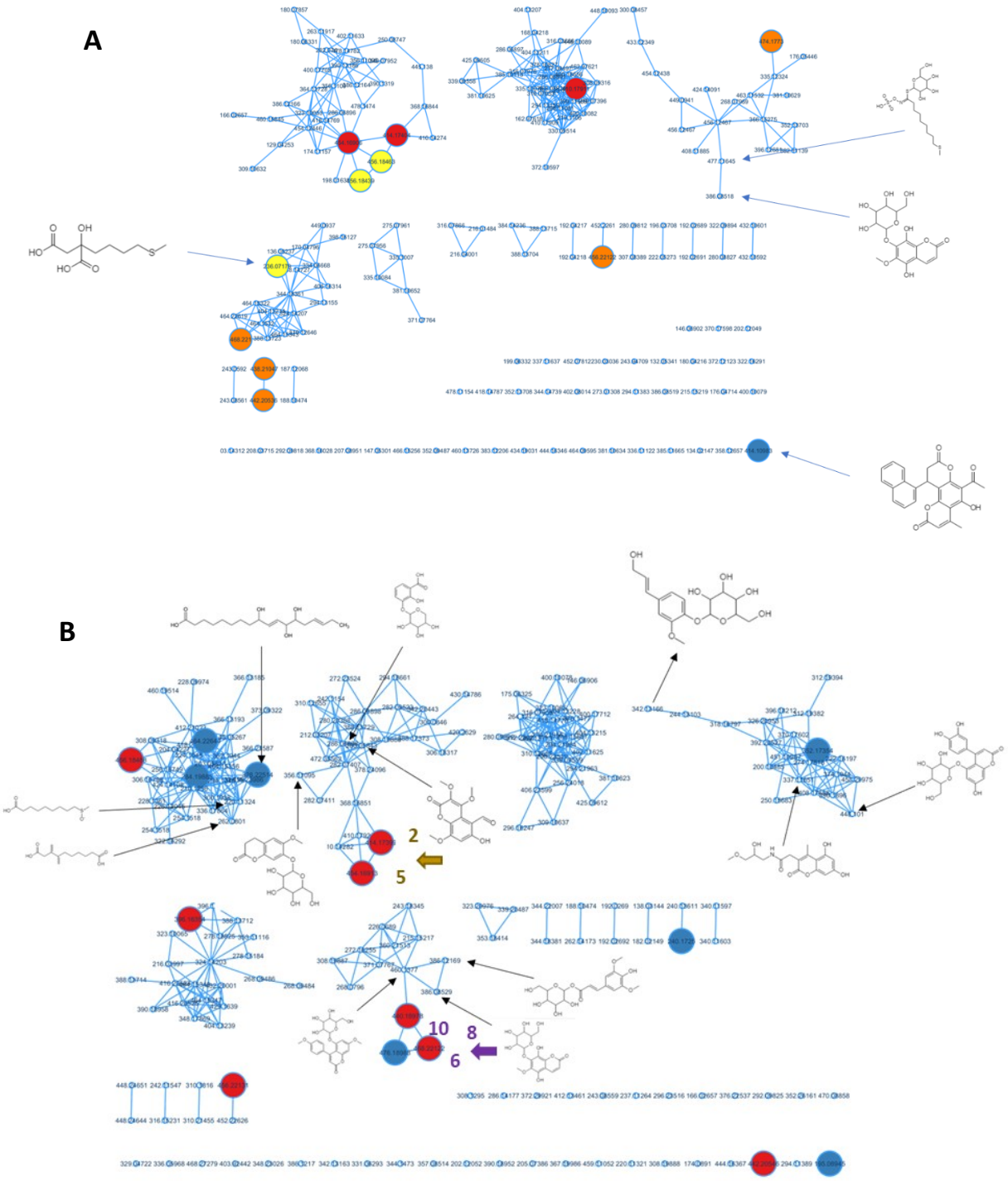


Figure 18. Metabolic network analysis: approximately 300 m/z collected in ESI-detection mode (untargeted metabolomics) were used to construct below metabolic clustering for root extracts **(A)** and exudates **(B).**

Drought experiment in soil

To investigate the functional role of UGT79B9 in drought stress responses, a soil-based experiment was conducted. Col-0, *ugt79b9*, overexpression lines 35S::*UGT799*/Col-0 and 35S::*UGT79B9*/*ugt79b9*, and Est-1 subjected to physiological drought conditions for 25 days (**Figure 19A)**. After that time plants were re-watered and the fresh and dry weight (FW and DW) was measured**.** Under control conditions, all genotypes displayed comparable growth and rosette biomass. Est-1 and Col-0 showed the most significant symptoms of drought in terms of FW when compared to well-watered control plants (**Figure 19B**). 35S::*UGT799*/Col-0 overexpression line had significantly higher FW than its genetic background Col-0. Although not significant, we also observed slightly higher FW in 35S::*UGT79B9*/*ugt79b9* overexpressing line when compared to *ugt79b9*. These differences were less pronounced in terms of DW (**Figure 19C**).

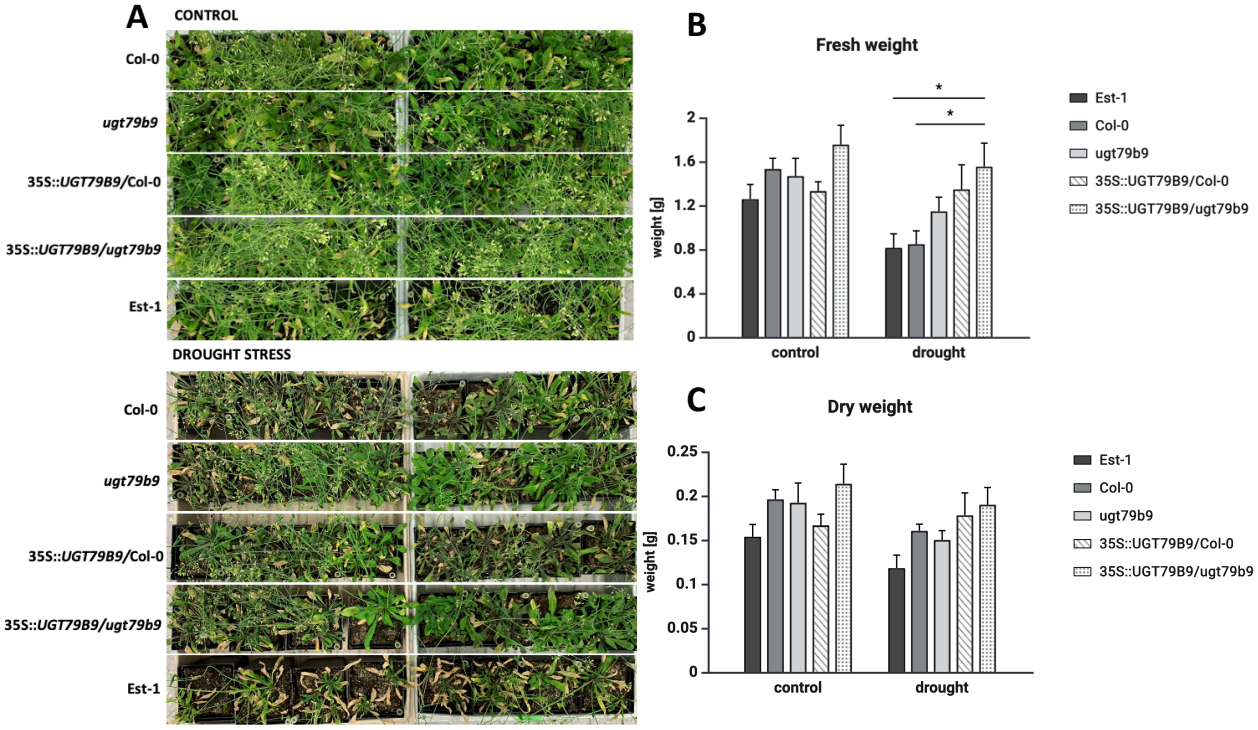


Figure 19. Impact of drought stress on Est-1, Col-0, *ugt79b9,* 35S::UGT799/Col-0 and 35S::UGT79B9/ugt79b9 grown in soil. **(A)** Representative images of tested genotypes grown under well-watered (control) and drought-stressed conditions. Plants were grown for 22 days under control conditions, followed by a 25-day drought period and re-watering phase. **(B)** Fresh weight (FW) and **(C)** dry weight (DW) of shoots were measured at the end of the recovery phase. Data is presented as mean ±SEM. **P*<0.05 (two-way ANOVA with Bonferroni multiple comparisons test, n=9).

To further evaluate the physiological response of UGT79B9-related genotypes to drought stress, we assessed membrane stability and tissue hydration by measuring electrolyte leakage (EL) and water content (WC) in rosette leaves. As shown in **Figure 20A**, under control conditions, EL was significantly

lower in Est-1 compared to Col-0 and mutant lines (*ugt79b9*, 35S::UGT79B9/Col-0, and 35S::UGT79B9/*ugt79b9*). This can be explained by the fact that Est-1 plants had a substantial amount of necrotic tissue after drought stress, which did not contribute to EL. Upon drought treatment, EL slightly increased in Col-0, *ugt79b9* and 35S::UGT79B9/*ugt79b9* in contrary to 35S::UGT79B9/Col-0 where EL decreased, however these differences were not statistically significant.

In parallel, WC measurements (**Figure 20B**) revealed that all genotypes maintained similar hydration levels under control conditions. However, under drought, Col-0 exhibited a significant reduction in WC, while both *UGT79B9* overexpression lines retained higher WC, comparable to the *ugt79b9*. These results imply a potential role of UGT79B9 in mitigating drought-induced cellular damage and maintaining tissue hydration.

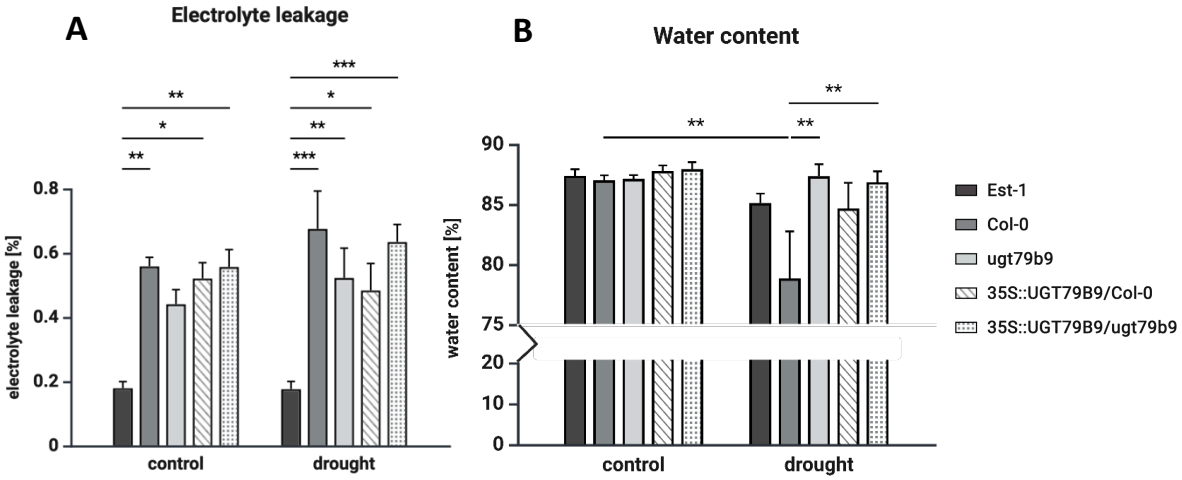


Figure 20. Electrolyte leakage and water content under drought stress conditions. **(A)** Electrolyte leakage in control and drought-treated rosettes. **(B)** Water content was calculated from fresh and dry weights using the formula: WC [%] = (FW − DW) / FW × 100. Asterisks indicate statistically significant differences between genotypes or treatments (two-way ANOVA with Bonferroni post hoc test; **P*<0.05, ***P*<0.01, ****P*<0.001). Bars represent mean values ± SEM (n=6–7 biological replicates).

DISCUSSION

We previously demonstrated that coumarins are widely distributed across various Arabidopsis natural populations and identified QTL regions underlying variation in scopoletin and scopolin content among seven accessions (Siwinska *et al.*, 2014). We selected several candidate genes that may influence this phenomenon. One gene of particular interest was *UGT79B9*, a member of the UDP-glucosyltransferase (UGT) family with an unknown biological function that is relatively highly expressed in roots.

Here, we conducted a comparative biochemical analysis using the *UGT79B9* coding sequences from the Arabidopsis Col-0 and Est-1 accessions, which differ in scopoletin and scopolin content and served as parental lines in the QTL mapping population. We transiently overexpressed *UGT79B9* in *N. benthamiana* and performed metabolic profiling using UPLC-MS. This analysis revealed

a significantly increased level of glycosylated scopoletin (scopolin) in the transformed leaves. Next, we conducted experiments using Arabidopsis UGT knock-out T-DNA insertional mutant line (*ugt79b9*) in the Col-0 genetic background, as well as lines overexpressing *UGT79B9*. Metabolic profiling of the *ugt79b9* mutant line revealed an increased level of scopoletin and a disrupted scopolin-to-scopoletin ratio compared to wild-type (WT) plants. *In silico* characterization of the *UGT79B9* gene suggested its potential involvement in osmotic stress response. To investigate this further, we cultivated *ugt79b9* mutant and WT lines under osmotic/drought stress conditions to compare their phenotypic appearance and overall plant performance. We found that plants with higher scopoletin content exhibited a more extensive root system and were better able to cope with osmotic stress caused by various factors. In addition to these *in vivo* experiments, we overproduced *UGT79B9* sequences from Col-0 and Est-1 *in vitro* and performed enzymatic assays.

The results strongly suggest that the UGT79B9 enzyme plays a key role in the plant's response to both osmotic stress and Fe deficiency, functioning as a metabolic switch. The regulatory role of scopoletin may have important physiological consequences, as unmodified scopoletin has been linked to enhanced stress defense but also cytotoxicity when excessively accumulated (Stringlis *et al*., 2019). Therefore, mutations in the gene encoding UGT79B9 may impact the accumulation of various secondary metabolites from different chemical groups, including coumarins.

The link between coumarins and plant tolerance to drought stress, as recently reported by other researchers (Zandalinas et al. 2017, Doll et al. 2018), along with the observed correlation between coumarin content in root exudates under Fe deficiency, aligns with our findings. This supports the hypothesis that UGT79B9 may play an important, albeit not yet fully understood, role in plant protection against stress factors. This is especially important due to the recent reports indicating the interconnection of Fe and osmotic stress signalling in plants (Kanwar *et al*., 2021; Quagliata *et al*., 2025). Such effects may be linked to the dual role of coumarins in redox homeostasis and modulation of the root microbiome, which contributes to plant fitness under water-limiting conditions (Voges *et al.,* 2019; Harbort *et al.,* 2020).

Our findings are important due to the fact that so far no biological function for Arabidopsis UGT79B9 has been described. Its potential role in Fe-deficiency response probably *via* coumarin biosynthesis and its involvement in plant responses to osmotic stress, including stress caused by water shortage, seems to be of particular interest at a time of climate change and the increasing occurrence of drought and nutrients-poor soils in the natural environments. Further characterization of UGT79B9, related glycosyltransferases and their potential substrates may contribute to the development of more resilient crops through targeted metabolic engineering.

MATERIALS AND METHODS

Plant material

The following *Arabidopsis thaliana* lines were used in this study: Columbia-0 (Col-0) as the wild-type (WT) reference, a T-DNA insertional knockout mutant of *UGT79B9* (GABI_048G01; hereafter referred to as *ugt79b9*) in the Col-0 background, and two transgenic overexpression lines: 35S::*UGT79B9*/ugt79b9 (complementation line) and 35S::*UGT79B9*/Col-0 (overexpression in WT background). Seeds of the *ugt79b9* mutant are publicly available from the Nottingham Arabidopsis Stock Centre (NASC; http://arabidopsis.info/).

Genotyping of the *ugt79b9* mutant and confirmation of transgene integration in the overexpression lines were performed using tiOptiTaq DNA Polymerase (EURx, Poland). Gene-specific primers for *UGT79B9* were 5′-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGCCAAAATTTTCAC-3′ and 5′-GGGGACCACTTTGTACAAGAAAGCTGGGTTTCATTCAAGATTTGTATCG-3′. As a housekeeping control, *ACTIN2* was amplified using primers 5′-TCCCAGTGTTGTTGGTAGGC-3′ and 5′-CAAGACGGAGGATGGCATGA-3′. The absence of *UGT79B9* transcript in the *ugt79b9* mutant was validated by RT-PCR using the same gene-specific primers (Supplementary Figure 1).

Nicotiana benthamiana seeds, used for transient expression assays, were kindly provided by Dr. Etienne Herbach (INRA, Colmar, France).

Growth conditions and treatments

Hydroponic and in vitro plate and liquid cultures

Arabidopsis plants under control conditions were grown in hydroponic cultures (10x Heeg hydroponic solutions) and *in vitro* cultures (0.5 MS) as described by Siwinska *et al.* (2018). Similarly, iron deficiency conditions (0 µM Fe2+) were obtained similarly to Siwinska *et al.* (2018), while osmotic stress conditions were obtained by transferring plants on day 10 to medium supplemented with 3% (w/v) PEG 8000.

Soil cultures

Arabidopsis seed were stratified prior to sowing for 4 days in 4°C. Plants were grown in 3:1 (v/v) mixture of a peat-free commercial substrate (COMPO SANA®, COMPO, Germany) and 3–6 mm vermiculite granules. Pots were placed in trays and maintained in growth chambers under a 16 h light/8 h dark cycle (120 µmol m^{-2} s^{-1}), at 22°C/20°C day/night temperatures.

Growth conditions of *N. benthamiana* plants

Two weeks after germination, *N. benthamiana* seeds were transplanted and grown independently for another 3 weeks in plant growth chambers with a photoperiod of 16 h of light (120 µmol m^{-2} s^{-1}) at 24°C and 8 h of darkness at 22°C and 70% humidity.

Drought experiment in soil

For drought experiments, each pot was adjusted to a uniform pre-sowing mass (120 ± 0.5 g) to standardize initial water content across replicates before sowing seeds. Drought stress was applied in the third week of growth by withholding irrigation for 25 days, followed by rehydration with 3 L of water per tray. Fresh weight (FW) of aerial tissues was measured three days post-rehydration; dry weight (DW) was recorded after drying at 85°C for 18h until constant mass (n=9). Water content (WC) was calculated as WC [%] = [(FW–DW)/FW] × 100. Membrane stability was assessed *via* electrolyte leakage: leaves were incubated overnight in 20 mL Milli-Q water (RT, 114 rpm), and conductivity of 6 mL aliquots was measured (HANNA EDGE™, USA) and normalized to leaf area (n=6).

Plasmid construction, stable transformation and generation of overexpression lines

To generate overexpression lines, the full-length CDS of *UGT79B9* was amplified from Col-0 cDNA using gene-specific primers containing Gateway® *attB* recombination sites (*att*B1: 5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGCCAAAATTTTCAC-3' and *att*B2: 5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTTCATTCAAGATTTGTATCG-3'). The PCR product was purified and cloned into the pDONR™ vector *via* BP recombination using BP Clonase™ II (Thermo Fisher Scientific, USA), following the manufacturer's instructions. Positive clones were verified by colony PCR and Sanger sequencing. The resulting entry clone (pENTRY:: *UGT79B9*) was subsequently recombined with the pK2GW7.0 destination vector using LR Clonase™ II to obtain the binary vector pDEST::*UGT79B9*, which confers kanamycin resistance in plants and spectinomycin resistance in bacteria. Electrocompetent *Agrobacterium tumefaciens* GV3101 cells (rifampicin- and gentamycin-resistant) were transformed with the pDEST::*UGT79B9* construct *via* electroshock, and positive transformants were selected on LB agar containing rifampicin (10 mg/L) and spectinomycin (50 mg/L). Colony PCR was performed to confirm transformation. For stable plant transformation, the floral dip method was applied to Arabidopsis Col-0 and *ugt79b9* mutant lines. Bacterial cultures were grown overnight in LB medium supplemented with rifampicin and spectinomycin at 28°C with shaking (180 rpm). The next day, bacteria were pelleted, resuspended in infiltration medium (2.2 g/L MS salts, 50 g/L sucrose, 1 mL/L Gamborg's vitamins, 0.5 g/L MES buffer, pH 5.7, and 0.005% Silwet L-77), and used for dipping. Flowering plants were submerged in the infiltration medium for 15 seconds and transferred to humidity chambers for 24 h in darkness. Plants were then returned to standard growth conditions and grown until seed harvest. Seeds were collected and sown on selective medium containing kanamycin to identify transgenic progeny. Resistant seedlings were transferred to soil, and transgene integration was confirmed by PCR. Homozygous T3 an T4 lines were selected for downstream phenotypic and molecular analyses.

***In silico* analysis of the *UGT79B9* expression**

In silico analysis of *UGT79B9* was conducted using publicly available AtGenExpress datasets generated with the Affymetrix ATH1 GeneChip microarray platform. Data were retrieved from the Arabidopsis eFP Browser (http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi) to assess tissue-specific and stress-induced expression patterns. Expression levels were visualized across various time points and in response to abiotic stresses.

Gene expression analysis by qPCR

Total RNA was extracted from Arabidopsis Col-0 and Est-1 root s using the GeneMATRIX Universal RNA Purification Kit (EURx, Poland), including on-column DNase. First-strand cDNA was synthesized from 1 µg of RNA using oligo(dT) primers and smart First Strand cDNA Synthesis Kit (EURx). qPCR was performed using Bio-Rad CFX96™ Real-Time PCR Detection System (Bio-Rad, USA) and Luminaris HiGreen qPCR Master Mix (Thermo Scientific, USA) under the following conditions: 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 95°C for 10 s and 60°C for 5 s with fluorescence acquisition. A final cooling step at 40°C for 30 s was included, and melt curve analysis was conducted from 60°C to 97°C. All reactions were performed in technical triplicates for each biological replicate. Relative transcript levels were calculated using the $2^{-\Delta\Delta Ct}$ method. We used the following gene-specific primers: for the reference gene *ACTIN2* (*ACT2*) 5'-CTTGCACCAAGCAGCATGAA-3' and 5'-CCGATCCAGACACTGTACTTCCTT-3' (Czechowski *et al.*, 2005); for *UGT79B9* 5'-GGAGGTTTTGGTTAGTCCTGG-3' and 5'-CGTTGACTATATTCTGCAAAGTTTC-3'. Primer specificity for *UGT79B* was confirmed by the analysis of melting curves

Heterologous expression in *N. benthamiana* leaves

The protocol for heterologous expression in *N. benthamiana* leaves and followed targeted metabolite profiling by UHPLC-MS was described in Siwinska et al. (2018).

Preparation of methanol extracts and metabolite profiling

Methanol extracts preparation from Arabidopsis roots, as well as Extraction of root exudates from nutrient solutions, and following targeted metabolic profiling (UHPLC-MS) was according to Siwinska et al. (2018).

Untargeted metabolic profiling (Thermo Scientific ORBITRAP)

Metabolite profiling was carried out using a Vanquish UHPLC system (ThermoFisher Scientific, USA) equipped with a quaternary pump, autosampler, and column oven. Chromatographic separation was achieved on a Kinetex XB-C18 core-shell column (150 × 2.1 mm, 2.6 µm; Phenomenex Inc., Torrance, CA, USA) maintained at 40 °C. The mobile phases consisted of water with 0.1% formic acid (A) and methanol with 0.1% formic acid (B), applied at a flow rate of 200 µL min^{-1}. The gradient elution started with 10% B for 2 minutes, ramped linearly to 30% B over 10 minutes, then to 95% B

over another 10 minutes. The column was held at 95% B for 5 minutes, followed by a 4-minute re-equilibration at initial conditions. Sample injections were randomized to avoid systematic bias.

High-resolution mass spectrometric detection was performed using an Orbitrap ID-X system (ThermoFisher Scientific, Bremen, Germany), operated in negative electrospray ionization (ESI) modes. The capillary voltage was set at 2.5 kV for negative mode. Sheath, auxiliary, and sweep gas flow rates were adjusted to 40, 8, and 1 arbitrary units, respectively, with the vaporizer temperature set at 320 °C. Full-scan MS data were acquired over an m/z range of 120–1200 at a resolving power of 60,000 (at m/z 200).

For compound characterization, MS/MS spectra were obtained from a pooled quality control (QC) sample using the AcquireX data-dependent acquisition workflow. Fragmentation was achieved through higher-energy collisional dissociation (HCD), and MS/MS spectra were recorded at a resolution of 15,000.

Correlation Heatmap and Metabolic Network analysis

To investigate metabolic co-regulation and compound associations in response to stress and genotypic variation, a heatmap and molecular network analysis were conducted on features derived from untargeted UHPLC-HRMS profiling. Differentially accumulated features between control and stress conditions across genotypes were first identified using Compound Discoverer 3.3 (Thermo Fisher Scientific, USA), applying standard workflows for peak detection, alignment, and normalization.

Correlation matrices were visualized using R (v4.2.1). For molecular network analysis, the processed feature table (m/z, retention time, area) was used to construct molecular networks based on mass spectral similarity. This was performed following the workflow described in Roumani *et al*. (2022).

***In vitro* enzymatic analyses**

The full-length coding sequence of Arabidopsis UGT79B9 was amplified from Col-0 and Est-1 cDNA and cloned into the pCR8®/GW/TOPO® entry vector (Invitrogen, USA). Amplification was performed using EmeraldAmp® GT PCR Master Mix (Takara, Japan) and gene-specific primers (with 5'-ATGGGCCAAAATTTTCACGCT-3' and 5'-TCATTCAAGATTTGTATCGTTGACT-3'). For bacterial expression, the ORF was subcloned into the pET28a(+) expression vector (Thermo Fisher Scientific, USA) *via Bam*HI and *Hind*III digestion or alternatively into the pCOLD™ vector system (Clontech, Japan) for cold-shock induced protein overexpression (Supplementary Figure 2). Constructs were verified by sequencing and introduced into *E. coli* Rosetta 2 competent cells by electroporation. Colonies were selected on LB agar supplemented with kanamycin (pET28a(+)) or ampicillin (pCOLD™) and chloramphenicol.

Starter cultures were expanded and grown at 37°C until OD_{600} reached 0.5. Expression was induced using 1 mM IPTG, either at 37°C overnight (pET system) and under cold-shock (4°C)

according to the manufacturer's protocol (pCOLD system). After induction, bacterial cells were harvested, resuspended in Tris-HCl (pH 8.0) lysis buffer with lysozyme and 0.1M NaCl, and disrupted by sonication. The resulting lysates were centrifuged, and insoluble proteins in the pellet were solubilized in 8 M urea.

Denatured proteins were refolded by 2 step dialysis with decreasing urea concentrations and 1 mM DTT at 4 °C (Olry *et al.*, 2005). Soluble fractions were concentrated using Amicon® Ultra filters (10 kDa cutoff; Merck), and His-tagged UGT79B9 proteins were partially purified by Ni-NTA affinity chromatography. *In vitro* activity assays were performed according to Matros *et al.* (2004).

Statistical analyses

Statistical analyses were performed in MS Excel, R programming or BioRender (https://www.biorender.com/).

FUNDING

This work was supported by the Polish National Science Centre grant UMO-2022/47/B/NZ2/01835 (AI) and the Polish National Agency for Academic Exchange grant BPN/BFR/2022/1/00039 (AI).

ACKNOWLEDGMENTS

We would like to Charles Clement for his valuable technical assistance.

An artificial intelligence (AI) tool (ChatGPT) was used solely as a tool to improve the style and grammar of some text parts previously written by non-English authors themselves (analogous to and an extension of word-processing and grammar-checking software).

AUTHOR CONTRIBUTIONS

I.P. designed the research, performed the experiments, analyzed data, contributed to the manuscript writing; J.S. performed the research, analyzed data.; A.D. analyzed data; F.M. analyzed data; J.G. performed the experiments; A.H. critically revised manuscript; A.O. designed and performed the experiments, analyzed data; E.L. supervised the research; A.I. designed the research, performed and supervised the research, analyzed data, wrote the manuscript, received funding for the project.

Conflict of interest statement: The authors declare that they have no conflict of interests.

REFERENCES

Gachon CM, Langlois-Meurinne M, Saindrenan P. Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends Plant Sci. 2005 Nov;10(11):542-9.

Fourcroy P, Sisó-Terraza P, Sudre D, Savirón M, Reyt G, Gaymard F, Briat JF. Involvement of the ABCG 37 transporter in secretion of scopoletin and derivatives by Arabidopsis roots in response to iron deficiency. New Phytol. 2014;201(1):155–67.

Harbort CJ, Hashimoto M, Inoue H, Niu Y, Guan R, Rombolà AD, Schulze-Lefert P. Root-secreted coumarins and the microbiota interact to improve iron nutrition in Arabidopsis. Cell Host Microbe. 2020;28(6):825–37.

Kai K, Mizutani M, Kawamura N, Yamamoto R, Tamai M, Yamaguchi H, Shimizu BI. Scopoletin is biosynthesized via Ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in Arabidopsis thaliana. Plant J. 2008;55(6):989–99.

Kai K, Shimizu BI, Mizutani M, Watanabe K, Sakata K. Accumulation of coumarins in Arabidopsis thaliana. Phytochemistry. 2006;67(4):379–86.

Kanwar P, Baby D, Bauer P. Interconnection of iron and osmotic stress signalling in plants: is FIT a regulatory hub to cross-connect abscisic acid responses? Plant Biol (Stuttg). 2021 May;23 Suppl 1:31-38.

Matros, A., & Mock, H. P. (2004). Ectopic expression of a UDP-glucose: phenylpropanoid glucosyltransferase leads to increased resistance of transgenic tobacco plants against infection with Potato Virus Y. Plant and cell physiology, 45(9), 1185-1193.

Mizutani M, Ohta D. Diversification of P450 genes during land plant evolution. Annu Rev Plant Biol. 2010;61:291-315.

Perkowska I, Potrykus M, Siwinska J, Siudem D, Lojkowska E, Ihnatowicz A. Interplay between coumarin accumulation, iron deficiency and plant resistance to Dickeya spp. Int J Mol Sci. 2021b;22(12):6449.

Perkowska I, Siwinska J, Olry A, Grosjean J, Hehn A, Bourgaud F, Ihnatowicz A. Identification and quantification of coumarins by UHPLC-MS in Arabidopsis thaliana natural populations. Molecules. 2021a;26(6):1804.

Olry, A., Boschi-Muller, S., Yu, H., Burnel, D., & Branlant, G. (2005). Insights into the role of the metal binding site in methionine-R-sulfoxide reductases B. Protein science, 14(11), 2828-2837.

Quagliata G, Molina MDG, Mannino G, Coppa E, Saidi MN, Palombieri S, Sestili F, Vigani G, Astolfi S. Drought affects Fe deficiency-induced responses in a purple durum wheat (Triticum turgidum subsp. durum) genotype. Plant Biol (Stuttg). 2025 Mar 17.

Rajniak J, Giehl RF, Chang E, Murgia I, von Wirén N, Sattely ES. Biosynthesis of redox-active metabolites in response to iron deficiency in plants. Nat Chem Biol. 2018;14(5):442–50. 10.1038/s41589-018-0019-2.

Rehman, H.M., Nawaz, M.A., Shah, Z.H. et al. Comparative genomic and transcriptomic analyses of Family-1 UDP glycosyltransferase in three Brassica species and Arabidopsis indicates stress-responsive regulation. Sci Rep 8, 1875 (2018).

Rohde A, Morreel K, Ralph J, Goeminne G, Hostyn V, De Rycke R, Boerjan W. Molecular phenotyping of the pal1 and pal2 mutants of Arabidopsis thaliana reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. Plant Cell. 2004;16(10):2749–71.

Roumani, M., Le Bot, J., Boisbrun, M., Magot, F., Péré, A., Robin, C., ... & Larbat, R. (2022). Transcriptomics and metabolomics analyses reveal high induction of the phenolamide pathway in tomato plants attacked by the leafminer Tuta absoluta. Metabolites, 12(6), 484.

Schmid NB, Giehl RF, Döll S, Mock HP, Strehmel N, Scheel D, von Wirén N. Feruloyl-CoA 6′-Hydroxylase1-dependent coumarins mediate iron acquisition from alkaline substrates in Arabidopsis. Plant Physiol. 2014;164(1):160–72.

Schmidt H, Günther C, Weber M, Spörlein C, Loscher S, Böttcher C, Clemens S. Metabolome analysis of Arabidopsis thaliana roots identifies a key metabolic pathway for iron acquisition. PLoS ONE. 2014;9(7):e102444. 10.1371/journal.pone.0102444.

Shimizu B. 2-Oxoglutarate-dependent dioxygenases in the biosynthesis of simple coumarins. Front Plant Sci. 2014 Nov 3;5:549.

Sisó-Terraza P, Luis-Villarroya A, Fourcroy P, Briat JF, Abadía A, Gaymard F, Abadía J, Álvarez-Fernández A. Accumulation and Secretion of coumarinolignans and other coumarins in Arabidopsis thaliana roots in response to Iron Deficiency at High pH. Front Plant Sci. 2016;7.

Siwinska J, Kadzinski L, Banasiuk R, Gwizdek-Wisniewska A, Olry A, Banecki B, Ihnatowicz A. Identification of QTLs affecting scopolin and scopoletin biosynthesis in Arabidopsis thaliana. BMC Plant Biol. 2014;14:1–14.

Siwinska J, Siatkowska K, Olry A, Grosjean J, Hehn A, Bourgaud F, Ihnatowicz A. Scopoletin 8-hydroxylase: a novel enzyme involved in coumarin biosynthesis and iron-deficiency responses in Arabidopsis. J Exp Bot. 2018;69(7):1735–48.

Stassen MJ, Hsu SH, Pieterse CM, Stringlis IA. Coumarin communication along the microbiome–root–shoot axis. Trends Plant Sci. 2021;26(2):169–83.

Stringlis IA, De Jonge R, Pieterse CM. The age of coumarins in plant–microbe interactions. Plant Cell Physiol. 2019;60(7):1405–19.

Stringlis IA, Yu K, Feussner K, de Jonge R, Van Bentum S, Van Verk MC, Pieterse CM. MYB72-dependent coumarin exudation shapes root microbiome assembly to promote plant health. Proc Natl Acad Sci. 2018;115(22):E5213–22.

Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22(22):4673–80.

Tsai HH, Rodríguez-Celma J, Lan P, Wu YC, Vélez-Bermúdez IC, Schmidt W. Scopoletin 8-hydroxylase-mediated fraxetin production is crucial for iron mobilization. Plant Physiol. 2018;177(1):194–207. 10.1104/pp.18.00178.

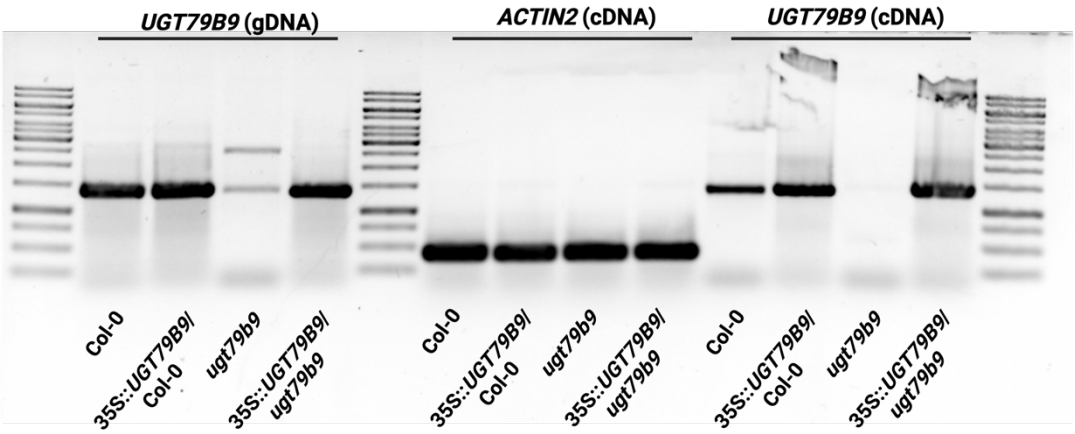
Voges MJ, Bai Y, Schulze-Lefert P, Sattely ES. (2019). Plant-derived coumarins shape the composition of an Arabidopsis synthetic root microbiome. Proc Natl Acad Sci USA. 2019;116(25):12558-65.

Wang M, Ji Q, Lai B, Liu Y, Mei K. Structure-function and engineering of plant UDP-glycosyltransferase. Comput Struct Biotechnol J. 2023 Oct 27;21:5358-5371.

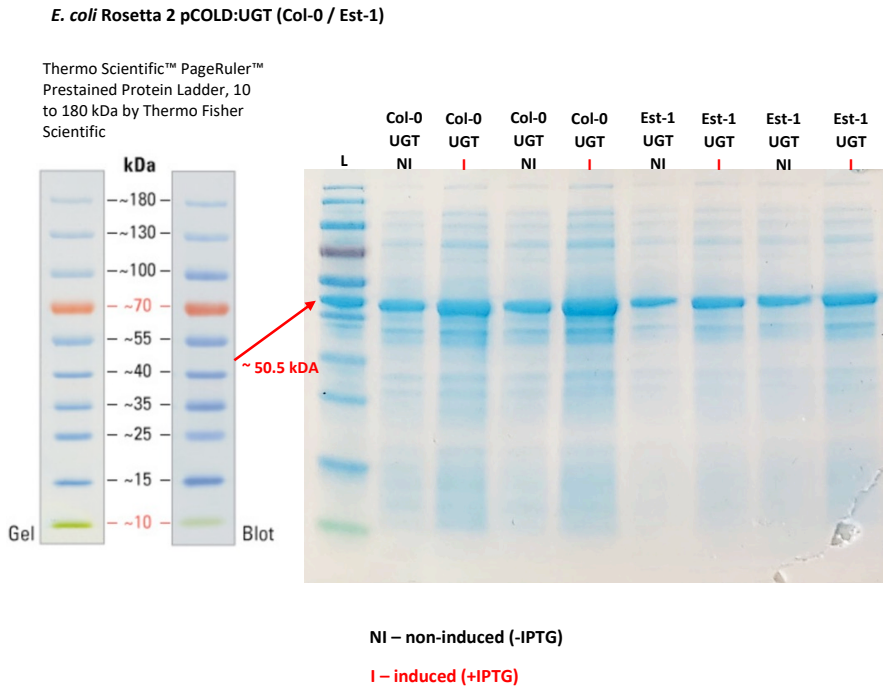
Werck-Reichhart D, Feyereisen R. Cytochromes P450: a success story. Genome Biol. 2000 ;1(6).

Zandalinas SI, Sales C, Beltrán J, Gómez-Cadenas A, Arbona V. Activation of Secondary Metabolism in Citrus Plants Is Associated to Sensitivity to Combined Drought and High Temperatures. Front Plant Sci. 2017 Jan 9;7:1954. doi: 10.3389/fpls.2016.01954. PMID: 28119698; PMCID: PMC5220112.

SUPPLEMENTRY DATA



Supplementary Figure S1. Genotyping and expression verification of Arabidopsis *UGT79B9* transgenic lines. PCR amplification was used to confirm the presence or absence of the *UGT79B9* gene in genomic DNA (gDNA), and to assess transcript levels in complementary DNA (cDNA) from Col-0 wild-type, 35S::*UGT79B9* overexpression lines, *ugt79b9* knockout mutants, and complemented lines (ugt79b9/35S::*UGT79B9*).



Supplementary Figure 2. SDS-PAGE analysis of recombinant UGT79B9 protein expression in *E. coli* Rosetta 2 cells using the pCOLD vector system. Protein extracts from bacterial cultures transformed with UGT79B9 constructs from Col-0 and Est-1 Arabidopsis accessions were analyzed under non-induced (NI; −IPTG) and induced (I; +IPTG) conditions. A distinct band corresponding to the predicted size of recombinant UGT79B9 (~50.5 kDa) was observed in IPTG-induced samples from both accessions, confirming successful overexpression of the target protein. L – molecular weight ladder (Thermo Scientific™ PageRuler™, 10–180 kDa).

Gdańsk, 16.05.2025

MSc Izabela Perkowska
Laboratory of Plant Protection and Biotechnology,
Intercollegiate Faculty of Biotechnology,
University of Gdansk and Medical University of Gdansk

AUTHOR CONTRIBUTION STATEMENT

I hereby declare my contribution to the following manuscripts:

1) Perkowska, I., Siwinska, J., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Ihnatowicz, A. (2021). Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. *Molecules*, 26(6), 1804

   validation, investigation, writing – original draft preparation

   Perkowska

2) Perkowska, I., Potrykus, M., Siwinska, J., Siudem, D., Lojkowska, E., & Ihnatowicz, A. (2021b). Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *International Journal of Molecular Sciences*, 22(12), 6449

   validation, formal analysis, investigation

   Perkowska

3) Ihnatowicz, A., Siwinska, J., Perkowska, I., Grosjean, J., Hehn, A., Bourgaud, F., ... & Olry, A. (2024). Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of *Arabidopsis thaliana*. *BMC Plant Biology*, 24(1), 806

   data analysis, figures preparation, validation, investigation, writing – review and editing

   Perkowska

4) Perkowska, I., Olry, A., Barrit, T., Dobek, A., Magot, F., Charles, C., Januchowska, A., Siwinska, J., Lojkowska, E., & Ihnatowicz, A. (2025a) Elucidating the role and regulatory mechanism of *Arabidopsis thaliana DOXC21-A*, its NAT and *DOXC21-B* in plant adaptation to terrestrial conditions (*manuscript in preparation*)

   conceptualization, methodology, investigation, validation, data analysis, writing – original draft preparation, writing – review and editing

   Perkowska

Gdańsk, 16.05.2025

MSc Izabela Perkowska

Laboratory of Plant Protection and Biotechnology,

Intercollegiate Faculty of Biotechnology,

University of Gdansk and Medical University of Gdansk

AUTHOR CONTRIBUTION STATEMENT

I hereby declare my contribution to the following manuscript:

5) Perkowska, I., Siwinska, J., Dobek, A, Magot, F., Grosjean, J., Hehn, A., Olry, A., Lojkowska, E., & Ihnatowicz, A. (2025b) Arabidopsis UGT79B9-mediated crosstalk between secondary metabolism, root exudation and plant responses to osmotic stress and iron deficiency (*manuscript in preparation*)

conceptualization, methodology, investigation, data analysis, writing – original draft preparation, writing – review and editing

IPerkowska

Gdańsk, 16.05.2025

Prof. Ewa Lojkowska

Laboratory of Plant Protection and Biotechnology,

Intercollegiate Faculty of Biotechnology,

University of Gdansk and Medical University of Gdansk

AUTHOR CONTRIBUTION STATEMENT

I hereby declare my contribution to the following manuscripts:

1) Perkowska, I., Siwinska, J., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Ihnatowicz, A. (2021). Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. *Molecules*, 26(6), 1804

   conceptualization, writing – review and editing, supervision

   ..... Ewa Łojkowska .....

2) Perkowska, I., Potrykus, M., Siwinska, J., Siudem, D., Lojkowska, E., & Ihnatowicz, A. (2021b). Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *International Journal of Molecular Sciences*, 22(12), 6449

   conceptualization, writing – review and editing, supervision

   ..... E. Łojkowska .....

3) Ihnatowicz, A., Siwinska, J., Perkowska, I., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska E. & Olry, A. (2024). Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of *Arabidopsis thaliana*. *BMC Plant Biology*, 24(1), 806

   writing – review and editing, funding acquisition, supervision

   ..... E. Łojkowska .....

4) Perkowska, I., Olry, A., Barrit, T., Dobek, A., Magot, F., Charles, C., Januchowska, A., Siwinska, J., Lojkowska, E., & Ihnatowicz, A. (2025a) Elucidating the role and regulatory mechanism of *Arabidopsis thaliana DOXC21-A*, its NAT and *DOXC21-B* in plant adaptation to terrestrial conditions (*manuscript in preparation*)

   conceptualization, supervision, writing – review and editing, funding acquisition

   ..... E. Łojkowska .....

5) Perkowska, I., Siwinska, J., Dobek, A, Magot, F., Grosjean, J., Hehn, A., Olry, A., Lojkowska, E., & Ihnatowicz, A. (2025b) Arabidopsis UGT79B9-mediated crosstalk between secondary metabolism, root exudation and plant responses to osmotic stress and iron deficiency (*manuscript in preparation*)

   supervised the research

   ..... E. Łojkowska .....

Ewa Łojkowska

Gdańsk, 16.05.2025

Dr. Anna Ihnatowicz

Laboratory of Plant Protection and Biotechnology,

Intercollegiate Faculty of Biotechnology,

University of Gdansk and Medical University of Gdansk

AUTHOR CONTRIBUTION STATEMENT

I hereby declare my contribution to the following manuscripts:

1) Perkowska, I., Siwinska, J., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Ihnatowicz, A. (2021). Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. *Molecules*, 26(6), 1804

   conceptualization, methodology, writing – original draft preparation, writing – review and editing, supervision, funding acquisition

   .....

2) Perkowska, I., Potrykus, M., Siwinska, J., Siudem, D., Lojkowska, E., & Ihnatowicz, A. (2021b). Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *International Journal of Molecular Sciences*, 22(12), 6449

   conceptualization, methodology, writing – original draft preparation, writing – review and editing, supervision, project administration, funding acquisition

   .....

3) Ihnatowicz, A., Siwinska, J., Perkowska, I., Grosjean, J., Hehn, A., Bourgaud, F., ... & Olry, A. (2024). Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of *Arabidopsis thaliana*. *BMC Plant Biology*, 24(1), 806

   conceptualization, data analysis, methodology, writing – original draft preparation, writing – review and editing, supervision, funding acquisition

   .....

Gdańsk, 16.05.2025

Dr. Anna Ihnatowicz
Laboratory of Plant Protection and Biotechnology,
Intercollegiate Faculty of Biotechnology,
University of Gdansk and Medical University of Gdansk

AUTHOR CONTRIBUTION STATEMENT

I hereby declare my contribution to the following manuscripts:

4) Perkowska, I., Olry, A., Barrit, T., Dobek, A., Magot, F., Charles, C., Januchowska, A., Siwinska, J., Lojkowska, E., & Ihnatowicz, A. (2025a) Elucidating the role and regulatory mechanism of *Arabidopsis thaliana DOXC21-A*, its NAT and *DOXC21-B* in plant adaptation to terrestrial conditions (*manuscript in preparation*)

   conceptualization, methodology, supervision, data analysis, validation, writing – review and editing, funding acquisition

   Anna Ihnatowicz

5) Perkowska, I., Siwinska, J., Dobek, A, Magot, F., Grosjean, J., Hehn, A., Olry, A., Lojkowska, E., & Ihnatowicz, A. (2025b) Arabidopsis UGT79B9-mediated crosstalk between secondary metabolism, root exudation and plant responses to osmotic stress and iron deficiency (*manuscript in preparation*)

   conceptualization, methodology, data analysis, validation, supervision, writing – original draft preparation, writing – review and editing, funding acquisition

   Anna Ihnatowicz

Gdańsk, 16.05.2025

Dr. Joanna Siwinska

2010 – 2015 PhD student at Laboratory of Plant Protection and Biotechnology,

Intercollegiate Faculty of Biotechnology,

University of Gdansk and Medical University of Gdansk

AUTHOR CONTRIBUTION STATEMENT

I hereby declare my contribution to the following manuscripts:

1) Perkowska, I., Siwinska, J., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Ihnatowicz, A. (2021). Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. *Molecules*, 26(6), 1804

   validation, investigation

   Signature

2) Perkowska, I., Potrykus, M., Siwinska, J., Siudem, D., Lojkowska, E., & Ihnatowicz, A. (2021b). Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *International Journal of Molecular Sciences*, 22(12), 6449

   methodology, validation, investigation

   Signature

3) Ihnatowicz, A., Siwinska, J., Perkowska, I., Grosjean, J., Hehn, A., Bourgaud, F., ... & Olry, A. (2024). Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of *Arabidopsis thaliana*. *BMC Plant Biology*, 24(1), 806

   validation, investigation, data analysis, writing – original draft preparation

   Signature

4) Perkowska, I., Olry, A., Barrit, T., Dobek, A., Magot, F., Charles, C., Januchowska, A., Siwinska, J., Lojkowska, E., & Ihnatowicz, A. (2025a) Elucidating the role and regulatory mechanism of *Arabidopsis thaliana DOXC21-A*, its NAT and *DOXC21-B* in plant adaptation to terrestrial conditions (*manuscript in preparation*)

   investigation

   Signature

5) Perkowska, I., Siwinska, J., Dobek, A, Magot, F., Grosjean, J., Hehn, A., Olry, A., Lojkowska, E., & Ihnatowicz, A. (2025b) Arabidopsis UGT79B9-mediated crosstalk between secondary metabolism, root exudation and plant responses to osmotic stress and iron deficiency (*manuscript in preparation*)

   investigation, data analysis

   Signature

Gdańsk, 16.05.2025

Dr. Marta Potrykus
Laboratory of Biologically Active Compounds,
Intercollegiate Faculty of Biotechnology,
University of Gdansk and Medical University of Gdansk

AUTHOR CONTRIBUTION STATEMENT

I hereby declare my contribution to the following manuscript:

2) Perkowska, I., Potrykus, M., Siwinska, J., Siudem, D., Lojkowska, E., & Ihnatowicz, A. (2021b). Interplay between coumarin accumulation, iron deficiency and plant resistance to *Dickeya* spp. *International Journal of Molecular Science*s, 22(12), 6449

conceptualization, methodology, validation, investigation, writing – original draft preparation

Marta Potrykus

.....

Vandoeuvre-lès-Nancy, 16.05.2025

Dr. Eng. Alexandre Olry
ENSAIA
Laboratoire Agronomie et Environnement UMR1121
Université de Lorraine

AUTHOR CONTRIBUTION STATEMENT

I hereby declare my contribution to the following manuscripts:

1) Perkowska, I., Siwinska, J., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Ihnatowicz, A. (2021). Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. *Molecules*, 26(6), 1804

conceptualization, methodology, validation, investigation, writing – original draft preparation, writing – review and editing, funding acquisition

A.OLRY

3) Ihnatowicz, A., Siwinska, J., Perkowska, I., Grosjean, J., Hehn, A., Bourgaud, F., ... & Olry, A. (2024). Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of *Arabidopsis thaliana*. *BMC Plant Biology*, 24(1), 806

conceptualization, methodology, validation, investigation, data analysis, writing – original draft preparation, writing – review and editing, project administration

A.OLRY

4) Perkowska, I., Olry, A., Barrit, T., Dobek, A., Magot, F., Charles, C., Januchowska, A., Siwinska, J., Lojkowska, E., & Ihnatowicz, A. (2025a) Elucidating the role and regulatory mechanism of *Arabidopsis thaliana DOXC21-A*, its NAT and *DOXC21-B* in plant adaptation to terrestrial conditions (*manuscript in preparation*)

investigation, methodology, supervision, writing – review and editing

A.OLRY

5) Perkowska, I., Siwinska, J., Dobek, A, Magot, F., Grosjean, J., Hehn, A., Olry, A., Lojkowska, E., & Ihnatowicz, A. (2025b) Arabidopsis UGT79B9-mediated crosstalk between secondary metabolism, root exudation and plant responses to osmotic stress and iron deficiency (*manuscript in preparation*)

investigation, methodology, supervision, writing – review and editing

A.OLRY

Curriculum vitae

Personal data

First name: **Izabela**
Last name: **Perkowska**
Date of birth: 16.03.1992
E-mail address: izabela.perkowska@phdstud.ug.edu.pl / perkowskaizabela@gmail.com
ORCID: 0000-0001-9094-9217

Education:

- 2020 – until now **Intercollegiate Biotechnology Doctoral School of University of Gdańsk and Medical University of Gdańsk (IBDS UG & MUG), Poland** PhD candidate, Discipline: Biotechnology
- 2017 – 2020 **Intercollegiate Faculty of Biotechnology of University of Gdańsk and Medical University of Gdańsk (IFB UG & MUG), Poland** Life Sciences and Mathematics Interdisciplinary Doctoral Studies, PhD candidate, Discipline: Biological Sciences
- 2011 – 2017 **Intercollegiate Faculty of Biotechnology of University of Gdańsk and Medical University of Gdańsk (IFB UG & MUG), Poland** bachelor's and master's degree, Discipline: Biotechnology

Publications:

1. Ihnatowicz, A., Siwinska, J., **Perkowska, I**., Grosjean, J., Hehn, A., Bourgaud, F., ... & Olry, A. (2024). Genes to specialized metabolites: accumulation of scopoletin, umbelliferone and their glycosides in natural populations of Arabidopsis thaliana. BMC Plant Biology, 24(1), 806. DOI: 10.1186/s12870-024-05491-w
2. **Perkowska, I.**, Potrykus, M., Siwinska, J., Siudem, D., Lojkowska, E., & Ihnatowicz, A. (2021). Interplay between Coumarin Accumulation, Iron Deficiency and Plant Resistance to *Dickeya* spp. International Journal of Molecular Sciences, 22(12), 6449. DOI: 10.3390/ijms22126449
3. **Perkowska, I**., Siwinska, J., Olry, A., Grosjean, J., Hehn, A., Bourgaud, F., Lojkowska, E., & Ihnatowicz, A. (2021). Identification and Quantification of Coumarins by UHPLC-MS in *Arabidopsis thaliana* Natural Populations. Molecules, 26(6), 1804. DOI:10.3390/molecules26061804
4. Potrykus, M., Decorosi, F., **Perkowska, I.,** Viti, C., Mengoni, A., Hugouvieux-Cotte-Pattat, N., & Lojkowska, E. (2020). The metabolic shift in highly and weakly virulent *Dickeya solani* strains is more affected by temperature than by mutations in genes encoding global virulence regulators. *FEMS Microbiology Ecology*, *96*(3). DOI: 10.1093/femsec/fiaa023

Research projects executor:

1. **OPUS 24 (2022/47/B/NZ2/01835), NCN, PI: dr Anna Ihnatowicz** (01.10.2024 – 31.04.2024, 7 months)
   Department of Plant Protection and Biotechnology, IFB UG & MUG, Gdańsk, Poland
   "*The battle for iron between plants and pathogens: genetic and molecular basis of Arabidopsis responses to Fe-deficiency and Dickeya spp. infection associated with the accumulation of coumarins"*
   Tasks: Assistance in conducting experiments, analyzing and interpreting results
2. **The Smart Growth Operational Programme (POIR.01.01.01-00-0689/21), NCBiR, PI: 'NIVISS' Leszek Łosin Sp. J.** (29.03.2023-25.04.2023, <1 month)
   "An innovative family of niPlants line luminaires designed for growing plants with the option of active PPFD stabilization and automatic adjustment of the illuminated area to the size of plants"
   Task: Preparation of a report on plant breeding
3. **PRELUDIUM BIS 1 (2019/35/O/NZ1/02751), NCN, PI: Prof. Ewa Łojkowska** (01.10.2020 – 30.09.2024, 4 years) Department of Plant Protection and Biotechnology, IFB UG & MUG, Gdańsk, Poland
   "Elucidating the role and mechanism of regulatory network of genes encoding dioxygenases in terms of plant adaptation to land conditions"
   Task: Preparation of doctoral dissertation
4. **PHC Polonium (PPN/BFR/2019/1/00050), NAWA, PI: Dr. Anna Ihnatowicz** (01.01.2020 – 31.12.2021)
   Department of Plant Protection and Biotechnology, IFB UG & MUG, Gdańsk, Poland
   Laboratoire Agronomie et Environnement, Universite de Lorraine INRA, Vandœuvre-lès-Nancy, France
   "Exploring functional diversity of selected 2-oxoglutarate-dependent dioxygenases using natural variation of Arabidopsis thaliana"
   Tasks: Research internship for the practice in the field of metabolic profiling and biochemical analysis
5. **OPUS 8 (2014/15/B/NZ2/01073), NCN, PI: Dr. Anna Ihnatowicz** (10.2017 – 02.2020)
   Department of Plant Protection and Biotechnology, IFB UG & MUG, Gdańsk, Poland

"The role of iron-dependent and oxoglutarate-dependent dioxygenases in the response of Arabidopsis thaliana to environmental stress and in the regulation of iron homeostasis"
Tasks: Functional analysis of Fe/2-OG dioxygenases and phenotyping of A. thaliana mutants.

6. **PRELUDIUM 7 (2014/13/N/NZ9/0108), NCN, PI: Dr. Marta Potrykus** (01.07.2018 – 30.08.2018)
Department of Plant Protection and Biotechnology, IFB UG & MUG, Gdańsk, Poland
"Evaluation of the relationship between the metabolic profile of selected bacterial strains of plant pathogens Dickeya solani and their virulence"
Task: Investigation of *D. solani* growth in media with varying nutrient compositions.

Internships:

1. **Laboratoire Agronomie et Environnement, Universite de Lorraine INRA, Vandœuvre-lès-Nancy, France**
(01.2018 – 03.2018; 04.2019; 01.2020; 11.2021; 09.2023 – 11.2023; 18. – 22.11.2024; >9 months in total)
Experience: Molecular cloning and vector construction, bacteria transformation, transient transformation, overproduction and purification of plant proteins in heterologous systems (*E. coli*, *N.*, SDS-PAGE, enzymatic activity and substrate specificity tests, UHPLC-LTQ-Orbitrap-ESI-MS analysis, metabolic profiling analysis, bioinformatic skills (Compound Discoverer, Cytoscape software)
2. **Department of Plant Protection and Biotechnology, IFB UG & MUG, Gdańsk, Poland**
(01.2014 – 06.2014; 11.2014 – 06.2015; 07.2016 – 10.2016; >16 months in total)
Experience: Plant cultivation in laboratory conditions (soil, hydroponics, *in vitro* cultures), molecular cloning, bacteria transformation, transient transformation, competent cells preparation, DNA and RNA isolation, cDNA synthesis, PCR, RT-PCR, colony PCR, qPCR, bioinformatic skills (R programming, MEGA, BioEdit)
3. **Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, Netherlands** (07.2015 – 10.2015; 3 months)
Experience: Examination of the interactions between the bacteria and fungi (fungal culture, isolation of DNA from bacteria and fungi, bacterial transformation, colony PCR, 16S DNA PCR, interpretation and analysis of obtained results including sequencing data, keeping detailed documentation)
4. **J.S. Hamilton Poland S.A., Gdynia, Poland** (07.2014 – 10.2014; 3 months)
Experience: Support in the determination of element concentrations in medicines, food, food packaging, cosmetics and environmental samples using techniques of mass spectrometry (mineralizing oven SPPEDWAVE and ERTEC, ICP-MS, ICP-OS, FAAS, mercury analyzer CV-AAS)
5. **Department of Genetics and Marine Biotechnology, Institute of Oceanology, Polish Academy of Sciences, Sopot, Poland** (09.2013; 1 month)
Experience: Identification of microorganisms colonizing marine organisms by PCR method, isolation of the genetic material of marine organisms using different methods, agarose gel electrophoresis

Conference oral and poster presentations:

1. **International Conference on Arabidopsis Research ICAR2025, Ghent, Belgium** (16.–20.06.2025)
Oral and poster presentation *"How does the mutation in Natural Antisense Transcript-encoding region shape Arabidopsis response to environmental stresses?"* **Izabela Perkowska**, Alexandre Olry, Thibault Barrit, Alicja Dobek, Florent Magot, Clement Charles, Alicja Januchowska, Joanna Siwinska, Ewa Lojkowska and Anna Ihnatowicz
2. **21st International Symposium on Iron Nutrition & Interactions in Plants ISINIP 2024, Düsseldorf, Germany** (08. –12.07.2024)
Co-authorship of a poster presentation *"From roots to shoots and back again: the interplay between iron nutrition, the coumarin pathway and infection with bacterial pathogens"* Alicja Dobek, Clément Charles, Marta Potrykus, **Izabela Perkowska**, Jeremy Grosjean, Weronika Babińska-Wensierska, Alain Hehn, Ewa Łojkowska, Alexandre Olry, Anna Ihnatowicz
3. **EMBO/EMBL Diversity in Plants: from genomes to metabolism, Heidelberg, Germany** (09–12.04.2024)
Co-authorship of an oral presentation "*Esculetin is biosynthesized via novel O-demethylation activity in Arabidopsis thaliana"* Alicja Dobek, Clément Charles, **Izabela Perkowska**, Jeremy Grosjean, Alain Hehn, Ewa Łojkowska, Anna Ihnatowicz, Alexandre Olry
4. **Iberian Plant Biology Congress, University of Minho, Braga, Portugal** (09 – 12.07.2023)
Poster presentation '*"Interplay between dioxygenases involved in plant secondary metabolism and drought stress"* **Izabela Perkowska**, Alicja Januchowska, Alexandre Olry, Alain Hehn, Ewa Łojkowska, Anna Ihnatowicz

Co-authorship of an oral presentation *"Arabidopsis F6'H3 dioxygenase plays a role in nutrient homeostasis affecting iron and sulfate deficiency response"* Alicja Dobek, **Izabela Perkowska**, Alexandre Olry, Clément Charles, Jeremy Grosjean, Alain Hehn, Ewa Łojkowska, Anna Ihnatowicz

5. **3rd PhD Meeting of Plant Science, University of Minho, Braga, Portugal** (09.07.2023)
   Oral presentation *"Role of dioxygenases in plant adaptation to oxidative and cold stress responses"* **Izabela Perkowska,** Alicja Januchowska, Alexandre Olry, Alain Hehn, Ewa Łojkowska, Anna Ihnatowicz
   Co-authorship of an oral presentation "*Exploring the role of selected Arabidopsis thaliana dioxygenase gene (DIOXY2) in abiotic and biotic stress responses"* Alicja Dobek, **Izabela Perkowska**, Alexandre Olry, Clément Charles, Jeremy Grosjean, Alain Hehn, Ewa Łojkowska, Anna Ihnatowicz
6. **International Scientific Conference on Plant Biodiversity & Sustainability, Università degli Studi "Gabriele d'Annunzio" (online)** (13-14.10.2022)
   Oral presentation *"The impact of Arabidopsis thaliana natural populations biodiversity on the content of simple coumarins"* **Izabela Perkowska**, Joanna Siwińska, Alexandre Orly, Jérémy Grosjean, Alain Hehn, Frédéric Bourgaud, Ewa Łojkowska, Anna Ihnatowicz
   Co-authorship of an oral presentation *"Arabidopsis thaliana F6'H3 gene encoding a 2-oxoglutarate/iron-dependent dioxygenase is involved in abiotic stress responses and nutrient homeostasis"* Alicja Dobek, **Izabela Perkowska**, Alexandre Olry, Clement Charles, Jérémy Grosjean, Alain Hehn, Ewa Łojkowska, Anna Ihnatowicz
7. **20th International Symposium on Iron Nutrition & Interactions in Plants ISINIP 2022, Reims, France** (04. –08.07.2022)
   Flash talk presentation *"Identification and characterization of Arabidopsis UDP-glucosyltransferase involved in scopolin biosynthesis"* **Izabela Perkowska**, Joanna Siwińska, Alexandre Orly, Jérémy Grosjean, Alain Hehn, Ewa Łojkowska, Anna Ihnatowicz
8. **4th International Erwinia Workshop (IEW), Assisi, Italy** (02.–03.07.2022)
   Co-authorship of a poster presentation "*Responses of Arabidopsis thaliana plants with disturbed coumarin accumulation to Dickeya dadantii and Dickeya solani infection"* **Izabela Perkowska**, Marta Potrykus, Joanna Siwinska, Dominika Siudem, Ewa Lojkowska, Anna Ihnatowicz
9. **8th Central European Congress of Life Sciences EUROBIOTECH 2022, Kraków, Poland** (20.-22.06.2022)
   Co-authorship of an oral presentation "*Functional characterization of Arabidopsis thaliana F6'H3 gene encoding a 2-oxoglutarate/iron-dependent dioxygenase involved in abiotic stress response"* Alicja Dobek, **Izabela Perkowsk**a, Alexandre Olry, Clement Charles, Jeremy Grosjean, Alain Hehn, Ewa Łojkowska, Anna Ihnatowicz
10. **Young for Fahrenheit – popular science conference, Gdańsk, Poland** (17.11.2021)
    Oral presentation *"W służbie roślinom i ludziom. Identyfikacja genów zaangażowanych w biosyntezę roślinnych związków farmakologicznie czynnych"* Izabela Perkowska
11. **Plant Genomes in a Changing Environment 2021 Virtual Conference, Hinxton, UK (online)** (18. – 20.10.2021)
    Flash talk presentation *"Variable profiles of simple coumarins between Arabidopsis thaliana accessions"* Izabela Perkowska, Joanna Siwińska, Alexandre Orly, Jérémy Grosjean, Alain Hehn, Frédéric Bourgaud, Ewa Łojkowska, Anna Ihnatowicz
12. **10th biennial Polish Society of Experimental Plant Biology Virtual Conference (online)** (20.-23.09.2021)
    Oral presentation "*Effect of combined environmental stresses on plant health: study using the model system of Arabidopsis thaliana/Dickeya spp. molecular interactions"* **Izabela Perkowska**, Marta Potrykus, Joanna Siwińska, Dominika Siudem, Ewa Łojkowska, Anna Ihnatowicz
13. **31st International Conference on Arabidopsis Research, ICAR2021 Virtual Conference (online)** (21.–25.06.2021)
    Poster presentation *"Effect of iron deficiency on infection symptoms caused by Dickeya spp. in Arabidopsis thaliana"* **Izabela Perkowska**, Marta Potrykus, Joanna Siwińska, Dominika Siudem, Ewa Łojkowska, Anna Ihnatowicz
14. **European Federation of Biotechnology, EFB2021 Virtual Conference (online)** (10.–14.10.2021)
    Flash poster presentation *"Profiling of coumarins in Arabidopsis thaliana accessions using UHPLC-MS"* **Izabela Perkowska**, Joanna Siwińska, Alexandre Orly, Jérémy Grosjean, Alain Hehn, Frédéric Bourgaud, Ewa Łojkowska, Anna Ihnatowicz
15. **Plant Genomes in a Changing Environment, Hinxton, United Kingdom** (16.–18.10.2019)
    Co-authorship of a poster presentation *"Arabidopsis thaliana responses to Fe deficiency and Dickeya spp. infection"* Anna Ihnatowicz, Dominika Siudem, Marta Potrykus, **Izabela Perkowska**, Joanna Siwińska, Ewa Łojkowska

16. **9th International conference organized by Polish Society of Experimental Plant Biology, Toruń, Poland** (09.–12.09.2019)
    Oral presentation *"Arabidopsis thaliana cation exchanger 4 (CAX4) is important for plant fitness and growth under Mn, Zn and Fe deficiencies"* **Izabela Perkowska**, Emilia Śledzińska, Ewa Łojkowska, Anna Ihnatowicz
17. **VIIIth Intercollegiate Biotechnology Symposium "Symbioza", Warsaw, Poland** (17.–19.05.2019)
    Oral presentation *"Revealing unknown. Discovering the biological function of one of the plant UDP-glucosyltransferases"* **Izabela Perkowska**, Ewa Łojkowska, Anna Ihnatowicz
18. **Keystone Symposia meeting on Climate Change-Linked Stress Tolerance in Plants, Hannover, Germany** (13.-16.05.2019)
    Poster presentation *"Scopoletin UDP-glucosyltransferase: a new player in plant response to osmotic stress"* **Izabela Perkowska**, Ewa Łojkowska, Anna Ihnatowicz
19. **29th OAK Scientific Presentations Workshops, Olsztynek, Poland** (21.–25.11.2018)
    Oral presentation *"Revealing unknown. Discovering plant enzymes from phenylopropanoid biosynthetic pathway"* **Izabela Perkowska**, Ewa Łojkowska, Anna Ihnatowicz
20. **3rd congress of Polish Biosciences BIO2018, Gdańsk, Poland** (18.–21.09.2018)
    Poster presentation *"Functional analysis of scopoletin UDP-glucosyltransferase in plant responses to environmental stresses"* **Izabela Perkowska**, Joanna Siwińska, Ewa Łojkowska, Anna Ihnatowicz

Certificates and Trainings:

- Student Business Academy (60 h), Institute of Public Debate
- Phylogenetic analysis (16 h), data2biology Ltd.
- High-Performance Liquid Chromatography UPLC/HPLC in Practice (20 h), EkotechLAB R&D Laboratory
- Introduction to RNA-Seq Data Analysis (16 h), data2biology Ltd. (formerly: ideas4biology)
- Requirements of PN-EN ISO 14001:2015-09 – Environmental Management Systems (8 h), ISOTOP S.c.
- Requirements of the New Standard PN-EN ISO/IEC 17025:2018-02 (8 h), ISOTOP S.c.
- Proteomic Analysis Using Mass Spectrometry (22 h), IFB UG&MUG
- Bioinformatics for Next-Generation Sequencing (16 h), data2biology Ltd. (formerly: ideas4biology)
- Working in an Isotope Laboratory (8 h), IFB UG&MUG
- The basics of statistical methods in the R program (8 h), Warsaw University of Life Sciences
- Presentation of Research Results and Scientific Writing in English (8 h), University of Gdańsk / Zatoka Nauki
- Public Presentations at Scientific Conferences (8 h), Young Researchers Academy, Institute of Public Debate
- Mendeley – Reference Management Software, UG Library
- Autoclave Operation and Safety (16 h), Vocational Development Centre
- General Laboratory Techniques (16 h), LifeSciences Training Centre
- Category B Driving Licence

Scholarships and awards:

- Foreign Doctoral Internships NAWA Preludium BIS 1 scholarship (2023)
- ERASMUS+ scholarships (2018, 2023)
- Pro-quality scholarship for the best PhD students (2018/2019)
- 2nd place for the best revised oral presentation during 29th OAK Scientific Presentation Workshops (2018)
- 3rd place for the best poster during the 29th OAK Scientific Presentation Workshop (2018)

Skills:

- **Laboratory techniques & molecular biology:** PCR, RT-PCR, qRT-PCR, molecular cloning, colony PCR, cDNA synthesis, gene expression analysis, bacterial transformation, plant transformation (transient and stable, *Nicotiana benthamiana, Arabidopsis thaliana*), protein overproduction purification, SDS-PAGE, enzymatic activity assays and substrate specificity testing
- **Analytical techniques:** HPLC and UHPLC-MS (LTQ-Orbitrap-ESI-MS), metabolic profiling, Compound Discoverer software, bioinformatics tools for sequence analysis (MEGA, BioEdit), network analysis (Cytoscape), statistical analysis and data visualization in R
- **Plant cultivation & experimental design:** cultivation of plants under controlled conditions (soil, hydroponics, *in vitro*), stress treatment setup (e.g., low pH, osmotic, drought, Fe deficiency), plant metabolite profiling

- **Science communication & teaching:** public presentation of research (oral and poster) at national and international conferences; conducted training on laboratory methods of plant cultivation and micropropagation, active organizer of science outreach events (e.g., Biotechnology Summer School, Biologists' Night, ART+SCIENCE MEETING)
- **Professional involvement & leadership:** member in the following doctoral student governance bodies (2021–2023): Doctoral Students Council of IBDS UG & MUG, Doctoral Students Council of the UG, UG Senate Committee on Science, UG Disciplinary Committee for Doctoral Students, and Faculty Scholarship Committee (2018–2020)

Languages:

- English – Advanced level (C1)
- French – Intermediate level (B1)