

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Gluza, Pawel

Title:

From nucleotide sugars to polysaccharides: How do plants control the delivery of substrates for cell wall biosynthesis and protein glycosylation?

Date:

2021

Persistent Link:

<https://hdl.handle.net/11343/274899>

Terms and Conditions:

Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.

From nucleotide sugars to polysaccharides:

**How do plants control the delivery of
substrates for cell wall biosynthesis and
protein glycosylation?**

Paweł Gluza

ORCID: 0000-0002-3568-4153

PhD thesis

February 2021

School of BioSciences

The University of Melbourne

Submitted in total fulfilment of the requirements
of the degree of Doctor of Philosophy

ABSTRACT

Plant cell walls constitute one of the most abundant raw biomaterials on Earth. The synthesis of long-chain polysaccharides, the main components of plant cell walls starts *ab ovo* in the cytoplasm where most of the building blocks for polysaccharide synthesis, so-called nucleotide sugars, are produced. The monosaccharide moieties of nucleotide sugars are incorporated into growing polysaccharide chains either directly at the plasma membrane by glycosyltransferases (GTs) that form cellulose synthase complexes or by those residing in the Golgi apparatus. In the latter case, nucleotide sugars have to pass the Golgi membranes with the help of nucleotide sugar transporters (NSTs). Once inside, they are used by Golgi GTs which assemble polysaccharide chains and decorate proteins and lipids with sugar residues. Recent evidence suggests that in plants nucleotide sugars can be guided to specific polysaccharides and/or glycan decorations, yet the molecular mechanisms of these processes are not fully understood.

This PhD research attempts to explore the phenomenon of the targeted substrate delivery by investigating two possible hypotheses: the spatiotemporal distribution of proteins within the Golgi apparatus and the occurrence of direct interactions between NSTs and GTs. The author of this dissertation has tested both of those hypotheses by investigating protein-protein interactions, localising the individual components to their specific sub-compartments within the cell and tracking changes in mutant plants where these processes are modulated. The bifunctional UDP-RHAMNOSE/UDP-GALACTOSE TRANSPORTER (URGT) family from the model plant *Arabidopsis thaliana* was selected for this study. While this family has six members which *in vitro* are capable of transporting the same substrates, plant mutant studies indicate substrate preference and targeted delivery to specific cell wall components *in vivo*.

This thesis presents the first evidence that members of the URG family localise to specific sub-Golgi compartments. The colocalisation studies undertaken as part of this thesis place URG1 and URG5 in the cis-Golgi, URG2, URG4 and URG6 – in the medial Golgi, while URG3 seems to localise to trans-Golgi stacks.

Protein-protein interactions studies have identified multiple interaction partners for the six URGs. Notably, many of those are known galactosyltransferases, which aligns with the transport function of the URGs. It is therefore highly likely that the identified candidates are true interactors, which use the proximity of the transporter to increase the process efficiency by substrate channelling. This finding is further supported by the fact that observed interactions

between URGF family members and GTs often localise to the same sub-Golgi compartment. The study identified new potential galactosyl- and rhamnosyltransferases, including two putative arabinogalactan protein galactosyltransferases.

The data obtained during this project suggest, that URGFs may determine the flow of substrate through both spatial separation within the sub-Golgi stacks and direct interactions with GTs. The study presents new insight into the sugar substrate delivery processes in plants, suggests that similar processes may take place in other NST families and that their specificity may be similarly tuned by localisation and interactions.

DECLARATION

I hereby declare that:

- (i) the thesis comprises only my original work towards the Doctor of Philosophy Degree except where indicated in the Preface;
- (ii) due acknowledgement has been made in the text to all other material used;
- (iii) the thesis is fewer than the maximum word limit (100,000) in length, exclusive of tables, maps, bibliographies and appendices.

Signed:

Paweł Gluza

Date: 04/02/2021

PREFACE

I hereby acknowledge the hard work of my colleagues, collaborators and sponsors that contributed to this study:

Preparation of samples for transmission electron microscopy was performed in collaboration with Dr Arun Sampathkumar's lab at Max Planck Institute for Plant Molecular Physiology in Potsdam-Golm, Germany. The work was conducted under the supervision of Anja Froehlich with the help of Alban Mariette. However, while this project has been originally part of this thesis, we refocused for the benefit of the more developed chapter on the characterisation of the GT candidates.

Dr Jennifer Bromley and Dr Berit Ebert performed the preliminary split-ubiquitin screen with URG1, URG2, URG3 and UXT3 at Joint BioEnergy Institute (JBEI) in Berkeley, California, USA under the supervision of Prof. Henrik V. Scheller. Dr Jennifer Bromley has also created the GT cDNA library used in this study. The research included in chapter 5.2 of this thesis is part of the collaboration with Prof. Wei Zeng (Zhejiang A&F University, China).

Generation of the RNAi construct and plant lines described in chapter 5.2.3. and qRT-PCR analysis of the expression of *At1g11940* and *At1g62305* genes described in chapter 5.2.2. were performed by Jacob Calabria, a research assistant in Ebert Lab.

The author would like to acknowledge funding from International Max Planck Research School and Melbourne Research Scholarship, as well as the following awards: Norma Hilda Schuster (nee Swift) Scholarship, John S. Turner Postgraduate Scholarship, Science Abroad Travelling Scholarship and Australian Society of Plant Scientists Student Travel Award. This author of this work has also received support from Botany Foundation and the University of Melbourne.

The author would like to thank Dr Berit Ebert for the proof-reading of this thesis and all the suggestions she has made and acknowledge the hard work of Joanna Kaptur, who had thoroughly read the first draft and polished it with a series of corrections and improvements.



SCHOOL OF BIOSCIENCES

COVID-19 IMPACT STATEMENT

Student name: Pawel Gluza

Date: 04.02.2021

Supervisor names: Berit Ebert, Joshua L. Heazlewood, Staffan Persson

Preamble/Synopsis

In response to COVID-19, the University of Melbourne transitioned to a 'working from home' model, closing all building and restricting the access for students and staff alike. Therefore, I had no access to the lab from 24.03.2020 to 2.11.2020, except for a short period before the second lockdown between 1.07.2020 and 3.08.2020, where I was allowed to work at a 33% capacity. From 2.11.2020 I was allowed back into the lab for 3 days a week, which gradually rose to full time by 01.01.2021.

The physical shutdown occurred 3 years and 4 months since I started my PhD. While at that time I was already starting to write my thesis, there were still several key experiments that needed to be undertaken. Most of those were already underway and had to be abandoned halfway through. The lockdown not only made it impossible for me to perform the planned experiments but also significantly postponed necessary repairs and maintenance of crucial equipment for these experiments. In the further sections, I will highlight those experiments and outline how the thesis could look different in anticipation of what those results might look like.

Summary:

Over the course of 2020, I have effectively lost about 8 months of lab work.

Laboratory experiments

(1) The plant work

Since my lab work requires the propagation of plants, each time I returned to the lab after a lockdown, I was forced to grow new plants and wait for about 6 weeks before I could perform assays just before they flowered (e.g. GUS-assisted staining of plant organs). For that reason, the actual period of not being able to conduct any assays was in my case longer than the lockdown periods itself and lasted roughly from February to December.

(2) The repairs and maintenance of lab equipment

Not only the pandemic influenced me personally and my ability to work, but also the undergoing repairs of three pieces of equipment that were crucial for the analysis of samples for my project – (I) Beckman Coulter Optima L-80 XP ultracentrifuge, (II) SpeedVac Vacuum Concentrator and (III) Dionex HPAEC-PAD system. Even though my samples were ready for processing and/or analysis, the delayed shipping of the required instrument parts and the difficulties of contracting service technicians still hampered my availability to finish experiments, even when appropriate lab access was restored. All of the above instruments have only been fixed throughout January 2021.

Consequences to research report:

For the reasons described above, the coimmunoprecipitation (Co-IP) experiment, which was originally planned to be a secondary method used to confirm the interactions between the NST and GT proteins, as described in chapter 4 could not be finalised. The repairs also significantly delayed the analysis of the monosaccharide samples of seed mucilage, and alcohol-insoluble residue from cell walls of Arabidopsis loss-of-function and overexpression mutants, as well as *N. benthamiana* leaves overexpressing the proteins of interest. Those results would comprise an additional subchapter in chapter 5 of this thesis. Lastly, the lockdown forced the author to put the scanning electron microscopy (SEM) on hold after the first batch of samples. The author declares a willingness to work towards the finalisation of those experiments during the thesis examination period.

Summary:

Co-IP and monosaccharide analyses are missing from the thesis. SEM experiment was performed only on some of the originally planned Arabidopsis lines.

Data analysis/Thesis writing**(1) Lack of office access**

Along with the lost lab access, I have also lost access to my office PC, with its two large screens and all of the licensed software. For the better part of 2020, I was forced to use my small laptop with its limited processing power and free software. This may have led to a lower quality of some figures.

(2) Lack of daily contact with supervisors

While I usually discuss the project with my PI during casual visits in her office a few times a day, during the lockdown our contact was limited to weekly or even fortnightly meetings.

Summary:

The quality of some of the images might be lower than it would be in normal circumstances.

Comments from supervisors

As the three supervisors of Pawel Gluza, we concur with Pawel's assessment of how the COVID-19 changes in workplace impacted his/her thesis.

ACKNOWLEDGEMENTS

Ask the green plants of the earth and they'll teach you

Job, 12:8

Every book is a journey. It needs voyagers and the voyagers need an adventure. However, dear reader, before we embark on it over the next hundred pages or so, let me introduce you to the *dramatis personae*; for I was not alone on this thorny road for the last four years.

First of all, there's **Berit**. She is the warmest and kindest lab head I have ever met in academia; one that doesn't mind being messaged in the middle of the night, just because you came up with a wild hypothesis to explain your data; one that tells you to have some rest when you work too hard. In her, I was blessed to have a supervisor that supports and fights for every member of her team and always stands in their corner. Encourages you to have a go at something that seems wildly out of your reach, yet somehow you end up succeeding.

Then there are **Josh and Staffan** – the people whose names I knew well years before moving to Australia and would never guess I will end up meeting them in person. Getting to know them not from the author list in a journal but from the perspective of a member of their labs was a real eye-opener – they're not only great scientists but also friendly people that I enjoyed hanging out with. I am especially deeply grateful to Josh for offering me a lift from the airport and a place to stay back in 2016 when I was fresh off the plane.

My lab colleagues and my PhD committee - they were always there to reach out with a helping hand and a piece of advice, both at the experiment planning stage and while elbows-deep in the wet lab work. All of those people created hands-down the best working environment a student can imagine; I'll miss it so much that lately, I start to devise plans on how to stick around the lab for just a bit longer.

Gosia and Marek – people who encouraged me at the start of my PhD journey and provided me with references to kickstart my scientific career. **Laurent and Gerard** – who selflessly mentored me through my PhD and shown me how to lead my career past this stage.

Finally, **my family, my friends, my loved ones** – you know who, you are. Thanks for believing in me and supporting my choice when I decided to pursue my PhD in a far-away land. Apologies for not being around in the moments of need. Thanks to **Joanna** for joining me in this endeavour and pulling her weight next to mine. Your wisdom, beauty and selflessness always inspired me to try and be a better scientist and human.

To all of those people, I offer my greatest thanks and eternal gratitude. This thesis would not exist without your help.

For me, writing those words is an end to a certain journey; one that consumed the best years of my life, leaving behind nothing but a few memories. But for you, dear reader, this is only a beginning. May it be a good one.

Does anything still move you since you're educated now?

And all grown up and travelled so well

Do you still hear the sound of the thunder while you lie up by yourself?

Brian Fallon

I dedicate this thesis to my friend Dawid, who tragically passed away in 2019.

TABLE OF CONTENTS

ABSTRACT	i
DECLARATION.....	iii
PREFACE.....	iv
COVID-19 IMPACT STATEMENT	v
ACKNOWLEDGEMENTS.....	vii
TABLE OF CONTENTS.....	viii
ABBREVIATIONS	xi
LIST OF FIGURES & TABLES	xvii
1. INTRODUCTION.....	1
1.1. The plant cell wall.....	1
1.2. Nucleotide sugar interconverting enzymes (NSEs)	5
1.3. Nucleotide sugar transporters (NSTs)	7
1.4. Glycosyltransferases (GTs)	10
1.5. The Arabidopsis UDP-rhamnose/UDP-galactose transporter (URGT) family	12
1.6. Possible substrate channelling occurring in the Golgi apparatus of <i>Arabidopsis thaliana</i>	13
1.7. Aims of the study	17
2. MATERIALS AND METHODS	19
2.1. Bioinformatic tools.....	19
2.1.1. Phylogeny	19
2.2. Chemicals, reagents and services.....	20
2.3. Molecular techniques and materials	20
2.3.1. Bacterial strains, growth media and conditions	20
2.3.2. Transformation of bacterial strains	21
2.3.3. DNA cloning	21
2.3.3.1. Polymerase chain reaction (PCR).....	21
2.3.3.2. Topoisomerase based (TOPO®) cloning.....	22
2.3.3.3. Gateway® cloning.....	22
2.3.3.4. Classical cloning approach – restriction digest and ligation	23
2.4. Yeast assays	23
2.4.1. Yeast growth media and conditions	23
2.4.2. Split-ubiquitin based interaction screen	24

2.4.3. Split-ubiquitin pairwise assay	25
2.5. Plant materials and methods	26
2.5.1. Plant material	26
2.5.2. Seed sterilisation	27
2.5.3. Growth conditions	27
2.5.3.1. <i>In vitro</i>	27
2.5.3.2. Cultivation on soil	28
2.5.4. Extraction of genomic DNA from <i>Arabidopsis</i> leaves	28
2.5.5. PCR for genotyping	29
2.5.6. Extraction of total RNA from <i>Arabidopsis</i> leaves	29
2.5.7. Analysis of <i>Arabidopsis</i> T-DNA insertional mutants by reverse transcription PCR (RT-PCR) ...	29
2.5.7.1. Analysis of <i>Arabidopsis</i> mutants by quantitative RT-PCR (qRT-PCR)	30
2.5.8. Stable <i>Arabidopsis</i> transformation using the floral dip	31
2.5.9. Generation of genetic crosses	31
2.5.9.1. Transgenic <i>Arabidopsis</i> lines used for genetic crosses	32
2.5.10. Transient transformation of <i>Nicotiana benthamiana</i> leaves	33
2.6. Microscopy methods	34
2.6.1. Light & fluorescence microscopy methods	34
2.6.2. Spinning disk microscopy	34
2.6.3. Colocalisation - imaging and analysis	35
2.6.4. Bimolecular Fluorescence Complementation (BiFC)	35
2.6.5. Acceptor photobleaching - Förster resonance energy transfer (APB-FRET)	36
2.6.6. Scanning Electron Microscopy (SEM)	36
2.7. Biochemical methods	37
2.7.1. Ruthenium red seed staining	37
2.7.2. Seed mucilage composition analysis	37
2.7.3. Cell wall monosaccharide composition analysis	38
2.7.4. Stress treatments	39
2.7.5. Co-immunoprecipitation	39
2.7.6. GUS reporter-aided detection of promoter activity	40
3. EXAMINING THE SUB-GOLGI LOCALISATIONS	41
3.1. Confirming Golgi localisation of URG1 family members	41

3.2. Colocalisation of URG1 family members with fluorescent markers specific for distinct Golgi subcompartments.....	43
3.3. Conclusion	48
4. IDENTIFICATION AND VALIDATION OF GTs	49
4.1. <i>In silico</i> analysis of publicly available interaction and co-expression data.....	49
4.1.1. Interaction data.....	49
4.1.2. Co-expression data.....	50
4.2. Identification of GTs interacting with NSTs using the split-ubiquitin screen	53
4.3. Confirmation of the interactions using a pairwise split-ubiquitin assay	57
4.4. Validation of protein-protein interactions using additional methods	64
4.5. Conclusion	66
5. FUNCTIONAL CHARACTERISATION OF GTs.....	67
5.1. GT candidates selected for characterisation	67
5.1.1. Subcellular distribution of GTs interacting with URG1 family	68
5.1.2. Analysis of GT loss-of-function mutant lines.....	70
5.2. Characterisation of AT1G11940 and AT1G62305, two novel putative galactosyltransferases	74
5.2.1. Colocalisation of AT1G11940 and AT1G62305 with fluorescent markers specific for distinct Golgi subcompartments.....	75
5.2.2. <i>At1g11940</i> and <i>At1g62305</i> expression pattern analysis through qRT-PCR and GUS staining	77
5.2.3. Phenotypic analysis of double mutants and knock-down lines affecting the functions of <i>At1g11940</i> and <i>At1g62305</i>	79
5.3. Conclusion	81
6. CONCLUSION	82
REFERENCES.....	84
SUPPLEMENTAL TABLES	94

ABBREVIATIONS

3-AT	3-aminotriazole
ABRC	Arabidopsis Biological Resource Center
Ace	Aceric acid
ACN	Acetonitrile
Ade	Adenine
AGP	Arabinogalactan protein
AINM	Arabidopsis Interactome Network Map
AIR	Alcohol insoluble residue
ALG	ASPARAGINE-LINKED GLYCOSYLATION
APB-FRET	Acceptor photobleaching - Förster resonance energy transfer
Api	Apiose
ARAD	ARABINAN DEFICIENT
Arap/Araf	Arabinopyranose/Arabinofuranose
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
BiFC	Bimolecular fluorescence complementation
CAZy	Carbohydrate-active enzymes
CCRC	Complex Carbohydrate Research Center
cDNA	Complementary deoxyribonucleic acid
CDS	Coding sequence
CESA	CELLULOSE SYNTHASE
CMP	Cytidine monophosphate
Co-IP	Co-immunoprecipitation
Col-0	Columbia-0 ecotype
CSL	CELLULOSE SYNTHASE-LIKE
Cub	C-terminus of ubiquitin
DCB	2,6-dichlorobenzonitrile
DGD	DIGALACTOSYL DIACYLGLYCEROL DEFICIENT
DGH	Dark-grown hypocotyl
Dha	3-deoxy-D-lyxo-heptulosaric
DNase	Deoxyribonuclease

dNTP	Deoxyribose nucleoside triphosphate
DTT	Dithiothreitol
DUF	DOMAIN OF UNKNOWN FUNCTION
EBS	EMS-MUTAGENIZED BRI1 SUPPRESSOR
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(β -aminoethyl ether)-N,N,N',N'-tetraacetic acid
EPC	ECTOPICALLY PARTING CELLS
ER	Endoplasmic reticulum
FA	Formic acid
FP	Fluorescent protein
FRET	Förster resonance energy transfer
Fuc	Fucose
FUT	FUCOSYLTRANSFERASE
Gal	Galactose
GalA	Galacturonic acid
GALS	GALACTAN SYNTHASE
GALT	GALACTOSYLTRANSFERASE
GATL	GALACTURONOSYLTRANSFERASE-LIKE
GAUT	GALACTURONOSYLTRANSFERASE
GAX	Glucuronoarabinoxylan
gDNA	Genomic deoxyribonucleic acid
GDP	Guanine diphosphate
GER	GDP-4-KETO-6-DEOXY-D-MANNOSE 3,5-EPIMERASE-4-REDUCTASE
GFP	Green fluorescent Protein
GFT	GDP-FUCOSE TRANSPORTER
Glc	Glucose
GlcA	Glucuronic acid
GlcAT	GLUCURONIC ACID TRANSFERASE
GlcNAc	N-ACETYLGLUCOSAMINE
GM	Glucomannan
GMD	GDP-D-MANNOSE-4,6-DEHYDRATASE
GME	GDP-D-MANNOSE 3,5-EPIMERASE
GM	Glucomannan

GmMan1	<i>Glycine max</i> α -1,2 mannosidase 1
GNT	N-ACETYLGLUCOSAMINYLTRANSFERASE
GONST	GOLGI-LOCALIZED NUCLEOTIDE SUGAR TRANSPORTER
GOLS	GALACTINOL SYNTHASE
GT	Glycosyltransferase
GUS	β -glucuronidase
GUX	GLUCURONIC ACID SUBSTITUTION OF XYLAN
GX	Glucuronoxylan
HA	Hemagglutinin
HDEL	ER-retention motif (histidine-aspartic acid-glutamic acid-leucine)
HG	Homogalacturonan
His	Histidine
HPAEC-PAD	High-performance anion-exchange chromatography with pulsed amperometric detection
HPGT	HYDROXYPROLINE O-GALACTOSYLTRANSFERASE
HPLC	High-performance liquid chromatography
IAA	Indole-3-acetic acid
IAGLU	INDOLE-3-ACETATE β -D-GLUCOSYLTRANSFERASE
IPUT	INOSITOL PHOSPHORYLCERAMIDE GLUCURONOSYLTRANSFERASE
IRX	IRREGULAR XYLEM
JACoP	Just Another Colocalization Plugin
JTT	Jones-Taylor-Thornton model
Kdo	Ketodeoxyoctonic acid
KOB	KOBITO
KT	Lysine-threonine motif
LB	Lysogeny broth
Leu	Leucine
LGT	LOS GLYCOSYLTRANSFERASE
LOPIT	Localisation of organelle proteins by isotope tagging
LPXB	LIPID X B
LTPG	GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHORED LIPID TRANSFER PROTEIN
Man	Mannose
MES	2-(N-morpholino)ethanesulfonic acid

MGP	MALE GAMETOPHYTE DEFECTIVE
MIND	Membrane-based Interactome Network Database
MLG	Mixed-linkage glycan
MOPS	3-(N-morpholino)propane sulfonic acid
MS	Murashige & Skoog medium
MSR	MANNAN SYNTHESIS-RELATED
MUM	MUCILAGE-MODIFIED
MUR	MURUS
NAG	<i>N</i> -ACETYLGLUCOSAMINYL TRANSFERASE I
NASC	Nottingham Arabidopsis Stock Centre
NDP	Nucleoside diphosphate
NP-40	Nonyl phenoxy polyethoxy ethanol
NRS/ER	NUCLEOTIDE-RHAMNOSE SYNTHASE/EPIMERASE-REDUCTASE
NSE	NDP-sugar interconversion enzymes
NST	Nucleotide Sugar Transporter
Nub (G/I)	N-terminus of ubiquitin (glycine/isoleucine)
OD	Optical density
PAGR	PECTIC ARABINO GALACTAN SYNTHESIS-RELATED
PCR	Polymerase chain reaction
PCW	Primary cell wall
PGSIP	PLANT GLYCOGENIN-LIKE STARCH INITIATION PROTEIN
PLV	Protein A-LexA-VP16 transcription factor
PM	Plasma Membrane
PNT	PEANUT
PP-60	Polyphenon-60
PPI	Protein-protein interaction
qRT-PCR	Quantitative Reverse Transcription-PCR
RAY	REDUCED ARABINOSE YARIV
RFP	Red fluorescent protein
RG (I/II)	Rhamnogalacturonan (I/II)
RGP	REVERSIBLY GLYCOSYLATED POLYPEPTIDE
RGXT	RHAMNOGALACTURONAN XYLOSYLTRANSFERASE
Rha	Rhamnose

RHM	UDP-L-RHAMNOSE SYNTHASE
RNase	Ribonuclease
ROCK	UDP- <i>N</i> -ACETYLGLUCOSAMINE / UDP- <i>N</i> -ACETYL GALACTOSAMINE TRANSPORTER
RRA	REDUCED RESIDUAL ARABINOSE
RRT	RG-I:RHAMNOSYLTRANSFERASE
RT-PCR	Reverse Transcription-PCR
SCW	Secondary cell wall
SD	Synthetic dextrose minimal medium
SEM	Scanning electron microscopy
S.E.M.	Standard error of the mean
Sia	Sialic acid
SOC	Super optimal broth with catabolite repression
SQD	SULFOQUINOVOSYLDIACYLGLYCEROL
ST	Rat α -2,3-sialyltransferase
STL	STELLO
SUBA	Subcellular Localisation Database for Arabidopsis Proteins
SUS	SUCROSE SYNTHASE
TAE	Tris/Acetate/EDTA buffer
TAIR	The Arabidopsis Information Resource
T-DNA	Transfer DNA of the tumour-inducing plasmid
TE	Tris/EDTA buffer
TEM	Transmission Electron Microscopy
TFA	Trifluoroacetic acid
TGN	Trans-Golgi network
TIRF	Total internal reflection fluorescence
TPS	TREHALOSE-6-PHOSPHATE SYNTHASE
TPT	Triose phosphate/phosphate translocator
Tris	Tris(hydroxymethyl)aminomethane
Trp	Tryptophane
UAFT	UDP-ARABINOFURANOSE TRANSPORTER
UAXS	UDP-D-APIOSE/UDP-D-XYLOSE SYNTHASE
UDP	URIDINE DIPHOSPHATE
UGD	UDP-D-GLUCOSE DEHYDROGENASE

UGE	UDP-D-GLUCOSE 4-EPIMERASE
UGlcAE	UDP-D-GLUCURONIC ACID 4-EPIMERASE
UGNT	UDP-N-ACETYLGALUCOSAMINE TRANSPORTER
UGT	UDP-GLUCOSYLTRANSFERASE
UNE	UNFERTILIZED EMBRYO SAC
URGT	UDP-RHAMNOSE/UDP-GALACTOSE TRANSPORTER
UTR	UDP-GLUCOSE/UDP-GALACTOSE TRANSPORTER
UUAT	UDP-URONIC ACID TRANSPORTER
UXE	UDP-D-XYLOSE 4-EPIMERASE
UXS	UDP-D-XYLOSE SYNTHASE
UXT	UDP-XYLOSE TRANSPORTER
v/v	Volume to volume
WAK	WALL-ASSOCIATED RECEPTOR KINASE
WT	Wild type
w/v	Weight to volume
XGA	Xylogalacturonan
XGD	XYLOGALACTURONAN DEFICIENT
X-Gluc	5-bromo-4-chloro-3-indolyl- β -D-glucuronic acid, cyclohexylammonium salt
XT	XYLOSYLTRANSFERASE
XXT	XYLOGLUCAN XYLOSYLTRANSFERASE
XyG	Xyloglucan
Xyl	Xylose
XYLT	β -1,2-XYLOSYLTRANSFERASE
YFP	Yellow fluorescent protein
YPDA	Yeast extract-peptone-dextrose-adenine complete medium

LIST OF FIGURES & TABLES

Figure 1.	Biosynthesis of cell wall polysaccharides in Golgi apparatus.....	4
Figure 2.	NDP-sugar interconversion reactions in plants.....	6
Figure 3.	Phylogenetic tree of the NST families in Arabidopsis.....	8
Figure 4.	UXT family activities in Arabidopsis.....	14
Figure 5.	Two models illustrating the putative substrate channelling mechanism.....	15
Figure 6.	The workflow of yeast split-ubiquitin interaction screen.....	25
Figure 7.	The workflow of yeast split-ubiquitin interaction pairwise assay.....	26
Figure 8.	The workflow for the generation of transgenic Arabidopsis plants.....	32
Figure 9.	Subcellular localisation of the URG T family members.....	42
Figure 10.	Colocalisation of the same Golgi cisternae markers fused to different FPs.....	43
Figure 11.	Fluorescence images of colocalisation of sub-Golgi markers and URG Ts.....	44-45
Figure 12.	Average colocalisation values of URG Ts.....	46
Figure 13.	Sub-Golgi localisations of the URG T family members.	47
Figure 14.	Split-ubiquitin pairwise assay results for URG Ts.....	59-60
Figure 15.	Split-ubiquitin pairwise assay results for UXTs.....	62
Figure 16.	Subcellular localisation of GT interactors of URG Ts.....	69
Figure 17.	Mucilage release phenotypes of seeds coming from the knockout lines.....	72
Figure 18.	SEM images of seeds and mucilage secretory cells.....	73
Figure 19.	Phylogenetic tree of GT14-L family.....	74
Figure 20.	Sub-Golgi localisations of three GTs from GT14-L family.....	76
Figure 21.	Average colocalisation values of three GTs from GT14-L family.....	77
Figure 22.	qRT-PCR analysis of At1g11940 and At1g62305 expression.....	78
Figure 23.	qRT-PCR measurements of At1g62305 downregulation	80
Figure 24.	RNAi At1g62305 insertion mutants in at1g11940 backgrounds.....	80
Table 1.	Co-expression networks between URG Ts, UXTs and other NSTs.....	51
Table 2.	Co-expression networks between URG Ts, UXTs and GTs.....	52
Table 3.	GT-NST interactions identified using a split-ubiquitin assay screen.....	55-56
Table 4.	GT-NST interactions confirmed through split-ubiquitin pairwise assay.....	63

Table 5.	Shortlist of characterised GT candidates.....	67
Table 6.	Arabidopsis T-DNA insertion lines used in the study.....	70
Table S1.	Nucleotide sugars found in plants.	94
Table S2.	NSTs characterised in Arabidopsis.....	95
Table S3.	Oligonucleotides used for cloning of NSTs/GTs with their native promoters.....	96
Table S4.	Oligonucleotides used for Sanger sequencing.....	96
Table S5.	Oligonucleotides used for the creation of new constructs.....	97
Table S6.	Oligonucleotides used for genotyping of plant lines	97
Table S7.	Oligonucleotides used for amplification of full sequence CDS.....	98
Table S8.	Interactions of members of URGT and UXT families presented in databases.....	99
Table S9.	Oligonucleotides used for qRT-PCR.....	99
Table S10.	Selected constructs created for the needs of experiments.....	99
Figure S1.	Secondary methods used for validation of protein-protein interactions.....	100
Figure S2.	Colocalisation of GTs interacting with URGTs with ER and Golgi markers...	101
Figure S3.	Loss-of-function plant lines – validation via RT-PCR.....	102
Figure S4.	Loss-of-function plant lines – phenotype analysis.....	103
Figure S4.	Loss-of-function plant lines – seedling phenotypes in stress conditions.....	104

1. INTRODUCTION

1.1. The plant cell wall

The presence of the cell wall is a unique feature, vital for the development and growth of plants. It functions as an exoskeleton; provides scaffolding for the cells and tissues, regulates cell growth, gives shape and allows cells to withstand high turgor pressure and low suction pressure without bursting or losing their shape. In addition, it is an intercellular boundary; as such, it performs a plethora of essential signalling functions, including the communication between adhering cells, response to pathogens and interactions between plant and microbe cells. Finally, the cell wall offers a barrier from the environment; it is the first and most important line of defence protecting the contents of the cell.

Plant cell walls have found a wide variety of applications in the industry. Since this energy-rich and versatile raw material is also the most abundant bioresource on Earth, it has quickly found its way into everyday human life in the form of biofuels, textiles, timber, paper, gums, biocomposites, pharmaceuticals or simply as fodder, dietary fibre and firewood (Jung & Allen, 1995, Chapple & Carpita, 1998, Billosta *et al.*, 2006, Himmel *et al.*, 2007, Blanch *et al.*, 2008, Pettolino *et al.*, 2012). Despite the overall usefulness of plant cell wall material, its further use in industry and medicine could greatly benefit from developing a way to produce biomass with improved properties, reduced heterogeneity (Mazzarino *et al.*, 2015) and increased susceptibility to enzyme degradation (Reiter, 2008). A detailed understanding of the processes underpinning the biosynthesis of cell wall polysaccharides is also an imperative for plant research and will have a considerable impact on modern biotechnological applications. Therefore, recent years have seen a growing interest in cell wall research, especially in studies focusing on creating tailored, custom-made cell walls, engineered to fit a purpose (Johnson *et al.*, 2018, Svagan *et al.*, 2010, Zhong & Ye, 2007).

Plant cell walls are generally divided into two types differing in structure, composition and function: primary cell walls (PCW) and secondary cell walls (SCW) (Cosgrove & Jarvis, 2012). PCWs are synthesised at the stage of intensive growth and consist of thin, plastic, highly hydrated layers forming structures surrounding single cells. Their purpose is to provide durable and robust support for the plant that would not hamper the physical enlargement of growing cells (Hamant & Traas, 2010). On the other hand, the role of SCWs is to provide stiffness and

reinforcement for tissues, that have ceased to grow. SCWs make up the vessels of terrestrial plants so that they can sustain the presence of negative suction pressure which makes it possible to transport water from the roots to the above-ground parts of the plant (Koch *et al.*, 2004, Meents *et al.*, 2018).

Plant cell walls can differ in structure and composition between different cell types, but also within specific microdomains within the same cell. Therefore, the cell wall will be specialised, look and behave differently in e.g. tracheary elements, endodermis and guard cells. Plant cells can be classified into three categories based on their composition: collenchyma, parenchyma and sclerenchyma. While the two first have PCWs exclusively, the sclerenchyma is combining both PCWs and SCWs (Zhong & Ye, 2007).

Structurally, cell walls are composed of multiple polysaccharide and glycoprotein components that are organised into a tangled, three-dimensional matrix. The most abundant component of this mesh is cellulose, a long fibril structure composed of single β -1,4-linked glucan chains, that are intertwined and glued together by the force of hydrogen bonds (McFarlane *et al.*, 2014a). Those fibrils provide the central scaffold, which in PCWs is filled with a gel-like matrix of pectins, including xylogalacturonan (XGA), homogalacturonan (HG), rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) (Harholt *et al.*, 2010); and hemicelluloses such as xylans, xyloglucans (XyG), glucans, mannans, glucomannan (GM) and mixed-linkage glucans (MLG) (Scheller & Ulvskov, 2010). The occurrence and abundance of these polysaccharides are heavily species-dependent and differs between different plants. Pectins, the most complex and diverse polysaccharides of the PCW, perform multiple roles in growth, wall penetrability, adhesion, strength, and mediation of intercellular signalling (Houston *et al.*, 2016). HG is composed solely of 1,4-linked galacturonic acid (GalA) residues. RG-I has a backbone combining (1,4) GalA and (1,2) rhamnose (Rha) residues with branches made of 1,4-linked galactose (Gal) or 1,5-linked arabinose (Ara) chains, occasionally terminated with (1,2) fucose (Fuc) and (1,4/1,6) glucuronic acid (GlcA) (Ridley *et al.*, 2001). Finally, RG-II is composed of heavily diversified sidechains containing Ara, Rha, Fuc, apiose (Api), GlcA, GalA, Gal, Xyl, ketodeoxyoctonic acid (Kdo), 3-deoxy-D-lyxo-heptulosaric (Dha) and aceric acid (Ace) on an HG backbone (Harholt *et al.*, 2010, Ndeh *et al.*, 2017). Hemicelluloses are yet another group of plant cell wall polysaccharides. They serve as the linkers connecting cellulose fibrils with the pectic matrix via hydrogen bonds, providing mechanical support (Scheller & Ulvskov, 2010). In dicots, the two most common hemicelluloses are XyG, consisting of 1,4-linked glucose (Glc) backbone with 1,6-linked

xylose (Xyl), Gal and Fuc sidechains, and glucuronoxylan (GX) consisting of a 1,4-linked Xyl backbone with 1,2- or 1,3-linked glucuronate residues. In contrast to cellulose, both pectins and hemicelluloses can be branched and comprised of multiple different types of monosaccharides (Houston *et al.*, 2016).

In *Arabidopsis*, PCWs are composed of five main polysaccharides: HG (23%), RG-I (11%), RG-II (8%), XyG (20%) and cellulose (14%) (Gigli-Bisceglia *et al.*, 2020). The remaining biomass consists of less abundant polysaccharides, a network of cross-linked structural glycoproteins, proteoglycans and small amounts of phenols and minerals (Pettolino *et al.*, 2012). PCWs are also heavily hydrated with water making up to 60% of the weight (Pettolino *et al.*, 2012).

Secondary cell walls, in turn, do not contain many proteins and pectic components but rather have large amounts of cellulose, xylan and GM. Xylan is composed of 1,4-linked Xyl residues, usually substituted with 2-*O*-linked GlcA, while GM is an unbranched strand of 1,4-linked Glc and mannose (Man) residues, with 2-4 times higher Man content (Zhong *et al.*, 2019). The network of those polysaccharides in SCW is impregnated with lignin, which gives this structure rigidity, mechanical strength and increases its hydrophobicity. Due to the presence of lignin, SCWs are also entirely dehydrated (Zhong *et al.*, 2019).

Apart from cellulose, callose and mixed-linkage glucan, the majority of sugar polymers is synthesised inside the Golgi apparatus and subsequently trafficked to the cell wall matrix via vesicular transport (Schneider *et al.*, 2016, Wilson *et al.*, 2015). The synthesis of plant cell wall polymers in the Golgi apparatus is exceptionally complex and involves thousands of proteins located across multiple cellular compartments. It relies on the concerted actions of nucleoside diphosphate sugar (NDP-sugar) interconversion enzymes (NSE, **Chapter 1.2.**), nucleotide sugar transporters (NST, **Chapter 1.3.**) and glycosyltransferases (GT, **Chapter 1.4.**). The process is presented in **Figure 1** and will be described in greater detail in the next subchapters of this thesis.

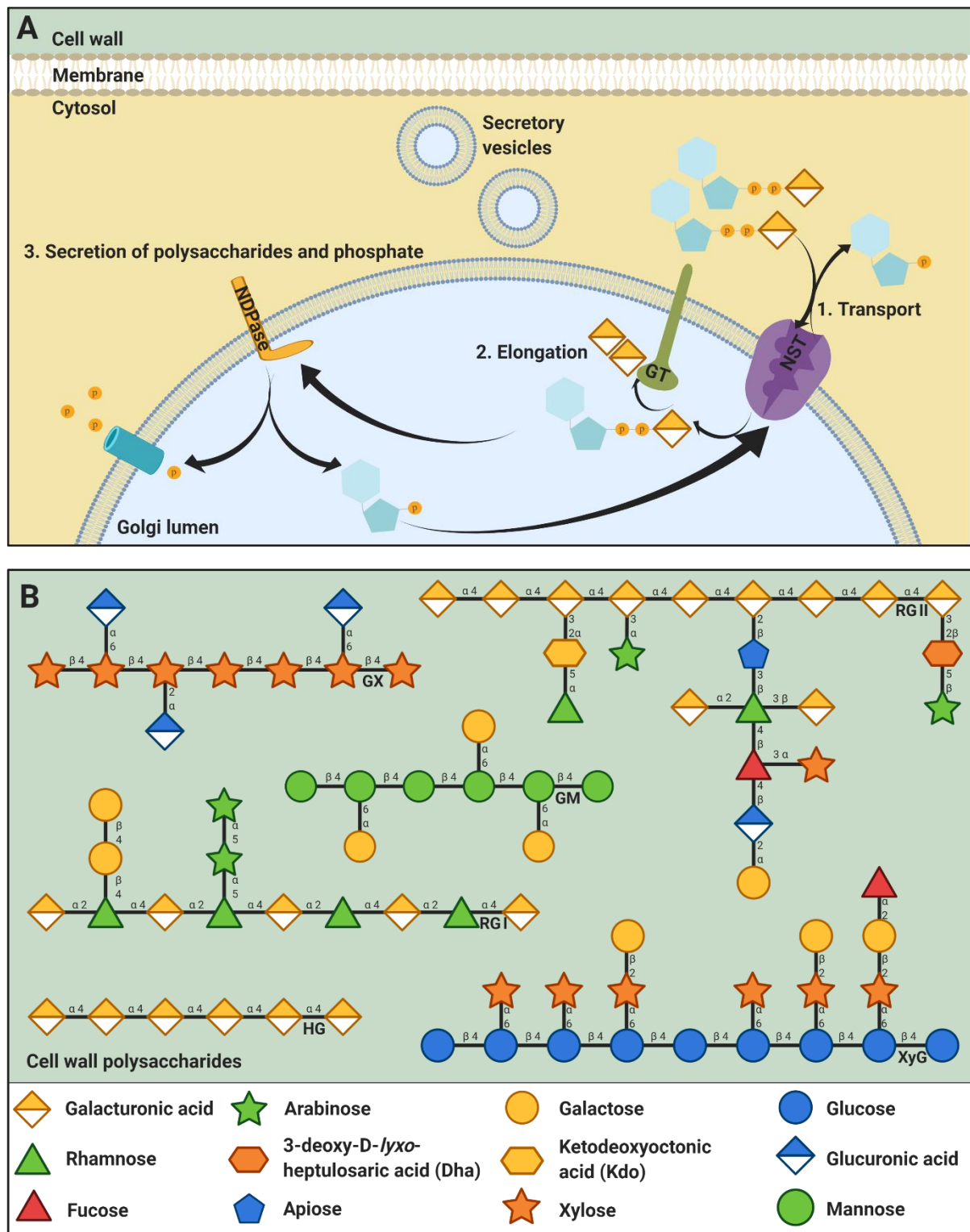


Figure 1. (A) A cartoon depicting the antiport transport of a nucleotide sugar coupled with the export of a nucleoside monophosphate, subsequent biosynthesis of cell wall polysaccharides and their transport to the cell wall via secretory vesicles. (B) Examples of pectins and hemicelluloses produced in the Golgi apparatus. Abbreviations: GM – glucomannan, GX – glucuronoxylan, HG – homogalacturonan, RG I/II – rhamnogalacturonan I/II, XyG – xyloglucan.

1.2. Nucleotide sugar interconverting enzymes (NSEs)

Nucleotide sugars are high-energy compounds serving as the sugar donors for polysaccharide biosynthesis. The energy released during the breakdown of the nucleotide sugar allows the formation of a new glycosidic bond. The majority of the 50 billion tonnes of carbon produced by plants annually during photosynthesis undergoes interconversion reactions and finally ends up in cell wall polymers. Only a relatively small portion gets incorporated into the remainder of oligosaccharides, glycolipids, glycoproteins, glycosides and storage polysaccharides (Field *et al.*, 1998, Verbančič *et al.*, 2018).

So far, over 30 different nucleotide sugars have been identified in plants (**Supplemental Table S1**); out of which 14 have been found in cell wall polysaccharides (Feingold, 1982; Mohnen, 2008). It has been established that nucleotide sugar synthesis can be achieved as a result of one of six processes: *de novo* synthesis via photosynthetic carbon incorporation, sucrose mobilisation, degradation of storage carbohydrates, the salvation of glycan residues, recycling of cell walls and sourcing sugar from plant-microbe interactions (Bar-Peled & O'Neill, 2011). After the synthesis of the initial precursor, D-fructose-6-P (Fru-6-P), the nucleotide sugars are interconverted by NSEs belonging to 11 families (see **Figure 2**). Those enzymes predominantly localise to the cytosol; however, some can also be found in the Golgi apparatus (Seifert, 2018). All NSE families are generally well conserved across algae and different plant species, but not nearly as diverse or abundant in other eukaryotic kingdoms. That suggests a distinct role of those enzymes in plants and algae that have a plant-specific cell wall (Yin *et al.*, 2011). However, it is worth noting that prokaryotes possess even more polysaccharide diversity in cell walls and capsules, which goes along with huge variability among bacterial NSEs (Herget *et al.*, 2008).

The interconverting pathways start with Fru-6-P, which is the first phosphorylated hexose produced during photosynthesis and the entry point for interconversion reactions. Fru-6-P is further converted by phosphosugar mutases and isomerases to produce Glc-1-phosphate and Man-1-phosphate (Reiter & Vanzin, 2001). From there, through activation by pyrophosphorylases, it can lead to one of two pivotal points in the pathway: GDP- α -D-Man, the precursor for most of the GDP-sugars, or UDP- α -D-Glc, the precursor of the UDP-sugars (Reiter, 2008). The combined action of hundreds of enzymes connects these pivotal points of plant sugar metabolism with their final products: UDP- α -D-Gal, UDP- α -D-GlcA, UDP- α -D-GalA, UDP- α -D-Xyl, UDP- α -D-Api, UDP- β -L-Rha, GDP- α -D-Glc, GDP- β -L-Gal, UDP- β -L-

Araf, UDP- β -L-Araf and GDP- β -L-Fuc, to name the most prominent ones. The synthesis of those building blocks is the very first step of cell wall biosynthesis.

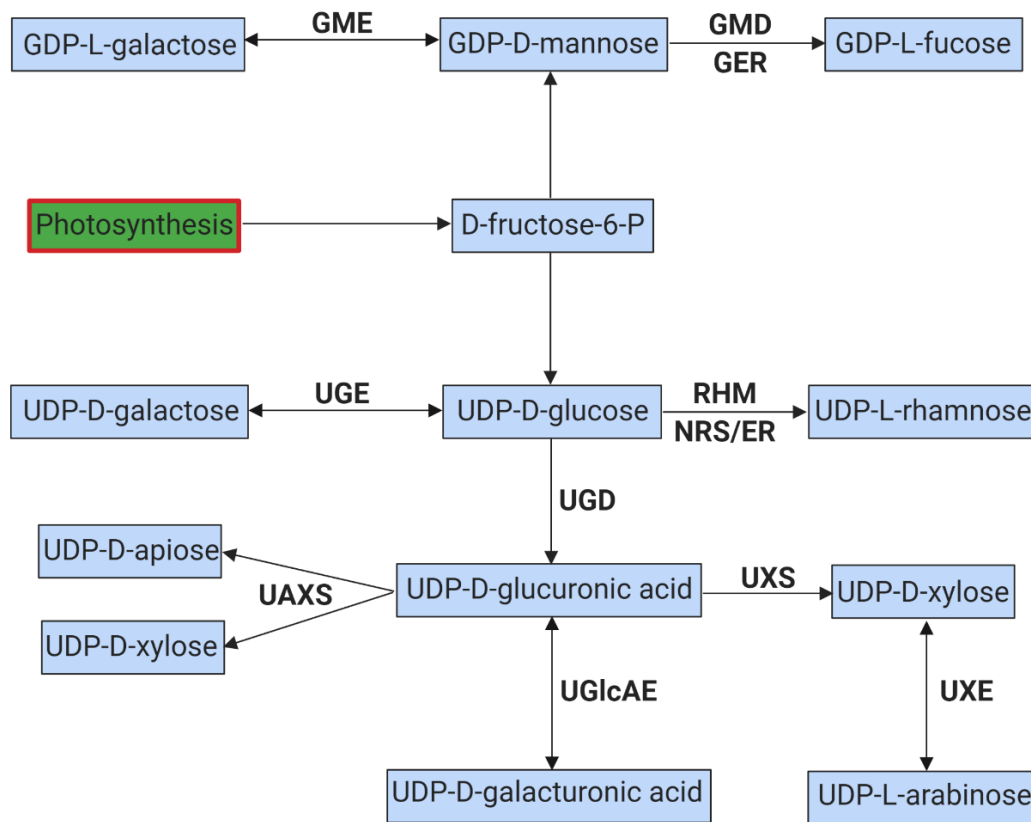


Figure 2. Reactions catalysed by 11 families of interconversion enzymes in plants. Abbreviations: GDP - guanine diphosphate; UDP - uridine diphosphate; GER - GDP-4-KETO-6-DEOXY-D-MANNOSE 3,5-EPIMERASE-4-REDUCTASE; GMD - GDP-D-MANNOSE-4,6-DEHYDRATASE; GME - GDP-D-MANNOSE 3,5-EPIMERASE; NRS/ER - NUCLEOTIDE-RHAMNOSE SYNTHASE/EPIMERASE-REDUCTASE, UER; RHM - UDP-L-RHAMNOSE SYNTHASE; UAXS - UDP-D-APIOSE/UDP-D-XYLOSE SYNTHASE, AXS; UGD - UDP-D-GLUCOSE DEHYDROGENASE; UGE - UDP-D-GLUCOSE 4-EPIMERASE; UGlcAE - UDP-D-GLUCURONIC ACID 4-EPIMERASE, GAE; UXE - UDP-D-XYLOSE 4-EPIMERASE; UXS - UDP-D-XYLOSE SYNTHASE. Adapted from Yin *et al.*, 2011.

Previous research led to a concept that the production of nucleotide sugars might not only be responsible for delivering the substrates for further reactions but also *de facto* be a regulatory step of cell wall biosynthesis. This type of control would be of crucial importance, as some highly exergonic interconversion reactions cannot be reversed *in vivo* and the pathway has to rely on the feedback inhibition of the NSEs (Reiter, 2008). Since the cell wall structure differs in diverse tissue types, it has been proposed by Seifert (2004), that its composition is regulated at the transcriptional level by differential expression of NSEs and then fine-tuned by redox control, feedback inhibition and the formation of multiprotein complexes. The same author also postulated that this regulation might be a response to stress factors, metabolic needs or plant development.

1.3. Nucleotide sugar transporters (NSTs)

NSTs are indispensable components of the plant glycosylation pathway, facilitating the transfer of nucleotide sugars across intracellular membranes and providing the full range of substrates that are a prerequisite for the action of GTs. The majority of them localise to the Golgi apparatus (Reyes & Orellana, 2008), where most polysaccharides and glycans are synthesised, although some of them have been observed to be localised in the endoplasmic reticulum (ER; Temple *et al.*, 2016) or both of those compartments (Ebert *et al.*, 2015). Their function generates a pool of nucleotide sugars inside the ER and Golgi apparatus and allows the synthesis of glycans and glycoconjugates.

NSTs are present across all eukaryotic kingdoms and despite differences, share common traits. Among them are: translocation of the nucleotide sugar in its entirety; translocation that can be saturated and leads to a 50-100-fold increase in the concentration of nucleotide sugars within ER or Golgi lumen over the external environment; translocation based on antiport mechanism, coupled to the export of nucleoside monophosphate, insensitive to ionophores and ATP concentration; translocation sensitive to competitive inhibition by mono- and diphosphates but not free sugars (Hadley *et al.*, 2014).

Structurally, NSTs are a class of integral membrane proteins present in all eukaryotes. These highly hydrophobic, homodimeric transporters consist of 300 to 350 amino acids that are organised into six to ten transmembrane helices, which as a result causes both termini to be exposed to the cytosol (Temple *et al.*, 2016, Parker & Newstead 2017). Even though some NSTs share similarities in their amino acid composition, their substrates seemed unpredictable solely on the basis of the sequence in the past; likely due to the fact that many transport activities had not been yet found and that specific motifs responsible for substrate recognition are just starting to be discovered (Parker and Newstead, 2019).

This unpredictability of NST substrate specificity based on substrate recognition motifs enforces the application of functional activity assays for their characterisation. However, in the past, this approach suffered largely from poor experimental setup, reliability issues and substrate availability. The situation changed recently when a novel platform for the high-throughput characterisation of biochemical properties and transport activities of NSTs was developed by Rautengarten *et al.* (2014). The development of this new assay resulted in the identification of novel *Arabidopsis* NST functions. Until not long ago it was also unclear how

NSTs, despite seemingly similar primary structure, recognise and transport the activated monosaccharides for which they are specific. However, in the last two years, structural studies shed light on NST specificity. Parker & Newstead (2017) observed that the substrates are bound by the yeast VRG-4 transporter in the extended conformation, with discrete pockets for the nucleotide and the sugar, allowing for separate recognition, while Hadley *et al.* (2019) discovered the importance of the formation of glycine-rich hydrophilic channels and serine and tyrosine residues in the translocation of cytidine-5'-monophospho-*N*-acetylneuraminic acid (CMP-Sia) in mammals.

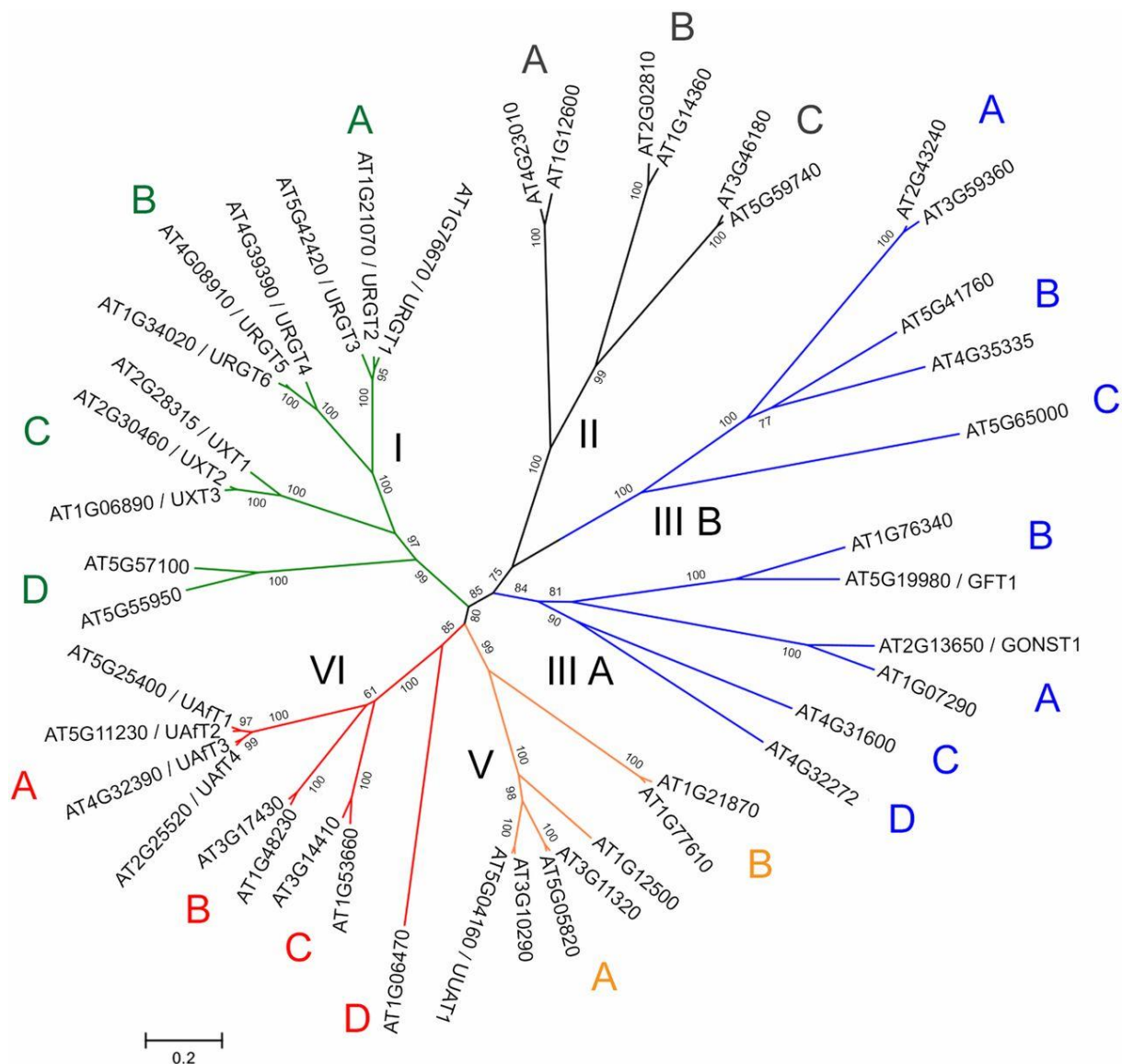


Figure 3. Phylogenetic tree of the nucleotide sugar transporter (NST) family in Arabidopsis. NSTs from clade I subclades A and B are the main focus of this thesis and will be discussed in more detail. Clade IV contains chloroplast-localised triose phosphate translocators (TPTs) and thus was omitted from the figure (image taken from Rautengarten *et al.*, 2017).

NSTs constitute one of the critical control points for cell wall composition and thus have attracted the interest of researchers for the past twenty years (Zhang *et al.*, 2011). Arabidopsis NSTs belong to the NST/triose-phosphate/phosphate translocator (TPT) superfamily consisting of 51 members, including six chloroplast-localised triose phosphate translocators (Knappe *et al.*, 2003). To date, 44 putative NSTs grouped into five subclades have been identified in Arabidopsis (**Figure 3**), out of which 18 have been functionally characterised using the transport activity assay developed by Rautengarten *et al.* (2014). This latter group includes UDP-RHAMNOSE/UDP-GALACTOSE TRANSPORTER 1-6 (URGT1-6, Rautengarten *et al.*, 2014), UDP-XYLOSE TRANSPORTER 1-3 (UXT1-3, Ebert *et al.*, 2015), GDP-FUCOSE TRANSPORTER 1 (GFT1, Rautengarten *et al.*, 2016), UDP-ARABINOFURANOSE TRANSPORTER 1-4 (UAfT1-4, Rautengarten *et al.*, 2017), UDP-URONIC ACID TRANSPORTER 1 (UUAT1, Saez-Aguayo *et al.*, 2017), UDP-*N*-ACETYLGLUCOSAMINE (GlcNAc) TRANSPORTER 1 (UGNT1, Ebert *et al.*, 2018a), and two GDP-Man transporters (GOLGI-LOCALISED NUCLEOTIDE SUGAR TRANSPORTER 1-2; GONST1-2, Baldwin *et al.*, 2001, Handford *et al.*, 2004, Mortimer *et al.*, 2013). Other transport functions have been reported for example for UDP-GLUCOSE/UDP-GALACTOSE TRANSPORTER 1-3 and 7 (UTr1-3 and 7, Norambuena *et al.*, 2002, 2005, Reyes *et al.*, 2010, Handford *et al.*, 2012), CMP-SIALIC ACID TRANSPORTER (CMP-Sia transporter, Bakker *et al.*, 2008) and UDP-*N*-ACETYLGLUCOSAMINE / UDP-*N*-ACETYLGLACTOSAMINE TRANSPORTER (ROCK1, Niemann *et al.*, 2015), however, the roles of those transporters have not been proven unambiguously and are still debated. (**Supplemental Table S2**).

Defects in the activity of NSTs are related to certain developmental disorders and various pathological phenotypes in many eukaryotic systems. In plants, phenotypical changes were described for the Arabidopsis *gft1*, *gonst1*, *uxt1 uxt2 uxt3* triple mutant, UAfT3 and UAfT4 overexpression mutants and *Osnst1*, a mutant of the putative rice UDP-Glc transporter. *gft1* RNAi lines had a large decrease of Fuc in their cell wall content, were smaller and growing slower and only ~50% survived the selection (Rautengarten *et al.*, 2016). Dwarfed phenotype and hypersensitivity to salicylic acid were reported for *gonst1* plants (Mortimer *et al.*, 2013). *uxt1 uxt3* double and *uxt1 uxt2 uxt3* triple mutants show delayed growth and have shorter stems, darker leaves, collapsed vessels and reduced SCW (Zhao *et al.*, 2018). UAfT3 and UAfT4 overexpression lead to growth retardation and early leaf senescence (Rautengarten *et al.*, 2017). Finally, *Osnst1* plants are brittle, with altered polysaccharide composition of the cell wall and irregular deposition of cellulose microfibrils (Song *et al.*, 2011).

1.4. Glycosyltransferases (GTs)

The final step of cell wall polysaccharide biosynthesis is catalysed by GTs. GTs facilitate the transfer of carbohydrate moieties to the acceptor molecules, such as polysaccharides, peptides, nucleic acids and lipids (Scheible & Pauly, 2004, Lao *et al.*, 2014). The high substrate and acceptor specificity in addition to a vast number of different GTs that have evolved in plants allows for catalysing a plethora of reactions (Stonebloom *et al.*, 2016). In contrast, some of the GTs have evolved to only provide structural support for complexes involving other active GTs, like IRREGULAR XYLEM (IRX) 9 and IRX14, which make up a complex and likely pass the substrate to IRX10 (Ren *et al.*, 2014). Another example is GALACTURONOSYLTRANSFERASE (GAUT) 7, which is suggested to be a membrane anchor for GAUT1, forming the GAUT1/GAUT7 complex (Atmodjo *et al.*, 2011). GTs that are involved in the synthesis of plant cell wall polysaccharides are predominantly type II membrane proteins localising to Golgi apparatus, with a short N-terminal tail exposed to the cytosol, a single transmembrane domain followed by a stem of inconstant length and an active site globular domain located in the lumen of Golgi apparatus. The only exceptions from this rule are GT2 and GT48 families that span the membranes multiple times and localise to either the Golgi apparatus or the plasma membrane (PM; Hansen *et al.*, 2012).

Based on their mode of action, GTs can be divided into two groups: inverting and retaining, depending on the inversion or retention of the anomeric carbon stereochemistry of the substrate. Alternatively, they can also be classified into a metal-dependent GT-A or metal-independent GT-B fold variant (Culbertson & Zabortina, 2015). Further classification into families may be done based on sequence similarity according to the CAZy (Carbohydrate Active enZyme) database (Lombard *et al.*, 2013), which as of November 2020 contains 108 GT families numbered from GT1 to GT111, with the exceptions of GT36 (now GH94), GT46 (deleted) and GT86 (merged with GT83) (www.cazy.org). To date, 463 different GTs categorised in 42 GT families and 100 GT-like genes have been identified in *Arabidopsis* (Stonebloom *et al.*, 2016; Ebert *et al.*, 2018b).

However, simple classification is not enough for a detailed prediction of GT functions. Even though due to their extreme substrate and acceptor specificity, up to hundreds of GTs participate in cell wall biosynthesis in *Arabidopsis*, only a few have been biochemically characterised to this day (Cicéron *et al.*, 2016). These include e.g. GAUT 1 (Sterling *et al.*, 2006, Atmodjo *et al.*, 2011), IRX10 and IRX10-L (Urbanowicz *et al.*, 2014), XYLOGALACTURONAN

DEFICIENT 1 (XGD1, Jensen *et al.*, 2008), FUCOSYLTRANSFERASE 1 (FUT1, Urbanowicz *et al.*, 2017), RHAMNOGALACTURONAN-1 RHAMNOSYLTRANSFERASE 1-4 (RRT1-4, Takenaka *et al.*, 2018) GLUCURONIC ACID SUBSTITUTION OF XYLAN 1-3 (GUX1, Rennie *et al.*, 2012), RHAMNOGALACTURONAN XYLOSYLTRANSFERASE 1-3 (RGXT1-3, Egelund *et al.*, 2006) and GALACTAN SYNTHASE 1-3 (GALS1-3, Liwanag *et al.*, 2012, Ebert *et al.*, 2018b). Much effort also went into the biochemical characterisation of other GTs such as GALACTURONOSYLTRANSFERASE-LIKE 5 (GATL5, Kong *et al.*, 2013), ARABINAN DEFICIENT 1 and 2 (ARAD1, 2, Harholt *et al.*, 2006, 2008) and PECTIC ARABINOGALACTAN SYNTHESIS-RELATED (PAGR, Stonebloom *et al.*, 2016); however, the activity of those enzymes could not be fully established yet.

Structural characterisation of plant GTs has remained a challenge for a long time and until 2015 no GT crystal structures had been resolved (Culbertson & Zabolina, 2015). Only recently a major breakthrough has been made when two structural analyses of FUT1 were published (Rocha *et al.*, 2016a, Urbanowicz *et al.*, 2017), which started a series of further studies in the following years.

The difficulty behind the functional characterisation of GTs is partly caused by the inability to overexpress the majority of these proteins in bacterial expression systems. With the exception of the GT1 subfamily, which is a group of soluble cytosolic proteins involved in secondary metabolite glycosylation, the majority of Arabidopsis GTs are poorly soluble, membrane-bound proteins that rely on the presence of post-translational modifications and proper folding to be active (Cicéron *et al.*, 2016). An additional problem is caused by the lack of suitable enzymatic assays mainly due to the small differences between the substrates, but also selectivity based on fine structural details of acceptors and numerous ways in which the sugars can link to the acceptor (Culbertson & Zabolina, 2015).

Finally, a vast number of GTs are represented by multiple genes encoding the same enzyme and even higher-order mutants sometimes do not display a phenotype that could hint at a specific function (Keegstra, 2010). For that reason, the structural and functional characterisation of GTs has been exceptionally challenging (Urbanowicz *et al.*, 2017). However, two approaches have expedited plant GT studies in recent years: the expression of recombinant GTs in human embryonic kidney cells, a method established at the Complex Carbohydrate Research Center (CCRC) (Subedi *et al.*, 2015, Urbanowicz *et al.*, 2017), and the development of microarrays printed with automatically assembled glycans as substrates for GT activity assays (Bartetzko & Pfrenge, 2019).

1.5. The Arabidopsis UDP-rhamnose/UDP-galactose transporter (URGT) family

Among the components of the cell wall biosynthesis machinery described in the previous three chapters, one NST family is of particular interest for the present study; the URG family first described by Rautengarten *et al.* (2014). Those bifunctional transporters belong to subclade I of the Arabidopsis NST/TPT superfamily, which has a distinct lysine-threonine motif (KT), hence the clade was named NST-KT (Knappe *et al.*, 2003). The URGs have been demonstrated to possess largely overlapping functions *in vitro*, with all six transporters being capable of transporting both UDP-Rha and UDP-Gal. However, a detailed assessment of the functions of URG1 and URG2 in a plant context, revealed that there might be additional regulatory mechanisms that are missed by the *in vitro* study. The *urg1* mutants were found to contain reduced amounts of cell wall Gal, while overexpression lines accumulated it, specifically in RG-I side chains. However, no changes in Rha content were observed. The situation is different for URG2, which has been implicated in UDP-Rha transport *in planta* but not UDP-Gal transport. The *urg2* mutants were observed to display a reduction in seed mucilage release. Rha is an essential component of the RG-I backbone, hence aberrations in seed mucilage composition could be expected as it is rich in RG-I. However, the data did not allow to conclude if URG2 is also transporting UDP-Gal *in vivo*. Interestingly, overexpression lines for all other URGs besides URG1 did also not reveal any changes to their cell wall composition. Regulation at the level of transcription is possible but seems unlikely, as publicly available expression data suggest a ubiquitous expression pattern across all plant tissues and developmental stages for those six proteins with only small variations in specific tissues and similar expression values except for URG2, which is expressed mainly in late flowers, pollen, seeds and siliques.

The results of *in vivo* studies of the URG family, indicating additional unknown regulatory factors present only in living cells, was the inspiration for this PhD project. Hence, this thesis aims to elucidate these mechanisms underlying the phenomenon of directed substrate delivery for cell wall biosynthesis.

1.6. Possible substrate channelling occurring in the Golgi apparatus of *Arabidopsis thaliana*

Most *Arabidopsis* NSTs are grouped into families with multiple members that *in vitro* use the same substrates and therefore members appear to be ostensibly redundant. However, increasing evidence from studies published in recent years links the activity of distinct members of NST subfamilies with the biosynthesis of particular sugar polymers. The observations made *in planta* hint at the separation of substrate streams coming from diverse NSTs and leading to different glycans (Ebert *et al.*, 2015; Rautengarten *et al.* 2014). Therefore, each plant NST, even the ones with identical substrate specificities *in vitro*, may play a distinct and specific role in polysaccharide synthesis or protein and lipid glycosylation rather than creating a shared pool of substrates within the Golgi apparatus.

In addition to URG1 and 2 - the two members of the URG1 family which were discussed earlier and are proposed to synthesise Gal sidechains and provide the Rha residues for the backbone of RG-I respectively - other examples suggesting substrate channelling have been reported. For example, Ebert *et al.* (2015) and Rautengarten *et al.* (2017) argue that Xyl utilised for GX biosynthesis in the Golgi is mainly derived from the cytosolic pool of UDP-Xyl which has been imported into Golgi by UXT1, while UDP-Xyl synthesised in the Golgi lumen seems to be predominantly used for UDP-Arap biosynthesis. Interestingly, Xyl-containing polysaccharides other than GX, like XyG, have been unaffected in *uxt1* mutants (**Figure 4**). It has been suggested that also UXT2 and UXT3 are involved in protein complex formation since their UDP-Xyl streams are directly utilised for xylosylation and never converted into UDP-Arap (Zhao *et al.*, 2018). Further, GONST1 has been shown to transport GDP-Man and, to a lesser extent, the rest of the GDP-sugars. However, according to the study by Mortimer *et al.* (2013), losing the transport activity of GONST1 has no impact on the cell wall composition and affects solely the mannosylation of glycosylated sphingolipids.

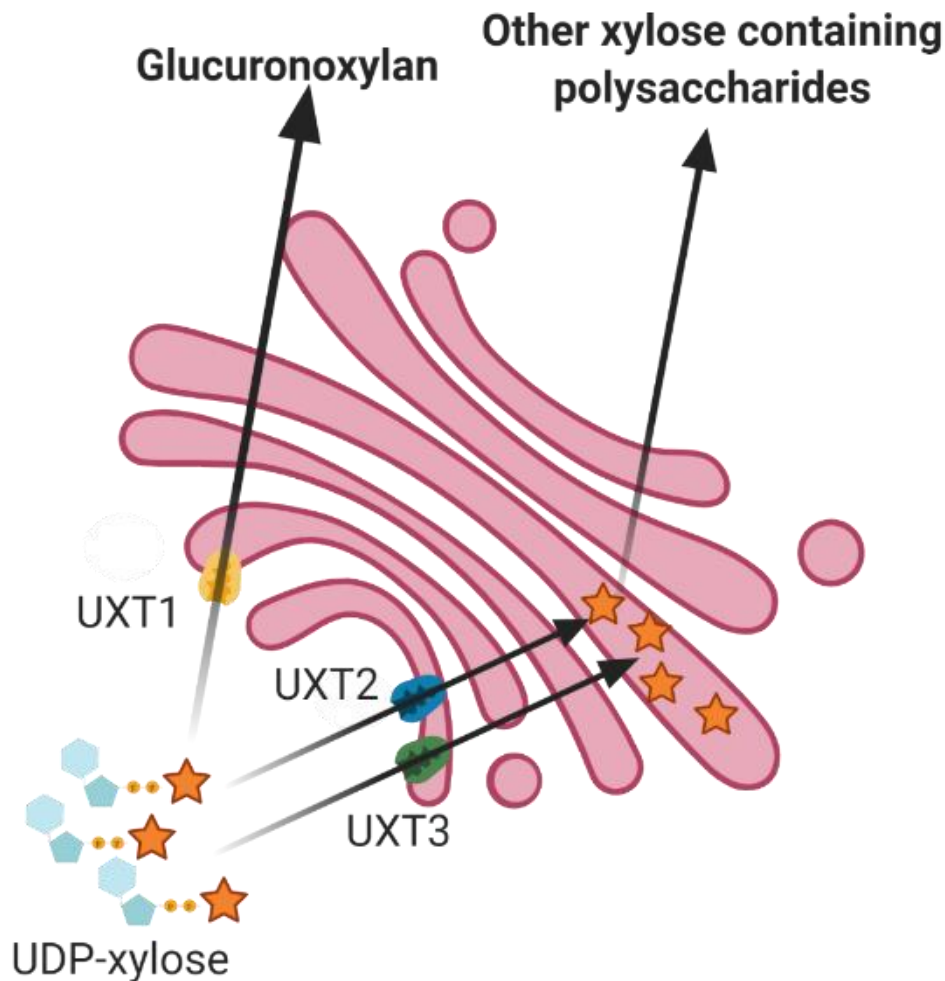


Figure 4. UXT1 seems to pass UDP-Xyl exclusively to glucuronoxylan, while UXT2 and UXT3 are currently believed to provide substrate mostly for the synthesis of other polysaccharides (Ebert *et al.*, 2015). It seems likely that those three transporters achieve their specificity through the formation of complexes with GTs.

The observations described so far indicate that there are additional factors present in the cellular context; a regulatory component or circumstances that make it possible to channel the substrate directly from the transporters to the final polysaccharide products. Such channelling could possibly be achieved through passing on the substrate directly between multiple active sites of numerous proteins in the same pathway without releasing it. Another possible way to direct substrate streams would be by enclosing the elements needed to perform all the reactions of a pathway inside a single, clearly defined compartment - in this case, Golgi cisternae - without the need for direct interactions between the components of the biosynthetic machinery (Seifert *et al.*, 2002; Jorgensen *et al.*, 2005) (**Figure 5**).

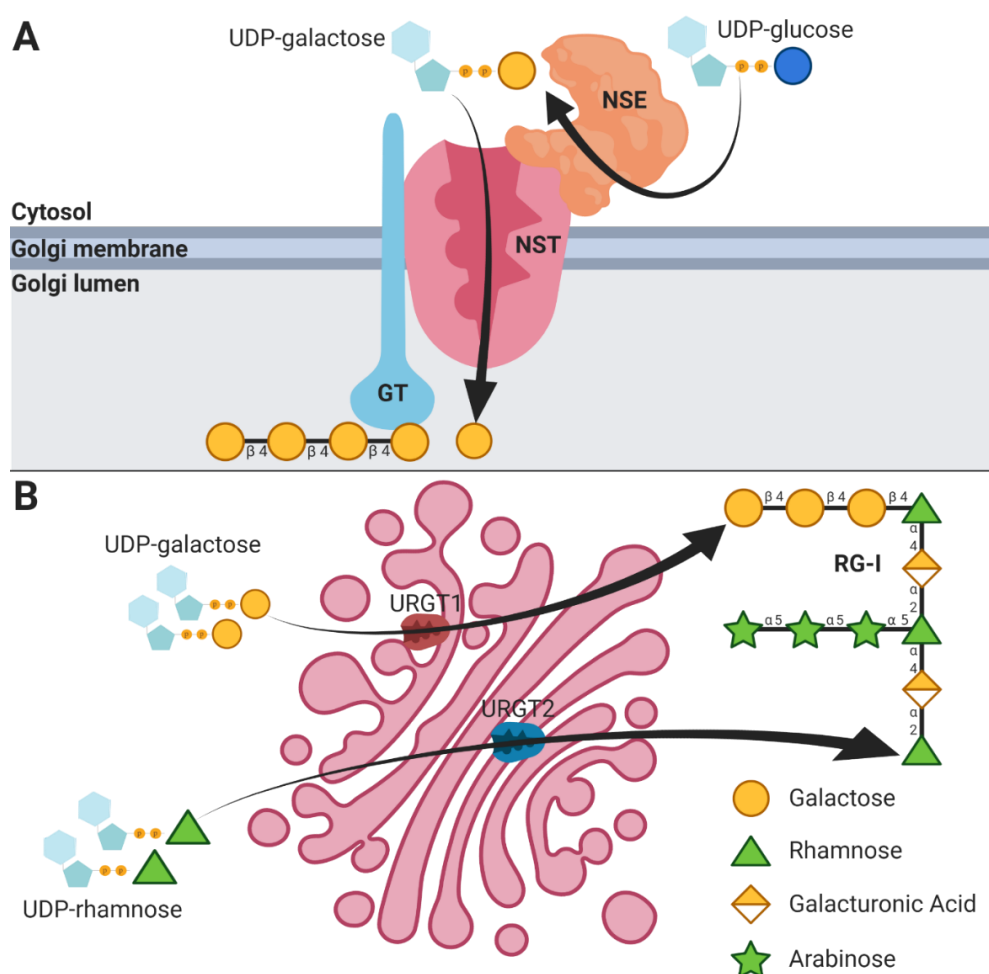


Figure 5. Two models illustrating the putative mechanisms underpinning substrate channelling in the Golgi apparatus: **(A)** metabolon formation; **(B)** spatiotemporal colocalisation of NSTs in distinct Golgi stacks with specific sets of GTs. URGT proteins have been used to illustrate the putative mechanism.

The assembly of large metabolic complexes, also called ‘metabolons’ is well established in plants. They are required for the synthesis pathways of isoprenoids, alkaloids, phenylpropanoids, flavonoids, isoflavonoids and cyanogenic glucosides (Miles *et al.*, 1999; Jørgensen *et al.*, 2005). The formation of metabolons can be transient, fragile or stable and their presence may improve the catalytic efficiency, reduces the substrate dilution, traps the toxic, inhibitory or unstable intermediates and allows for more natural control over the whole synthesis pathway (Jørgensen *et al.*, 2005). The concept involving the formation of large metabolic complexes in the Golgi apparatus based on protein-protein interactions (PPI) has been described previously by Seifert (2004, 2018). The author postulates the existence of complexes at the Golgi membrane composed of NSTs, GTs and NSEs that are guiding the substrate from its formation in the cytosol to its incorporation into specific polysaccharides in the lumen of the endomembrane system. The hypothesis was based on earlier work by the same

author (Seifert *et al.*, 2002), which had demonstrated that the flux of UDP-Gal is compartmentalised at the level of the NSEs as it is processed by diverse enzymes. Another paper by the same author established that different isoforms of UDP-GLUCOSE 4-EPIMERASE (UGEs) are involved in the production of different end products (Rösti *et al.*, 2007).

The theory that NSTs participate in complex formation with GTs is further supported by examples coming from mammalian systems, where a UDP-GlcNAc transporter seems to form a heterologous complex with a UDP-Gal transporter and mannosidic acetylglucosaminyltransferases (Maszczak-Seneczko *et al.*, 2012, Maszczak-Seneczko *et al.*, 2015) and SLC35A5 transporter homodimerises and likely interacts with other members of solute carrier family 35 (SLC35A) family (Sosicka *et al.*, 2019).

It should also be noted that over the past decade, many interactions between GTs have been described in Arabidopsis. For example, the work of Atmodjo *et al.* (2011) has shown the interaction between GAUT1 and GAUT7. In the GAUT1:GAUT7 complex, the GAUT7 domain has a purely structural function, anchoring the GAUT1 homodimer to the membrane and thus forming an active trimeric complex. Interactions between MURUS 3 (MUR3), XYLOGLUCAN L-SIDE CHAIN GALACTOSYLTRANSFERASE POSITION 2 (XLT2), FUT1, XYLOGLUCAN XYLOSYLTRANSFERASE (XXT) 2 and 5 and CELLULOSE-SYNTHASE LIKE C4 (CSLC4) have been described by Chou *et al.*, (2012, 2015) providing evidence for the formation of a XyG biosynthesis complex in the Golgi apparatus. Other examples include ARABINAN DEFICIENT 1/2 (ARAD1/ARAD2) homo- and heterocomplexes (Harholt *et al.*, 2012), GALACTOSYLTRANSFERASE 29A/31A (GALT29A/GALT31A) heterocomplex (Dilokpimol & Geshi, 2014) and heterocomplexes involved in *N*-glycan processing including α -MANNOSIDASE 1 (MNS1), ACETYLGLUCOSAMINYLTRANSFERASE 1 (GnTI), GOLGI α -MANNOSIDASE II (GMII), β -1,2-XYLOSYLTRANSFERASE (XYLT), GALT1 and FUT13 (Schoberer *et al.*, 2013). Furthermore, wheat research showed an interaction between *Triticum aestivum* glycosyltransferases TaGT43-4, TaGT47-13 and TaGT75-4, which together participate in the synthesis of glucuronoarabinoxylan (GAX; Zeng *et al.*, 2010).

Another possible explanation for the targeted substrate delivery for cell wall biosynthesis may be the spatiotemporal colocalisation of proteins delivering and utilising substrates for the synthesis of a specific polymer. The glycosylation processes in which NSTs and GTs are involved are often tightly regulated and divided into certain steps that are highly

compartmentalised and take place in specific regions of ER and Golgi stacks (Parker & Newstead, 2017). Therefore, proteins from the same family and displaying the same activity *in vitro*, placed in the surroundings of a specific subset of proteins may redirect their substrate and irreversibly connect it to the synthesis of a specific polysaccharide. This hypothesis is supported by studies showing that some cell wall polysaccharides are synthesised in stages, as they progress through the Golgi stacks. For example, HG in its unesterified form is synthesised in cis-Golgi, esterified in medial-Golgi and substituted in trans-Golgi (Zhang and Staehelin, 1992; Staehelin and Moore, 1995; Sterling *et al.*, 2001). Similarly, side chains of RG-I are added on the trans side of the Golgi apparatus while a backbone-synthesising GT colocalises with the cis Golgi marker (Zhang and Staehelin, 1992, Parsons *et al.*, 2019). Lastly, *N*-glycosylation has been well established to happen step-by-step over the course of protein passage through the ER and Golgi stacks. Glycan decorations on proteins that arrive in the cis-Golgi consist of two GlcNAc and eight or nine Man residues and are gradually altered as they move through the secretory pathway. While the glycosidases can remove existing residues, GTs residing in the cis-Golgi can add GlcNAc, in the medial – GlcNAc and Fuc, and those in the trans-Golgi and trans-Golgi network (TGN) – GlcNAc, Xyl, Fuc and Gal (Nagashima *et al.*, 2018).

1.7. Aims of the study

This study aims to identify the mechanisms that control the targeted delivery of activated sugars in plants from the cytosol to the specific end products. Since this phenomenon is linked to the step of nucleotide sugar import into the Golgi apparatus mediated by NSTs, these proteins have become the initial focus of the research project. The main emphasis was placed on the URGH family, which revealed interesting characteristics in a recent study (Rautengarten *et al.*, 2014), as described in **Chapter 1.5**. Furthermore, discovering GTs interacting with members of the URGH family can plausibly help to uncover the novel GT functions on the basis of their association with NSTs of known function.

The overarching question of this project was:

How is the channelling of the nucleotide sugars to specific glycans controlled in plants?

To tackle it, the following questions have guided the research:

- 1) Does the channelling of nucleotide sugars in plants involve the formation of stable complexes and direct protein interactions or indirect protein association through spatial colocalisation within the Golgi cisternae?
- 2) If the interactions are occurring, then which GTs are involved?
- 3) Are the members of the URGH family differentially distributed across the Golgi apparatus and where exactly do they reside?
- 4) Are transporters and their interactors localised to the same distinct Golgi subcompartments?
- 5) Can novel GTs associated with the assembly of Gal- and/or Rha-containing glycans be identified?
- 6) How would mutations to these GTs affect the plant phenotype?

To answer the questions listed above, yeast-based split-ubiquitin assays were performed to identify URGH-GT interactions and fluorescence-based methods were used to validate them. Advanced live-cell imaging and colocalisation of URGH proteins with sub-Golgi markers were undertaken to localise the URGHs and selected GTs within the specific Golgi stacks. Finally, for the interacting GTs, both overexpression and loss-of-function mutant plant lines were generated and analysed in detail.

2. MATERIALS AND METHODS

2.1. Bioinformatic tools

For Arabidopsis genome analyses, the Integrated Genome Browser (bioviz.org; Freese *et al.*, 2016) and The Arabidopsis Information Resource (TAIR) (arabidopsis.org; Huala *et al.*, 2001) were used. Primer3 (bioinfo.ut.ee; Koressaar & Remm, 2007) was used for primer design and Benchling (benchling.com) for DNA editing and analyses (including restriction digests, translation and sequence alignment) as well as a DNA sequence repository and lab notebook. Images were analysed using FIJI distribution (Schindelin *et al.*, 2012) of ImageJ2 tool (Rueden *et al.*, 2017). Gene expression levels across different tissues and under stress conditions were assessed using the Arabidopsis eFP Browser (bar.utoronto.ca; Winter *et al.*, 2007).

Protein localisations were obtained from the Subcellular Localisation Database for Arabidopsis proteins (SUBA4; suba.live; Hooper *et al.*, 2017), co-expression network analysis was conducted via PlaNet (aranet.mpimp-golm.mpg.de; Mutwil *et al.* 2011), predicted interactions were sourced from the Membrane-based Interactome Network Database (MIND; biodb.lumc.edu/mind; Chen *et al.* 2012) and the Arabidopsis Interactome Network Map (AINM; interactome.dfci.harvard.edu; Dreze *et al.*, 2011) databases.

2.1.1. Phylogeny

DUF266 proteins were identified using the Phytozome v13.0 database (Goodstein *et al.*, 2012). The closest homologs of AT1G11940 were established in five selected species: spike moss (*Selaginella moellendorffii*, Banks *et al.*, 2011), rice (*Oryza sativa*, Ouyang *et al.*, 2007), mouse-ear cress (*Arabidopsis thaliana*, Lamesch *et al.*, 2012), poplar (*Populus trichocarpa*, Tuskan *et al.*, 2006) and rose gum (*Eucalyptus grandis*, Myburg *et al.*, 2014) and used to identify other DUF266 proteins in their respective species based on sequence similarity. The identity of all candidates was confirmed by the presence of the Core-2/I-Branching enzyme domain (Pfam PF02485, Finn *et al.*, 2016). Alignment of the sequences and construction of phylogenetic trees was conducted using the MEGA-X software (Kumar *et al.*, 2018). A maximum likelihood tree was constructed using amino acid sequences aligned using the MUSCLE software with a maximum number of 16 iterations. The best-fitting model was

calculated using the built-in function of MEGA-X. The Jones-Taylor-Thornton (JTT) model was chosen as the most appropriate based on its lowest Bayesian Information Criterion score. An initial tree for the heuristic search was obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model and then selecting the topology with superior log-likelihood values. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site.

2.2. Chemicals, reagents and services

Chemicals used in this study were purchased from Sigma-Aldrich, Merck Millipore, Thermo Fisher Scientific, Ajax, Oxoid & BioRad. Amino acid supplements for yeast media were purchased from Clontech and Murashige & Skoog (MS) medium including vitamins for the plant *in vitro* culture was manufactured by Duchefa.

Plasmid DNA extraction kits were purchased from Qiagen, Macherey-Nagel and Promega. Plant RNA extraction kits and gel extraction kits were made by Qiagen. MyTaq™ polymerase and Tetro reverse transcriptase were supplied by Bioline, while restriction enzymes, Phusion® High-Fidelity DNA Polymerase and T4 ligase were acquired from New England Biolabs. TOPO cloning kits and LR clonase II were bought from Thermo Fisher Scientific. JiffyPots® for plant growth were purchased from Jiffy Products International. VectoBac WG was supplied by Sumitomo Chemical Australia.

Sanger sequencing services were provided by the Australian Genome Research Facility (AGRF) and Garvan Institute for single samples and by Macrogen for the large-scale yeast interaction screen. Primer synthesis was conducted by Sigma-Aldrich. A full list of primers used can be found in **Supplemental Tables S3** through **S7**.

2.3. Molecular techniques and materials

2.3.1. Bacterial strains, growth media and conditions

Both *Escherichia coli* (DH5α, DH10B, TOP10, DB3.1) and *Agrobacterium tumefaciens* (GV3101) strains were grown on classical lysogeny broth (LB) supplemented with 1.5% (w/v) agar for solid medium and suitable antibiotics. Super Optimal broth with Catabolite repression (SOC) medium was used for the recovery step after transformation. *E. coli* competent cells were

either bought commercially from New England Biolabs or prepared in-house using the Inoue method (Im, 2011). Electrocompetent *A. tumefaciens* cells were made according to the protocol by Höfgen & Wilmitzer (1988), with the minor modification of resuspending the cells in ice-cold water.

2.3.2. Transformation of bacterial strains

E. coli was transformed using the heat shock method, where a 50 µL aliquot of competent *E. coli* cells was combined with 100 ng of plasmid DNA, incubated on ice for 30 min, submerged in a 42°C water bath for 30 s and incubated on ice for another 2 min. The reaction mix was subsequently supplemented with 500 µL of SOC and incubated at 37 °C for 1 h under constant shaking (220 rpm) before plating half of the volume on a selective medium containing appropriate antibiotics. Bacteria were grown for 16 h at 37 °C in a plate incubator. A single colony was picked with a sterile tip and placed in a 12 mL Greiner tube with 6 mL of LB with an antibiotic. Such culture was grown in the 37°C shaker for 16-24 h followed by plasmid extraction using the Plasmid Mini Kit (Qiagen) according to the manufacturer's instructions.

A. tumefaciens cells were transformed by electroporation at 1.8 kV for 5-6 ms, using a BioRad MicroPulser and 1 mm cuvettes. After the electroporation, the cells were supplemented with 500 µL of SOC and recovered in a 28°C shaker for 2 h. Recovered cells were plated on LB plates supplemented with appropriate antibiotics, wrapped in parafilm and grown at 28°C for 48 h. Single selected colonies were used to start liquid cultures for floral dipping of *Arabidopsis* (**Chapter 2.5.8.**) and *N. benthamiana* leaf infiltration (**Chapter 2.5.10.**).

2.3.3. DNA cloning

2.3.3.1. Polymerase chain reaction (PCR)

PCRs for cloning purposes were conducted using Phusion[®] high fidelity DNA polymerase, according to the protocol recommended by the manufacturer using gDNA or cDNA as templates. Typically, a 50 µL PCR mix containing 10 µL of 5x Phusion[®] HF buffer, 1 µL of 10 mM dNTP mix, 2.5 µL of primer mix (10 µM each) 1 U of Phusion[®] polymerase and 1 µL of template DNA was combined in PCR tube strips and amplified in BioRad or Applied Biosystems thermal cyclers. Reaction mixes were typically cycled through the following

conditions: initial denaturation at 98°C for 30 s, 35 cycles of denaturation at 98°C for 10 s, annealing at 55-60°C for 30 s and elongation at 72°C for 30 s, ending with a final extension at 72°C for 10 min. Products were resolved on average for 20 min at 120 V on a 1% (w/v) agarose gel containing ethidium bromide in Tris/Acetate/EDTA (TAE) buffer and then imaged using BioRad GelDoc and Image Lab software.

2.3.3.2. Topoisomerase based (TOPO®) cloning

DNA fragments obtained by PCR served as templates for the generation of entry clones using either D-TOPO® or TOPO-TA® cloning (Invitrogen™); methods that do not involve traditional restriction/ligation approach but instead rely on spontaneous hydrogen bond formation between complementary bases. In the case of D-TOPO® cloning, primers were designed to introduce a CACC extension upstream of the PCR product, thus ensuring directional cloning, while TOPO-TA® cloning relied on additional A-tailing with Taq polymerase after the fragment was amplified with Phusion® polymerase. The cloning steps were performed according to the suggestions of the manufacturer, using half of the recommended enzyme volume. A-tailing was performed by mixing 1 pmol of purified PCR product, 1 µL of 10 mM dATP, 5 µL of 10x PCR buffer with Mg, 1 U of Taq DNA polymerase and Milli-Q water up to a reaction volume of 50 µL and incubating the reaction mix at 72°C for 30 min. Typical TOPO® reactions were done for up to 30 min at room temperature and consisted of 1 µL of salt solution, 2 µL of fresh PCR product, 2 µL of Milli-Q water and 0.5 µL of TOPO® vector (pENTR™/SD/D-TOPO® or pCR™8/GW/TOPO®, depending on the method used). 1 µL of those reaction mixes were used for amplification in *E. coli* and subsequent selection of positive transformants.

2.3.3.3. Gateway® cloning

Entry clones (as described in **Chapter 2.3.2.2.**) were used to create multiple destination clones using the Gateway® technology (Invitrogen™), a cloning system that enables the transfer of DNA fragments between plasmids through recombination of the corresponding Gateway “att” sites (Katzen, 2007). Recombination reactions were performed with LR Clonase™ II (Invitrogen™) according to the manufacturer’s instructions, using half of the recommended

enzyme volume. In cases where the resistance of entry and destination plasmid were matching, the LR reaction was conducted using PCR-amplified recombination cassette of entry clones instead of the full plasmid, as described in Kumar *et al.* (2013). A typical reaction consisted of 150 ng of entry clone, 150 ng of destination vector, TE buffer, pH 8.0 up to 8 μ L and 1 μ L of LR Clonase II. The reaction was conducted at room temperature for 3 h or overnight and then stopped by incubation with 1 μ L of Proteinase K for 10 min at 37°C. After Proteinase K treatment, 1 μ L of the reaction mix was used for transformation, plasmid selection and amplification in *E. coli*.

2.3.3.4. Classical cloning approach – restriction digest and ligation

Some of the plasmids generated in this study were generated using a traditional cloning approach, beginning with double restriction enzyme digestion. Only single cutting enzymes leaving non-compatible, sticky ends were selected, which simplified the ligation step and removed the need for vector dephosphorylation. A typical digestion reaction was composed of 1 μ L of DNA, 1 U of each restriction enzyme, 1x concentrated buffer diluted with Milli-Q water to the final volume of 20 μ L. The restriction digest was conducted for up to 4 h at 37°C and visualised on a 1% (w/v) agarose gel. The digested product was purified through a DNA clean-up procedure using a QIAprep Spin Miniprep Kit (Qiagen) or via gel extraction using QIAquick Gel Extraction Kit (Qiagen), where necessary.

A typical DNA ligation reaction was prepared by mixing digested vector backbone and insert at a 1:3 molar ratio supplemented with 1 U of T4 phage ligase and 1x T4 ligation buffer in 20 μ L of reaction volume. The reaction was incubated for 12 h or overnight in an icy water bath.

2.4. Yeast assays

2.4.1. Yeast growth media and conditions

All yeast experiments were done using *Saccharomyces cerevisiae* strain NMY51 (Lentze & Auerbach, 2008). The genotype of the strain consists of: MAT α his3 Δ 200 trp1-901 leu2-3,112 ade2 LYS2::(*lexAop*)₄-HIS3 ura3::(*lexAop*)₈-lacZ ade2::(*lexAop*)₈-ADE2 GAL4. Cells not harbouring any plasmids were grown using standard liquid and solid YPDA (Yeast extract, Peptone, Dextrose, Adenine) medium to achieve optimal growth conditions. Minimal media

(synthetic defined; SD) without leucine (SD-Leu) or both leucine and tryptophan (SD-Leu-Trp) were used to apply selective pressure during culturing of yeast possessing the bait vector alone or both bait and prey vectors respectively. SD without leucine, tryptophan, adenine and histidine (SD-Leu-Trp-Ade-His) was used to visualise the PPI through the growth of autotrophic yeast. All of the yeast liquid cultures were grown in a 30°C shaker set to 200 rpm overnight while yeast plates were kept at the same temperature for 2-4 days until colonies became visible.

2.4.2. Split-ubiquitin based interaction screen

Cells from a frozen glycerol stock of *S. cerevisiae* were plated on a fresh YPDA plate and incubated for 24 h at 30°C, thus creating a working plate, the starting point for all the subsequent yeast assays. Five independent colonies were then picked from the working plate and used to start a liquid preculture in 5 mL of YPDA. After overnight incubation in a 30°C shaker, the preculture was diluted to an optical density (OD) = 0.1 with 100 mL YPDA in a sterile, baffled flask and grown until OD₆₀₀ reached the value of 0.4-0.6 (log phase). Competent yeast cells were prepared using the LiAc heat-shock method, as described in Hill *et al.* (1991). In brief, this method involves washing yeast pellets with water to remove the residual medium, resuspending them in LiAc solution and then proceeding to heat-shock based transformation using 50% (v/v) polyethylene glycol. Such competent cells were transformed with bait vectors (pBT3-N, DualSystems Biotech) containing the URG1 and UXT genes, respectively. Cells that accepted the bait were selected on SD-Leu plates and used to start liquid SD-Leu cultures. Cultures were subsequently used for the self-activation test with 3-AT (at 0, 2.5 mM and 10 mM) and then transformed with a GT library, based on the JBEI GT collection (Lao *et al.*, 2014), cloned into the prey vector (pPR3-N, DualSystems Biotech) prepared by Dr Jennifer Bromley. Transformed cells were plated on a series of SD-Leu-Trp-Ade-His plates in 1x, 10x and 100x dilutions to select positive interactions. After 4 days of growth, the colonies served as a starter for new SD-Leu-Trp liquid cultures for plasmid purification using lyticase (Sigma-Aldrich) pre-treatment to digest the cell walls and a QIAprep Spin Miniprep Kit (Qiagen) to extract the plasmid from resulting spheroplasts. Prey plasmids were then rescued from the mix consisting of bait and prey plasmids by amplification in *E. coli* cells and selection on LB plates with ampicillin (Sigma-Aldrich). Plasmid purification from *E. coli* was followed by large scale

Sanger sequencing (Macrogen, Korea). A detailed summary of the split-ubiquitin based interaction screen is illustrated in **Figure 6**.

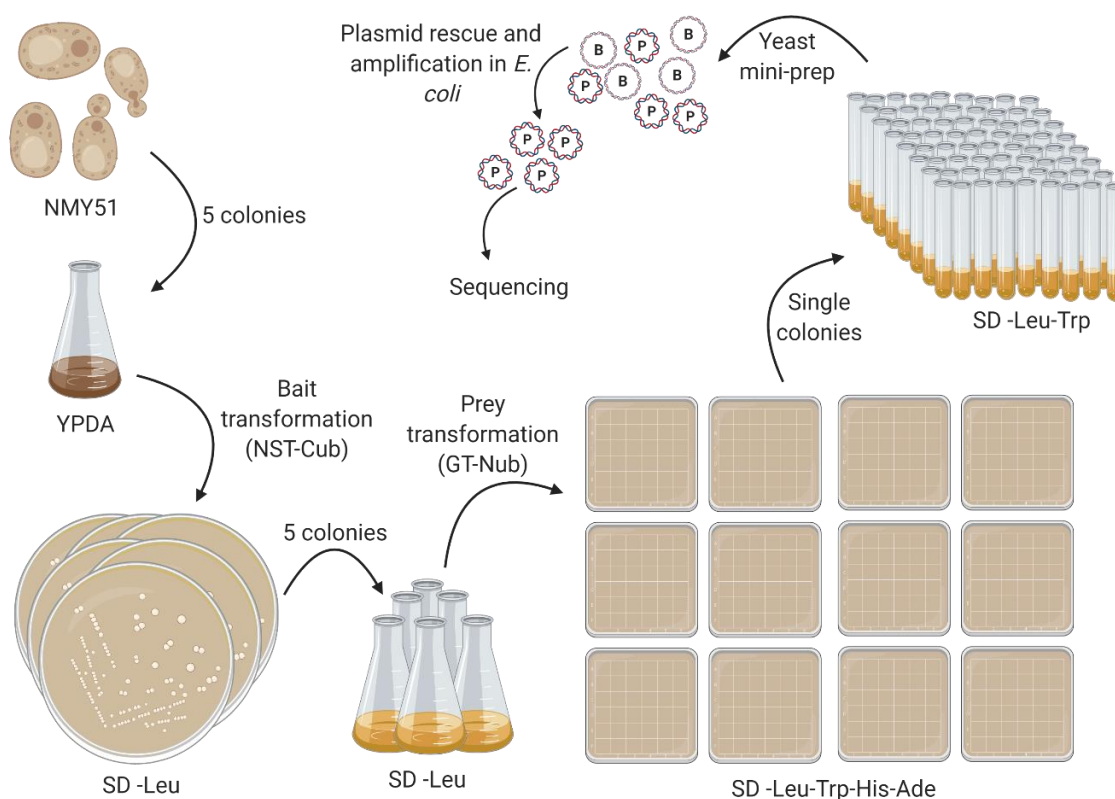


Figure 6. A cartoon depicting the workflow of the large-scale yeast split-ubiquitin interaction screen, including the transformation of NMY51 strain with bait constructs, subsequent transformation with prey library, the selection on quadruple dropout plates, preparation of liquid cultures, rescue and amplification of prey plasmids and Sanger sequencing to establish interactor identities. B indicates bait plasmid and P denotes prey plasmid.

2.4.3. Split-ubiquitin pairwise assay

Pairwise assays further tested prey proteins identified as putative interactors in the screening phase. NMY51 with bait vectors were cultured again in liquid SD-Leu medium and subsequently transformed with the respective prey vectors extracted from yeast and amplified in *E. coli*, as described in **Chapter 2.4.2**. After the transformation, the reaction mixes were plated on SD-Leu-Trp solid medium and grown for 2 days (see **Chapter 2.4.1** for details). Next, three separate colonies from each plate served as starters for 5 mL liquid cultures grown overnight. On the following morning, the liquid cultures were normalised to an $OD_{600} = 0.1$ and 10 μ L of 1x, 10x, 100x and 1000x dilutions were plated onto SD-Leu-Trp-Ade-His plates. Plates were wrapped in parafilm and kept in the yeast incubator for up to 5 days, with pictures

being taken after 48 h, 72 h, 96 h and 120 h with Canon EOS 700D on a tripod. **Figure 7** shows a summary of the procedure.

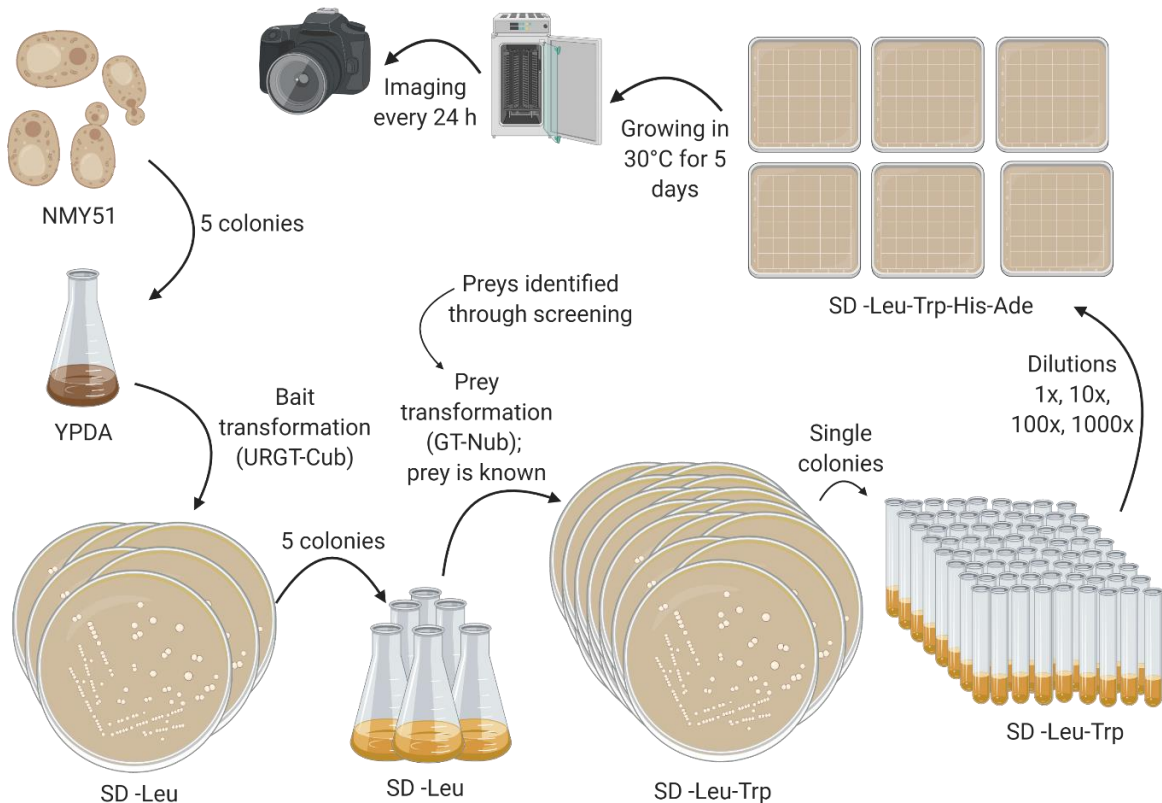


Figure 7. A cartoon depicting the workflow of the pairwise yeast split-ubiquitin interaction assay including the transformation of NMY51 strain with bait constructs, subsequent transformation with known prey plasmids, the selection on double dropout plates, preparation of liquid cultures, plating on quadruple dropout plates in series of dilutions and daily imaging of yeast growth.

2.5. Plant materials and methods

2.5.1. Plant material

Wild type *Arabidopsis thaliana* used in this study was of the Columbia-0 (Col-0) background. The T-DNA insertion lines were obtained from the Arabidopsis Biological Resource Center (ABRC; abrc.osu.edu) and the Nottingham Arabidopsis Stock Centre (NASC; arabidopsis.info). The lines investigated in this study include loss-of-function mutants for At3g11420 (SALK_036197, SAIL_237_A01, SAIL_842_C07), At2g37730 (SALK_050983), At5g23850 (SALK_067277, SALK_145225, SALK_031660C), At1g78800 (SALK_004198, SALK_014293), At3g27540 (GK-616G02), At1g68470 (WiscDsLox437E03), At5g40900 (SALK_076243), At5g44820 (SALK_007904, SALK_042539), At4g19970 (SALK_055999,

SALK_096621, SALK_139634), At1g11940 (SALK_127740, SALK_076175), At1g62305 (SALK_017871) and At5g16170 (SALK_126261, SAIL_245_B01, SALK_130475C).

For transient expression, a standard *Nicotiana benthamiana* laboratory strain was used.

2.5.2. Seed sterilisation

Arabidopsis seeds were surface sterilised by either ethanol wash or vapour-phase sterilisation before sowing on the MS plates. No sterilisation was conducted prior to planting on the soil.

For large seed amounts of single lines, the seeds were washed in 70% (v/v) ethanol for 2 min, followed by a second wash with 1.5% (v/v) bleach with 0.02% (v/v) Triton X-100 for 10 min under shaking. The bleach was decanted, and the seeds were planted after four washes with sterile Milli-Q water.

For small amounts (<100) of seeds, the vapour-phase sterilisation procedure was adapted. Seeds were placed in open Eppendorf tubes inside of the desiccator. Along with the seeds, the desiccator contained a 250 mL flask filled with 100 mL of 12.5% (w/v) sodium hypochlorite, to which 3 mL of concentrated HCl was added before sealing the lid. The sterilisation procedure was conducted for 3 h.

2.5.3. Growth conditions

2.5.3.1. *In vitro*

Sterilised seeds were grown on ½ strength MS medium, buffered with 5 mM 2-(N-morpholino)ethanesulfonic acid (MES), pH adjusted to 5.8 with KOH, with the addition of 0.7% (w/v) agar and 1% (w/v) sucrose, except where noted otherwise.

Arabidopsis seedlings were stratified for 2 days in the dark at 4°C and grown under a 24 h light regime in a Conviron growth cabinet (70% relative humidity, 21°C and 120 $\mu\text{E}/\text{m}^2\cdot\text{s}^{-1}$ light intensity) for either 3 days (root tip imaging), 7 days (root and hypocotyl measurements) or 10 days (glufosinate, kanamycin or fluorescence screens). Seeds for dark-grown hypocotyl (DGH) measurements were light-activated for 4 h prior to wrapping them up in a double layer of aluminium foil, while seeds for all other experimental approaches were grown in constant light.

2.5.3.2. Cultivation on soil

Non-sterile seeds were grown on JiffyPots[®] (Jiffy Products International). Pots were hydrated with water supplemented with VectoBac WG pesticide (Sumitomo Chemical Australia) - to minimise the damage to plants caused by insects - then covered with plastic wrap to maintain high humidity during the early growth phases. As with the plate-grown *Arabidopsis*, trays with pots were placed at 4°C in the dark for 2 days for cold stratification before moving them to a growth room or chamber.

Arabidopsis seedlings on soil were cultivated under a long-day photoperiod for crossing, selection of homozygous loss-of-function mutants and seed amplification or a short-day photoperiod for increased biomass production, which is beneficial for cell wall analyses. For long-day conditions, plants were grown in a growth room with 60% relative humidity, 20°C and 16 h of daily light with intensity ranging from 95 to 120 $\mu\text{E}/\text{m}^2\cdot\text{s}^{-1}$. The short-day regime was obtained using a Conviron growth chamber set to 10 h daily light with an intensity of 120 $\mu\text{E}/\text{m}^2\cdot\text{s}^{-1}$, 70% humidity, 18°C during the ‘night’ and 21°C during the ‘day’.

2.5.4. Extraction of genomic DNA from *Arabidopsis* leaves

Genomic DNA extraction was conducted according to the protocol from Edwards *et al.* (1991). A single young *Arabidopsis* leaf from a 21-day-old plant was harvested and ground using a Potter homogeniser in a 1.7 mL Eppendorf tube with 400 μL of extraction buffer (200 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% (w/v) SDS). Samples were vortexed and then centrifuged at 11,000 $\times g$ for 3 min in a tabletop centrifuge. Upon transfer of the supernatant into a fresh tube, 300 μL of isopropanol was added and the mixture was combined by inversion. After an 8 min incubation at room temperature, the samples were centrifuged at 11,000 $\times g$ for 5 min. Pelleted nucleic acids were washed with 70% (v/v) ethanol and centrifuged at 11,000 $\times g$ for 2 min. After removal of the ethanol, pellets were air-dried overnight in the fume hood and resuspended in 20 μL of Milli-Q water. DNA solutions were kept at -20°C until further use and 2 μL was typically used in a PCR reaction.

2.5.5. PCR for genotyping

PCR for the purpose of plant genotyping was conducted using a highly robust commercial DNA polymerase, MyTaq™ (Bioline) according to the protocol recommended by the manufacturer using genomic DNA as a template. 20 µL of the reaction mix (4 µL of 5x MyTaq™ Reaction Buffer, 0.3 µL of MyTaq™ DNA Polymerase, 2 µL of genomic DNA template, 2 µL of 10 µM primer mix, water up to 20 µL) were pipetted into the wells of a 96-well PCR plate, covered with plastic seal and then DNA was amplified in a thermal cycler. The products were visualised on a 1% (w/v) agarose gel with ethidium bromide in TAE buffer, resolved typically for 20 min at 120 V, then imaged using a BioRad GelDoc and Image Lab software. Primers used for genotyping have been presented in **Supplemental Table S6**.

2.5.6. Extraction of total RNA from Arabidopsis leaves

Total RNA was extracted from 21-day-old leaves as source material using an RNeasy Plant Mini kit (Qiagen) with DNase I (Qiagen) on-column treatment. In brief, samples were shock-frozen in liquid nitrogen in 2 mL Eppendorf tubes with tungsten beads and then ground using plant tissue homogeniser at a frequency of 18.0 min⁻¹ for 30 s. Subsequently, the beads were removed and the powdered plant material was resuspended in RLT buffer supplied with the kit, which was followed by further steps outlined in the manufacturer's protocol. RNA was eluted in 30 µL of RNase-free water and the concentration was measured using a NanoDrop microvolume spectrophotometer. All the steps were completed quickly and conducted on ice whenever possible, to avoid RNA degradation.

2.5.7. Analysis of Arabidopsis T-DNA insertional mutants by reverse transcription PCR (RT-PCR)

To confirm the total loss of transcript in T-DNA insertional lines via RT-PCR, first-strand cDNA synthesis was performed using 1 µg of total RNA extracted as described in **Chapter 2.5.6**. The synthesis step was performed directly after the RNA extraction to minimise RNA degradation with the use of a Tetro cDNA synthesis kit (Bioline) and according to the

instructions supplied with the kit. After the cDNA synthesis, cDNA and RNA were stored at -80°C until used.

For RT-PCR, MyTaq™ polymerase (Bioline) was used. 2 µL of cDNA served as a template and primers used were complementary to the full cDNA sequence of the gene (**Supplemental Table S7**). The rest of the conditions were the same as outlined in **Chapter 2.5.5**. For each line, two reactions were performed: amplification of the gene of interest and amplification of *ACTIN2* (At3g18780), which was chosen as the reference gene. WT cDNA was used as a positive control for the amplification of the full-length gene sequences. Afterwards, the PCR products were resolved on a 1% (w/v) agarose gel in TAE and visualised using a BioRad GelDoc and Image Lab software.

2.5.7.1. Analysis of Arabidopsis mutants by quantitative RT-PCR (qRT-PCR)

qRT-PCR was used for the analysis of the expression pattern of *At1g11940* and *At1g62305*, and the identification of their respective transcript levels in RNAi mutant plants.

To analyse the expression pattern, samples for RNA extraction have been harvested from a range of various tissues and subsequently reverse-transcribed into cDNA using Tetro cDNA synthesis kit. To determine transcript abundance in plants harbouring hairpin constructs, leaves from 21-day-old plants were harvested and reverse-transcribed using SuperScript IV Reverse Transcriptase. In both cases, 500 ng of RNA was used as a template for the first strand synthesis. RT-PCR was done in triplicates using SsoAdvanced Universal SYBR Green Supermix and the primers listed in **Supplemental Table S9** on a Bio-Rad CFX384™ Real-Time System equipped with the CFX Maestro Software. Relative expression levels were determined using the geNorm algorithm (Vandesompele *et al.*, 2002), which uses the geometric mean of the housekeeping genes (*UBIQUITIN 10*, At4g05320; *PROTEIN PHOSPHATASE 2A*, At1g10430; and *SAND FAMILY PROTEIN*, At2g28390). The PCR was performed with the default SYBR program recommended by the manufacturer (95°C for 30 s, 40 cycles of 95°C for 15 s and 60°C for 30 s) and the melt curve was recorded after cycle 40 by heating from 65°C to 95°C with an increment of 0.5°C per each 5 s. Relative expression was calculated using the $\Delta\Delta C_q$ method (Schmittgen & Livak, 2008).

2.5.8. Stable *Arabidopsis* transformation using the floral dip

Transformation of mature *Arabidopsis* plants was mediated by *A. tumefaciens* strain GV3101 and accomplished using the floral dip procedure described by Clough and Bent (1998). Plasmids were transferred to GV3101 *Agrobacterium* cells through electroporation (see **Chapter 2.3.2.**). After that, 5-10 colonies were used to start 100 mL liquid cultures that were grown in Erlenmeyer flasks overnight at 28°C under constant shaking. The following day, cultures were spun down at 3000 x g for 20 min and pelleted cells were resuspended in 5% (w/v) sucrose with 0.05% (v/v) Silwet L-77. The suspension was poured into a fresh wide-mouth flask and healthy *Arabidopsis* plants with young flowers and mature buds were dipped in the liquid thrice, then the excess liquid was removed, and the plants were placed horizontally on wet paper towels in a plastic tray covered with plastic wrap to maintain humidity. After the transformation, plants were grown until the siliques matured (~3-4 weeks), then the T1 seeds were harvested and screened on ½ MS plates supplemented with 15 mg/mL glufosinate (Sigma-Aldrich) or 50 mg/mL kanamycin (Sigma-Aldrich) for 10 days before transplanting the seedling to the soil.

2.5.9. Generation of genetic crosses

For the generation of genetic crosses, the earliest emerging flower buds of healthy plants were used. Before crossing, a pair of fine tweezers was used to remove all the siliques, flowers, open flower buds and apical meristem from the mother plant. Then 2-3 of the biggest mature buds were stripped of sepals, petals and anthers, leaving the pistil unprotected. Flowers from the father plant were gently squeezed to open them and then used to pollinate the pistil by touching the anthers on the stigma. The pistil was then wrapped in plastic wrap to maintain humidity, labelled and repollinated after 48 h. Swelling and elongation of the pistil was the indication for a successful cross, in which case the seeds were harvested about 25 days after the pollination and dried in paper bags.

Genetic crosses were obtained mainly to establish the subcellular localisation of proteins of interest by generating lines expressing both proteins of interest and cellular marker proteins localising to specific organelles and sub-organelle compartments (see **Chapter 2.5.9.1.** for details). Plants expressing both proteins were examined under the microscope and the

localisations of the proteins were compared in the acquired images. **Figure 8** summarises the procedure applied, from the generation of transgenic lines to image acquisition.

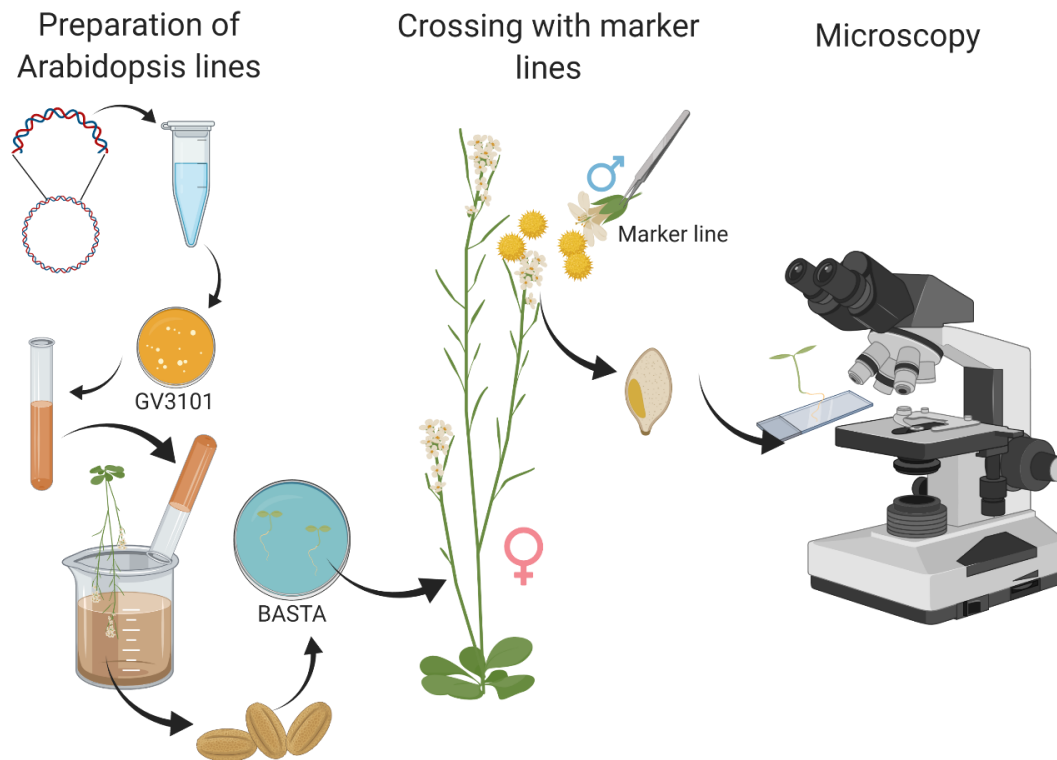


Figure 8. A cartoon depicting the workflow for the generation of transgenic Arabidopsis plants (**Chapters 2.3.2.** and **2.5.8.**) harbouring fluorescent fusion proteins of interest and crossing them with fluorescent organelle marker lines.

2.5.9.1. Transgenic Arabidopsis lines used for genetic crosses

Four different organelle marker lines have been used in this study. The GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHORED LIPID TRANSFER PROTEIN – red fluorescent protein – HDEL motif (LTPG-RFP-HDEL) served as the ER marker for colocalisation experiments. To generate the marker line, the fusion protein was cloned according to Batoko *et al.* (2000), using the first 102 amino acids from the Arabidopsis LTPG as a signal peptide (DeBono *et al.*, 2009), while RFP and the plasmid backbone were derived from pUBC-RFP-Dest (Grefen *et al.*, 2010). The HDEL ER retention signal (Napier *et al.*, 1992) was added through a flanking sequence on a primer used to amplify the LTPG-RFP cassette. Arabidopsis *N*-ACETYLGLUCOSAMINYL TRANSFERASE I - RFP (NAG-RFP) fusion served as a cis-Golgi marker and the marker line was prepared according to Grebe *et al.* (2003), also using the pUBC-RFP-Dest backbone. XYLT-RFP and ST-RFP lines (Schoberer & Strasser, 2011) served as medial- and trans-Golgi marker lines, respectively.

Genomic copies of members of the URG1 and UXT families including the ~1 kilobase pair (Kb) sequence upstream of the start of the gene were cloned into pEarleyGate301 (Earley *et al.*, 2006) modified by Dr Carsten Rautengarten. Originally, pEarleyGate301 contained a promoterless Gateway cassette fused to a hemagglutinin (HA) tag, which was replaced with a yellow fluorescent protein (YFP). This vector was used to express the proteins of interest from the URG1 and UXT families, fused to YFP under the control of their respective native promoters (**Supplemental Table S3**). Where this could not be achieved, a CDS copy of the gene (provided kindly by Dr Carsten Rautengarten) was expressed under the control of the ubiquitin-10 promoter using the pUBC-GFP-Dest vector (Grefen *et al.*, 2010).

2.5.10. Transient transformation of *Nicotiana benthamiana* leaves

Infiltration of 28 days old *N. benthamiana* plants was mediated by *A. tumefaciens* strain GV3101 according to Wydro *et al.* (2006). Plasmids were transferred into bacterial cells through electroporation (as described in **Chapter 2.3.2.**). After that, 5-10 Agrobacterium colonies were used to start 10 mL of liquid cultures that were grown in 50 mL Falcon tubes overnight at 28°C under shaking. The following day, cultures were spun down at 3000 x g for 20 min and pelleted cells were washed in infiltration buffer containing 10 mM MgCl₂, 10 mM MES pH 5.7 and 100 µM acetosyringone. The centrifugation step was repeated and the cells were resuspended in the same buffer. After measuring the OD₆₀₀ using a cuvette spectrophotometer, the suspensions were diluted to a final OD₆₀₀ between 0.03 and 0.3. Generally, a lower OD₆₀₀ was preferred in experiments, where genes were placed under strong promoters or when false positives should be avoided i.e. in interaction assays, while higher concentrations were used when genes were expressed under their native promoters. The cell suspensions were subsequently injected with a needleless syringe into the abaxial side of the *N. benthamiana* leaves.

The plants were prepared for infiltration by draining the tray 5 days before the infiltration and adding 1 L of water per tray 4 h before the infiltration in order to allow for the stomata to open and increase the solution intake. After the infiltration, the leaves were carefully labelled and plants were placed for 24 h in low-light conditions before transferring them back to standard conditions, as described in **Chapter 2.5.3.2.** After 3-4 days, 1 cm² leaf sections were cut out and imaged.

2.6. Microscopy methods

2.6.1. Light & fluorescence microscopy methods

Light microscopy was used for the imaging of seed mucilage stained with ruthenium red (see **Chapter 2.7.1.**) Stained seeds were imaged with the Leica M205FA Stereomicroscope and using the Leica LAS V4.9 software for imaging and documenting the observations.

Fluorescence microscopy observations that relied merely on the presence or absence of a signal (e.g. screening the seedlings for fluorescence) and those not requiring high resolution were also conducted using the Leica M205FA Stereomicroscope.

2.6.2. Spinning disk microscopy

More advanced fluorescence microscopy experiments, e.g. imaging of protein colocalisation, where the resolution of resulting images was crucial, were conducted using Nikon Eclipse Ti-E inverted microscope with optical autofocus system paired with an Andor Revolution CSU-W1 spinning disk and an optical autofocus system.

To prepare samples for spinning disk microscopy, different approaches were undertaken for *N. benthamiana* leaves and Arabidopsis roots. For *N. benthamiana*, a square fragment of a leaf blade was cut out, then wetted with water on both sides and mounted onto microscope slides in water with the abaxial side facing upwards. Next, a sample was topped with a coverslip and secured using a small piece of sticky tape. For Arabidopsis, 100 μ L of melted 0.8% (w/v) agarose was placed in between round coverslips, forming a sandwich. A rectangular coverslip was affixed to a steel holder with silicon grease and then three 3-day-old Arabidopsis seedlings were mounted in the middle of the holder in 5 μ L of water each, facing their root tips to the centre of the holder. One of the round coverslips was then removed from the sandwich and the seedlings were gently immobilised in the agarose disc. Slides prepared this way were subsequently immobilised in a slide holder on the microscope stage and used for imaging.

Leaf sections and root tips were observed using 60x (objective Apo TIRF MRD01691, NA 1.49) and 100x (objective Apo TIRF MRD01991, NA 1.49) magnifications and simultaneous detection in two channels using an Andor Ixon Ultra (EM-CCD) Dual-camera. Images were acquired with the help of the Metamorph software.

2.6.3. Colocalisation - imaging and analysis

Plant lines expressing both proteins of interest fused to either yellow or green fluorescent proteins (GFP) and the respective ER/Golgi markers fused to RFP, described in **Chapter 2.4.9.1.**, were grown on ½ MS plates. After 3 days, the seedlings were screened for both types of fluorescence (see **Chapter 2.6.1.**) using a stereomicroscope, before imaging was done on the spinning disk microscope. For each seedling expressing both proteins, three z-stacks were taken from the root tip area using 488 nm and 561 nm lasers for excitation and 525/50 and 610/40 filter wheels for acquisition and saved for further analysis.

Dual camera z-stacks were split into single channels using a calibration image of TetraSpeck™ Microspheres (0.1 µm, Thermo Fisher) and DualCam FIJI plugin developed by Dr Jacqueline Nowak (github.com/jnowak90/DualCam). Images were then converted to 8-bit and single cells expressing both proteins at similar levels were cropped out from them. After that, the JACoP plugin (Bolte & Cordelieres, 2006) was used to assess the Pearson correlation coefficient and to perform an object-based colocalisation analysis. The standard error of the mean (S.E.M.) was calculated by dividing the standard deviation of the sample by the square root of the sample count using Microsoft Excel.

Initial colocalisation using only ER and Golgi markers was tested using transient expression in *N. benthamiana* leaves. GV3101 Agrobacterium strains carrying the appropriate vectors, pEarleyGate101 (35S-Gateway-YFP-HA) and pEarleyGate104 (35S-YFP-Gateway) (Earley *et al.*, 2006) with genes of interest, were mixed with cultures containing agrobacteria harbouring constructs expressing markers for the ER (signal peptide from AtWAK2 fused to RFP and HDEL) or the Golgi apparatus (transmembrane domain from soybean endo-β-mannanase 1 fused to RFP) (Nelson *et al.*, 2007). The cultures were mixed in a 1:1 ratio, as described in **Chapter 2.5.10.** and imaged after 3 days.

2.6.4. Bimolecular Fluorescence Complementation (BiFC)

BiFC constructs were generated by cloning the CDS of the genes of interest into the pVYCE(R) and PVYNE vectors (Waadt *et al.*, 2008, Gehl *et al.*, 2009), encoding halves of the VENUS fluorescent protein on the N-terminus of GTs and C-terminus of NSTs respectively. Transformation and infiltration of *N. benthamiana* leaves were mediated by the GV3101

Agrobacterium strain (as described in **Chapters 2.3.2.** and **2.5.10.**). The final volume of the infiltration mix was 1 mL and the OD₆₀₀ was equal to 0.03. Imaging was conducted 3 days post infiltration using a spinning disk microscope. The fluorescence observed was compared against a known positive and negative control consisting of proteins localising to the Golgi apparatus and already known to interact, GAUT1/GAUT7 (Atmodjo *et al.*, 2011) or predicted as unlikely to interact INOSITOL PHOSPHORYLCERAMIDE GLUCURONOSYLTRANSFERASE 1 (IPUT1)/URGT1.

2.6.5. Acceptor photobleaching - Förster resonance energy transfer (APB-FRET)

APB-FRET was conducted using a Zeiss LSM880 Airyscan Fast confocal microscope and the pABindGFP, pABindmCherry and pABindFRET set of vectors (Bleckmann *et al.*, 2010). The pABindGFP vector was modified to enable GFP tag fusion on the N-terminus of GTs, while pABindmCherry was modified to allow mCherry fusion at the N-terminus of NSTs. URG1 in pABindFRET was used as a positive control for the method. Infiltration and induction were performed according to the protocol published by Somssich and Simon (2017). Protein expression was induced by spraying the abaxial side of the infiltrated *N. benthamiana* leaf with 20 µM β-estradiol 14 h and 3 h prior to imaging. Leaves were harvested right before the microscopy session and samples were prepared as described in **Chapter 2.6.2.** Cells expressing both fluorophores were selected and subsequently bleached while monitoring the fluorescence signal from both donor and acceptor.

2.6.6. Scanning Electron Microscopy (SEM)

SEM was used to visualise the morphology of seeds of selected loss-of-function mutants presumed to be involved in mucilage secretion. Samples for SEM experiments were prepared by carefully placing 15-20 mature Arabidopsis seeds on sticky carbon tape and coating them in a 1-5 nm thick gold layer using a Xenosput XE200 Sputter Coater. Imaging was conducted using an FEI/Philips XL-30 Field Emission scanning electron microscope at 5.00 kV using 500x, 1500x and 5000x magnifications.

2.7. Biochemical methods

2.7.1. Ruthenium red seed staining

Ruthenium red staining was performed to assay the defects in mucilage release in *Arabidopsis* mutants according to the protocol published by McFarlane *et al.* (2014b). About 20 seeds from a given line were placed in 800 μL of water and vigorously mixed (400 rpm) in an orbital shaker for 2 h to hydrate the seeds. Then, the solution was replaced with 800 μL of 0.01% (w/v) ruthenium red (Sigma-Aldrich) solution and incubated for 1 h. The seeds were subsequently washed once in the water, mounted in water on a glass slide and covered with a coverslip. Imaging was performed using a Leica M205FA microscope in a brightfield view. Images were acquired and later visually examined.

2.7.2. Seed mucilage composition analysis

To quantitatively analyse mucilage released from seeds, it was extracted, hydrolysed with trifluoroacetic acid (TFA) and measured by High-Performance Anion Exchange Chromatography with Pulsed Amperometric Detection system (HPAEC-PAD) as described in Voiniciuc *et al.* (2015). 5 mg of seeds from the line of interest were hydrated in 800 μL of water and mixed for 2 h in a horizontal shaker with vigorous agitation (400 rpm). Next, the samples were centrifuged and 700 μL of supernatant was transferred to a fresh screw-cap tube and dried overnight in a SpeedVacTM vacuum concentrator. Dry mucilage pellets were resuspended in 850 μL of water and vortexed well.

Subsequently, TFA hydrolysis was performed. Upon addition of 150 μL of TFA, samples were immediately thoroughly vortexed and transferred to a 120°C heat block and incubated for 1 h. Next, the samples were allowed to cool to ambient temperature, spun down and dried in a SpeedVacTM vacuum concentrator overnight. Samples were stored at -20°C until the analysis using HPAEC-PAD.

The monosaccharide composition of mucilage was analysed on an ICS3000 HPAEC-PAD (Dionex) equipped with a CarboPac PA20 column according to the protocol detailed by Rautengarten *et al.* (2016). The column was equilibrated with 10 mM NaOH for 10 min and then samples were injected and eluted with 10 mM NaOH for 15 min to separate neutral

monosaccharides. To separate acidic sugars, the concentration of NaOH was increased to 450 mM. Samples were quantitated by linear regression between the sample and a standard mix consisting of L-Fuc, L-Rha, L-Ara, D-Gal, D-Glc, D-Xyl, D-GalA and D-GlcA. Results were acquired using the Chromeleon software (Dionex).

2.7.3. Cell wall monosaccharide composition analysis

Leaf material for cell wall analyses was collected from plants grown in short-day conditions to maximise rosette growth. Plants were subjected to dark treatment for 24 h before the harvest to minimise starch content. Leaf samples were harvested from 8 mature, 4 weeks old rosettes per line, placed in screw-cap Eppendorf tubes with tungsten beads, snap-frozen in liquid nitrogen and either processed directly or stored at -80°C until further use. Plant material was ground using a TissueLyser II (Qiagen) at 18.0 min⁻¹ for 90 s. Next, the beads were removed and the powdered material was boiled in 1 mL of 96% (v/v) ethanol for 30 min, including a vortexing step after 15 min. Finally, samples were spun down in a tabletop centrifuge for 5 min at 12.000 x g. Upon ethanol removal, the pellets were washed two more times with 1 mL 70% (v/v) ethanol until the pellets became white. The last wash was done with 1 mL of acetone, followed by drying the alcohol-insoluble residue (AIR) in a SpeedVacTM vacuum concentrator overnight.

The resultant AIR was subsequently destarched. In brief, the samples were incubated at 85°C for 10 min with 150 µL of 50 mM MOPS pH=7.0 supplemented with 3U/mL of amylase. Then, 200 µL of 200 mM sodium acetate pH=4.5 with 3U/mL of amyloglucosidase and 1U/mL of pullulanase were added to the mix and the tubes were incubated at 50°C for 2 h. After that, the pellets were washed once with 96% (v/v) and once with 70% (v/v) ethanol. The pellets were dried in a SpeedVacTM vacuum concentrator overnight.

After the destarching step, AIR was resuspended in 850 µL of Milli-Q water and vortexed thoroughly. Hydrolysis was started by the addition of 150 µL of TFA, quickly screwing the cap back on, vortexing and incubating the samples at 120°C for exactly 1 h. Samples were then centrifuged for 1 min, transferred to a SpeedVacTM vacuum concentrator and dried overnight. Next, the pellets were resuspended in 1 mL of sterile Milli-Q water, vortexed and incubated at 30°C with shaking for 1 h. Samples were stored at -20°C until the analysis using HPAEC-PAD.

2.7.4. Stress treatments

Root and DGH phenotypes were evaluated by growing at least 15 seedlings of each loss-of-function mutant on ½ MS plates, as described in **Chapter 2.5.3.1**. Treatments used in these studies include: heat (30°C; applied to DGH and roots), cold (10°C; applied to DGH and roots), polyphenon-60 (PP60; green tea extract, Sigma-Aldrich, 0.0375 mg/mL applied to roots), 2,6-dichlorobenzonitrile (DCB; Sigma-Aldrich, 500 nM applied to DGH), isoxaben (Sigma-Aldrich, 2 nM applied to DGH), latrunculin-B (Sigma-Aldrich, 25 mM applied to DGH), 4.5% (w/v) sucrose (applied to DGH and roots), ionic osmotic stress (sodium chloride, 130 mg/mL applied to DGH and roots), and non-ionic osmotic stress (sorbitol, 260 mg/mL applied to DGH and roots).

2.7.5. Co-immunoprecipitation

The protocol was based on the publication by Luzarowski *et al.* (2017). Plants expressing proteins of interest fused to either GFP (pUBC-GFP-DEST, Grefen *et al.*, 2010) or HA tags (pEarleyGate301, Earley *et al.*, 2006) were grown on soil for 3 weeks before harvesting 5-6 leaves for protein extraction. Leaves were snap-frozen in liquid nitrogen and ground into a fine powder using an ice-cold mortar and a pestle. 1 volume of the powder was dissolved in 3 volumes of extraction buffer (50 mM MOPS pH=7.0 with 2 mM EGTA and 2 mM EDTA with EDTA-free protease inhibitor cocktail) by vortexing. Samples were then transferred to ultracentrifuge tubes and spun down using an Optima L-80 XP ultracentrifuge (Beckman Coulter) for 15 min at 10,000 x g at 4°C. The supernatant was collected and centrifuged again for 70 min at 100,000 x g at 4°C to collect the membrane fraction in the pellet. Pelleted proteins were then dissolved in 500 µL of the ice-cold lysis buffer (10 mM Tris-HCl, 150 mM NaCl, 0.5 mM EDTA, 0.5% (v/v) NP-40) for GFP trap or HA trap (Nano-Trap, Chromotek) application in a 1.7 mL Eppendorf tube. Then, an additional centrifugation step for 10 min at 10,000 x g at 4°C was performed using a tabletop centrifuge to get rid of the remaining cell debris. The supernatant was then transferred to fresh tubes and stored at -20°C until Co-IP was performed. The procedure was conducted on ice to prevent protein degradation. Extracted protein was quantified using a bicinchoninic acid assay (BCA) kit (Pierce™) in order to keep the amounts the same across all the samples and replicates. The resultant extracts were stored at -80°C until further analysis. At this stage, the experiment had to be stopped due to COVID-19.

2.7.6. GUS reporter-aided detection of promoter activity

Native promoter sequences upstream of the ATG start codon from genes At1g11940 and At1g62305, consisting of respectively 0.6 and 2.0 kilobase pairs, were amplified using a proof-reading polymerase. Subsequently, promoters were cloned into the pBI101.3 vector (Jefferson *et al.*, 1987) using HindIII and BamHI restriction sites upstream of the β -Glucuronidase (GUS) reporter gene. The clones were transformed into *Agrobacterium* strain GV3101 and introduced into wild type *Arabidopsis* plants.

Positive transformants were selected on MS plates supplemented with kanamycin and transplanted to the soil after the selection stage. All major plant tissues and organs, including leaves, stems, siliques, seedlings and flowers, were harvested at different developmental stages and incubated with 5-bromo-4-chloro-3-indolyl- β -D-glucuronic acid (X-Gluc, X-Gluc DIRECT) dissolved in GUS incubation solution (50 mM phosphate buffer pH 7.2, 330 mg/L K_3FeCn_6 , 0.1% (v/v) Tween-20, 0.1% (v/v) Triton-X100, 10 mM EDTA, 0.5 g/L X-Gluc) for 1-16 h at 37°C. Next, tissues were washed with 80% (v/v) ethanol until all chlorophyll was removed and stored in 80% (v/v) ethanol until imaged.

3. EXAMINING THE SUB-GOLGI LOCALISATIONS OF MEMBERS OF THE URGF FAMILY IN ARABIDOPSIS

As mentioned in the introduction chapter, an interaction between proteins or a formation of a stable protein complex is not an absolute necessity for substrate channelling to occur. Therefore, another possible explanation for the apparent redundancy between URGF family members is that they might be distinctly distributed across the different Golgi subcompartments. Differential localisation of those transporters in specific subcompartments would place them together with specialised subsets of enzymes (Parker & Newstead, 2017). This could be contributing to different substrate flows, which in turn could lead to different end products. To test this hypothesis, the specific localisation of the URGFs within the Golgi was assessed and is outlined in this chapter. For this purpose, a variety of protein markers localising to specific organelles or the distinct cisterns of the Golgi apparatus and their localisations in relation to the URGFs was examined. An identification of the exact protein localisations within the Golgi will allow determining whether distinct spatial distribution might contribute to the observed substrate channelling.

3.1. Confirming Golgi localisation of URGF family members

To confirm the Golgi localisation of the URGFs before probing their sub-Golgi distribution in more details, $P_{URGF}::URGF-YFP$ and $P_{UBQ10}::URGF-GFP$ constructs were generated (as described in **Chapter 2.5.9.1.**). While the use of native promoter constructs is optimal for the experiments described in this chapter, P_{UBQ10} constructs were created as a backup and were used both in preliminary subcellular localisation and in the case of URGF6, where the native promoter could not be cloned. Both sets of constructs were transiently transformed and expressed in *Nicotiana benthamiana* leaves. This experiment was aiming to confirm that all the proteins are properly expressed in the plant cell environment and confirm their previously shown Golgi localisation (Rautengarten *et al.*, 2014).

The spatial distribution of the $P_{UBQ10}::URGF-GFP$ constructs was assessed alongside the ER marker $P_{35S}::RFP-HDEL$ and the Golgi marker $P_{35S}::TMD(GmMan1)-RFP$ (Nelson *et al.*, 2007) to confirm the Golgi localisation of URGF proteins and is presented in **Figure 9**. Comparable results were observed with $P_{URGF}::URGF-YFP$ constructs.

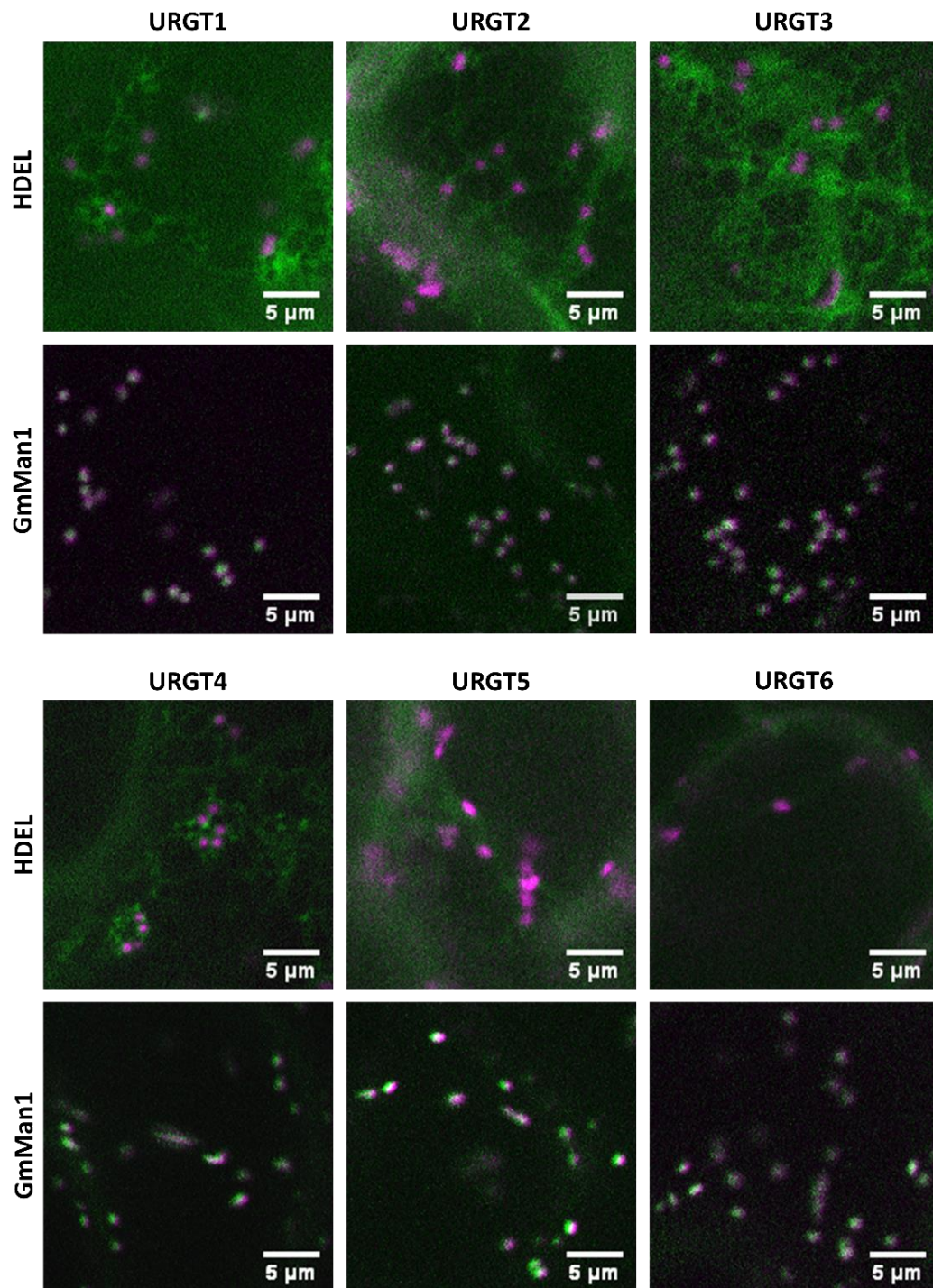


Figure 9. Subcellular localisation of the members of the URG1 family. Co-expression of *P_{UBQ10}::URG1-GFP* (magenta) and *P_{35S}::SP(AtWAK2)-RFP-HDEL* or *P_{35S}::TMD(GmMan1)-RFP* (green). All URG1s co-localise with the Golgi marker GmMan1.

Transient expression in tobacco leaves has shown that all URG1s can be observed as fast-moving, punctae-like structures, that colocalise with the Golgi marker. No signal localised to a web-like structure observed for the ER marker, thereby confirming that all URG1s localise to the Golgi apparatus.

3.2. Colocalisation of URGT family members with fluorescent markers specific for distinct Golgi subcompartments

Having confirmed that all of the URGT transporters localise to the Golgi apparatus using transient expression in tobacco leaves, stable Arabidopsis lines were generated. Those were then crossed with Golgi marker lines to create double fluorescent lines, expressing both a transporter and a selected protein marker for each of the different Golgi stacks (NAG for cis-Golgi, XYLT for medial Golgi and ST for trans-Golgi), while a transporter coupled to an ER marker (RFP-HDEL) served as a negative control. Because of the lag between the two channels of the microscope and the rapid streaming of the Golgi apparatus in the cytoplasm (Neberführ *et al.*, 1999), prior to the actual experiment, a measurement of two copies of the same marker fused to two different fluorophores was undertaken. NAG-RFP was imaged along with NAG-GFP and ST-RFP along with ST-GFP. Upon the imaging, the colocalisation of those pairs was measured (**Figure 10**). This theoretically perfect colocalisation has shown values ranging around 90%, which were used for the normalization of the data acquired from the experiment. Merged fluorescence signals generated by fusion proteins are shown in **Figure 11**.

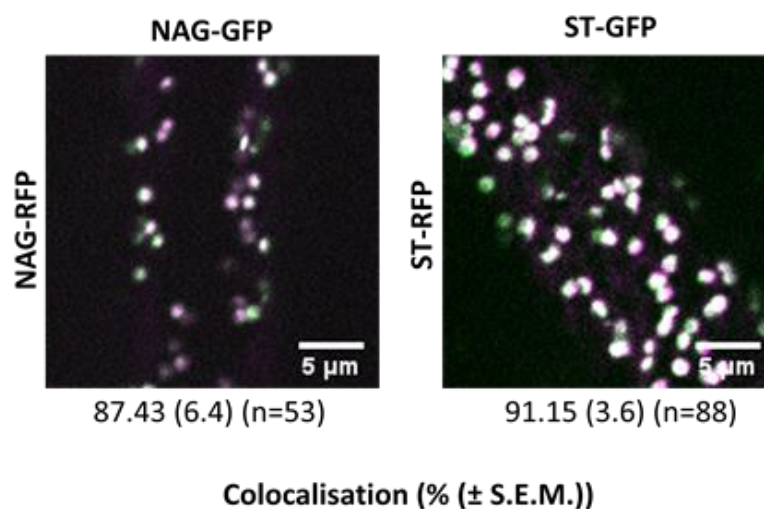


Figure 10. Colocalisation of the same Golgi cisternae marker fused to two different fluorescent proteins. Images were taken from root tips of stably transformed Arabidopsis lines with $P_{35S}::(NAG/ST)-GFP$ (magenta) and $P_{35S}::(NAG/ST)-RFP$ (green). n = number of centres of mass; S.E.M. = standard error of the mean.

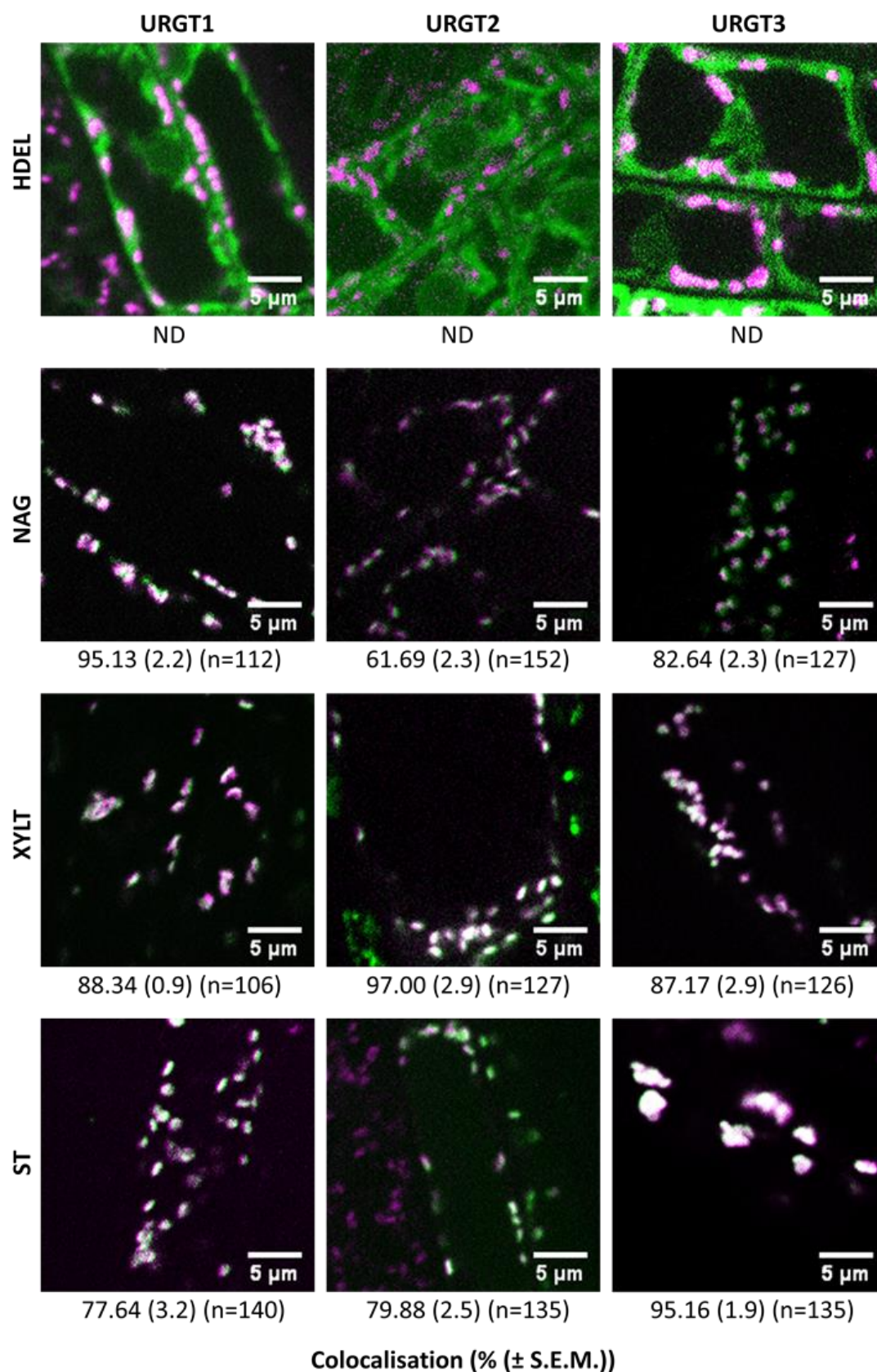


Figure 11. Fluorescence images showing the colocalisation of sub-Golgi markers and URGT proteins. Images of Arabidopsis root tips from lines transformed with *P_{URGT}::URGT-YFP* (magenta) and *P_{UBQ10/35S}::(NAG/XYLT/ST)-RFP* (green).

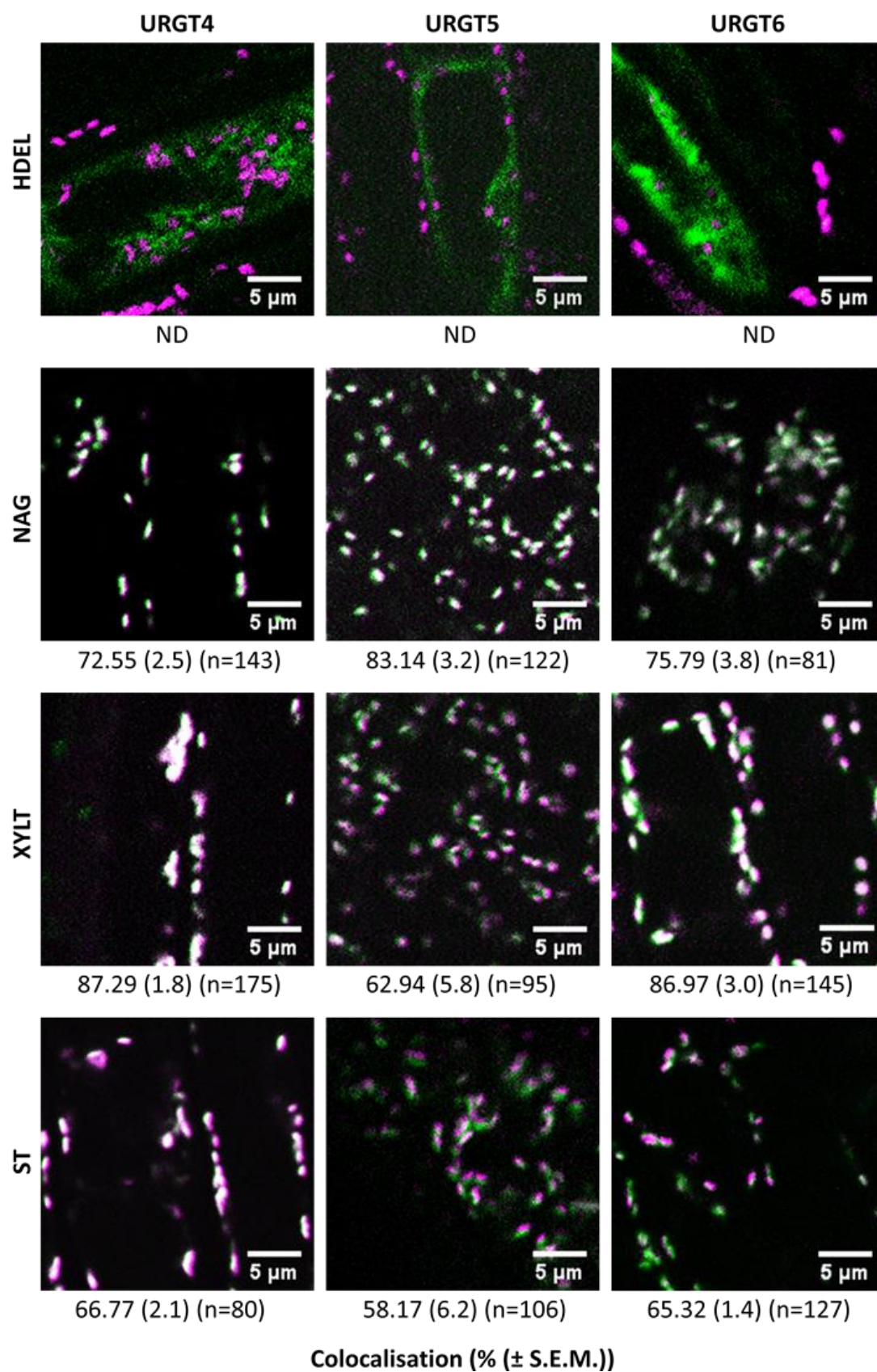


Figure 11. (continued).

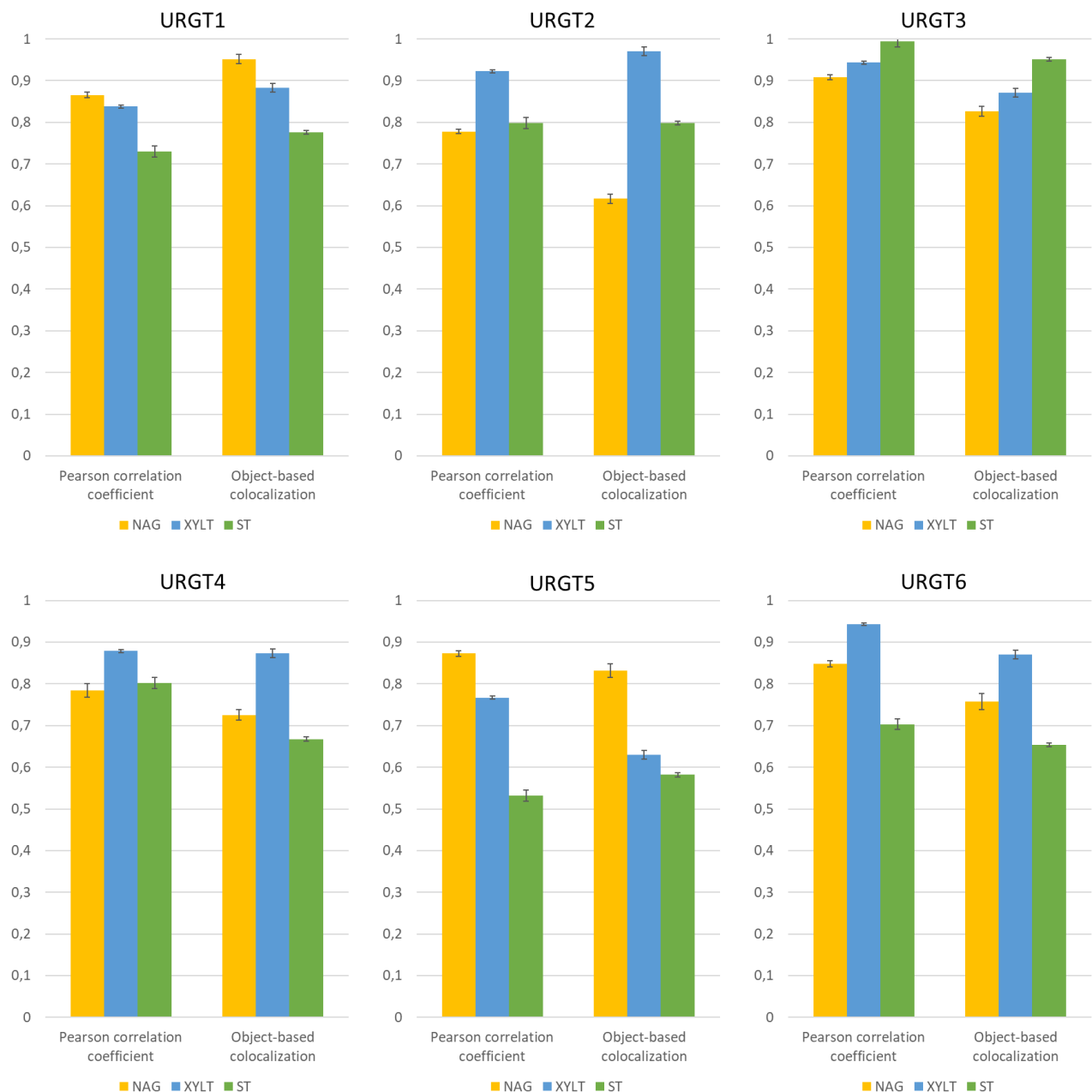


Figure 12. Average Pearson correlation coefficient and object-based colocalisation values for colocalisation of URGT family members with protein markers of respective Golgi stacks. NAG represents cis, XYLT medial and ST trans cisterns of Golgi apparatus.

Figure 12 is presenting a summary of the URGT/sub-Golgi markers colocalisation values obtained via calculation of the Pearson correlation coefficients (Manders *et al.*, 1992) and the object-based approach (Jaskolski *et al.*, 2005). The results indicate a differential distribution of the transporters across the Golgi stacks. URGT1 and 5 tend to localise mostly in the early cisterns of the cis stacks, URGT2, 4 and 6 in the medial Golgi, while URGT 3 is most often found in the trans-Golgi.

While the Pearson coefficient analyses the location and the intensity of pixels in an image, the object-based approach compares the placement of objects in the image (in this case, the Golgi bodies) using their centres of mass. Both methods led to similar results, with the object-based approach being slightly more sensitive and showing starker contrast between the main stack and the neighbouring ones.

The observed sub-Golgi distribution of the members of the URG T family (**Figure 13**) leads to several interesting conclusions. Firstly, URG T2, URG T4 and URG T6, which were proposed to be involved in UDP-Rha delivery for mucilage biosynthesis (A. Orellana, personal communication) and belong to the same co-expression networks (Mutwil *et al.*, 2011), tend to colocalise in the medial Golgi. URG T1 and URG T5, which are two out of three URG Ts found to be interacting with known GalTs from GT family 8 and GTs from the novel family GT14-L, localise in the early cisternae on the cis face of the Golgi apparatus. This observation might suggest that galactosylation reactions, for which they provide the substrate occurs in the early stages of polysaccharide biosynthesis. Finally, URG T3 is the only URG T, which localisation is shifted towards the trans-face of the Golgi apparatus and is possibly feeding its substrate to a completely different set of enzymes.

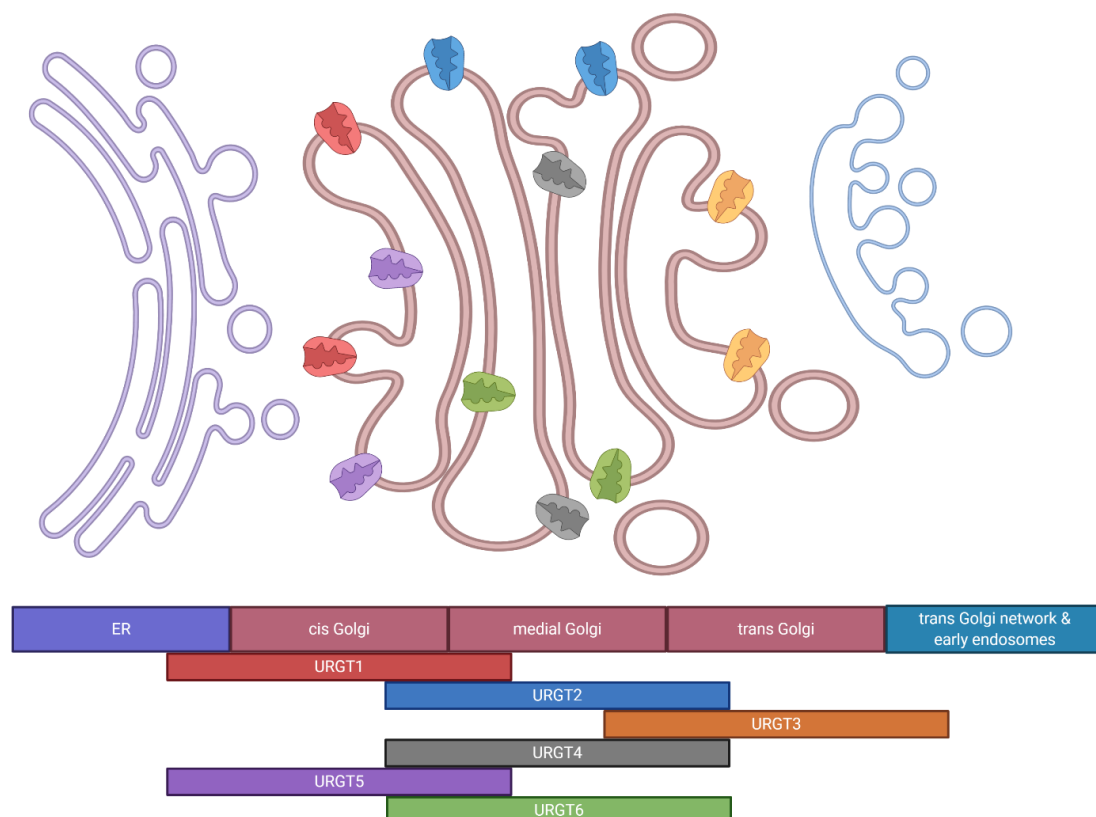


Figure 13. Illustration summarising the observed sub-Golgi localisations of the members of the URG T family in this study. Indicating distinct sub-Golgi localisation for the six URG Ts.

As a result, the data obtained in this chapter suggest that URGTs are functioning in the environment of different enzymes sharing the same substrates, and thus, their substrate specificity may be affected.

3.3. Conclusion

In conclusion, the data presented in this chapter suggest that the members of the URG T family are indeed localising to different cisterns of the Golgi apparatus, which may have an influence on the substrate of the URG Ts independently from, or in addition to, the complex formation and possible PPIs. The observed colocalisation values generally varied little between replicates of the same URG T – marker couple. Additional evidence that confirms the observed localisation of some of the URG T members in respective Golgi stacks will be presented in **Chapter 4**. However, to get the ultimate confirmation, further studies, potentially involving transmission electron microscopy (TEM) experiments, would be useful due to their high resolution. Unfortunately, TEM imaging was outside the scope of this thesis and is thus a recommendation for further work.

4. IDENTIFICATION AND VALIDATION OF GTs INTERACTING WITH ARABIDOPSIS URG1 AND UXT FAMILIES

As mentioned previously, a direct interaction between NSTs and GTs is one of the likely scenarios explaining the *in planta* regulation of nucleotide sugar flow to specific cell wall polymers; however, tangible evidence to support this hypothesis is still missing. Thus, this chapter seeks to identify and validate Arabidopsis GTs potentially interacting with members of the URG1 (URG1, At1g76670; URG2, At1g21070; URG3, At5g42420; URG4, At4g39390; URG5, At1g34020; URG6, At4g09810) and UXT (UXT1, At2g28315; UXT2, At2g30460; UXT3, At1g06890) families to collect evidence supporting the substrate channelling theory and to establish a list of candidate GTs possibly involved in Rha, Gal, and Xyl turnover.

4.1. *In silico* analysis of publicly available interaction and co-expression data

An *in silico* approach was used to survey publicly available data and to generate a list of published proteins potentially interacting with members of the URG1 and UXT families and to identify genes co-expressed with them. In all of these analyses, the focus was to identify links between GTs and URG1s/UXTs.

4.1.1. Interaction data

Interaction data was gathered from three databases: SUBA4 (literature-based; Hooper *et al.*, 2017), Arabidopsis Interactome Network Map (AINM; yeast 2-hybrid analysis-based; Dreze *et al.*, 2011) and Membrane-based Interactome Network Database (MIND; split-ubiquitin analysis-based; Chen *et al.*, 2012). SUBA4 and AINM present only information for URG1, assigning only 6 and 8 interactors respectively. However, none of the predicted interactors is a GT or NST.

MIND, a database summarising an interactome-wide split-ubiquitin screen of Arabidopsis protein interactions, provides the most extensive source of putative interactions, naming 101

possible interaction partners for URG3, 8 for URG5, 417 for UXT1 and 71 for UXT3. URG3 is shown to interact with two other NSTs, UDP-URONIC ACID TRANSPORTER UUAT3 (At5g05820) and UUAT5 (At1g12500; Saez-Aguayo *et al.*, 2017). The only GT found to be potentially interacting with URG3 is CELLULOSE SYNTHASE 3 (CESA3, At5g05170; Burn *et al.*, 2002), which likely represents an artefact, as members of the GT2 family are typically located in the PM. UXT1 is predicted to interact with the same two NSTs as reported for URG3, but also with itself (At2g28315), UTR2 (At4g23010) and UTR3 (At1g14360) – putative UDP-Gal transporters (Norambuena *et al.* 2002). In addition, UXT1 has been shown to pair with five GTs: CESA2 (At4g39350; Burn *et al.*, 2002), CESA3, PLANT GLYCOGENIN-LIKE STARCH INITIATION PROTEIN 7 (PGSIP7; At2g35710; Yin *et al.* 2010), GT14-L protein (At1g68380) and PAGR (At3g26370; Stonebloom *et al.*, 2016). While both CESAs seem to be implausible interactors *in planta* due to their PM localisation and the substrates used, the remaining three are interesting in relation to the split-ubiquitin results discussed later in this thesis. Lastly, UXT3 does not bind any other NSTs and – similarly to UXT1 – seems to interact with PAGR. A complete set of results can be accessed under biodb.lumc.edu/mind, while the interactions discussed in this paragraph can be found in **Supplemental Table S8**.

4.1.2. Co-expression data

The investigation of co-expression networks of genes related to cell wall synthesis has proven to be a valuable tool in the past (Persson *et al.*, 2005). It can be used to predict the functions of genes based on their links to co-expression networks and detect functional homologues, as well as conserved pathways across plant species (Mutwil *et al.*, 2011). In many cases, distinct subunits of protein complexes tend to be expressed together to preserve stoichiometric ratios and assure proper assembly (Liu *et al.*, 2009). For that reason, available information from co-expression networks of the *URG*T and *UXT* families was reviewed and the correlation between them, other NST (see **Table 1**) and GT genes (**Table 2**) was summarised. Information on *Arabidopsis thaliana* co-expression networks was obtained from the PlaNet database (Mutwil *et al.*, 2011). No co-expression results were obtained for UXT1 since the ATH1 array (Applied Biosystems™) used for expression studies does not contain a probe for At2g28315.

Table 1. Summary of co-expression network data and published putative interactions between NSTs and members of the URG T and UXT families in *Arabidopsis thaliana* according to PlaNet and MIND databases (Chen *et al.*, 2012, Mutwil *et al.*, 2011).

Sub-clade	AGI	Description	NST name	Co-expressed with	Published putative interactors
IA	AT1G21070	UDP-RHAMNOSE/UDP-GALACTOSE TRANSPORTER 2	URGT2	URGT4	-
IB	AT4G39390	UDP-RHAMNOSE/UDP-GALACTOSE TRANSPORTER 4	URGT4	URGT2, UXT3	-
	AT4G08910	UDP-RHAMNOSE/UDP-GALACTOSE TRANSPORTER 5	URGT5	UXT2	-
IC	AT2G30460	UDP-XYLOSE TRANSPORTER 2	UXT2	URGT5	-
	AT1G06890	UDP-XYLOSE TRANSPORTER 3	UXT3	URGT4	-
IIA	AT4G23010	UDP-GALACTOSE TRANSPORTER 2	UTR2	-	UXT1
IIB	AT1G14360	UDP-GALACTOSE TRANSPORTER 3	UTR3	URGT4	UXT1
IIIAB	AT1G76340	GOLGI NUCLEOTIDE SUGAR TRANSPORTER 3	GGLT1	URGT3, URG T4	-
	AT5G19980	GDP-FUCOSE TRANSPORTER 1	GFT1	URGT3, URG T4	-
IIIAC	AT4G31600	UDP-GALACTOSE TRANSPORTER 7	UTR7	URGT5, UXT2	-
IIIBB	AT4G35335	-	-	URGT3	-
VA	AT5G04160	UDP-URONIC ACID TRANSPORTER 1	UUAT1	URGT2, URG T4	-
	AT5G05820	UDP-URONIC ACID TRANSPORTER 3	UUAT3	URGT2, URG T4	URGT3, UXT1
	AT3G11320	UDP-URONIC ACID TRANSPORTER 4	UUAT4	URGT4	-
	AT1G12500	UDP-URONIC ACID TRANSPORTER 5	UUAT5	URGT1	URGT3, UXT1
VIA	AT5G11230	UDP-ARABINOFURANOSE TRANSPORTER 2	UAFT2	URGT2, URG T4	-
VIB	AT3G17430	-	-	URGT6	-
	AT1G48230	-	-	URGT6	-
VID	AT1G06470	-	-	URGT4	-

Based on the co-expression analysis, two groups seem to form within the clade I of the *Arabidopsis* NST-TPT family. The first one consists of *URGT2-URGT4-UXT3* and the second one of *URGT5-UXT2*. The members of those groups also appear to often share co-expression partners, such as *UUAT1*, *UUAT3* and *UAFT2* are co-expressed with members of the *URGT2-URGT4-UXT3* group and *UTR7* with both *URGT5-UXT2*. Interestingly, the co-expression cluster including *URGT2*, *URGT4* and *UXT3* was also supported by tissue expression maps available in the eFP browser (Winter *et al.*, 2007). Those three transporters are expressed ubiquitously across all tissues, but the transcript abundance is twice as high in flowers, petals and root apex. *URGT2* and *UXT3* are also both upregulated in response to salt treatment. The expression pattern of *UXT2* is in turn comparable to that of *URGT5*. A recent study by Parra-Rojas *et al.* (2019) further confirmed those observations, by showing that in the *urgt2-2* loss-of-function mutant transcripts of *URGT4*, *UXT3*, *UUAT1*, *UUAT3* and *UAFT2* were all significantly upregulated to compensate for lost *URGT2* activity. In the same study, *URGT2*, *URGT4*, *UUAT1* and *UUAT3* have been implicated in mucilage synthesis, providing Rha and GalA for the synthesis of RG-I, the main component of *Arabidopsis* seed mucilage. At the same time, *URGT1* seems to have few co-expression partners among NSTs and the network of *URGT6* appears to be very specific, with its partners coming exclusively from the poorly characterised VI B subclade and are not being shared with other *URGTs*.

Table 2. Summary of co-expression network data and published putative interactions between GTs and members of the URGT and UXT families in *Arabidopsis thaliana*, according to PlaNet and MIND databases (Chen *et al.*, 2012, Mutwil *et al.*, 2011). GT NC denotes non-classified, a group of GTs that are not yet categorised in the CAZy database into any of the known GT families.

GT family	AGI	Description	GT name	Co-expresses with	Published putative interactors
GT1	AT4G18230	UDP-N-ACETYLGLUCOSAMINE TRANSFERASE SUBUNIT ALG14-LIKE PROTEIN	ALG14-L	URGT1	-
GT2	AT2G39630	PUTATIVE DOLICHYL-PHOSPHATE B-GLUCOSYLTRANSFERASE	-	URGT3, UXT3	-
	AT3G07330	CELLULOSE SYNTHASE-LIKE C6	CSLC6	URGT4, UXT3	-
	AT4G32410	CELLULOSE SYNTHASE 1	CESA1	URGT1	-
	AT4G39350	CELLULOSE SYNTHASE A2	CESA2	URGT1	UXT1
	AT5G03760	CELLULOSE SYNTHASE-LIKE A9	CSLA9	URGT1	-
	AT5G05170	CELLULOSE SYNTHASE 3	CESA3	URGT1	URGT3, UXT1
	AT5G09870	CELLULOSE SYNTHASE 5	CESA5	URGT1	-
	AT5G64740	CELLULOSE SYNTHASE 6	CESA6	URGT1	-
GT4	AT1G19710	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	-	URGT6	-
GT8	AT1G02720	GALACTURONOSYLTRANSFERASE-LIKE 5	GATL5	URGT2, URG4	-
	AT2G20810	GALACTURONOSYLTRANSFERASE 10	GAUT10	URGT6	-
	AT2G35710	PLANT GLYCOGENIN-LIKE STARCH INITIATION PROTEIN 7	PGSIP7	-	UXT1
	AT2G38650	GALACTURONOSYLTRANSFERASE 7	GAUT7	URGT6	-
	AT3G25140	GALACTURONOSYLTRANSFERASE 8	GAUT8	URGT6	-
	AT5G15470	GALACTURONOSYLTRANSFERASE 14	GAUT14	URGT2	-
	AT5G47780	GALACTURONOSYLTRANSFERASE 4	GAUT4	URGT6	-
GT14	AT1G71070	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	-	URGT6	-
GT29	AT2G37585	GLUCURONIC ACID TRANSFERASE 14C	GLCAT14C	URGT4, URG6	-
	AT1G08280	GLYCOSYLTRANSFERASE 29A	GALT29A	URGT6	-
GT31	AT1G53290	GALACTOSYLTRANSFERASE FAMILY PROTEIN	-	URGT4, URG6	-
GT34	AT5G53340	HYDROXYPROLINE O-GALACTOSYLTRANSFERASE 1	HPGT1	URGT6	-
	AT4G02500	UDP-XYLOSYLTRANSFERASE 2	XT2	URGT1	-
GT57	AT5G38460	ASPARAGINE-LINKED GLYCOSYLATION 6	ALG6	URGT5	-
GT58	AT2G47760	ASPARAGINE-LINKED GLYCOSYLATION 3	ALG3	URGT4, URG5	-
GT59	AT5G02410	ASPARAGINE-LINKED GLYCOSYLATION 10	ALG10	URGT3, URG4, URG5, UXT2	-
GT76	AT1G11880	-	-	URGT3, UXT2	-
GT77	AT4G01220	MALE GAMETOPHYTE DEFECTIVE 4	MGP4	URGT3, URG5, UXT2, UXT3	-
GT92	AT2G33570	GALACTAN SYNTHASE 1	GALS1	URGT1	-
GT NC	AT1G11940	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	-	URGT2	-
	AT1G14020	O-FUCOSYLTRANSFERASE FAMILY PROTEIN	RRT4	URGT2, URG4	-
	AT1G68380	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	-	-	UXT1
	AT2G04280	CALCIUM-ION BINDING PROTEIN	-	URGT6	-
	AT2G41770	STELLO 1	STL1	URGT1, URG4, URG6	-
	AT3G02250	O-FUCOSYLTRANSFERASE FAMILY PROTEIN	RRT2	URGT1	-
	AT3G08550	KOBITO	KOB1	URGT4	-
	AT3G21190	MANNAN SYNTHESIS RELATED 1	MSR1	URGT6	-
	AT3G26370	PECTIC ARABIBOGALACTAN SYNTHESIS-RELATED	PAGR	-	UXT1, UXT3
	AT4G25870	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	-	URGT6	-
	AT4G38500	-	-	URGT4	-
	AT3G57220	-	-	URGT3	-
	AT5G50420	O-FUCOSYLTRANSFERASE FAMILY PROTEIN	-	URGT6	-

Co-expression relations between *URGT* and *UXT* families and GTs seem to show some patterns. *URGT1* is expressed predominantly with members of the GT2 family, which consists of CESAs: a group of cellulose synthases localising mostly to the PM. Although surprising, a hypothesis explaining this observation might be the temporal convergence of cellulose and RG-I synthesis, as *URGT1* was proposed to be involved in the synthesis of the latter. This theory might be supported by the fact, that despite the same co-expression network, no interaction has been observed between *URGT1* and CESAs, and their activity focuses on different substrates. Hence, they are not likely to be involved in the same processes. Another GT of the *URGT1* network is *GALS1*, a UDP-Gal-dependent GT known to elongate galactan sidechains of RG-I (Liwanag *et al.*, 2012). The specific involvement of *URGT1* in the biosynthesis of galactan sidechains has been shown by Rautengarten *et al.* (2014, see **Chapter 1.5.** for details), which highlights a functional link between these two. *GALS1* has also been found to interact with

URGT1 by Dr Jennifer Bromley (personal communication, see **Table 3.**), making the interaction or even complex formation between them highly likely.

GTs belonging to families 4, 8, 14, 29 & 31 could only be found in the networks of *URGT2*, *URGT4* and *URGT6*, which, as mentioned previously, are presumed to be involved in RG-I synthesis, which is a component of seed coat mucilage. Interestingly, members of the GT8 family – GATL5 (Kong *et al.*, 2013) and GAUT11 (Voiniciuc *et al.*, 2018) – are known galacturonosyltransferases, and both have been shown to also play a role in seed mucilage synthesis, while GT14, 29 & 31 are all involved in the synthesis of arabinogalactan protein (AGP) (Knoch *et al.*, 2014). Lastly, *URGT5* is expressed alongside genes encoding members of GT57, GT58 and GT59, which are involved in *N*-glycosylation (Farid *et al.*, 2011).

Unfortunately, the available data is scarce for the *UXT* gene family. Among the few genes co-expressed with members of the *UXT* family, one is encoding the MALE GAMETOPHYTE DEFECTIVE 4 (MGP4) protein, a well-characterised RG-II xylosyltransferase (Liu *et al.*, 2011). This co-expression could hint at an interaction between the two, as both share UDP-Xyl as substrate.

4.2. Identification of GTs interacting with NSTs using the split-ubiquitin screen

To identify interaction partners of the *Arabidopsis* URG and UXT families experimentally, the split-ubiquitin system (Stagljar *et al.*, 1998) was applied. This yeast-based approach allows for a high-throughput screen of large libraries containing proteins that interact with a bait of choice. Bait and prey proteins are fused to halves of ubiquitin, and upon their binding functional ubiquitin is reconstituted. Cleavage of reassembled ubiquitin releases a transcription factor fused downstream of ubiquitin C-terminus, protein A-LexA-VP16 (PLV). The method is similar to the yeast-2-hybrid system (Fields & Song, 1989); however, the fact that the bait protein has to be insoluble and membrane-anchored, makes it a perfect system for analysis of membrane protein interactions.

The split-ubiquitin assay was performed as described in **Chapter 2.4.2.**, using nine bait proteins – URG1-6 and UXT1-3 – and a prey library containing a broad set of 403 GTs based on the JBEI GT Collection (Lao *et al.* 2014), which was prepared by Dr Jennifer Bromley at JBEI. N-terminal tags for both GTs and NSTs were used, as GTs are in the vast majority type-

II membrane proteins, exposing the N-terminus to the cytosol while the active C-terminus faces the lumen (Oikawa *et al.*, 2013). NSTs on the other hand have 6 to 10 transmembrane helices, presenting both termini to the cytosolic side (Parker & Newstead 2017).

The split-ubiquitin screen has detected 253 interactions between tested NSTs and GTs from the library. The results obtained in this experiment are summarised in **Table 3**, along with results obtained by Dr Jennifer Bromley (JBEI, USA) in an initial screen of URGT1, 2 and 3 with the same library undertaken prior to the abovementioned experiment. The screen served as a tool to select the most promising candidates from the library, however without assessing the strength of the interaction. To validate the observed putative interactions, a pairwise split-ubiquitin assay has also been conducted and is outlined in **Chapter 4.4**.

Most interestingly, five protein pairs have been found to be linked both in the performed interaction screens and in co-expression networks established by Mutwil *et al.* (2011). These include GALS1-URGT1, GATL5-URGT4, MGP4-URGT5, GT29 family galactosyltransferase 1 (GalT29A)-URGT6 and cellulose-synthase like C4 (CSLC6)-UXT3. The results hint at interlinked functions of GATL5 and URGT4 in seed mucilage production. The engagement of both those proteins in mucilage synthesis is supported by previous reports, which indicate that GATL5 regulates the size of mucilage pectin polymers and that URGT4 is overexpressed in *urgt2* loss-of-function mutants to minimise the disruptions in mucilage synthesis (Kong *et al.*, 2013; Voiniciuc *et al.*, 2018, Parra-Rojas *et al.*, 2019). Taken together, that information could suggest complex formation between those two proteins. GalT29A is involved in adding Gal to type-II arabinogalactan (Dilokpimol *et al.*, 2014); a process, for which UDP-Gal could be provided by co-expressed URGT6. Similarly, GALS1-URGT1 could be working together in the production of RG-I galactan sidechains (Ebert *et al.*, 2018b), as URGT1 overexpression plants have been shown to accumulate excess UDP-Gal exclusively as galactan (Rautengarten *et al.*, 2014).

Table 3. (pages 55-56). GT-NST interactions identified using a split-ubiquitin assay screen. The numbers in the chart correspond to the number of times a GT was identified in the screen. Pre-screen data contained only qualitative and not quantitative data. Multiple attempts to screen UXT1 yielded zero colonies and hence this bait was excluded from the results presented.

			Initial screen performed by Dr Jennifer Bromley at JBEI (USA)				Screen undertaken as a part of this thesis							
GT family	GT name & description	GT AGI	URGT1	URGT2	URGT3	UXT3	URGT1	URGT2	URGT3	URGT4	URGT5	URGT6	UXT2	UXT3
GT1	UDP-GLUCOSYLTRANSFERASE 71C4 (UGT71C4)	AT1G07250							1					
	UDP-GLUCOSYLTRANSFERASE 85A5 (UGT85A5)	AT1G22370									1			
	UDP-GLUCOSYLTRANSFERASE 78D1 (UGT78D1)	AT1G30530									1			
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT1G50580						1						
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT2G36780										2		
	UDP-GLUCOSYLTRANSFERASE 74F2 (UGT74F2)	AT2G43820							1		4			
	UDP-GLYCOSYLTRANSFERASE 76B1 (UGT76B1)	AT3G11340										1		
	UDP-GLUCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT3G22250						1						
	UDP-GLUCOSYLTRANSFERASE 76E11 (UGT76E11)	AT3G46670							1					
	UDP-GLUCOSYLTRANSFERASE (UGT84A3)	AT4G15490				+						1	1	
	INDOLE-3-ACETATE BETA-D-GLUCOSYLTRANSFERASE (IAGLU)	AT4G15550							2			1		
	UDP-N-ACETYLGLUCOSAMINE TRANSFERASE ALG14-LIKE (ALG14-L)	AT4G18230								1				
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT4G27560							1	1		1		
	UDP-GLUCOSYLTRANSFERASE 73B2 (UGT73B2)	AT4G34135										1		
GT2	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT5G53990									1			
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT5G65550											1	
	CELLULOSE SYNTHASE-LIKE C6 (CSLC6)	AT3G07330				+								
	CELLULOSE SYNTHASE-LIKE A1 (CSLA1)	AT4G16590				+								
GT4	CELLULOSE SYNTHASE-LIKE G1 (CSLG1)	AT4G24010												
	CELLULOSE SYNTHASE-LIKE A11 (CSLA11)	AT5G16190									1			
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT1G78800	+				1	1	7	2		1		
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT3G10630					5		4					
	DIGALACTOSYL DIACYLGLYCEROL DEFICIENT 2 (DGD2)	AT4G00550		+	+		1		3	6	1	1	1	
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT4G19460							1					
	SULFOQUINOVOXYLDIACYLGLYCEROL 2 (SQD2)	AT5G01220							1				1	
GT8	SUCROSE SYNTHASE 5 (SUS5)	AT5G37180				+								
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT5G59070												
	GALACTURONOSYLTRANSFERASE-LIKE 5 (GATL5)	AT1G02720				+				1	1			
	GALACTINOL SYNTHASE 3 (GOLS3)	AT1G09350				+			1					
	GALACTURONOSYLTRANSFERASE-LIKE 3 (GATL3)	AT1G13250										2		
	LOS GLYCOSYLTRANSFERASE 5 (LGT5)	AT2G30575							3					
	GALACTURONOSYLTRANSFERASE 7 (GAUT7)	AT2G38650								2				
	GALACTINOL SYNTHASE 1 (GOLS1)	AT2G47180				+								
	GALACTURONOSYLTRANSFERASE-LIKE 4 (GATL4)	AT3G06260			+					1				
	GLUCURONIC ACID SUBSTITUTION OF XLAN 1 (GUX1)	AT3G18660					1							
	GALACTURONOSYLTRANSFERASE-LIKE 10 (GATL10)	AT3G28340				+				1		2		
	GALACTURONOSYLTRANSFERASE-LIKE 7 (GATL7)	AT3G62660										1		
	GALACTURONOSYLTRANSFERASE-LIKE 6 (GATL6)	AT4G02130								1				
	GALACTINOL SYNTHASE 6 (GOLS6)	AT4G26250										1		
GT13	GALACTINOL SYNTHASE 5 (GOLS5)	AT5G23790										1		
	GALACTURONOSYLTRANSFERASE 12 (GAUT12)	AT5G54690				+						1	2	
GT14	N-ACETYLGLUCOSAMINYLTRANSFERASE I (GNTI)	AT4G38240											1	
GT17	GLUCURONIC ACID TRANSFERASE 14C (GLCAT14C)	AT2G37585				+								
	UNFERTILIZED EMBRYO SAC 7 (UNE7)	AT3G03690				+								
GT17	BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	AT1G12990				+	2							
	BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	AT3G01620										1		1
	BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	AT3G27540			+				2					
	BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	AT5G14480							4	1			1	

			Initial screen performed by Dr Jennifer Bromley at JBEI (USA)				Screen undertaken as a part of this thesis							
GT family	GT name & description	GT AGI	URGT1	URGT2	URGT3	UXT3	URGT1	URGT2	URGT3	URGT4	URGT5	URGT6	UXT2	UXT3
GT19	LIPID X B (LPXB)	AT2G04560										1		
GT20	TREHALOSE-6-PHOSPHATASE SYNTHASE S6 (TPS6)	AT1G68020				+								
GT22	EMS-MUTAGENIZED BRI1 SUPPRESSOR 3 (EBS3)	AT1G16900		+										
GT28	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT1G73740										1		
GT29	GLYCOSYLTRANSFERASE 29A (GALT29A)	AT1G08280										1		
GT31	GALACTOSYLTRANSFERASE FAMILY PROTEIN	AT1G11730				+								
	GALACTOSYLTRANSFERASE FAMILY PROTEIN	AT1G77810												1
	BETA-1,3-N-ACETYLGLUCOSAMINYLTRANSFERASE LUNATIC PROTEIN	AT3G11420			+									
	GALACTOSYLTRANSFERASE FAMILY PROTEIN	AT3G14960				+	3							
	-	AT4G11350			+									
	FRINGE-LIKE PROTEIN	AT4G23490											1	
	HYDROXYPROLINE O-GALACTOSYLTRANSFERASE 1 (HPGT1)	AT5G53340					3	3	3				1	
GT32	ALPHA 1,4-GLYCOSYLTRANSFERASE FAMILY PROTEIN	AT2G38150								1				
GT34	XYLOGLUCAN XYLOSYLTRANSFERASE 4 (XXT4)	AT1G18690							1				1	
	UDP-XYLOSYLTRANSFERASE 2 (XT2)	AT4G02500							1					
	GALACTOSYL TRANSFERASE GMA12/MNN10 FAMILY PROTEIN	AT4G38310									1			
GT47	EXOSTOSIN FAMILY PROTEIN	AT1G68470	+	+		+			1			1		
	EXOSTOSIN FAMILY PROTEIN	AT1G74680							3			1		
	MURUS 3 (MUR3)	AT2G20370				+	2		3				3	
	EXOSTOSIN FAMILY PROTEIN	AT5G16890				+							1	
	EXOSTOSIN FAMILY PROTEIN	AT5G41250							2					
	IRREGULAR XYLEM 10 (IRX10-L)	AT5G61840								1				
GT50	PEANUT 1 (PNT1)	AT5G22130				+								
GT58	ASPARAGINE-LINKED GLYCOSYLATION 3 (ALG3)	AT2G47760			+		2		3			1		
GT61	-	AT2G41640			+		4		3	1				1
	MUCILAGE-MODIFIED 5 (MUM5)	AT3G10320			+	+								
GT64	ECTOPICALLY PARTING CELLS (EPC1)	AT3G55830				+				2				
GT75	REVERSIBLY GLYCOSYLATED POLYPEPTIDE 1 (RGP1)	AT3G02230									1			
	REVERSIBLY GLYCOSYLATED POLYPEPTIDE 2 (RGP2)	AT5G15650										1		
GT77	REDUCED RESIDUAL ARABINOSE 3 (RRA3)	AT1G19360					1							
	NUCLEOTIDE-DIPHOSPHO-SUGAR TRANSFERASE FAMILY PROTEIN	AT1G28710				+				1				
	REDUCED ARABINOSE YARIV 1 (RAY1)	AT1G70630				+								
	REDUCED RESIDUAL ARABINOSE 1 (RRA1)	AT1G75120		+										
	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	AT2G02060						1			3			
	MALE GAMETOPHYTE DEFECTIVE 4 (MGP4)	AT4G01220							2					
	-	AT5G40900		+										
	NUCLEOTIDE-DIPHOSPHO-SUGAR TRANSFERASE FAMILY PROTEIN	AT5G44820		+					4	1				
GT90	O-GLUCOSYLTRANSFERASE RUMI-LIKE PROTEIN (DUF821)	AT3G61270				+								
	O-GLUCOSYLTRANSFERASE RUMI-LIKE PROTEIN (DUF821)	AT3G61290									1			
	O-GLUCOSYLTRANSFERASE RUMI-LIKE PROTEIN (DUF821)	AT5G23850		+						1				
GT92	GALACTAN SYNTHASE 1 (GALS1)	AT2G33570	+											
NC	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT1G10280										2		
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT1G11940		+		+	1		2	1	2		1	
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT1G68380					1							
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT5G16170			+	+	5		9		1	1	3	2
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT5G25970											1	
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT5G57270					1		2	1				

4.3. Confirmation of the interactions using a pairwise split-ubiquitin assay

To validate the protein-protein interaction identified in the large-scale URG-T split-ubiquitin screen, they were subsequently also tested in a pairwise manner. For the split-ubiquitin-based pairwise assay, yeast cells harbouring a bait were transformed with the respective prey identified during the split-ubiquitin screen, where URG-T baits were probed against a GT prey library.

The UGNT1 protein was chosen as control as it localises to Golgi apparatus, yet is unlikely to interact with URG-T and UXT proteins, since a recent study has shown its involvement in the delivery of UDP-GlcNAc for *N*-glycans and sphingolipids (Ebert *et al.*, 2018a). The negative and positive controls used in the pairwise screen were composed of UGNT1 fused to, respectively, a mutated (NubG) or the wild-type (NubI) N-terminus of ubiquitin. The difference between NubG and NubI is a mutation of the Ile-13 residue to glycine, which weakens the affinity of the ubiquitin halves and enables reconstitution of a functional protein only when brought to close proximity. Therefore fusing NubG to a protein completely unrelated to URG-T and UXT activity, yet residing in the same compartment, should ideally yield no interaction, thus serving as a negative control, while fusing the same protein to NubI should work as a positive control, since the interaction will always happen, as long as both of the proteins are properly expressed, independently of the identity of the fused protein.

Pairwise assays were preceded by a self-activation test, which aimed to exclude the possibility of bait autoactivation resulting in detection of false positives in the absence of the bait and to determine the suitable concentration of 3-amino-1,2,4-triazole (3-AT). 3-AT is a histidine synthesis inhibitor that helps to remove false positives from yeast-based interaction assays. The autoactivation test was performed by growing yeast with bait constructs, UGNT1-NubG negative control and UGNT1-NubI positive control, paired with empty prey vector on SD-Leu-Trp plates supplemented with a gradient of 3-AT (0, 2.5 mM, 10 mM). In the absence of the inhibitor, no growth has been observed for cells containing only bait constructs accompanied by an empty prey vector. Additionally, only the cells expressing URG-T6-Cub have grown in the presence of the negative control UGNT1-NubG (see **Figure 14**). However, it is worth noting that for URG-T6-Cub, the growth inhibition with 2.5 mM and 10 mM concentrations of 3-AT was similar for both positive and negative controls, thus rendering further use of 3-AT pointless. It should be also noted that URG-T6 cannot be considered as an

autoactivator, since the yeast growth is in this case dependent on the presence of the prey. Together, these observations indicate that the interaction between URGT6 and the control construct could be dependent on the UGNT1 and provide an indication for possible URGT6-UGNT1 interaction.

Yeast cultures with baits and preys were diluted to $OD_{600} = 0.1$ and plated in 1x, 10x, 100x and 1000x dilutions. Yeast growth was recorded with a Canon EOS 700D camera at 48, 72, 96 and 120 h of incubation and the interactions were assessed after 72 h (**Figures 14 & 15**).

Out of 105 protein interactions observed in the initial screen (**Table 3**), the pairwise assay confirmed in total 55 unique interaction pairs between the six members of the URGT family and the GTs present in the prey library. 11 out of 17 interactions have been confirmed for two members of the UXT family.

For URGT1, 15 identified putative interaction partners belonging to GT families 4, 8, 17, 31, 47, 58, 61, 77 and 14-L (**Table 3**) were tested in the pairwise assay. Out of those 15, 9 could be confirmed as interactors in the pairwise assay, namely: AT1G78800, AT3G10630, GLUCURONIC ACID SUBSTITUTION OF XYLAN 1 (GUX1), AT1G12990, AT3G14960, AT2G41640, AT1G11940, AT1G68380 and AT5G16170 (**Figure 14A**). Notably, the majority of these proteins has not been functionally characterised yet. This includes mostly proteins belonging to the relatively new GT14-L family, AT1G11940, AT1G68390 and AT5G16170 (Hansen *et al.*, 2009), which have been speculated to be putative GalTs. Interestingly, the pairwise split-ubiquitin assay confirmed interactions of GT14-Ls not only with URGT1 but also URGT5 and URGT6; the same three URGTs that have been shown to interact with proteins from the GT8 family, including known GalTs GALACTINOL SYNTHASE GolS5 and GolS6 (Taji *et al.*, 2002) (**Table 4**). This could hint at the possibility that UDP-Gal is directly channelled from these URGTs to GTs from those two classes. So far the only functionally described putative URGT1 interaction partner, GUX1, has been shown to add GlcA residues to GX (Mortimer *et al.*, 2010, Rennie *et al.*, 2012), which, unfortunately, cannot be readily connected with URGT1 activity; one possible explanation might be however the interaction between GX and the products of URGT1 activity. The interactions identified for URGT1 are quite exciting, given that URGT1 has been directly linked to UDP-Gal, but not UDP-Rha, transport *in planta* and overexpression of URGT1 led to an accumulation of Gal in the leaf cell wall (Rautengarten *et al.*, 2014). Presented data suggest that URGT1 can form a complex with GalTs, however, this is a concept that needs to be confirmed by a second, independent method.

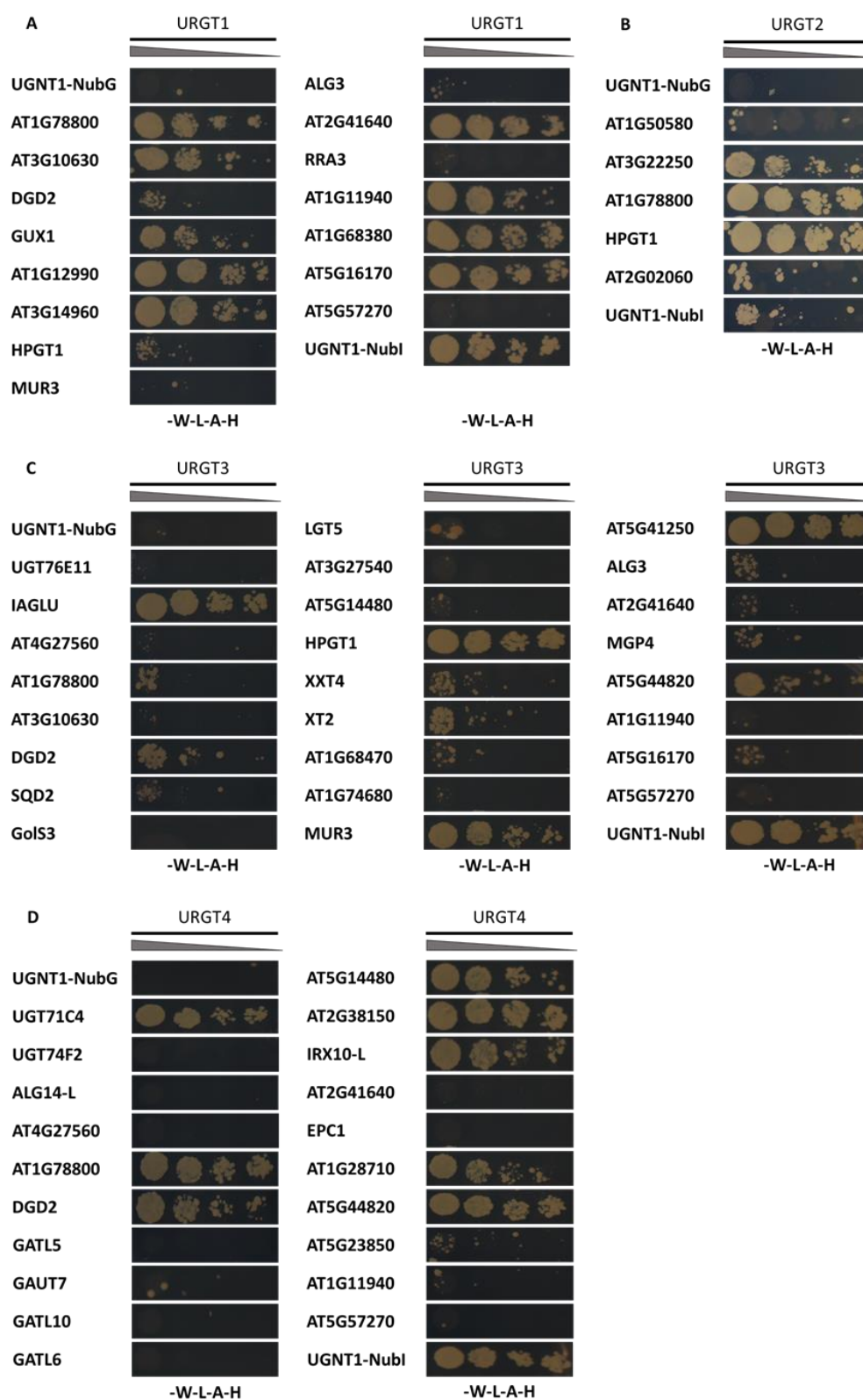


Figure 14. Validation of interactions between members of the URG family (baits) and their respective interactors identified using a GT library (preys) using the split-ubiquitin pairwise assay. URGs were fused to the Cub bait domain and GTs to the Nub prey domain. The assay was conducted on a quadruple dropout medium (SD-Trp-Leu-Ade-His). Photos show yeast growth after 72 h at 30°C. No 3-AT was used due to the complete lack of growth of negative controls, with the exception of **G**, where URG6 interaction with UGNT1-NubG and UGNT1-NubI was tested across a gradient of 3-AT inhibitor, to verify whether growth is related to a version of Nub used.

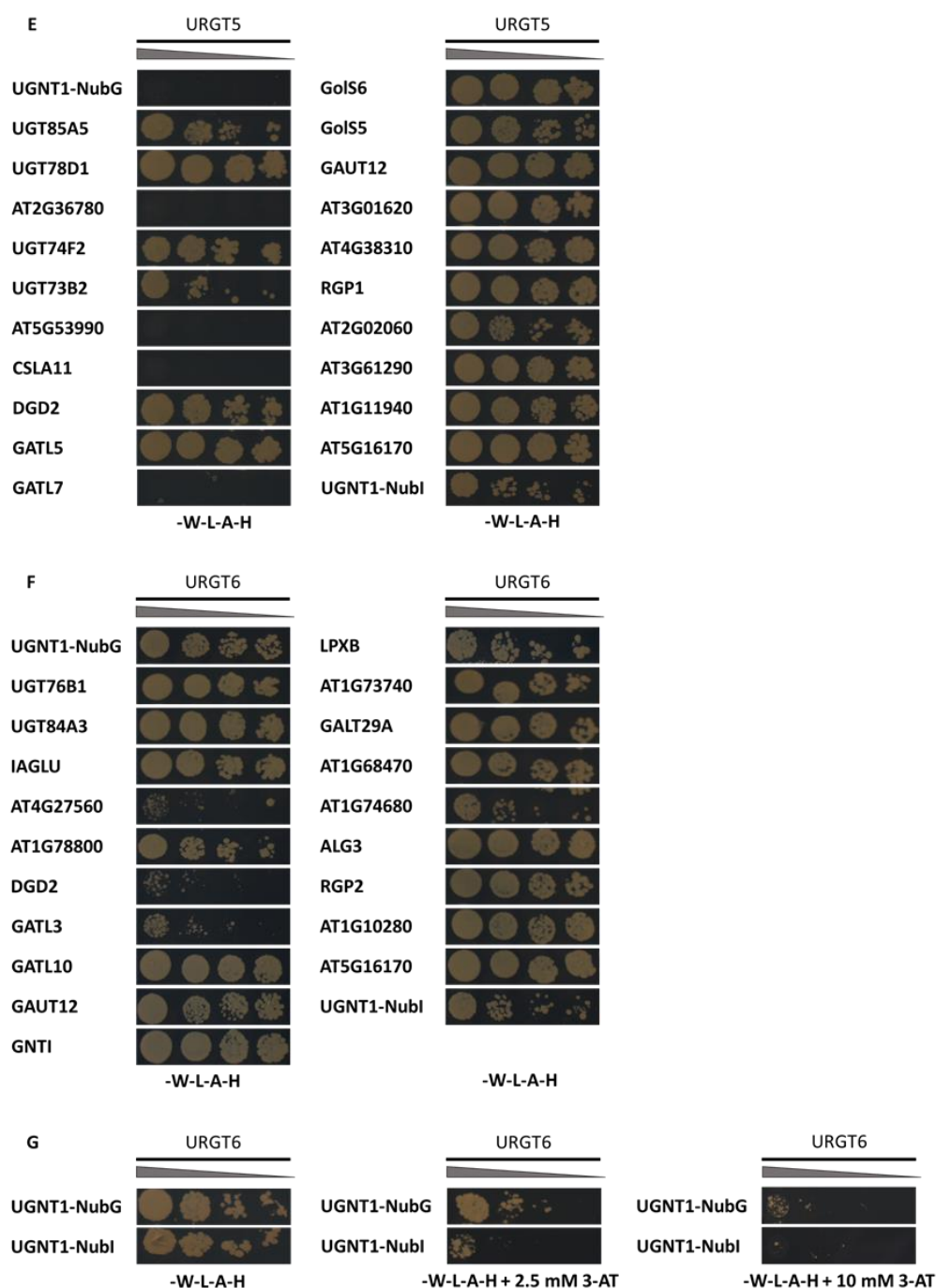


Figure 14. (continued).

URGTS2 was tested with 5 candidates previously identified in the screen from families GT1, 4, 31 and 77 (**Table 3**), out of which 3 were tested positive also in the pairwise screen, namely AT3G22250, AT1G78800 and HYDROXYPROLINE O-GALACTOSYL-TRANSFERASE 1 (HPGT1) (**Figure 14B**). Apart from uncharacterised AT3G22250 and AT1G78800, URGTS2, but also URGTS3, have shown strong interactions with HPGT1 (Ogawa-Ohnishi & Matsubayashi, 2015). Since the main substrate of HPGT1 is UDP-Gal, a substrate transported by URGTSs, this interaction is likely and should be followed up in future studies.

For URG3 the split-ubiquitin screen identified the most (26) putative interactors. The putative interactors belong to the GT1, 4, 8, 17, 31, 34, 47, 58, 61, 77 and 14-L families (**Table 3**). However, only 5 out of those 26 could be confirmed in a pairwise assay, namely: INDOLE-3-ACETIC ACID GLUCOSYLTRANSFERASE (IAGLU), HPGT1, MUR3, AT5G41250 and AT5G44820 (**Figure 14C**). In addition to the two uncharacterised proteins and already discussed HPGT1, URG3 binds to IAGLU (Jackson *et al.*, 2001) and MUR3. MUR3 is a bifunctional XyG GalT (Li *et al.*, 2013) that has been shown to function as part of a multiprotein XyG biosynthetic complex together with XLT2, FUT1, XXT2, XXT5 and CSLC4 (Chou *et al.*, 2015). As of now, there is no evidence suggesting that URG3 is involved in this complex; however, supplying the substrate directly to the complex assembling the side chains of XyG could be the elusive step of *in planta* regulation of this URG.

URGT4 was tested with 20 putative interactors from the original screen, belonging to GT families 1, 4, 8, 17, 32, 47, 61, 64, 77, 90 and 14-L (**Table 3**); 8 of those putative interactions could be confirmed in the pairwise assay, specifically UGT71C4, AT1G78800, DIGALACTOSYL DIACYLGLYCEROL DEFICIENT 2 (DGD2), AT5G14480, AT2G38150, IRX10-L, AT1G28710 and AT5G44820 (**Figure 14D**). URG4 shows a strong interaction (also seen for URG5) with DGD2, a protein required for galactosylation of monogalactosyldiacylglycerol (Rocha *et al.*, 2016b), which supports the possible complex formation of these two proteins. An interaction of URG4 with IRX10-L was also observed and confirmed; a fact that is rather surprising given that IRX10-L has been shown to synthesise GX (Urbanowicz *et al.*, 2014). This interaction might be also unspecific since IRXs belong to the same family as already mentioned MUR3, with which they share some sequence similarity.

Using URG5 as a bait in the screen yielded possible interactions with 20 GTs from families 1, 2, 4, 8, 17, 34, 75, 77, 90 and 14-L (**Table 3**). In the pairwise assay, the interaction was confirmed for 15 of them: UGT85A5, UGT78D1, UGT74F2, DGD2, GATL5, GoLS6, GoLS5, GAUT12, AT3G01620, AT4G38310, RGP1, AT2G02060, AT3G61290, AT1G11940 and AT5G16170 (**Figure 14E**). Apart from the uncharacterised proteins and those, that were already discussed, URG5 interacted with several members of the GT8 family, including GALT5 (Kong *et al.*, 2013), GolS5 and GolS6 (Taji *et al.*, 2002) and GAUT12, involved in GX and pectin biosynthesis (Hao *et al.*, 2014). Also, AT4G38310 was described as related to XyG synthesis by Faik *et al.* (2002) was found as an interacting partner. However, this publication suggested that AT4G38310 is not a functional protein but a polypeptide fragment, since it does not contain a transmembrane domain, nor a DXD motif.

Finally, URG6 was tested with 19 GTs from families GT1, 4, 8, 13, 19, 28, 29, 47, 58, 75 and 14-L (**Table 3**), which were identified in the screen as putative interaction partners. Out of those 19 interactions, for 15 were confirmed: UGT76B1, UGT84A3, IAGLU, AT1G78800, GATL10, GAUT12, GNTI, LIPID X B (LPXB), AT1G73740, GALT29A, AT1G68470, ASPARAGINE-LINKED GLYCOSYLATION 3 (ALG3), RGP2, AT1G10280 and AT5G16170 (**Figure 14F**). URG6 has been identified as an interactor of both previously mentioned IAGLU and GAUT12, as well as GnTI and ALG3). GnTI and ALG3 collaborate to synthesise *N*-glycans, utilising *N*-acetyl-D-glucosamine (Ebert *et al.*, 2018a) and Man (Veit *et al.*, 2018). Lastly, GALT29A is a UDP-Gal-dependent GalT (Dilokpimol *et al.*, 2014). Overall, URG6 was pretty promiscuous with its interactions, implying that this might be due to its ‘stickiness’ rather than specificity. This could also explain its affinity to the negative control. Therefore, even though some of the interactions identified for URG6 align well with its activity, they should be placed under additional scrutiny.

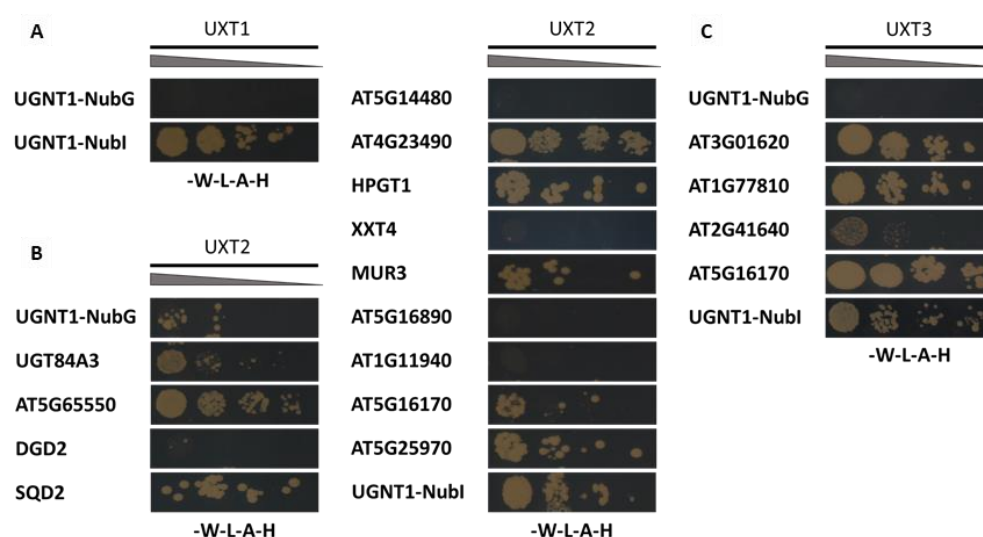


Figure 15. Validation of interactions between members of the UXT family (baits) and their respective interactors identified using a GT library (preys) using the split-ubiquitin pairwise assay. UXTs were fused to the Cub bait domain and GTs to the Nub prey domain. The assay was conducted on a quadruple dropout medium (SD-Trp-Leu-Ade-His). Photos show yeast growth after 72 h at 30°C. No 3-AT was used due to the complete lack of growth of negative controls. Names were given where applicable, while proteins not yet named were described by their locus.

No interactions could be identified in the initial screen using UXT1 as bait, although the pairwise test using the controls (**Figure 15A**) revealed that the UXT1 bait is expressed and functional, at least in regard to its role in the split-ubiquitin assay. In contrast, UXT2 came up in the screen with 13 potentially interacting GTs from families GT1, 4, 17, 31, 34, 47 and GT14-L (**Table 3**) and in the pairwise tests, the interactions were confirmed for 8: UGT84A3,

AT5G65550, SQD2, AT4G23490, HPGT1, MUR3, AT5G16170 and AT5G25970 (**Figure 15B**). Two of them were previously functionally characterised: HPGT and MUR3; both have also been identified as putative interaction partners of URGT3. MUR3, a GalT, is a likely interactor of UXT2, as they are both involved in XyG biosynthesis (Tedman-Jones *et al.*, 2008, Zhao *et al.*, 2018). Finally, UXT3 was tested with four GTs from families 17, 31, 61 and GT14-L (**Table 3**), out of which three interaction partners were confirmed in the pairwise assay: AT3G01620, AT1G77810 and AT5G16170 (**Figure 15C**). None of them has been previously described in the literature, making them interesting candidates for further follow-up studies.

A detailed summary of the identified interacting proteins and their family affiliations can be found in **Table 4**.

Table 4. GT-NST interactions established in the split-ubiquitin screen and confirmed through pairwise interactions in yeast cells.

GT family	GT name & description	GT AGI	URGT1	URGT2	URGT3	URGT4	URGT5	URGT6	UXT2	UXT3
GT1	UDP-GLUCOSYLTRANSFERASE 71C4 (UGT71C4)	AT1G07250				+				
	UDP-GLUCOSYLTRANSFERASE 85A5 (UGT85A5)	AT1G22370					+			
	UDP-GLUCOSYLTRANSFERASE 78D1 (UGT78D1)	AT1G30530					+			
	UDP-GLUCOSYLTRANSFERASE 74F2	AT2G43820					+			
	UDP-GLYCOSYLTRANSFERASE 76B1 (UGT76B1)	AT3G11340						+		
	UDP-GLUCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT3G22250		+						
	UDP-GLUCOSYLTRANSFERASE (UGT84A3)	AT4G15490						+		
	INDOLE-3-ACETATE BETA-D-GLUCOSYLTRANSFERASE (IAGLU)	AT4G15550			+			+		
GT4	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT5G65550								+
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT1G78800	+	+		+		+		
	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT3G10630	+							
	DIGALACTOSYL DIACYLGLYCEROL DEFICIENT 2 (DGD2)	AT4G00550				+	+			
GT8	SULFOQUINOVOSYLDIACYLGLYCEROL 2 (SQD2)	AT5G01220								+
	GALACTURONOSYLTRANSFERASE-LIKE 5 (GATL5)	AT1G02720					+			
	GLUCURONIC ACID SUBSTITUTION OF XYLAN 1 (GUX1)	AT3G18660	+							
	GALACTURONOSYLTRANSFERASE-LIKE 10 (GATL10)	AT3G28340						+		
	GALACTINOL SYNTHASE 6 (GOLS6)	AT4G26250					+			
	GALACTINOL SYNTHASE 5 (GOLS5)	AT5G23790					+			
	GALACTURONOSYLTRANSFERASE 12 (GAUT12)	AT5G54690					+	+		
GT13	N-ACETYLGLUCOSAMINYLTRANSFERASE I (GNTI)	AT4G38240						+		
GT17	BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	AT1G12990	+							
	BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	AT3G01620					+			+
	BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	AT5G14480				+				
GT19	LIPID X B (LPXB)	AT2G04560						+		
GT28	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT1G73740						+		
GT29	GLYCOSYLTRANSFERASE 29A (GALT29A)	AT1G08280						+		
GT31	GALACTOSYLTRANSFERASE FAMILY PROTEIN	AT1G77810								+
	GALACTOSYLTRANSFERASE FAMILY PROTEIN	AT3G14960	+							
	FRINGE-LIKE PROTEIN	AT4G23490								+
	HYDROXYPROLINE O-GALACTOSYLTRANSFERASE 1 (HPGT1)	AT5G53340		+	+					+
GT32	ALPHA 1,4-GLYCOSYLTRANSFERASE FAMILY PROTEIN	AT2G38150				+				
GT34	GALACTOSYL TRANSFERASE GMA12/MNN10 FAMILY PROTEIN	AT4G38310					+			
GT47	EXOSTOSIN FAMILY PROTEIN	AT1G68470						+		
	MURUS 3 (MUR3)	AT2G20370			+					+
	EXOSTOSIN FAMILY PROTEIN	AT5G41250			+					
	IRREGULAR XYLEM 10 (IRX10-L)	AT5G61840				+				
GT58	ASPARAGINE-LINKED GLYCOSYLATION 3 (ALG3)	AT2G47760						+		
GT61	-	AT2G41640	+							
GT75	REVERSIBLY GLYCOSYLATED POLYPEPTIDE 1 (RGP1)	AT3G02230					+			
	REVERSIBLY GLYCOSYLATED POLYPEPTIDE 2 (RGP2)	AT5G15650						+		
GT77	NUCLEOTIDE-DIPHOSPHO-SUGAR TRANSFERASE FAMILY PROTEIN	AT1G28710				+				
	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	AT2G02060					+			
	NUCLEOTIDE-DIPHOSPHO-SUGAR TRANSFERASE FAMILY PROTEIN	AT5G44820			+	+				
GT90	O-GLUCOSYLTRANSFERASE RUMI-LIKE PROTEIN (DUF821)	AT3G61290					+			
NC	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT1G10280						+		
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT1G11940	+				+			
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT1G68380	+							
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT5G16170	+				+	+		+
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT5G25970							+	

Notably, the obtained protein-protein interaction data can also be indirectly used as a means to support the observed sub-Golgi localisations of the URGs, described in **Chapter 3**. For

example, Parsons *et al.* (2019) published a dataset that classifies some GTs to certain Golgi stacks using a method called Localisation of organelle proteins by isotope tagging (LOPIT). In their experiments, HPGT1, GAUT7, GNTI and IRX10-L were listed as localised to the medial Golgi, while the interaction data described above indicate that they interact with URGT2, URGT4 and URGT6 – URGTs which have also been localised to the medial Golgi in this thesis. MGP4 and MUR3 are recognised in the dataset as trans-Golgi proteins and the interaction data shows that they interact with URGT3 – another trans-Golgi protein. However, URGT3 seems to interact also with proteins that Parsons classify as medial – HPGT1 and XXT1. This could hint at the fact that cis, medial and trans-Golgi stacks are still only abstract concepts, while cells are full of fluctuations and devoid of borders separating proteins from interacting with each other. Further studies are required to provide evidence validating both the localisation and the interaction data, as well as testing the hypotheses for how matrix polysaccharide synthesis is organised in the Golgi apparatus.

4.4. Validation of protein-protein interactions using additional methods

Since all of the currently widely applied PPI investigation methods are prone to create false negatives, false positives or both, it is always crucial to analyse putative protein pairs using at least two methods (Xing *et al.*, 2016). The methods chosen for validation of the possible interactions observed in the split-ubiquitin screen (described in **Chapter 4.4.**) were bimolecular fluorescence complementation (BiFC) and acceptor photobleaching - Förster resonance energy transfer (APB-FRET). A full list of constructs generated for the experiments described in this chapter can be found in **Supplemental Table S10**. For all fluorescence and CoIP constructs, as well as the GT overexpression constructs, stable Arabidopsis lines have also been generated.

BiFC is a method based on the reconstitution of the functional fluorophore from halves fused to proteins of interest (Kerppola, 2006). Due to its simplicity and widespread use in plant sciences, it was selected as the method of choice. However, over the course of the experiments, promiscuous interactions in every possible combination of Golgi-localising NST and GT were observed, including an unrelated GT used as a negative control, INOSITOL PHOSPHORYLCERAMIDE GLUCURONOSYLTRANSFERASE 1 (IPUT1). Several different strategies were followed to optimise the BiFC assays; e.g., using an OD₆₀₀ gradient to infiltrate the leaves and thresholding the signal to filter out the false positives. However, none

of these approaches helped to differentiate between real interactions and a random reassembly of the fluorophore (**Supplemental Figure S1A**). Therefore, despite immense efforts, it was concluded that this method might not be suitable to validate the split-ubiquitin data. The nature of Golgi protein biosynthesis is complex and involves vesicular transport between the ER and Golgi apparatus and between the individual Golgi stacks, during which the cargo proteins are packed tightly in small compartments. This may facilitate getting proteins of interest into close proximity and lead to irreversible spontaneous self-assembly of the fluorophore, rather than being caused by the actual interactions. The problem of spontaneous functional reconstitution of a fluorescent protein (FP) caused by exceeding a certain concentration threshold in a confined cellular localisation has been known for almost twenty years and most recently brought into the spotlight in a review by Kudla & Bock (2016).

Another approach explored for the validation of the split-ubiquitin assay results was APB-FRET (Wouters *et al.*, 1998). This method is based on the principle, that in a FRET pair upon bleaching of the fluorescence acceptor with a high-power laser, the fluorescence of the donor increases. The use of APB-FRET should have solved the problems that were encountered during the BiFC experiments since in this method the interaction between the two proteins of interest is the only force that holds the fluorophores in proximity and the interaction is not artificially stabilised by reconstitution of an FP. However, in this case, APB-FRET failed to produce any increase in the fluorescent signal coming from the donor, both in the case of the tested interactions (URGT1, URGT3 and URGT5 with AT1G11940 and AT1G62305), as well as for the positive control used (GAUT1 and GAUT7). The only exception from that was the control containing GFP and RFP covalently fused together in a single polypeptide, which responded to acceptor bleaching with an initial 88% increase of donor fluorescence, which quickly decayed to become stable at 39% (**Supplemental Figure S1B**). This construct, however, does not constitute an interaction and thus could have not been used as a meaningful positive control. To summarise, our APB-FRET assay could not confirm interactions even in the case of the positive control, which is well documented in the literature (Atmodjo *et al.*, 2011). This result, however, is not unexpected. In the review of PPI investigation methods, Xing *et al.* (2016) have classified this technique as one that is prone to create false negatives just as much as BiFC is prone to create false positives.

The last method chosen to validate the split-ubiquitin results was co-immunoprecipitation (Co-IP, Phizicky & Fields, 1995), where a tagged protein of interest is pulled down from the suspension by a tag-specific antibody together with its interacting partners. For this purpose, a

set of transgenic *Arabidopsis* lines expressing GT interactors fused to a hemagglutinin (HA) tag or GFP tag were generated and were, where possible, expressed in the respective loss-of-function mutant background. However, the outbreak of the COVID-19 pandemic forced the cessation of the experiments. Even though the plant samples were harvested and proteins extracted, the final steps of experiments could not be undertaken before the submission of this thesis. The author of this thesis plans to include that information in a subsequent paper based on the thesis that is set to be published later.

4.5. Conclusion

This chapter summed up the knowledge from publicly available databases of interactions and co-expression networks of URG and UXT families and then built further upon that information by probing members of those two families against an extensive library of *Arabidopsis* GTs in a split-ubiquitin assay. In total, out of 253 interactions detected in the large-scale screen, 95 unique putative interactors among the GTs were identified (**Table 3**), out of which 51 were confirmed in pairwise interactions assays in this study (**Table 4**). While many of them are not yet functionally characterised, the list contains some of the already known GalTs and others, whose activity might be readily linked with that of URGs.

While additional verification of these interactions was attempted multiple times, it was either unsuccessful on the grounds of methodology or delayed by the COVID-19 pandemic. Since the set of possible interactors obtained during this study was too big to conduct an in-depth characterization of all interactors, 15 GTs were selected for further analysis. The results of these analyses are presented in **Chapter 5**.

5. FUNCTIONAL CHARACTERISATION OF GTs PUTATIVELY INVOLVED IN THE BIOSYNTHESIS OF RHAMNOSE AND/OR GALACTOSE CONTAINING GLYCANS IN ARABIDOPSIS

The fifth chapter of this study aims to characterise some of the GT candidates identified in **Chapter 4**. To achieve that goal, a reverse genetics approach employing loss-of-function mutants was selected. The aim was to investigate the impact of loss-of-function mutations on plant morphology, cell wall composition and glycan biosynthesis.

5.1. GT candidates selected for characterisation

As mentioned in **Chapter 4**., a shortlist of 15 GTs from a range of GT families was generated. Candidates were chosen to represent as many families as possible and GTs that were not yet characterised. The shortlist of candidates selected for further analysis consisted exclusively of URGF family interactors and included: AT4G15490 (GT1), AT4G16590 (GT2), AT1G78800 (GT4), AT3G06260 (GT8), AT3G27540 (GT17), AT1G16900 (GT22), AT3G11420 & AT4G11350 (GT31), AT1G68470 (GT47), AT2G41640 (GT61), AT5G40900, AT5G44820 (BOTH GT77), AT5G23850 (GT90), AT1G11940 and AT5G16170 (both non-classified; NC). For those GTs, their subcellular localisation information was sourced from SUBA4 – SUBAcon Bayes Consensus Classifier output (Hooper *et al.*, 2017) (**Table 5**).

Table 5. Shortlist of GT candidates characterised in this chapter.

GT family	GT name & description	GT AGI	Localisation (SUBAcon)
GT1	UDP-GLUCOSYLTRANSFERASE (UGT84A3)	AT4G15490	Golgi, cytosol
GT2	CELLULOSE SYNTHASE-LIKE A1 (CSLA1)	AT4G16590	PM
GT4	UDP-GLYCOSYLTRANSFERASE SUPERFAMILY PROTEIN	AT1G78800	cytosol
GT8	GALACTURONOSYLTRANSFERASE-LIKE 4 (GATL4)	AT3G06260	extracellular
GT17	BETA-1,4-N-ACETYLGLUCOSAMINYLTRANSFERASE FAMILY PROTEIN	AT3G27540	PM, vacuole
GT22	EMS-MUTAGENIZED BRI1 SUPPRESSOR 3 (EBS3)	AT1G16900	ER
GT31	BETA-1,3-N-ACETYLGLUCOSAMINYLTRANSFERASE LUNATIC PROTEIN	AT3G11420	vacuole
	-	AT4G11350	PM
GT47	EXOSTOSIN FAMILY PROTEIN	AT1G68470	Golgi
GT61	-	AT2G41640	PM
GT77	-	AT5G40900	extracellular
	NUCLEOTIDE-DIPHOSPHO-SUGAR TRANSFERASE FAMILY PROTEIN	AT5G44820	PM
GT90	O-GLUCOSYLTRANSFERASE RUMI-LIKE PROTEIN (DUF821)	AT5G23850	Golgi
NC	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT1G11940	mitochondrion
	CORE-2/I-BRANCHING BETA-1,6-N-ACETYLGLUCOSAMINYLTRANSFERASE	AT5G16170	PM

Since the localisations obtained from SUBAcon (listed above) were based entirely on information from localisation predictors and not experimental data, these results are likely not reliable. For that reason, the subcellular localisation of the candidates was determined experimentally (**Chapter 5.1.1.**). In cases, where Golgi localisation could be confirmed, loss-of-function and overexpression mutant lines were generated and their respective phenotypes analysed in more detail. The morphological phenotypes of these mutant lines were also tested using phenotype-inducing stress factors, including non-optimal growth conditions and the application of cell wall synthesis inhibitors. Since the selected GTs could be involved in mucilage production, the seed and mucilage morphology were also investigated (**Chapter 5.1.2.**). While cell wall and mucilage monosaccharide composition were originally planned to comprise the last part of this characterisation, the outbreak of the COVID-19 pandemic delayed the processing of the samples and thus those cannot be presented at this time.

5.1.1. Subcellular distribution of GTs interacting with URG family

To investigate the subcellular localisation of the GTs identified as URG interactors, a set of 15 overexpression constructs was generated using the pEarleyGate101 vector. Those constructs contained the putative interactors fused to YFP under the control of the 35S promoter and were used to transiently transform tobacco to analyse their subcellular localisation. The fluorescence images of protein localisations in tobacco leaves were collected after 3 days (**Figure 16**). As a follow up to this screen, all 15 proteins of interest have also been colocalised with Golgi and ER marker (**Supplemental Figure S2**).

Out of the 15 lines, four proteins: AT4G15490 (GT1), AT4G16590 (GT2), AT1G78800 (GT4) and AT1G16900 (GT22) have been observed to colocalise with the ER marker. This is not unexpected as GT1s are generally cytosolic soluble enzymes, while the GT2 family contains enzymes involved in cellulose synthesis, cellulose synthases and cellulose synthase-like proteins, that reside mainly in the PM. AT1G16900, known also as EMS-MUTAGENIZED BRI1 SUPPRESSOR 3 (EBS3), is an α -1,2-mannosyltransferase participating in dolichol-PP-linked Glc₃Man₉GlcNAc₂ N-glycan in the ER (Liu *et al.*, 2018), while its localisation in the image also resembles the web-like ER structure.

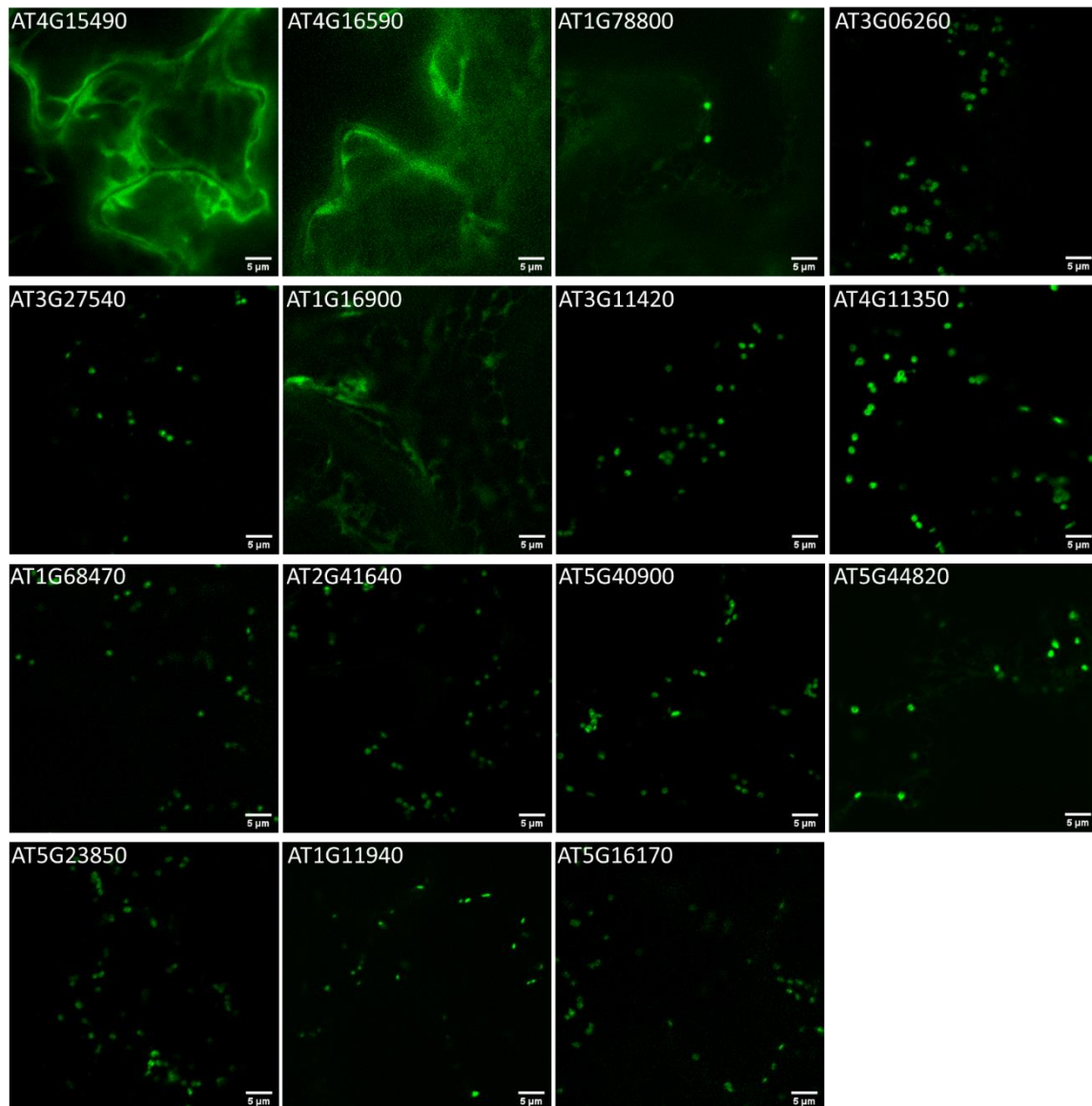


Figure 16. Subcellular localisations of the GTs identified to be interacting with members of the URGF family. Images of tobacco leaves expressing the *P_{35S}::GT-YFP* constructs. Four of the GTs seem to be expressed exclusively in the ER (AT4G15490, AT4G16590, AT1G78800 and AT1G16900), while the remaining 11 are localised exclusively to the Golgi apparatus. A special case has been made for AT1G78800, which seemed to localise to both ER and Golgi in some of the replicates and therefore was included in the further studies.

Interestingly, AT1G78800 has also been detected in fast-moving punctae-like Golgi structures, indicating dual localisation, which could be either naturally occurring or likely be attributed to the fact that they are overexpressed under a strong constitutive promoter, as Golgi proteins have to pass through the ER during their synthesis and hence was also taken into consideration as one of the potential URGF interactors. At the same time, the remaining 11 proteins (AT3G06260 (GT8), AT3G27540 (GT17), AT3G11420 & AT4G11350 (GT31), AT1G68470 (GT47), AT2G41640 (GT61), AT5G40900 (GT77), AT5G23850 (GT90),

AT1G11940, AT5G16170 and AT5G44820 (GT NC)) were observed exclusively in the Golgi apparatus. Since URGs localise to Golgi, the three GTs which were not detected in the Golgi have been excluded from the further analysis, reducing the shortlist to 12 proteins.

5.1.2. Analysis of GT loss-of-function mutant lines

To fully understand the role of these 12 GTs in relation to the URG family and Arabidopsis cell wall synthesis, a set of T-DNA insertion lines has been ordered from the ABRC seed repository (abrc.osu.edu). Depending on the T-DNA insertion site and line availability, between one and three independent lines have been ordered per gene (**Table 6**). All lines were grown through multiple generations until homozygous loss-of-function mutant lines were obtained. The lack of functional alleles was assessed by genotyping, followed by a confirmation of the absence of a transcript by RT-PCR. In two of the cases, SALK_076243 and SALK_007904, the presence of the T-DNA insertion was not observed and therefore, those two lines were excluded from further studies. In another two cases, SALK_004198 and SAIL_245_B01, homozygous mutants were not observed throughout three generations, hinting at possible lethality of the mutation, hence those lines were also eliminated. For SALK_126261 and SALK_130475C minor amounts of residual transcript in form of faint bands were observed, yet those lines were still analysed (**Supplemental Figure S3A**).

Table 6. Arabidopsis T-DNA insertion lines used in the study.

Family	AGI	Mutant name	Germplasm	Family	AGI	Mutant name	Germplasm
GT4	At1g78800	at1g78800-1	SALK_004198	GT90	At5g23850	at5g23850-1	SALK_067277
		at1g78800-2	SALK_014293			at5g23850-2	SALK_145225
GT17	At3g27540	at3g27540-1	GK-616G02			at5g23850-3	SALK_031660C
GT31	At3g11420	at3g11420-1	SALK_036197		At4g19970	at4g19970-1	SALK_055999
		at3g11420-2	SAIL_237_A01			at4g19970-2	SALK_096621
		at3g11420-3	SAIL_842_C07	GT NC	At1g11940	at1g11940-1	SALK_127740
	At2g37730	at2g37730-1	SALK_050983			at1g11940-2	SALK_076175
GT47	At1g68470	at1g68470-1	WiscDsLox437E03		At1g62305	at1g6230-1	SALK_017871
GT77	At5g40900	at5g40900-1	SALK_076243		At5g16170	at5g16170-1	SALK_126261
	At5g44820	at5g44820-1	SALK_007904			at5g16170-2	SAIL_245_B01
		at5g44820-2	SALK_042539			at5g16170-3	SALK_130475C

When cultivated on the soil in optimal growth conditions, none of the loss-of-function lines described in **Table 6** displayed any obvious morphological phenotype, which is most often the case with single GT loss-of-function mutants due to significant redundancy among GTs (Lee *et al.*, 2020). The presence of less-obvious differences was assessed by measuring the rosette diameters and stem lengths, which have not shown any statistically significant differences between the lines (**Supplemental Figure S3B & S3C**).

However, often phenotypes can be induced under specific growth conditions or in the presence of abiotic stress factors (Smakowska *et al.*, 2016). Thus, a range of various stress conditions as well as polysaccharide synthesis inhibitors has been applied to test their effects on seedling growth of the loss-of-function mutants. 7-day-old seedlings were exposed to heat (30°C), cold (10°C), ionic osmotic stress (sodium chloride, 130 mg/mL) and non-ionic osmotic stress (sorbitol, 260 mg/mL). Additionally, the effect of growth inhibitors such as polyphenon-60 (PP60; green tea extract, 0.0375 mg/mL), 2,6-dichlorobenzonitrile (DCB, 500nM), isoxaben (2 nM), latrunculin-B (25 mM) oryzalin (500 nM) and 4.5% (w/v) sucrose was assessed. Even though all of the treatments that were applied caused stress to the plants and induced slower growth, the rate of the slowdown was comparable for WT and the mutant lines (**Supplemental Figure S4**). Therefore, it can be concluded that it is unlikely that any of those 12 interactors have a stress-inducible phenotype under the conditions tested in this thesis.

Given the possible involvement of some of these GTs in the turnover of UDP-Rha for mucilage synthesis, a survey of the mutant mucilage release was undertaken. To that end, the mucilage was released by imbibing the seeds in water with vigorous shaking. The seeds were then stained with ruthenium red, a polycationic stain widely used to visualise negatively charged polysaccharides (**Figure 17**). Along with the GT loss-of-function lines, three *urgt2* lines were analysed since URG2 is mainly expressed in seeds and has been proposed to be involved in mucilage synthesis with a mucilage phenotype described previously (Rautengarten *et al.*, 2014, Parra-Rojas *et al.*, 2019). The mucilage staining experiment served as a screening tool before the monosaccharide analysis, and therefore a simple visual inspection was conducted in lieu of software-assisted quantification. Those initial results indicated that the mucilage release could possibly be affected in *at3g27540-1*, *at3g11420-2*, *at5g44820-1*, *at5g16170-2* and *urgt3-1* mutant lines in comparison to WT seeds.

In addition, scanning electron microscopy (SEM) has been used to visualise potential abnormalities in mucilage secretory cells on the seed surface (**Figure 18**). The experiment was conducted with WT, *urgt2-1* and *urgt2-3* lines for URG2, a transporter previously implicated in mucilage synthesis (Rautengarten *et al.*, 2014, Parra-Rojas *et al.*, 2019), and two loss-of-function mutants for *At3g11420*, a gene encoding one of the GTs for which a mucilage phenotype was observed in the initial mucilage staining experiment, as these were the most likely to have any visible phenotypes. However, the four lines tested by SEM showed comparable characteristics resembling WT both on the level of seed and cell morphology. Due to the Covid-19 pandemic, this experiment could not be expanded further to include all mutants.

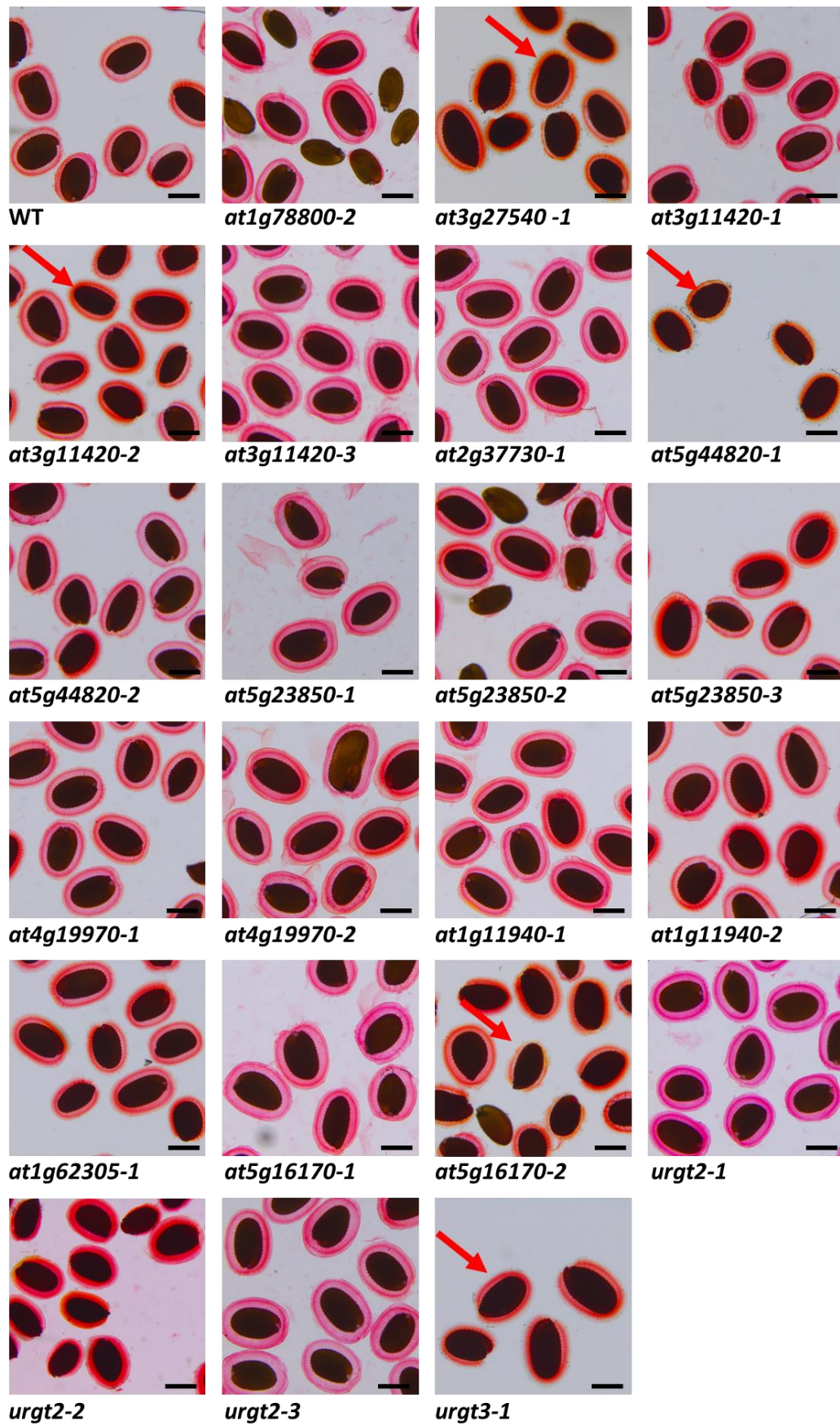


Figure 17. Mucilage release from seeds from mutant lines deficient in selected GTs and URGs that are putatively involved in the mucilage synthesis. Red arrows indicate lines, where the mucilage release is potentially affected.

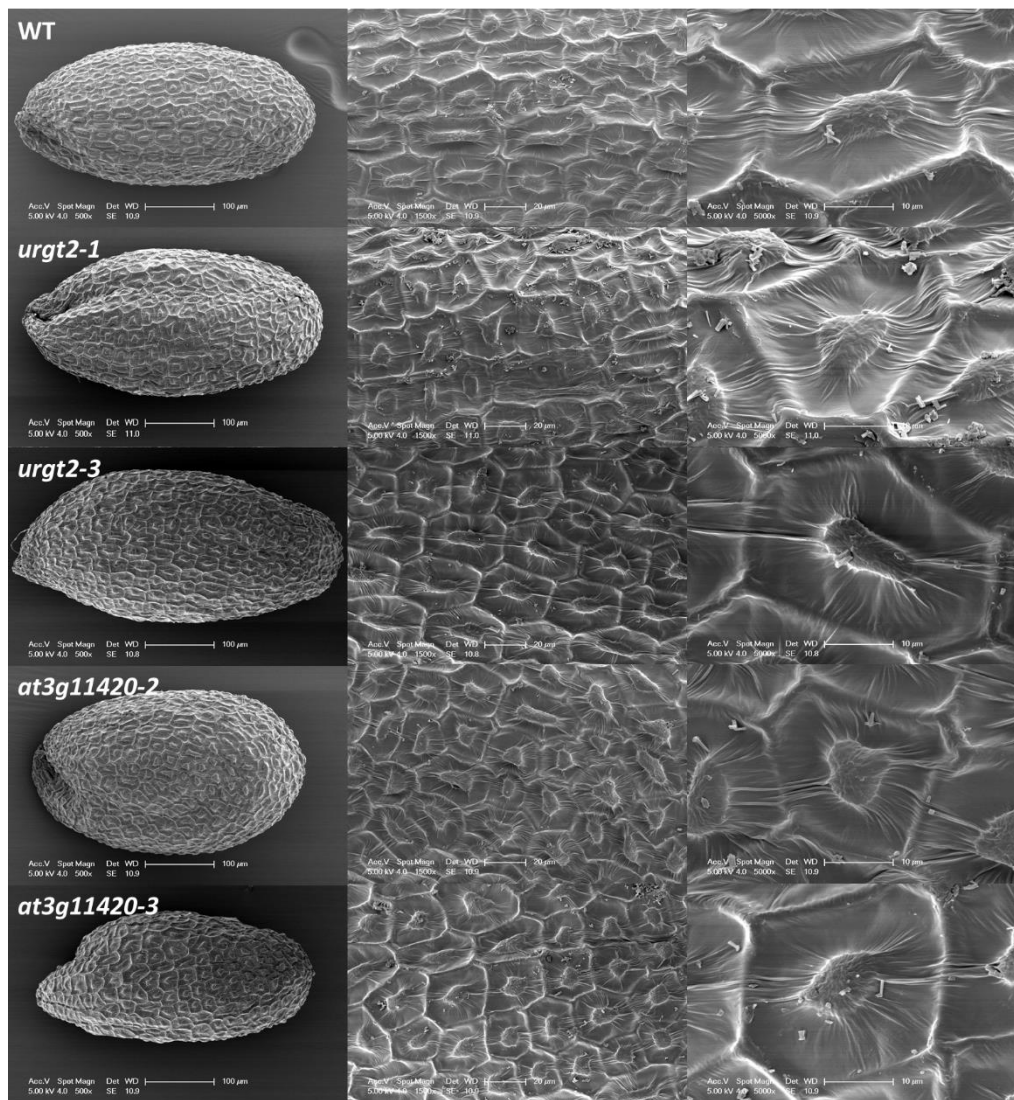


Figure 18. The seed coat of *URGT* and *GT* mutants, as imaged by SEM. The left column presents a whole seed under 500x magnification, the middle one – 1500x and the right one displays a single mucilage secretory cell under 5000x magnification. All of the samples analysed showed a wild type-like appearance.

Despite the lack of apparent differences in seed cell morphology, an analysis of the monosaccharide composition of the mucilage was attempted to explain the differences in mucilage release. Similarly, the cell wall monosaccharide composition of both loss-of-function mutants and plants overexpressing the proteins of interest were supposed to be assessed as part of this thesis. However, while mucilage extracts and cell wall alcohol insoluble residue (AIR) from mutant plants have been prepared and subjected to TFA hydrolysis to yield monosaccharide samples, due to COVID-19-related problems with repairs of the required instrumentation, the author of this thesis was not able to finalise these experiments before thesis submission. The experiments are currently being undertaken and will be possibly included in a subsequent publication based on the findings of this thesis.

5.2. Characterisation of AT1G11940 and AT1G62305, two novel putative galactosyltransferases

AT1G11940 which was identified in the screen and its close homolog, AT1G62305 (not detected in the screen), were taken under special consideration and in addition to the above-mentioned experiments were further investigated regarding their double mutant phenotype, sub-Golgi localisation and expression pattern. These two GTs were selected since data obtained by our collaborator, Prof. Wei Zeng (Zhejiang A&F University, China), suggested that they are involved in the addition of Gal residues from UDP-Gal as donor onto an AGP acceptor. In a biochemical assay performed by Prof. Zeng, both AT1G11940 and AT1G62305 catalysed the reaction of the addition of a Gal residue onto a galactan chain composed of 6 to 8 Gal residues. (W. Zeng, personal communication). This function may be possibly shared with other members of the GT14-L family. **Figure 19** depicts a phylogenetic tree of GT14-L proteins from five organisms: moss (*Selaginella moellendorffii*), rice (*Oryza sativa*), mouse-ear cress (*Arabidopsis thaliana*), poplar (*Populus trichocarpa*) and rose gum (*Eucalyptus grandis*).

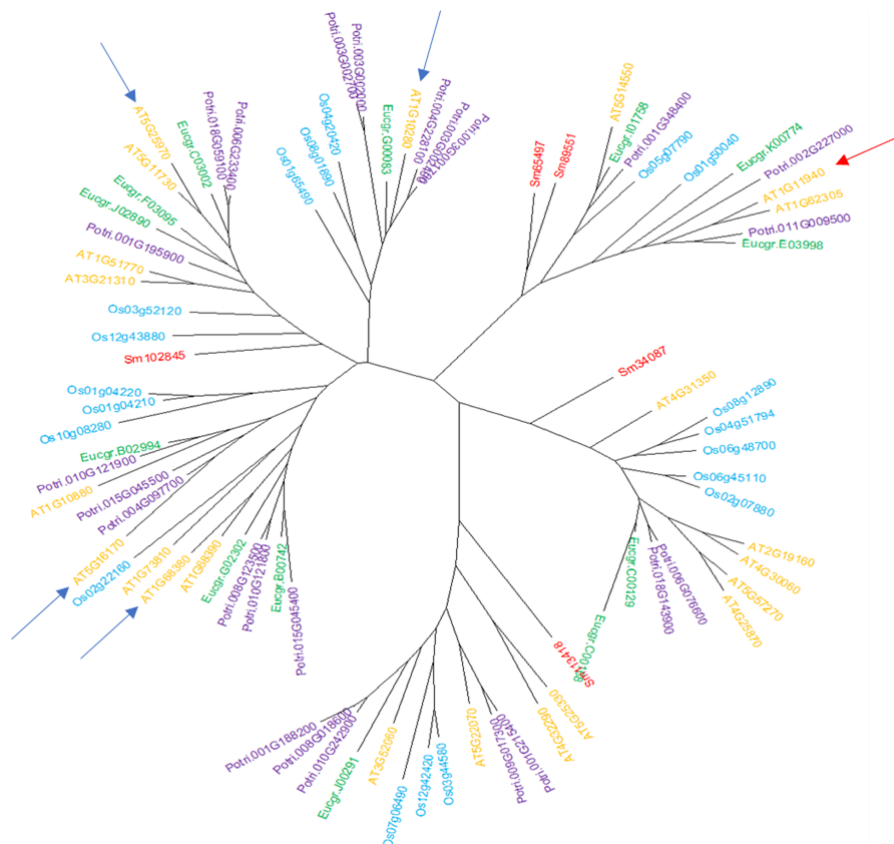


Figure 19. Phylogenetic tree of DUF266 containing genes (GT14-L family) from *Arabidopsis thaliana* and its homologs from four other land plants. Moss - red, rice - cyan, Arabidopsis – yellow, poplar – purple, rose gum – green. Genes for which protein products have been identified in the split-ubiquitin pairwise assay have been indicated with blue arrows. Red arrow indicates At1g11940.

5.2.1. Colocalisation of AT1G11940 and AT1G62305 with fluorescent markers specific for distinct Golgi subcompartments

For the two GT14-Ls, one identified in the split-ubiquitin assay and its homolog, an additional colocalisation experiment was performed using the same markers as in the URG T colocalisation experiment (**Figure 11**). The resulting merged images are presented in **Figure 20**, while the colocalisation values are shown in **Figure 21**.

Interestingly, the sub-Golgi localisation of these two GTs seems to be less defined than the spatial distribution observed for the URG Ts, for which one could easily pinpoint the localisation in one of the three subcompartments of the Golgi apparatus. AT1G11940 seems to be present in the medial and trans-Golgi, while its close homolog AT1G62305 localises to the cis and medial Golgi. AT1G62305 has been also shown by Parsons *et al.* (2019) as localising to the medial Golgi in LOPIT experiments. Given the ubiquitous localisation of these proteins across the Golgi apparatus, it is difficult to confirm or exclude interactions between them and the URG Ts basing on this experiment. In addition, the lower levels of expression deriving from GT expressed under their native promoters make the signal to noise ratio lower than in the case of the URG T measurements.

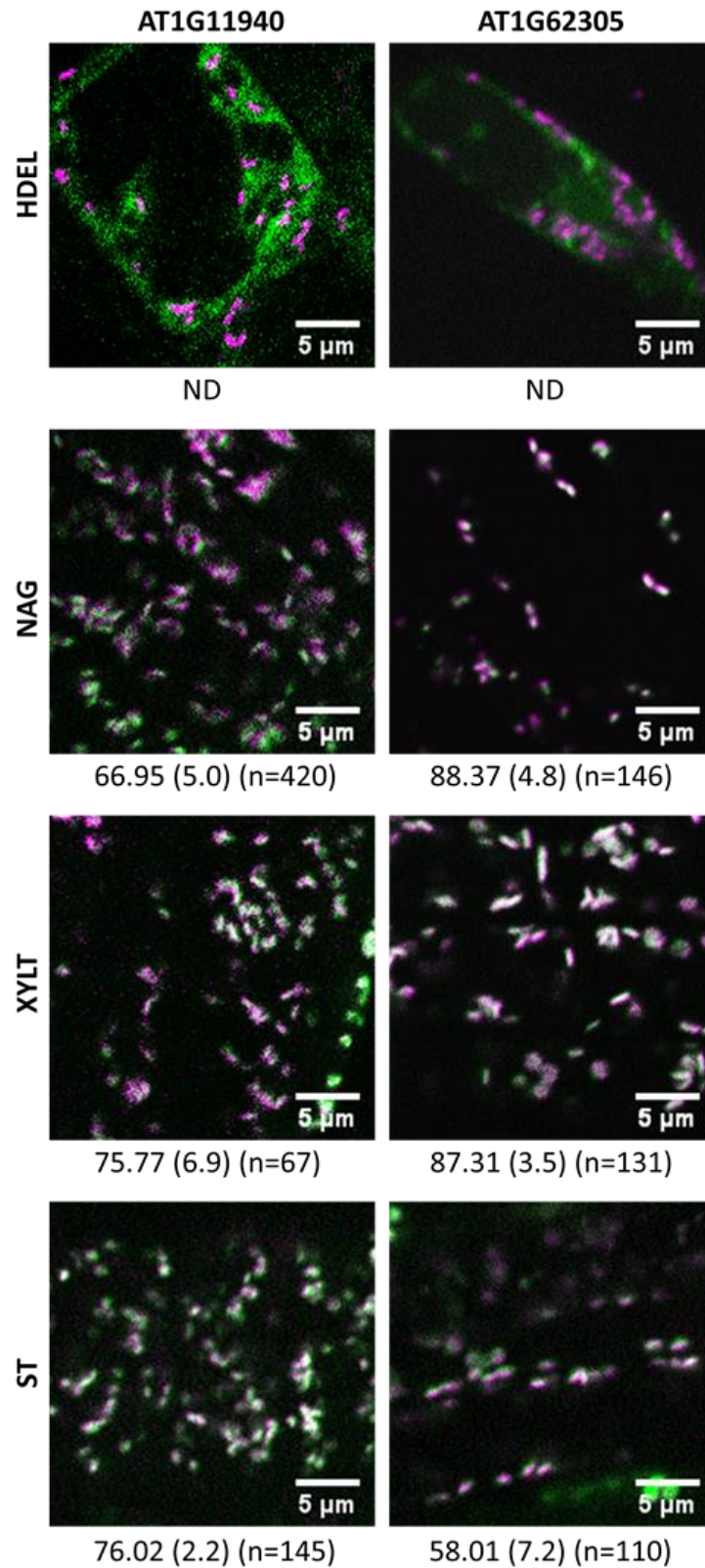


Figure 20. Sub-Golgi localisation of two GTs from the GT14-L family with sub-Golgi markers. Images of Arabidopsis root tips from lines stably expressing $P_{GT}::GT-YFP$ (magenta) and $P_{UBQ10}::(NAG/XYLT/ST)-RFP$ (green).

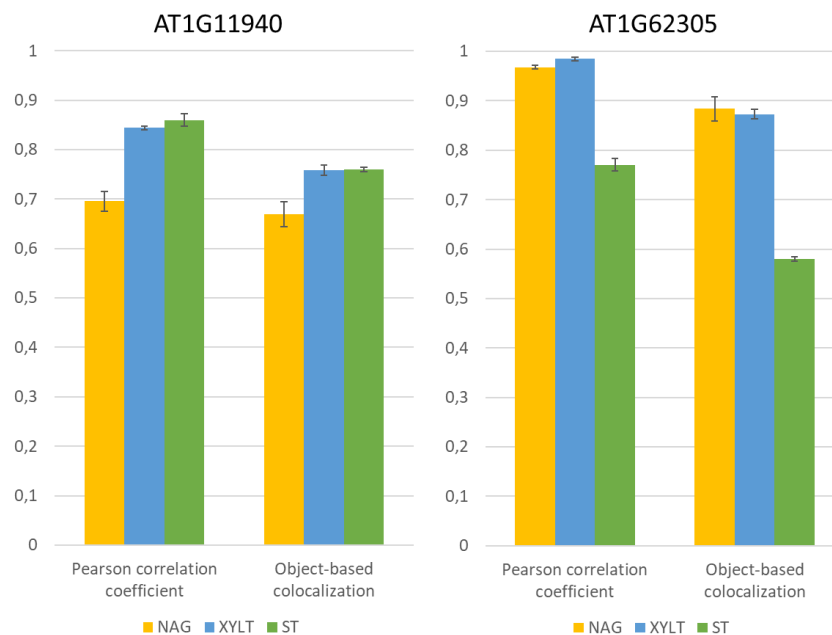


Figure 21. Average Pearson correlation coefficient and object-based colocalisation values of two GTs belonging to the GT14-L family with protein markers of respective sub-Golgi cisternae. NAG represents cis, XYLT medial and ST trans cisterns of the Golgi apparatus.

5.2.2. *At1g11940* and *At1g62305* expression pattern analysis through qRT-PCR and GUS staining

To analyse the expression of *At1g11940* and *At1g62305*, online databases were surveyed first. Based on the data from Klepikova *et al.* (2016) *At1g11940* is expressed mainly in mature flowers and siliques (7.0-8.0, 0.5-3.5 elsewhere), while *At1g62305* seems to be expressed exclusively in the anthers of the young flower (53.8, 1.0-4.0 elsewhere). The microarray expression data from a Developmental Map by Schmid *et al.* (2005, Winter *et al.*, 2007) showed ubiquitous expression for *At1g11940* throughout plant development, that peaked in flowers and inflorescences stages. Since *At1g62305* was not present on the ATH1 array (Applied Biosystems™), the relative expression levels for both GT were assessed using qRT-PCR.

To further investigate the expression pattern of *At1g11940* and *At1g62305* by qRT-PCR, samples from 10 tissue types (roots, leaves, seedlings, bottom, middle and top of the stem, flowers and bottom, middle and top siliques) were harvested and RNA was extracted. After the reverse transcription, the cDNA samples were tested using the primers listed in **Supplemental Table S9** for *At1g11940* and *At1g62305*, as well as primers for ubiquitin 10, protein

phosphatase 2A and SAND family protein, which served as a reference. The qRT-PCR was performed by Jacob Calabria. The expression of the different tissues was calculated using the $2^{-\Delta\Delta C_t}$ method (Rao *et al.*, 2013) (**Figure 22**).

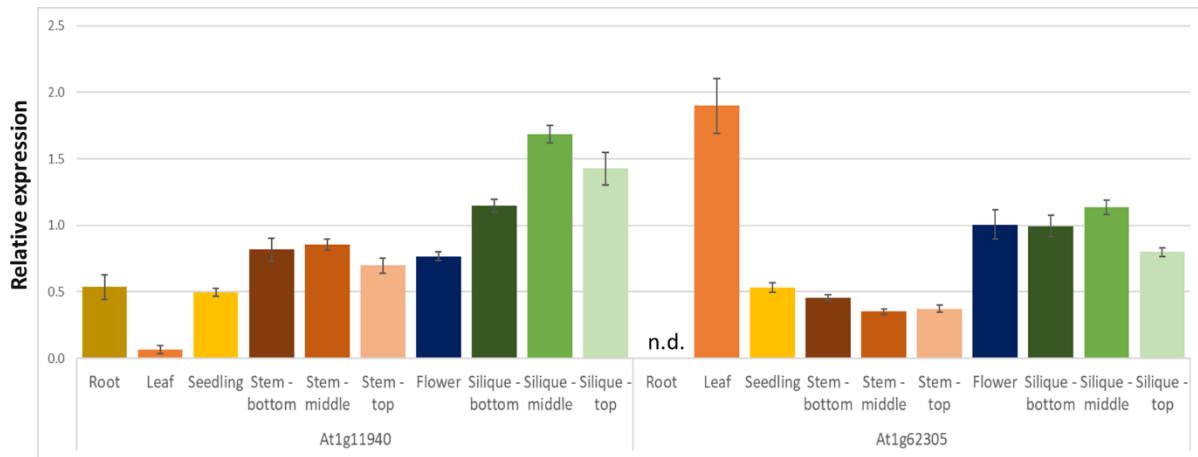


Figure 22. qRT-PCR analysis of *At1g11940* and *At1g62305* expression of organs and across developmental stages. Both genes are expressed throughout the plant with *At1g11940* peaking in siliques and *At1g62305* in leaves, flowers and siliques, while the expression of the latter in the roots is below the detection level.

The qRT-PCR analysis showed that the expression of *At1g11940* is low in leaves, moderate in other tissues and highest in siliques. At the same time, the expression of *At1g62305* is low in seedlings and stems, moderate in flowers and siliques and highest in leaves. Those results seem to differ from the publicly available RNASeq and microarray data described earlier in this chapter. Hence, in addition to the qRT-PCR analysis, the promoter activities of those two genes were tested using a GUS reporter-aided analysis. 626 and 1840 bp upstream of ATG codon of *At1g11940* and *At1g62305* genes respectively were amplified and cloned into the pBI101.3 plasmid (Jefferson *et al.*, 1987). Constructs were then introduced into *Arabidopsis* through *Agrobacterium*-mediated transformation. After selection with kanamycin, tissue samples (leaf, flower, stem, silique) were harvested from 4-week-old, mature plants and subjected to GUS staining. However, only minimal promoter activity could be observed in the young anthers for both *At1g11940* and *At1g62305* and thus the results remained inconclusive.

5.2.3. Phenotypic analysis of double mutants and knock-down lines affecting the functions of *At1g11940* and *At1g62305*

Due to the close homology of *At1g11940* and *At1g62305* (72% amino acid identity) and a lack of obvious morphological phenotypes for the single loss-of-function mutants, both in optimal and phenotype-inducing conditions, double loss-of-function mutants were generated. This was done by crossing the homozygous lines SALK_127740 and SALK_076175 for *At1g11940* with SALK_017871, a homozygous line identified for *At1g62305*. However, extensive genotyping of the generated progeny of over 200 plants revealed that the F2 generation for both of those crosses did not generate homozygous double mutants. These data suggest that a simultaneous loss of both of these genes could be embryo-lethal. However, both *at1g11940 At1g62305*^{+/-} and *At1g11940*^{+/-} *at1g62305* plants showed no obvious morphological phenotypes, indicating that seemingly a single functional copy of one of the homologs is enough for plant survival.

As a consequence, a hairpin construct specifically targeting *At1g62305* was generated by Jacob Calabria using pHANNIBAL/pART27 system (Wesley *et al.*, 2001) and subsequently introduced into wild type and the two homozygous *at1g11940* loss-of-function lines. The following primers, LP: ACGTCTCTCACATGTCAAATCCTTACG and RP: TGCAAGCAACAATCTTTCTGCTGCA were used to generate a 590 bp long amplicon for *At1g62305*, that was later cloned into pHANNIBAL in both sense and antisense orientation to create a hairpin. The expression cassette was then sub-cloned using *NotI* restriction enzyme into pART27 and transformed into Arabidopsis WT and *at1g11940* loss-of-function mutant plants via *Agrobacterium*-mediated transformation.

Interestingly, the reductions in the amount of *At1g62305* transcript reached only 40% in the *at1g11940* mutant background but over 60% in plants where *At1g11940* transcript was still present (data not shown). In the latter case, the increased reduction of *At1g62305* transcript was compensated by upregulation of the remaining *At1g11940* homolog in comparison with the empty vector control (**Figure 23**).

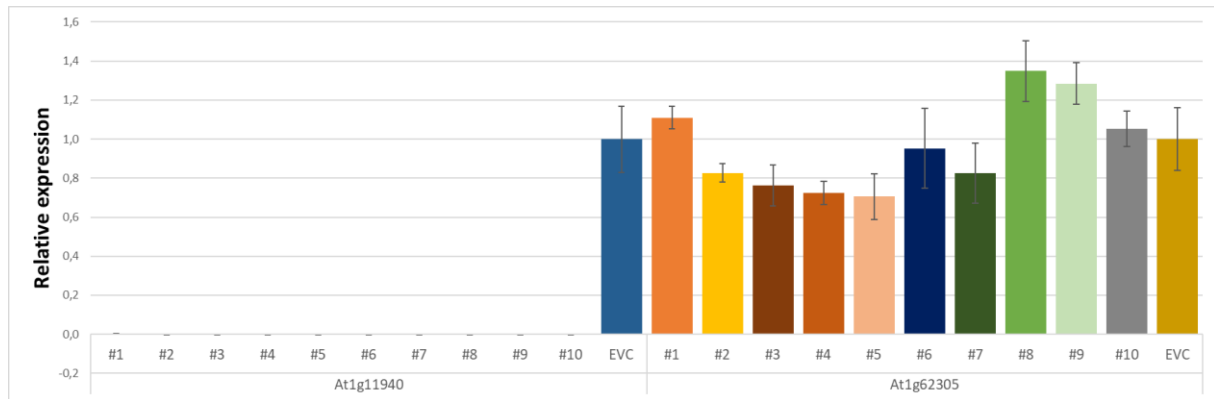


Figure 23. qRT-PCR confirmation of loss of *At1g11940* transcript and measurement of *At1g62305* expression across 10 different *at1g11940-2* plants harbouring the *At1g62305* hairpin construct. EVC = empty vector control.

Independent hairpin lines in both *at1g11940* mutant backgrounds showed significantly delayed growth and smaller rosette diameters in comparison to the empty vector control (**Figure 24**) suggesting a crucial function of these genes for plant development. Further analysis is required to link the observed morphological phenotype to a specific function of these GTs and in relation to the URGs.



Figure 24. Representative morphological phenotypes of *At1g62305* RNAi-mediated knock-down in the background of *at1g11940* loss-of-function mutants. Resultant plants show severe growth defects.

5.3. Conclusion

This chapter probed the subcellular localisations of 15 URGT interactors and described the approaches undertaken aiming at the characterisation of 12 of them that localised to the Golgi apparatus. Single loss-of-function mutants identified for those 12 GTs did not reveal obvious morphological phenotypes and maintained a similar size of rosettes and stem length when compared to the WT in both optimal and stress-inducing conditions. The seed and mucilage phenotype was also analysed, leading to preliminary evidence that in some of the mutants mucilage release could be impaired. However, due to malfunctioning instrumentation, this observation could not be further supported with data from a monosaccharide composition analysis. Similarly, the monosaccharide composition of cell wall polysaccharides also suffered from this problem. Further, studies including the generation of multiple mutants between close homologs but also between URGTs and GTs, analysis of monosaccharides and cell wall fractions, glycosidic linkage analysis and immunolabelling studies will be required to establish further links between these proteins and help to assign their function in cell wall-related processes.

After the broader screen of 12 GTs, two GT-14Ls, namely At1g11940 and At1g62305, were selected for a thorough characterisation as part of this thesis. For these two GTs, GalT activity could be detected using a structure resembling AGP backbone as an acceptor and UDP-Gal as a donor (Wei Zeng, personal communication). Notably, in rice a close homolog of At1g11940, OsBC10 has been described in the past as a gene with influence on cellulose and AGP synthesis (Zhou *et al.*, 2009). This information supports both the proposed GalT activity and the function in AGP synthesis for At1g11940.

A double mutant lacking those putative GalTs is seemingly lethal and a knock-down of one of them in the mutant background leads to plants with stunted growth and a dwarf phenotype, highlighting the importance of their function for plant growth and development. Summing up, the author of this thesis postulates, that At1g11940 and At1g62305 could be AGP GalTs and at least one functional copy of one of those genes is necessary for plant survival. However, further studies are required to unequivocally establish the function of these proteins (presumably as AGP GalTs) and to verify their links to the URGT family.

6. CONCLUSIONS & OUTLOOK

To conclude, both the spatiotemporal colocalisation, as well as the protein interaction theories that have been postulated in the introduction have proven to be true for the URG T transporters. The data generated as part of this thesis clearly shown that while the URG Ts are distinctly distributed across the cis, medial and trans-Golgi apparatus, they also seem to interact with a number of GTs, including known GalTs and potentially novel RhaTs and GalTs, in a similar manner as it happens in the Golgi apparatus of mammalian cells (Maszczak-Seneczko *et al.*, 2012, Maszczak-Seneczko *et al.*, 2015, Sosicka *et al.*, 2019). One could hypothesise that by placing NST in a pivotal point of the polysaccharide synthesis pathways and making their activity subject to regulation by both PPI and localisation, plants found a way to manipulate the polysaccharide content of their cell walls by expressing only the specific family member to direct the flow of the substrate. This manipulation might be a response to changing environmental conditions, biotic and abiotic stresses or simply plant growth and development and is tightly governed by the cell wall integrity sensing mechanism (Rui & Dinnyeny, 2020). It is also tempting to speculate that similar activity regulation via PPI and differential localisation might be occurring in other nucleotide sugar transporter families that have multiple members participating in the generation of different polysaccharides, such as the UXTs or the UAFTs. However, more studies are needed to confirm this, as another explanation might be also based on the regulation of the level of the transcript and thus currently both UXTs and UAFTs are a subject of further research in the author's research group.

The identification of numerous known GalTs in this study supports the hypothesis that novel GT functions can be identified using the URG T-GT interactions database established in this thesis. While a typical reverse genetics approach would not be suitable for establishing the function of GTs due to their large redundancy in plant cells, an interaction with a member of the NST family of known function might hint at the possible substrate specificity of the GT in question, or their involvement in protein complexes that synthesise heterogenous cell wall polysaccharides. However, this indication still merely works as a selection tool for a candidate of interest, that will be subjected to further functional analysis.

For the two GTs that were studied in more detail as part of this thesis, putative AGP GalT activity has been proposed. AGPs are a family of highly glycosylated glycoproteins found on the surface of the plant cells in both cell wall and PM that are composed of hydroxyproline-rich protein and long and diverse arabinogalactan chains. The latter is synthesised through the

concerted action of eight GT31s (Basu *et al.*, 2015, Ogawa-ohnishi & Matsubayashi, 2015) that glycosylate the hydroxyproline, AT1G77810 (Qu *et al.*, 2008) that builds the backbone on that initial sugar residue, two GT29s that initiate the galactan sidechains and GALT31A that elongate those sidechains up to five residues in most of plant species (Geshe *et al.*, 2013, Dilokpimol *et al.*, 2014). However, those sidechains are even longer in Arabidopsis leaves, which may be related to the activity of GT14-Ls described in this thesis, that elongate galactan chains consisting of 6 to 8 Gal residues (Wei Zeng, personal communication). Plant mutant analysis revealed that they are crucial for plant function but further experiments are necessary to confirm their GalT activity and link them to specific tissues and biological processes *in planta*. Therefore, the author and his collaborators intend to perform more experiments to establish the exact function of the two GalTs of interest in the context of AGP structure and analyse in detail the composition of the cell wall polysaccharides in the mutant plant lines on the microscopic level. When completed, this research will add a valuable piece of information to the current state of knowledge about the AGP synthesis, which in future might be of use in the context of applied science, since AGPs have multiple applications in industry and health sciences as adhesives, emulsifiers and water-holding agents (Saha & Bhattacharya, 2010; Barclay *et al.*, 2019)

Additional experiments will be conducted after the submission of this thesis to extend and validate for example the interaction results presented here using Co-IP experiments and monosaccharide composition analysis of mucilage and cell wall from GT loss-of-function and overexpression lines. In the long term, future studies may involve the full biochemical and *in planta* characterisation of all interactors identified in this study, guiding the potential discovery of novel GalTs and RhaTs. Additionally, analysis of localisation and interactions of UXT and UafT families is also underway and will be published as part of theses of other PhD students.

REFERENCES

- Atmodjo, M.A., Sakuragi, Y., Zhu, X., Burrell, A.J., Mohanty, S.S., Atwood, J.A., Orland, R., Scheller, H.V. & Mohnen, D. (2011). Galacturonosyltransferase (GAUT)1 and GAUT7 are the core of a plant cell wall pectin biosynthetic homogalacturonan: galacturonosyltransferase complex. *Proceedings of the National Academy of Sciences*, 108(50), 20225–20230.
- Bakker, H., Routier, F., Ashikov, A., Neumann, D., Bosch, D. & Gerardy-Schahn, R.A. (2008). CMP-sialic acid transporter cloned from *Arabidopsis thaliana*. *Carbohydrate Research*, 343(12), 2148–52.
- Baldwin, T.C., Handford, M.G., Yuseff, M.I., Orellana, A. & Dupree, P. (2001). Identification and characterization of GONST1, a Golgi-localised GDP-mannose transporter in *Arabidopsis*. *Plant Cell*, 13(10), 2283–95.
- Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., dePamphilis, C., (...) & Grigoriev, I.V. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science (New York, N.Y.)*, 20, 332 6032 960–3.
- Barclay, T. G., Day, C. M., Petrovsky, N. & Garg, S. (2019). Review of polysaccharide particle-based functional drug delivery. *Carbohydrate polymers*, 221, 94–112.
- Bar-Peled, M. & O'Neill, M.A. (2011). Plant nucleotide sugar formation, interconversion and salvage by sugar recycling. *Annual Review of Plant Biology*, 62, 127–55.
- Bartetzko, M.P. & Pfrengle, F. (2019). Automated glycan assembly of plant oligosaccharides and their application in cell-wall biology. *ChemBioChem*, 20, 877–885.
- Basu, D., Tian, L., Wang, W., Bobbs, S., Herock, H., Travers, A. & Showalter, A.M. (2015). A small multigene hydroxyproline-O-galactosyltransferase family functions in arabinogalactan-protein glycosylation, growth and development in *Arabidopsis*. *BMC Plant Biology*, 15, 295.
- Batoko, H., Zheng, H.-Q., Hawes, C. & Moore, I. (2000). A rab1 GTPase is required for transport between the endoplasmic reticulum and Golgi apparatus and for normal Golgi movement in plants. *The Plant Cell*, 12 (11), 2201–2217.
- Benchling [Biology Software]. (2017).
- Billosta, V., Brändström, J., Cochaux, Alain, Joseleau, J.P. & Ruel, K. (2006). Ultrastructural organisation of the wood cell wall may explain some modifications undergone by fibers during pulping process. *Cellulose Chemistry and Technology*, 40, 231–237.
- Blanch, H.W., Adams, P.D., Andrews-Cramer, K.M., Frommer, W.B., Simmons, B.A. & Keasling, J.D. (2008). Addressing the need for alternative transportation fuels: the Joint BioEnergy Institute. *ACS Chemical Biology*, 3, 17–20.
- Bleckmann, A., Weidtkamp-Peters, S., Seidel, C.A. & Simon, R. (2010). Stem cell signaling in *Arabidopsis* requires CRN to localise CLV2 to the plasma membrane. *Plant Physiology*, 152(1), 166–176.
- Bolte, S. & Cordelières, F.P. (2006). A guided tour into subcellular colocalisation analysis in light microscopy. *Journal of Microscopy*, 224 (3), 213–232.
- Burn, J.E., Hocart, C.H., Birch, R.J., Cork, A.C. & Williamson, R.E. (2002). Functional analysis of the cellulose synthase genes *CesA1*, *CesA2*, and *CesA3* in *Arabidopsis*. *Plant Physiology*, 129(2), 797–807.
- Chapple, C. & Carpita, N. (1998). Plant cell walls as targets for biotechnology. *Current Opinion in Plant Biology*, 1(2), 179–185.
- Chen J., Lalonde S., Obrdlik P., Noorani Vatani A., Parsa S.A., Vilarino C., Revuelta J.L., Frommer W.B. & Rhee S.Y. (2012). Uncovering *Arabidopsis* membrane protein interactome enriched in transporters using mating-based split ubiquitin assays and classification models. *Frontiers in Plant Science*, 3, 124.
- Chevalier, L., Bernard, S., Ramdani, Y., Lamour, R., Bardor, M., Lerouge, P., Follet-Gueye, M.L. & Driouich, A. (2010). Subcompartment localisation of the side chain xyloglucan-synthesizing enzymes within Golgi stacks of tobacco suspension-cultured cells. *The Plant Journal*, 64(6), 977–989.
- Chou, Y.H., Pogorelko, G. & Zabolina, O.A. (2012). Xyloglucan xylosyltransferases XXT1, XXT2 and XXT5 and the glucan synthase CSLC4 form Golgi-localised multiprotein complexes. *Plant Physiology* 159, 1355–1366.
- Chou, Y.H., Pogorelko, G., Young, Z.T. & Zabolina, O.A. (2015). Protein-protein interactions among xyloglucan-synthesizing enzymes and formation of Golgi-localised multiprotein complexes. *Plant Cell Physiology*, 56(2), 255–67.

- Cicéron, F., Rocha, J., Kousar, S., Hansen, S.F., Chazalet, V., Gillon, E., Breton, C. & Lerouxel, O. (2016). Expression, purification and biochemical characterization of AtFUT1, a xyloglucan-specific fucosyltransferase from *Arabidopsis thaliana*. *Biochimie*, 128–129, 183–192.
- Clough, S.J. & Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant Journal*, 16, 735–43.
- Colleoni, C., Linka, M., Deschamps, P., Handford, M.G., Dupree, P., Weber, A.P.M. & Ball, S.G. (2010). Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis. *Molecular Biology and Evolution*, 27, (12), 2691–2701.
- Cosgrove, D.J. & Jarvis, M.C. (2012). Comparative structure and biomechanics of plant primary and secondary cell walls. *Frontiers in Plant Science*, 3(August), 1–7.
- Culbertson, A.T. & Zabotina, O.A. (2015). The glycosyltransferases involved in the synthesis of plant cell wall polysaccharides: present and future. *JSM Enzymology and Protein Science*, 1, 1–5.
- DeBono, A., Yeats, T.H., Rose, J.K., Bird, D., Jetter, R., Kunst, L. & Samuels, L. (2009). *Arabidopsis* LTPG is a glycosylphosphatidylinositol-anchored lipid transfer protein required for export of lipids to the plant surface. *The Plant cell*, 21(4), 1230–1238.
- Dikeman, C.L. & Fahey, G.C. (2006). Viscosity as related to dietary fiber: a review. *Critical Reviews in Food Science and Nutrition*, 46(8), 649–63.
- Dilokpimol, A. & Geshi, N. (2014). *Arabidopsis thaliana* glucuronosyltransferase in family GT14. *Plant signaling & behavior*, 9(6), e28891.
- Dilokpimol, A., Poulsen, C.P., Vereb, G., Kaneko, S., Schulz, A. & Geshi, N. (2014). Galactosyltransferases from *Arabidopsis thaliana* in the biosynthesis of type II arabinogalactan: molecular interaction enhances enzyme activity. *BMC plant biology*, 14, 90.
- Dreze, M., Carvunis, A.R., Charleatoux, B., Galli, M., Pevzner, S.J., Tasan, M., Braun, P., Vidal, M., Ecker, J.R. & Hill D.E. (2011). Evidence for network evolution in an *Arabidopsis* interactome map. *Science*, 333, 601–7.
- Driouich, A., Follet-Gueye, M.L., Bernard, S., Kousar, S., Chevalier, L., Vicré-gibouin, M. & Lerouxel, O. (2012). Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. *Frontiers in Plant Science*, 3(April), 1–15.
- Earley, K., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K. & Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. *The Plant Journal*, 45, 616–629.
- Ebert, B., Rautengarten, C., Guo, X., Xiong, G., Stonebloom, S., Smith-Moritz, A.M., Herter, T., Chan, L.J.G., Adams, P.D., Petzold, C.J., Pauly, M., Willats, W.G.T., Heazlewood, J.L. & Scheller, H.V. (2015). Identification and characterization of a Golgi-localised UDP-xylose transporter family from *Arabidopsis*. *Plant Cell*, 27, 1218–27.
- Ebert, B., Rautengarten, C., McFarlane, H.E., Rupasinghe, T., Zeng, W., Ford, K., Scheller, H.V., Bacic, A., Roessner, U., Persson, S. & Heazlewood, J.L. (2018a). A Golgi UDP-GlcNAc transporter delivers substrates for N-linked glycans and sphingolipids. *Nature Plants*, 4(10), 792–801.
- Ebert, B., Birdseye, D., Liwanag, A.J.M., Laursen, T., Rennie, E.A., Guo, X., Catena, M., Rautengarten, C., Stonebloom, S.H., Gluza, P., Pidatala, V.R., Andersen M.C.F., Cheetamun, R., Mortimer, J.C., Heazlewood, J.L., Bacic, A., Clausen, M.H., Willats, W.G.T. & Scheller, H.V. (2018b). The three members of the *Arabidopsis* glycosyltransferase family 92 are functional β -1,4-galactan synthases. *Plant & Cell Physiology*, 59(12), 2624–2636.
- Edwards, K., Johnstone, C. & Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, 19(6), 1349.
- Egelund, J., Petersen, B.L., Motawia, M.S., Damager, I., Faik, A., Olsen, C.E., Ishii, T., Clausen, H., Ulvskov, P. & Geshi, N. (2006). *Arabidopsis thaliana* RGXT1 and RGXT2 encode Golgi-localised (1,3)- α -D-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II. *The Plant Cell*, 18 (10), 2593–2607.
- Egelund, J., Damager, I., Faber, K., Olsen, C.E., Ulvskov, P. & Petersen, B.L. (2008) Functional characterisation of a putative rhamnogalacturonan II specific xylosyltransferase. *FEBS Letters*, 582, 3217–3222.
- Faik, A., Price, N.J., Raikhel, N.V. & Keegstra, K. (2002). An *Arabidopsis* gene encoding α -xylosyltransferase involved in xyloglucan biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 99(11), 7797.
- Farid, A., Pabst, M., Schoberer, J., Altmann, F., Glössl, J. & Strasser, R. (2011). *Arabidopsis thaliana* α 1,2-glucosyltransferase (ALG10) is required for efficient N-glycosylation and leaf growth. *The Plant Journal*, 68, 314–325.
- Feingold, D.S. (1982). Aldo (and keto) hexoses and uronic acids. In: *Plant Carbohydrates I, Intracellular Carbohydrates*, ed. Loewus, F.A. & Tanner, W. 13A:3–76. Berlin: Springer-Verlag.

- Field, C.B., Behrenfeld, M.J., Randerson, J.T. & Falkowski P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science*, 281, 237-240.
- Fields, S. & Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature*, 340(6230), 245-246.
- Finn, R.D., Coghill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, E.J., Mitchell, A.L., Potter, S.C., Punta, P.M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J. & Bateman, A. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research*, 44, 279–85.
- Freese, N.H., Norris, D.C. & Loraine, A.E. (2016). Integrated Genome Browser: Visual analytics platform for genomics. *Bioinformatics*, 32(14), 2089-95.
- Gehl, C., Waadt, R., Kudla, J., Mendel, R-R. & Hänsch, R. (2009). New GATEWAY vectors for high throughput analyses of protein-protein interactions by bimolecular fluorescence complementation. *Molecular Plant*, 2 (5), 1051 – 1058.
- Geshi, N., Johansen, J.N., Dilokpimol, A., Rolland, A., Belcram, K., Verger, S., Kotake, T., Tsumuraya, Y., Kaneko, S., Tryfona, T., Dupree, P., Scheller, H.V., Höfte, H. & Mouille, G. (2013). A galactosyltransferase acting on arabinogalactan protein glycans is essential for embryo development in *Arabidopsis*. *Plant Journal*, 76, 128–137.
- Gigli-Bisceglia, N., Engelsdorf, T. & Hamann, T. (2020). Plant cell wall integrity maintenance in model plants and crop species-relevant cell wall components and underlying guiding principles. *Cellular and molecular life sciences: CMLS*, 77(11), 2049–2077.
- Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. & Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40 (D1): D1178-1186.
- Grebe, M., Xu, J., Möbius, W., Ueda, T., Nakano, A., Geuze, H.J., Rook, M.B. & Scheres, B. (2003). *Arabidopsis* sterol endocytosis involves actin-mediated trafficking via ARA6-positive early endosomes. *Current Biology*, 13 (16), 1378-1387.
- Grefen, C., Donald, N., Hashimoto, K., Kudla, J., Schumacher, K. & Blatt, M.R. (2010). A ubiquitin-10 promoter-based vector set for fluorescent protein tagging facilitates temporal stability and native protein distribution in transient and stable expression studies. *The Plant Journal*, 64, 355-365.
- Hadley, B., Maggioni, A., Ashikov, A., Day, C.J., Haselhorst, T. & Tiralongo, J. (2014). Structure and function of nucleotide sugar transporters: Current progress. *Computational and Structural Biotechnology Journal*, 10(16), 23–32.
- Hadley, B., Litfin, T., Day, C.J., Haselhorst, T., Zhou, Y. & Tiralongo, J. (2019). Nucleotide sugar transporter SLC35 family structure and function. *Computational and structural biotechnology journal*, 17, 1123–1134.
- Hamant, O. & Traas, J. (2010). The mechanics behind plant development. *The New Phytologist*, 185(2), 369-85.
- Handford, M.G., Sicilia, F., Brandizzi, F. & Chung, J.H. (2004). *Arabidopsis thaliana* expresses multiple Golgi-localised nucleotide-sugar transporters related to GONST1. *Mol Genet Genomics*, 272, 397-410.
- Handford, M.G., Rodríguez-Furlán, C., Marchant, L., Segura, M., Gómez, D., Alvarez-Buylla, E., Xiong, G.-Y., Pauly, M. & Orellana, A. (2012). *Arabidopsis thaliana* AtUTr7 encodes a golgi-localised udp–glucose/udp–galactose transporter that affects lateral root emergence. *Molecular Plant*, 5 (6), 1263-1280.
- Hansen, S.F., Bettler, E., Wimmerová, M., Imberty, A., Lerouxel, O. & Breton, C. (2009). Combination of several bioinformatics approaches for the identification of new putative glycosyltransferases in *Arabidopsis*. *Journal of Proteome Research*, 8(2), 743-753.
- Hansen, S.F., Harholt, J., Oikawa, A. & Scheller, H.V. (2012). Plant glycosyltransferases beyond CAZy: a perspective on DUF families. *Frontiers in Plant Science*, 3(March), 1–10.
- Hao, Z., Avci, U., Tan, L., Zhu, X., Glushka, J., Pattathil, S., Eberhard, S., Sholes, T., Rothstein, G. E., Lukowitz, W., Orlando, R., Hahn, M.G. & Mohnen, D. (2014). Loss of *Arabidopsis* GAUT12/IRX8 causes anther indehiscence and leads to reduced G lignin associated with altered matrix polysaccharide deposition. *Frontiers in Plant Science*, 5, 357.
- Harholt, J., Jensen, J.K., Sorensen, S.O., Orfila, C., Pauly, M., Scheller & H.V. (2006). ARABINAN DEFICIENT 1 is a putative arabinosyltransferase involved in biosynthesis of pectic arabinan in *Arabidopsis*. *Plant Physiology*, 140, 49–58.
- Harholt, J., Suttangkakul, A. & Scheller, H.V. (2010). Biosynthesis of Pectin. *Plant Physiology*, 153(2), 384–395.
- Harholt, J., Jensen, J.K., Verhertbruggen, Y., Søgaard, C., Bernard, S., Nafisi, M., Poulsen, C.P., Geshi, N., Sakuragi, Y., Driouich, A., Knox, J.P. & Scheller, H.V. (2012). ARAD proteins associated with pectic Arabinan biosynthesis form complexes when transiently overexpressed *in planta*. *Planta*, 236(1), 115–128.
- Herget, S., Ranzinger, R., Maass, K. & Lieth, C.-W.v.D. (2008). GlycoCT—a unifying sequence format for carbohydrates. *Carbohydrate Research*, 343 (12), 2162-2171.

- Hill, J., Donald, K.A. & Griffiths, D.E. (1991). DMSO-enhanced whole cell yeast transformation. *Nucleic Acids Research*, 19, 5791.
- Himmel, M.E., Ding, S.Y., Johnson, D.K., Adney, W.S., Nimlos, M.R., Brady, J.W. & Foust, T.D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*, 315(5813), 804-7.
- Hofgen, R. & Willmitzer, L. (1998). Storage of competent cells for *Agrobacterium* transformation. *Nucleic Acids Research*, 16(20), 9877.
- Hooper, C.M., Castleden, I.R., Tanz, S.K., Aryamanesh, N. & Millar, A.H. (2017). SUBA4: the interactive data analysis centre for *Arabidopsis* subcellular protein locations. *Nucleic Acids Research*, 45(D1), D1064-D1074.
- Houston, K., Tucker, M.R., Chowdhury, J., Shirley, N. & Little, A. (2016). The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. *Frontiers in plant science*, 7, 984.
- Huala, E., Dickerman, A., Garcia-Hernandez, M., Weems, D., Reiser, L., LaFond, F., Hanley, D., Kiphart, D., Zhuang, J., Huang, W., Mueller, L., Bhattacharyya, D., Bhaya, D., Sobral, B., Beavis, B., Somerville, C. & Rhee, S.Y. (2001). The *Arabidopsis* Information Resource (TAIR): A comprehensive database and web-based information retrieval, analysis and visualization system for a model plant. *Nucleic Acids Research*, 29(1), 102-5.
- Im, H. (2011). The Inoue method for preparation and transformation of competent *E. coli*: "ultra-competent" cells. *Bio-101*, e143.
- Jackson, R.G., Lim, E.K., Li, Y., Kowalczyk, M., Sandberg, G., Hoggett, J., Ashford, D.A. & Bowles, D.J. (2001). Identification and biochemical characterization of an *Arabidopsis* indole-3-acetic acid glucosyltransferase. *Journal of Biological Chemistry*, 276(6), 4350-6.
- Jacobs, D.R.Jr. & Gallaher, D.D. (2004). Whole grain intake and cardiovascular disease: a review. *Current Artherosclerosis Reports*, 6(6), 415-23.
- Jaskolski, F., Mulle, C. & Manzoni, O.J. (2005). An automated method to quantify and visualize colocalized fluorescent signals. *Journal of Neuroscience Methods*, 146, 42–49.
- Jefferson, R.A., Kavanagh, T.A. & Bevan, M.W. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO journal*, 6(13), 3901–3907.
- Jensen, J.K., Sørensen, S.O., Harholt, J., Geshi, N., Sakuragi, Y., Møller, I., Zandleven, J., Bernal, A.J., Jensen, N.B., Sørensen, C., Pauly, M., Beldman, G., Willats, W.G. & Scheller, H.V. (2008). Identification of a xylogalacturonan xylosyltransferase involved in pectin biosynthesis in *Arabidopsis*. *Plant Cell*, 20, 1289–1302.
- Jing, B., Ishikawa, T., Soltis, N., Inada, N., Liang, Y., Murawska, G., Andeberhan, F., Pidatalla, R., Yu, X., Baidoo, E., Kawai-Yamada, M., Loque, D., Kliebenstein, D.J., Dupree, P. & Mortimer, J.C. (2018). GONST2 transports GDP-Mannose for sphingolipid glycosylation in the Golgi apparatus of *Arabidopsis*. *bioRxiv*, 346775.
- Johnson, K.L., Gidley, M.J., Bacic, A. & Doblin, M.S. (2018). Cell wall biomechanics: a tractable challenge in manipulating plant cell walls 'fit for purpose!' *Current Opinion in Biotechnology*, 49, 163-171.
- Jørgensen, K., Rasmussen, A.V., Morant, M., Bak, S. & Møller, B.L. (2005). Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. *Current Opinions in Plant Biology*, 8(3), 280-91.
- Jung, H.G. & Allen, M.S. (1995). Characteristics of plant cell walls affecting intake and digestibility of forages by ruminants. *Journal of Animal Science*, 73(9), 2774-90.
- Katzen, F. (2007). Gateway[®] recombinational cloning: a biological operating system. *Expert Opinion on Drug Discovery*, 2(4), 571-589.
- Keegstra, K. (2010). Plant Cell Walls. *Plant Physiology*, 154(October), 483–486.
- Kerppola T.K. (2006). Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nature protocols*, 1(3), 1278–1286.
- Klepikova, A.V., Kasianov, A.S., Gerasimov, E.S., Logacheva, M.D. & Penin, A.A. (2016). A high-resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *Plant Journal*, 88, 1058-1070.
- Knappe, S., Flügge, U.I. & Fischer K. (2003). Analysis of the plastidic phosphate translocator gene family in *Arabidopsis* and identification of new phosphate translocator -homologous transporters, classified by their putative substrate-binding site. *Plant Physiology*, 131(March), 1178–1190.
- Knoch, E., Dilokpimol, A. & Geshi, N. (2014). Arabinogalactan proteins: focus on carbohydrate-active enzymes. *Frontiers in Plant Science*, 11, 198.
- Koch G.W., Sillett S.C., Jennings G.M. & Davis S.D. (2004). The limits to tree height. *Nature*, 428, 851–854.

- Kong, Y., Zhou, G., Abdeen, A.A., Schafhauser, J., Richardson, B., Atmodjo, M.A., Jung, J., Wicker, L., Mohnen, D., Western, T. & Hahn, M.G. (2013). GALACTURONOSYLTRANSFERASE-LIKE5 is involved in the production of Arabidopsis seed coat mucilage. *Plant Physiology*, 163(3), 1203–1217.
- Koressaar, T. & Remm, M. (2007). Enhancements and modifications of primer design program Primer3. *Bioinformatics*, 23(10), 1289–91.
- Kudla, J. & Bock, R. (2016). Lighting the Way to Protein-Protein Interactions: Recommendations on Best Practices for Bimolecular Fluorescence Complementation Analyses. *The Plant cell*, 28(5), 1002–1008.
- Kumar, K., Yadav, S., Purayannur, S. & Verma, P.K. (2013). An alternative approach in Gateway[®] cloning when the bacterial antibiotic selection cassettes of the entry clone and destination vector are the same. *Molecular Biotechnology*, 54, 133.
- Kumar S, Stecher G, Li M, Knyaz C and Tamura K (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547–1549.
- Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., Karthikeyan, A.S., Lee, C.H., Nelson, W.D., Ploetz, L., Singh, S., Wensel, A. & Huala, E. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic acids research*, 40, D1202–10.
- Lao, J., Oikawa, A., Bromley, J.R., Mcinerney, P., Suttangkakul, A., Smith-moritz, A.M., Plahar, H., Chiu, T.Y., González Fernández-Niño, S.M., Ebert, B., Yang, F., Christiansen, K.M., Hansen, S.F., Stonebloom, S., Adams, P.D., Ronald, P.C., Hillson, N.J., Hadi, M.Z., Vega-Sánchez, M.E., Loqué, D., Scheller, H.V. & Heazlewood, J.L. (2014). The plant glycosyltransferase clone collection for functional genomics. *The Plant Journal*, 79(3), 517–29.
- Lee, S., Kim, M.-H., Lee, J. H., Jeon, J., Kwak J.M. & Kim, Y.J. (2020) Glycosyltransferase-Like RSE1 Negatively Regulates Leaf Senescence Through Salicylic Acid Signaling in Arabidopsis. *Frontiers in Plant Science*, 11, 551.
- Lentze, N. & Auerbach, D. (2008). Membrane-based yeast two-hybrid system to detect protein interactions. *Current Protocols in Protein Science*, 19(17), 1–28.
- Li, W., Guan, Q., Wang, Z.Y., Wang, Y. & Zhu, J. (2013). A bi-functional xyloglucan galactosyltransferase is an indispensable salt stress tolerance determinant in Arabidopsis. *Molecular Plant*, 6(4), 1344–54.
- Liu, C.T., Yuan, S. & Li, K.C. (2009). Patterns of co-expression for protein complexes by size in *Saccharomyces cerevisiae*. *Nucleic acids research*, 37(2), 526–532.
- Liu, X.-L., Liu, L., Niu, Q.-K., Xia, C., Yang, K.-Z., Li, R., Chen, L.-Q., Zhang, X.-Q., Zhou, Y. & Ye, D. (2011). MALE GAMETOPHYTE DEFECTIVE 4 encodes a rhamnogalacturonan II xylosyltransferase and is important for growth of pollen tubes and roots in Arabidopsis. *The Plant Journal*, 65, 647–660.
- Liu, C., Niu, G., Zhang, H., Sun, Y., Sun, S., Yu, F., Lu, S., Yang, Y., Li, J. & Hong, Z. (2018). Trimming of N-glycans by the Golgi-localised α -1,2-mannosidases, MNS1 and MNS2, is crucial for maintaining rsw2 protein abundance during salt stress in Arabidopsis. *Molecular Plant*, 11 (5), 678–690.
- Liwanag, A.J.M., Ebert, B., Verherbruggen, Y., Rennie, E.A., Rautengarten, C., Oikawa, A., Andersen, M.C.F., Clausen, M.H. & Scheller, H.V. (2012). Pectin biosynthesis: GALS1 in Arabidopsis thaliana is a β -1,4-galactan β -1,4-galactosyltransferase. *The Plant Cell*, 24(December), 5024–5036.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M. & Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy). *Nucleic Acids Research*, 1 (42), 490–495.
- Luzarowski, M., Kosmacz, M., Sokolowska, E., Jasinska, W., Willmitzer, L., Veyel, D. & Skirycz, A. (2017). Affinity purification with metabolomic and proteomic analysis unravels diverse roles of nucleoside diphosphate kinases. *Journal of Experimental Botany*, 68(13), 3487–3499.
- Manders, E., Stap, J., Brakenhoff, G., van Driel, R. & Aten, J. (1992). Dynamics of three-dimensional replication patterns during the S phase, analysed by double labelling of DNA and confocal microscopy. *Journal of Cell Science*, 103, 857–862.
- Maszcak-Seneczko, D., Sosicka, P., Majkowski, Olczak, T. & Olczak, M. (2012). UDP-N-acetylglucosamine transporter and UDP-galactose transporter form heterologous complexes in the Golgi membrane. *FEBS Letters*, 586, 4082–7.
- Maszcak-Seneczko, D., Sosicka, P., Kaczmarek, B., Majkowski, M., Luzarowski, M., Olczak, T. & Olczak, M. (2015). UDP-galactose (SLC35A2) and UDP-N-acetylglucosamine (SLC35A3) transporters form glycosylation-related complexes with mannoside acetylglucosaminyltransferases (MGATs). *Journal of Biological Chemistry*, 290(25), 15475–15486.
- Mazzarino, L., Loch-Neckel, G., Dos Santos Bubniak, L., Ourique, F., Otsuka, I., Halila, S., Curi Pedrosa, R., Santos-Silva, M.C., Lemos-Senna, E. Curti Muniz, E. & Borsali, R (2015). Nanoparticles made from xyloglucan-block-polycaprolactone copolymers: safety assessment for drug delivery, *Toxicology Sciences*, 147, 104–115.
- McFarlane, H.E., Döring, A. & Persson, S. (2014a). The cell biology of cellulose synthesis. *Annual Review of Plant Biology*, 65, 69–94.

- McFarlane, H.E., Gendre, D. & Western, T.L. (2014b). Seed coat ruthenium red staining assay. *Bio-protocol*, 4(7): e1096.
- Meents, M.J., Watanabe, Y. & Samuels, A.L. (2018). The cell biology of secondary cell wall biosynthesis. *Annals of Botany*, 121(6), 1107–1125.
- Miles, E.W., Rhee, S. & Davies, D.R. (1999). The molecular basis of substrate channelling. *Journal of Biological Chemistry*, 274, 12193–6.
- Mohnen, D. (2008). Pectin structure and biosynthesis. *Current Opinion in Plant Biology*, 11 (3), 266–77.
- Mortimer, J.C., Miles, G.P., Brown, D.M., Zhang, Z., Segura, M.P., Weimar, T., Yu, X., Seffen, K.A., Stephens, E., Turner, S.R. & Dupree, P. (2010). Absence of branches from xylan in *Arabidopsis* gux mutants reveals potential for simplification of lignocellulosic biomass. *Proceedings of the National Academy of Sciences of the United States of America*, 107(40), 17409–17414.
- Mortimer, J.C., Yu, X., Albrecht, S., Sicilia, F., Huichalaf, M., Ampuero, D., Michaelson, L.V., Murphy, A.M., Matsunaga, T., Kurz, S., Stephens, E., Baldwin, T.C., Ishii, T., Napier, J.A., Weber, A.P.M., Handford, M.G. & Dupree, P. (2013). Abnormal glycosphingolipid mannosylation triggers salicylic acid-mediated responses in *Arabidopsis*. *The Plant Cell*, 25(5), 1881–1894.
- Myburg, A.A., Grattapaglia, D., Tuskan, G.A., Hellsten, U., Hayes, R.D., Grimwood, J., Jenkins, J., Lindquist, E., Tice, H., Bauer, D., (...) & Schmutz, J. (2014). The genome of *Eucalyptus grandis*. *Nature*, 510 7505 356–62.
- Mutwil, M., Klie, S., Tohge, T., Giorgi, F.M., Wilkins, O., Campbell, M.M., Fernie, A.R., Usadel, B., Nikoloski, Z. & Persson, S. (2011). PlaNet: combined sequence and expression comparisons across plant networks derived from seven species. *The Plant Cell*, 23(3), 895–910.
- Nagashima, Y., von Schaewen, A. & Koiwa, H. (2018). Function of N-glycosylation in plants. *Plant Science*, 274, 70–79.
- Napier, R.M., Fowke, L.C., Hawes, C., Lewis, M. & Pelham, H.R. (1992). Immunological evidence that plants use both HDEL and KDEL for targeting proteins to the endoplasmic reticulum. *Journal of Cell Science*, 102, 261–271.
- Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J., Venditto, I., Briggs, J., Zhang, X., Labourel, A., Terrapon, N., Buffetto, F., Nepogodiev, S., Xiao, Y., Field, R.A., Zhu, Y., O'Neil, M.A., Urbanowicz, B.R., York, W.S., Davies, G.J., Abbott, D.W., Ralet, M.C., Martens, E.C., Henrissat, B. & Gilbert, H.J. (2017). Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature*, 544(7648), 65–70.
- Nebenführ, A., Gallagher, L.A., Dunahay, T.G., Frohlick, J.A., Mazurkiewicz, A.M., Meehl, J.B. & Stachelin, L.A. (1999). Stop-and-Go Movements of Plant Golgi Stacks Are Mediated by the Acto-Myosin System. *Plant Physiology*, 121 (4), 1127–1141.
- Nelson, B.K., Cai, X. & Nebenführ, A. (2007). A multicolored set of *in vivo* organelle markers for co-localisation studies in *Arabidopsis* and other plants. *The Plant Journal*, 51, 1126–1136.
- Niemann, M.C., Bartrina, I., Ashikov, A., Weber, H., Novák, O., Spíchal, L., Strnad, M., Strasser, R., Bakker, H., Schmülling, T. & Werner, T. (2015). *Arabidopsis* ROCK1 transports UDP-GlcNAc/UDP-GalNAc and regulates ER protein quality control and cytokinin activity. *Proceedings of the National Academy of Sciences of the United States of America*, 112(1), 291–6.
- Norambuena, L., Marchant, L., Berninsone, P., Hirschberg C.B., Sliva, H. & Orellana, A. (2002). Transport of UDP-galactose in plants. *The Journal of Biological Chemistry*, 277(36), 32923–32929.
- Norambuena, L., Nilo, R. & Handford, M. (2005). AtUTr2 is an *Arabidopsis thaliana* nucleotide sugar transporter located in the Golgi apparatus capable of transporting UDP-galactose. *Planta*, 222(3), 521–9.
- Ogawa-Ohnishi, M. & Matsubayashi, Y. (2015). Identification of three potent hydroxyproline O-galactosyltransferases in *Arabidopsis*. *Plant Journal*, 81, 736–746.
- Oikawa, A., Lund, C.H., Sakuragi, Y. & Scheller, H.V. (2013). Golgi-localized enzyme complexes for plant cell wall biosynthesis. *Trends in Plant Science*, 18 (1), 1360–1385.
- Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K., Thibaud-Nissen, F., Malek, R.L., Lee, Y., Zheng, L., Orvis, J., Haas, B., Wortman, J. & Buell, C.R. (2007). The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic acids research*, 35, D883–7.
- Parker, J.L. & Newstead, S. (2017). Structural basis of nucleotide sugar transport across the Golgi membrane. *Nature*, 551(7681), 521–524.
- Parker, J.L. & Newstead, S. (2019). Gateway to the Golgi: molecular mechanisms of nucleotide sugar transporters. *Current Opinion in Structure Biology*, 57, 127–134.
- Parra-Rojas, J.P., Largo-Gosens, A., Carrasco, T., Celiz-Balboa, J., Arenas-Morales, V., Sepúlveda-Orellana, P., Temple, H., Sanhueza, D., Reyes, F.C., Meneses, C., Saez-Aguayo, S. & Orellana, A. (2019). New steps in mucilage biosynthesis revealed

- by analysis of the transcriptome of the UDP-rhamnose/UDP-galactose transporter 2 mutant. *Journal of experimental botany*, 70(19), 5071–5088.
- Parsons, H.T., Stevens, T.J., McFarlane, H.E., Vidal-Melgosa, S., Griss, J., Lawrence, N., Butler, R., Sousa, M.M.L., Salemi, M., Willats, W.G.T., Petzold, C.J., Heazlewood, J.L. & Lilley, K.S. (2019). Separating Golgi proteins from cis to trans reveals underlying properties of cisternal localisation. *The Plant Cell*, 31(9), 2010–2034.
- Persson, S., Wei, H., Milne, J., Page, G.P. & Somerville, C.R. (2005). Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proceedings of the National Academy of Sciences of the United States of America*, 102(24), 8633–8638.
- Pettolino, F.A., Walsh, C., Fincher, G.B. & Bacic, A. (2012). Determining the polysaccharide composition of plant cell walls. *Nature Protocols*, 7(9), 1590–1607.
- Phizicky, E.M., & Fields, S. (1995). Protein-protein interactions: methods for detection and analysis. *Microbiology and Molecular Biology Reviews*, 59(1), 94–123.
- Qu, Y., Egelund, J., Gilson, P.R., Houghton, F., Gleeson, P.A., Schultz, C.J. & Bacic, A. (2008). Identification of a novel group of putative *Arabidopsis thaliana* β (1,3)-galactosyltransferases. *Plant Molecular Biology*, 68, 43–59.
- Rao, X., Huang, X., Zhou, Z., & Lin, X. (2013). An improvement of the $2^{-\Delta\Delta CT}$ method for quantitative real-time polymerase chain reaction data analysis. *Biostatistics, bioinformatics and biomathematics*, 3(3), 71–85.
- Rautengarten, C., Ebert, B., Moreno, I., Temple, H., Herter, T., Link, B., Doñas-Cofré, D., Moreno, A., Saéz-Aguayo, S., Blanco, F., Mortimer, J.C., Schultink, A., Reiter, W.D., Dupree, P., Pauly, M., Heazlewood, J.L., Scheller, H.V. & Orellana, A. (2014). The Golgi localised bifunctional UDP-rhamnose/UDP-galactose transporter family of *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 119, 1–6.
- Rautengarten, C., Ebert, B., Liu, L., Stonebloom, S., Smith-Moritz, A.M., Pauly, M., Orellana, A., Scheller, H.V. & Heazlewood, J.L. (2016). The *Arabidopsis* Golgi-localised GDP-fucose transporter is required for plant development. *Nature Communication*, 7, 1–10.
- Rautengarten, C., Birdseye, D., Pattathil, S., McFarlane, H.E., Saez-Aguayo, S., Orellana, A., Persson, S., Hahn, M.G., Scheller, H.V., Heazlewood, J.L. & Ebert, B. (2017). The elaborate route for UDP-arabinose delivery into the Golgi of plants. *Proceedings of the National Academy of Science*, 114 (16), 4261–4266.
- Reiter, W.D. (2008). Biochemical genetics of nucleotide sugar interconversion reactions. *Current Opinion in Plant Biology*, 11 (3), 236–243.
- Reiter, W.D. & Vanzin G.F. (2001). Molecular genetics of nucleotide sugar interconversion pathways in plants. *Plant Molecular Biology*, 47(1-2), 95–113.
- Ren, Y., Hansen, S.F., Ebert, B., Lau, J. & Scheller, H.V. (2014). Site-directed mutagenesis of IRX9, IRX9L and IRX14 proteins involved in xylan biosynthesis: glycosyltransferase activity is not required for IRX9 function in *Arabidopsis*. *PLOS ONE* 9(8), e105014.
- Rennie, E.A., Hansen, S.F., Baidoo, E.E., Hadi, M.Z., Keasling, J.D. & Scheller, H.V. (2012). Three members of the *Arabidopsis* glycosyltransferase family 8 are xylan glucuronosyltransferases. *Plant Physiology*, 159(4), 1408–1417.
- Reyes, F., Marchant, L., Norambuena, L., Nilo, R., Silva, H. & Orellana, A. (2006). AtUTr1, a UDP-glucose/UDP-galactose transporter from *Arabidopsis thaliana*, is located in the endoplasmic reticulum and up-regulated by the unfolded protein response. *Journal of Biological Chemistry*, 281, 9145–51.
- Reyes, F. & Orellana, A. (2008). Golgi transporters: opening the gate to cell wall polysaccharide biosynthesis. *Current Opinions in Plant Biology*, 11(3), 244–51.
- Reyes, F., León, G., Donoso, M., Brandizzi, F., Weber, A.P. & Orellana, A. (2010). The nucleotide sugar transporters AtUTr1 and AtUTr3 are required for the incorporation of UDP-glucose into the endoplasmic reticulum, are essential for pollen development and are needed for embryo sac progress in *Arabidopsis thaliana*. *The Plant Journal*, 61, 423–435.
- Ridley, B.L., O'Neill, M.A. & Mohnen, D. (2001). Pectins: structure, biosynthesis, and oligogalacturonide-related signalling. *Phytochemistry*, 57(6), 929–67.
- Rocha, J., Cicéron, F., de Sanctis, D., Lelimosin, M., Chazalet, V., Lerouxel, O. & Breton, C. (2016a). Structure of *Arabidopsis thaliana* FUT1 reveals a variant of the GT-B class fold and provides insight into xyloglucan fucosylation. *The Plant Cell*, 28(10), 2352–2364.
- Rocha, J., Sarkis, J., Thomas, A., Pitou, L., Radzimanowski, J., Audry, M., Chazalet, V., de Sanctis, D., Palcic, M.M., Block, M.A., Girard-Egrot, A., Maréchal, E. & Breton, C. (2016b). Structural insights and membrane binding properties of MGD1, the major galactolipid synthase in plants. *Plant Journal*, 85, 622–633.

- Rollwitz, I., Santaella, M., Hille, D. & Flu, U. (2006). Characterization of AtNST-KT1, a novel UDP-galactose transporter from *Arabidopsis thaliana*. *FEBS Letters*, 580, 4246–4251.
- Rösti, J., Barton, C. J., Albrecht, S., Dupree, P., Pauly, M., Findlay, K., Roberts, K., & Seifert, G. J. (2007). UDP-glucose 4-epimerase isoforms UGE2 and UGE4 cooperate in providing UDP-galactose for cell wall biosynthesis and growth of *Arabidopsis thaliana*. *The Plant cell*, 19(5), 1565–1579.
- Rueden, C.T., Schindelin, J., Hiner, M.C., DeZonia, B.E., Walter, A.E., Arena, E.T. & Eliceiri, K.W. (2017). ImageJ2: ImageJ for the next generation of scientific image data, *BMC Bioinformatics*, 18, 529.
- Rui, Y. & Dinneny, J.R. (2020). A wall with integrity: surveillance and maintenance of the plant cell wall under stress. *New Phytologist*, 225, 1428–1439.
- Saez-Aguayo, S., Rautengarten, C., Temple, H., Sanhueza, D., Ejsmentewicz, T., Sandoval-Ibañez, O., Doñas, D., Parra-Rojas, J.P., Ebert, B., Lehner, A., Mollet, J.C., Dupree, P., Scheller, H.V., Heazlewood, J.L., Reyes, F.C. & Orellana, A. (2017). UUAT1 Is a Golgi-localised UDP-uronic acid transporter that modulates the polysaccharide composition of *Arabidopsis* seed mucilage. *The Plant Cell*, 29(1), 129–143.
- Saha, D. & Bhattacharya, S. (2010). Hydrocolloids as thickening and gelling agents in food: a critical review. *Journal of food science and technology*, 47(6), 587–597.
- Scheible, W.R. & Pauly, M. (2004). Glycosyltransferases and cell wall biosynthesis: Novel players and insights. *Current Opinion in Plant Biology*, 7(3), 285–295.
- Scheller, H.V. & Ulyskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology*, 61, 263–289.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9, 676–68.
- Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D. & Lohmann, J.U. (2005). A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics*, 37(5), 501–6.
- Schmittgen, T.D. & Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nature Protocols*, 3, 1101–1108.
- Schneider, R., Hanak, T., Persson, S. & Voigt, C.A. (2016). Cellulose and callose synthesis and organization in focus, what's new?. *Current Opinion in Plant Biology*, 34, 9–16.
- Schoberer, J. & Strasser, R. (2011). Sub-compartmental organization of Golgi-resident N-glycan processing enzymes in plants. *Molecular Plant*, 4 (2), 220–228.
- Schoberer, J., Liebminger, E., Botchway, S.W., Strasser, R. & Hawes, C. (2013). Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the Golgi stack *in planta*. *Plant Physiology*, 161 (4), 1737–1754.
- Sechet, J., Htwe, S., Urbanowicz, B., Agyeman, A., Feng, W., Ishikawa, T., Colomes, M., Kumar, K.S., Kawai-Yamada, M., Dinneny, J.R., O'Neill, M.A. & Mortimer, J.C. (2018), Suppression of *Arabidopsis* GGLT1 affects growth by reducing the L-galactose content and borate cross-linking of rhamnogalacturonan-II. *The Plant Journal*, 96, 1036–1050.
- Seifert, G.J., Barber, C., Wells, B., Dolan, L. & Roberts, K. (2002). Galactose biosynthesis in *Arabidopsis*: Genetic evidence for substrate channeling from UDP-D-galactose into cell wall polymers. *Current Biology*, 12(02), 1840–1845.
- Seifert, G.J. (2004). Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. *Current Opinions in Plant Biology*, 7(3), 277–84.
- Seifert, G.J. (2018). Mad moves of the building blocks - Nucleotide sugars find unexpected paths into cell walls. *Journal of Experimental Botany*, 69(5), 905–907.
- Smakowska, E., Skibior-Blaszczyk, R., Czarna, M., Kolodziejczak, M., Kwasniak-Owczarek, M., Parys, K., Funk, C. & Janska, H. (2016). Lack of FTSH4 Protease Affects Protein Carbonylation, Mitochondrial Morphology, and Phospholipid Content in Mitochondria of *Arabidopsis*: New Insights into a Complex Interplay. *Plant Physiology*, 171(4), 2516–35.
- Somssich, M. & Simon R. (2017). Studying protein-protein interactions *in planta* using advanced fluorescence microscopy. In: Busch W. (eds) *Plant Genomics. Methods in Molecular Biology*, vol 1610. Humana Press, New York, NY 2.
- Song, X., Zhang, B. & Zhou, Y. (2011). Golgi-localised UDP-glucose transporter is required for cell wall integrity in rice. *Plant signaling & behavior*, 6(8), 1097–1100.
- Sosicka, P., Bazan, B., Maszczak-Seneczko, D., Shauchuk, Y., Olczak, T. & Olczak, M. (2019). SLC35A5 protein — a Golgi complex member with putative nucleotide sugar transport activity. *International Journal of Molecular Sciences*, 20, 276.

- Stagljar, I., Korostensky, C., Johnsson, N. & te Heesen, S. (1998). A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 95(9), 5187–5192.
- Stonebloom, S., Ebert, B., Xiong, G., Pattathil, S., Birdseye, D., Lao, J., Pauly, M., Hahn, M.G., Heazlewood, J.L. & Scheller, H.V. (2016). A DUF-246 family glycosyltransferase-like gene affects male fertility and the biosynthesis of pectic arabinogalactans. *BMC Plant Biology*, 16, 90.
- Subedi, G.P., Johnson, R.W., Moniz, H.A., Moremen, K.W. & Barb, A. (2015). High Yield Expression of Recombinant Human Proteins with the Transient Transfection of HEK293 Cells in Suspension. *Journal of Visual Experiments*, (106), e53568.
- Svagan, A.J., Jensen, P., Dvinskikh, S.V., Furó, I. & Berglund L.A. (2010). Towards tailored hierarchical structures in cellulose nanocomposite biofoams prepared by freezing/freezing-drying. *Journal of Materials Chemistry*, 20, 6646–6654.
- Staehelin, L.A. & Moore, I. (1995). The plant Golgi apparatus: structure, functional organization and trafficking mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology*, 46, 261–288.
- Sterling, J.D., Quigley, H.F., Orellana, A. & Mohnen, D. (2001). The catalytic site of the pectin biosynthetic enzyme alpha-1,4-galacturonosyltransferase is located in the lumen of the Golgi. *Plant Physiology*, 127, 360–371.
- Sterling, J.D., Atmodjo, M.A., Inwood, S.E., Kumar Kolli, V.S., Quigley, H.F., Hahn, M.G. & Mohnen, D. (2006). Functional identification of an Arabidopsis pectin biosynthetic homogalacturonan galacturonosyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 5236–5241.
- Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., Yamaguchi-Shinozaki, K. & Shinozaki, K. (2002). Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana. *The Plant Journal*, 29, 417–426.
- Takenaka, Y., Kato, K., Ogawa-Ohnishi, M., Tsuruhama, K., Kajiura, H., Yagyu, K., Takeda, A., Takeda, Y., Kunieda, T., Hara-Nishimura, I., Kuroha, T., Nishitani, K., Matsubayashi Y. & Ishimizu T. (2018). Pectin RG-I rhamnosyltransferases represent a novel plant-specific glycosyltransferase family. *Nature Plants* 4, 669–676.
- Tedman-Jones, J.D., Lei, R., Jay, F., Fabro, G., Li, X., Reiter, W.D., Brearley, C. & Jones, J.D. (2008). Characterization of Arabidopsis mur3 mutations that result in constitutive activation of defence in petioles, but not leaves. *Plant J* 56:691–703.
- Temple, H., Saez-Aguayo, S., Reyes, F.C. & Orellana, A. (2016). The inside and outside: Topological issues in plant cell wall biosynthesis and the roles of nucleotide sugar transporters. *Glycobiology*, 26(9), 913–925.
- Tuskan, G.A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., (...) & Rokhsar D. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science (New York, N.Y.)*. 313 5793 1596–604.
- Urbanowicz, B.R., Peña, M.J., Moniz, H.A., Moremen, K.W. & York, W.S. (2014). Two Arabidopsis proteins synthesize acetylated xylan in vitro. *The Plant journal for cell and molecular biology*, 80(2), 197–206.
- Urbanowicz, B.R., Bharadwaj, V.S., Alahuhta, M., Peña, M.J., Lunin, V.V., Bomble, Y.J., Wang, S., Yang, J.Y., Tuomivaara, S.T., Himmel, M.E., Moremen, K.W., York, W.S. & Crowley, M.F. (2017). Structural, mutagenic and in silico studies of xyloglucan fucosylation in Arabidopsis thaliana suggest a water-mediated mechanism. *The Plant Journal*, 91(6), 931–949.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3, 34.1.
- Veit, C., König, J., Altmann, F. & Strasser, R. (2018). Processing of the terminal alpha-1,2-linked mannose residues from oligomannosidic N-glycans is critical for proper root growth. *Frontiers in Plant Science*, 9, 1807.
- Verbančič, J., Lunn, J.E., Stitt, M. & Persson, S. (2018). Carbon supply and the regulation of cell wall synthesis. *Molecular Plant*, 11 (1), 75–94.
- Voiniciuc, C., Schmidt, M.H.-W., Berger, A., Yang, B., Ebert, B., Scheller, H.V., North, H.M., Usadel, B. & Günl, M. (2015). MUCILAGE-RELATED10 produces galactoglucomannan that maintains pectin and cellulose architecture in Arabidopsis seed mucilage. *Plant Physiology*, 169 (1), 403–420.
- Voiniciuc, C., Engle, K.A., Günl, M., Dieluweit, S., Schmidt, M.H.-W., Yang, J.-Y., Moremen, K.W., Mohnen, D. & Usadel, B. (2018). Identification of key enzymes for pectin synthesis in seed mucilage. *Plant Physiology*, 178 (3), 1045–1064.
- Waadt, R., Schmidt, L.K., Lohse, M., Hashimoto, K., Bock, R. & Kudla, J. (2008). Multicolor bimolecular fluorescence complementation reveals simultaneous formation of alternative CBL/CIPK complexes *in planta*. *The Plant Journal*, 56, 505–516.

- Wesley, S.V., Helliwell, C.A., Smith, N.A., Wang, M.B., Rouse, D.T., Liu, Q., Gooding, P.S., Singh, S.P., Abbott, D., Stoutjesdijk, P.A., Robinson, S.P., Gleave, A.P., Green, A.G. & Waterhouse, P.M. (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant Journal*, 27(6), 581-90.
- Wilson, S.M., Ho, Y.Y., Lampugnani, E.R., Van de Meene, A.M.L., Bain, M.P., Bacic, A. & Doblin, M.S. (2015). Determining the subcellular location of synthesis and assembly of the cell wall polysaccharide (1,3; 1,4)- β -d-glucan in grasses. *The Plant Cell*, 27(3), 754-771.
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V. & Provart, N.J. (2007). An “electronic fluorescent pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS ONE*, 2(8), e718.
- Wouters, F.S., Bastiaens, P.I., Wirtz, K.W. & Jovin T.M. (1998). FRET microscopy demonstrates molecular association of non-specific lipid transfer protein (nsL-TP) with fatty acid oxidation enzymes in peroxisomes. *The EMBO Journal*, 17(24), 7179-89.
- Wydro, M., Kozubek, E. & Lehmann, P. (2006). Optimization of transient *Agrobacterium*-mediated gene expression system in leaves of *Nicotiana benthamiana*. *Acta Biochimica Polonica*, 53(2), 289-298.
- Xing, S., Wallmeroth, N., Berendzen, K. W. & Grefen, C. (2016). Techniques for the Analysis of Protein-Protein Interactions in Vivo. *Plant Physiology*, 171(2), 727-758.
- Yin, Y., Chen, H., Hahn, M.G., Mohnen, D., Xu, Y. (2010). Evolution and function of the plant cell wall synthesis-related glycosyltransferase family 8. *Plant Physiology*, 153, 1729-1746.
- Yin, Y., Huang, J., Gu, X., Bar-Peled, M. & Xu, Y. (2011). Evolution of plant nucleotide-sugar interconversion enzymes. *PLOS ONE*, 6(11), e27995.
- Zeng, W., Jiang, N., Nadella, R., Killen, T.L., Nadella, V. & Faik, A. (2010). A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47 and GT75 families and functions cooperatively. *Plant Physiology*, 154 (1), 78-97.
- Zhang, G.F. & Staehelin, L.A. (1992). Functional compartmentation of the Golgi apparatus of plant cells: immunocytochemical analysis of high-pressure frozen- and freeze-substituted sycamore maple suspension culture cells. *Plant Physiology*, 99, 1070-1083.
- Zhang, B., Liu, X., Qian, Q., Liu, L., Dong, G., Xiong, G. & Zeng, D. (2011). Golgi nucleotide sugar transporter modulates cell wall biosynthesis and plant growth in rice. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (12), 5110-5.
- Zhao, X., Liu, N., Shang, N., Zeng, W., Ebert, B., Rautengarten, C., Zeng, Q.Y., Li, H., Chen, X., Beahan, C., Bacic, A., Heazlewood, J.L. & Wu, A.M. (2018). Three UDP-xylose transporters participate in xylan biosynthesis by conveying cytosolic UDP-xylose into the Golgi lumen in *Arabidopsis*. *Journal of Experimental Botany*, 69(5), 1125-1134.
- Zhong, R. & Ye, Z.H. (2007). Regulation of cell wall biosynthesis. *Current Opinions in Plant Biology*, 10 (6), 564-72.
- Zhong, R., Cui, D. & Ye, Z.-H. (2019). Secondary cell wall biosynthesis. *New Phytologist*, 221, 1703-1723.
- Zhou, Y., Li, S., Qian, Q., Zeng, D., Zhang, M., Guo, L., Liu, X., Zhang, B., Deng, L., Liu, X., Luo, G., Wang, X. & Li, J. (2009). BC10, a DUF266-containing and Golgi-located type II membrane protein, is required for cell-wall biosynthesis in rice (*Oryza sativa* L.). *Plant Journal*, 57, 446-462.

SUPPLEMENTAL TABLES

Supplemental Table S1. Nucleotide sugars found in plants. See Feingold (1982), for a full list of sources cited.

Sugar residue	Nucleotides
L-arabinose	UDP, ADP, GDP
D-ribose	ADP
D-xylose	UDP, GDP
D-galactose	UDP, ADP, GDP
D-glucose	UDP, GDP, TDP, ADP
D-mannose	GDP, ADP, UDP
L-galactose	GDP, ADP, UDP
D-fructose	UDP, ADP, GDP
L-fucose	GDP
L-rhamnose	UDP
2-deoxy-D-glucose	UDP
D-digitoxose	UDP
D-apiose	UDP
D-glucuronic acid	UDP
D-galacturonic acid	UDP, TDP
D-mannuronic acid	GDP
L-guluronic acid	GDP
2-deoxy-2-acetamido-D-glucose	UDP, ADP
2-deoxy-2-acetamido-D-galactose	UDP
Oligosaccharides cellobiose	UDP
D-glucose + L-arabinose	UDP
D-glucose + L-mannose	GDP

Supplemental Table S2. NSTs characterized in Arabidopsis. ND – not determined,

* – putative substrates, where the biochemical analyses were not conducted.

Transporter name	AGI	Localization	Substrate	Reference
AtUTr1	At2g02810	ER	UDP-Glc*	Norambuena et al., 2002; Reyes et al., 2006, 2010
AtUTr2	At4g23010	Golgi	UDP-Gal*	Norambuena et al., 2005
AtUTr3	At1g14360	Golgi & ER	UDP-Glc*, UDP-Gal*	Norambuena et al., 2002; Reyes et al., 2010
AtUTr7	At4g31600	Golgi	UDP-Gal*, UDP-Glc*	Handford et al., 2012
CMP-SiaT	At5g41760	ND	CMP-Sia*	Bakker et al., 2008
GONST1	At2g13650	Golgi	GDP-Man, GDP-Glc, GDP-Fuc, GDP-Gal	Baldwin et al., 2001; Colleoni et al., 2010; Mortimer et al., 2013, Rautengarten et al., 2016
GONST2	At1g07290	ND	GDP-Man, GDP-Glc, GDP-Fuc, GDP-Gal	Handford et al., 2004, Rautengarten et al., 2016, Jing et al., 2018
GGLT1	At1g76340	ND	GDP-Gal	Sechet et al., 2018
GFT1	At5g19980	Golgi	GDP-Fuc	Rautengarten et al., 2016
GONST5	At1g21870	Golgi	GDP-Man*	Handford et al., 2004
ROCK1	At5g65000	ER	UDP-GlcNAc, UDP-GalNAc	Niemann et al., 2015; Ebert et al., 2018a
UAft1	At5g25400	Golgi	UDP-Araf	Rautengarten et al., 2017
UAft2	At5g11230	Golgi	UDP-Araf	Rautengarten et al., 2017
UAft3	At4g32390	Golgi	UDP-Araf	Rautengarten et al., 2017
UAft4	At2g25520	Golgi	UDP-Araf	Rautengarten et al., 2017
UUAT1	At5g04160	Golgi	UDP-GlcA, UDP-GalA	Saez-Aguayo et al., 2017
UUAT2	At3g1090	Golgi	UDP-GlcA, UDP-GalA	Saez-Aguayo et al., 2017
UUAT3	At5g05820	Golgi	UDP-GlcA, UDP-GalA	Saez-Aguayo et al., 2017
UUAT4	At3g11320	Golgi	UDP-GlcA, UDP-GalA	Saez-Aguayo et al., 2017
UUAT5	At1g12500	Golgi	UDP-GlcA, UDP-GalA	Saez-Aguayo et al., 2017
UGNT1	At4g32272	Golgi	UDP-GlcNAc	Ebert et al., 2018a
URGT1	At1g76670	Golgi	UDP-Rha,UDP-Gal	Bakker et al., 2008; Rautengarten et al., 2014
URGT2	At1g21070	Golgi	UDP-Rha,UDP-Gal	Rautengarten et al., 2014
URGT3	At5g42420	Golgi	UDP-Rha,UDP-Gal	Rautengarten et al., 2014
URGT4	At4g39390	Golgi	UDP-Rha,UDP-Gal	Rollwitz et al., 2006; Rautengarten et al., 2014
URGT5	At4g09810	Golgi	UDP-Rha,UDP-Gal	Rautengarten et al., 2014
URGT6	At4g34020	Golgi	UDP-Rha,UDP-Gal	Rautengarten et al., 2014
UXT1	At2g28315	Golgi	UDP-Xyl	Ebert et al., 2015
UXT2	At2g30460	Golgi	UDP-Xyl	Ebert et al., 2015
UXT3	At1g06890	Golgi	UDP-Xyl	Ebert et al., 2015

Supplemental Table S3. Oligonucleotides used for cloning of NSTs and GTs along with their respective native promoters

Gene	Primer	Sequence (5'-3')
At1g76670	FW_URGT1_NP_Mk2	CACCGGCGCTTATACAGAACATAACCAG
	RV_URGT1_NP	TGGTTTAGTGTCACCGAGTTCA
At1g21070	FW_URGT2_NP_1000bp	CACCACTGGCAATGGATTGATTCAC
	RV_URGT2_NP_1000bp	TGCTTTATTATTTCCAAGCTCCA
At1g21070	FW_URGT2_NP	CACCATACGGGGTGGAAATGTTGA
	RV_URGT2_NP	TGCTTTATTATTTCCAAGCTCCA
At5g42420	FW_URGT3_NP_Mk2	CACCATCACCATGTATGTCTGCG
	RV_URGT3_NP	ATCTTTTGTGCGGCCAAGTTC
At4g39390	FW_URGT4_NP_Mk2	CACCGTTGGTGGATTGCGGTTT
	RV_URGT4_NP	CACCTTTTCATCAGATTCGTTT
At4g09810	FW_URGT5_NP_Mk3	AGCATATAATGCTGCATTGGATTGAC
	RV_URGT5_NP_Mk3	CGCTTTATCTCGATGTTTCATCAGAATC
At1g34020	FW_URGT6_NP	CACCTATCAAACCGACTATCGTTTC
	RV_URGT6_NP	GGCTTTGTCTTCGTTGTCGT
At2g28315	FW_UXT1_NP	CACCTATTTATATGGGATGAAGATCTTGAG
	RV_UXT1_NP	TACTGGAGAATGCTTGCTAATTTTC
At2g30460	FW_UXT2_NP_Mk3	CACCCATGGTACTTCGCTCTTG
	RV_UXT2_NP	CGCTTGAAAATCTTTGTTTG
At1g06890	FW_UXT3_NP	CACCGAAGAGCCATTGTTCAATCAGTT
	RV_UXT3_NP	GGCTTGAAAATCTTTATTGAGTTC
At1g11940	At1g11940 NP FW	CGGTCTGTTTTGGCATATGAATG
	At1g11940 NP RV	TAAGGTTTTGTTGCTGATGATCC
At1g62305	At1g62305 NP FW	CGTGCAAGATTCCAAAGTTG
	At1g62305 NP RV	AAAGGTTGTTGTATCTATAGAGGATTC
At5g16170	At5g16170 NP FW	CCTTTTGTGGGGGATTAAATACC
	At5g16170 NP RV	TGGACCAAAACCAAGAACTATTGG

Supplemental Table S4. Oligonucleotides used for Sanger sequencing

Complementary to:	Primer	Sequence (5'-3')
Ubiquitin-10 promoter	pUBQ10_FW	GTGCGATCGAATTTGTAGAT
35S terminator	T35S_RV	CTTATCTGGGAACACTACACAC
BASTA resistance gene	BastaR FW	TCGATACTTACGTCACGTC
OCS terminator	OCS term RV	GGCGGTAAGGATCTGAGCTA
M13 universal vector	M13 FW	CGCCAGGGTTTTCCAGTCACGAC
	M13 RV	CAGGAAACAGCTATGAC
ADH1 promoter	pADH1 primer FW	GTTCTCGTTCCCTTTCTTC
LexA/35S promoter	FW lexA/p35S	GTCGAGGTAAGATTAGATATGG
Methionine-25 promoter	FW anti-PMet25	CGTGTAATACAGGGTCGTCAG
N-terminal Yellow Fluorescent Protein	YFP RV seq primer	GTCCAGCTCGACCAGGATG
	nYFP FW	CAGCCACAACGTCTATATCATGG
C-terminal Yellow Fluorescent Protein	cYFP FW	GGCATGGACGAGCTGTAC
	cYFP RV	GCCGTTCTTCTGCTTGTC
C-terminal ubiquitin	Cub seq primer	GATCCCTCCAGATCAACAAAG
	RV anti-Cub	GACAGCGTTCTACCGTCTTC
N-terminal ubiquitin	Nub FW	CAAGACAAGGAAGGAATCCCTG
	Nub RV	TTTACCGGTCAAAGTCTTGACG
GAL1 promoter	GAL1 primer FW	AATATACCTCTATACTTTAACGTC

Supplemental Table S5. Oligonucleotides used for the creation of new constructs

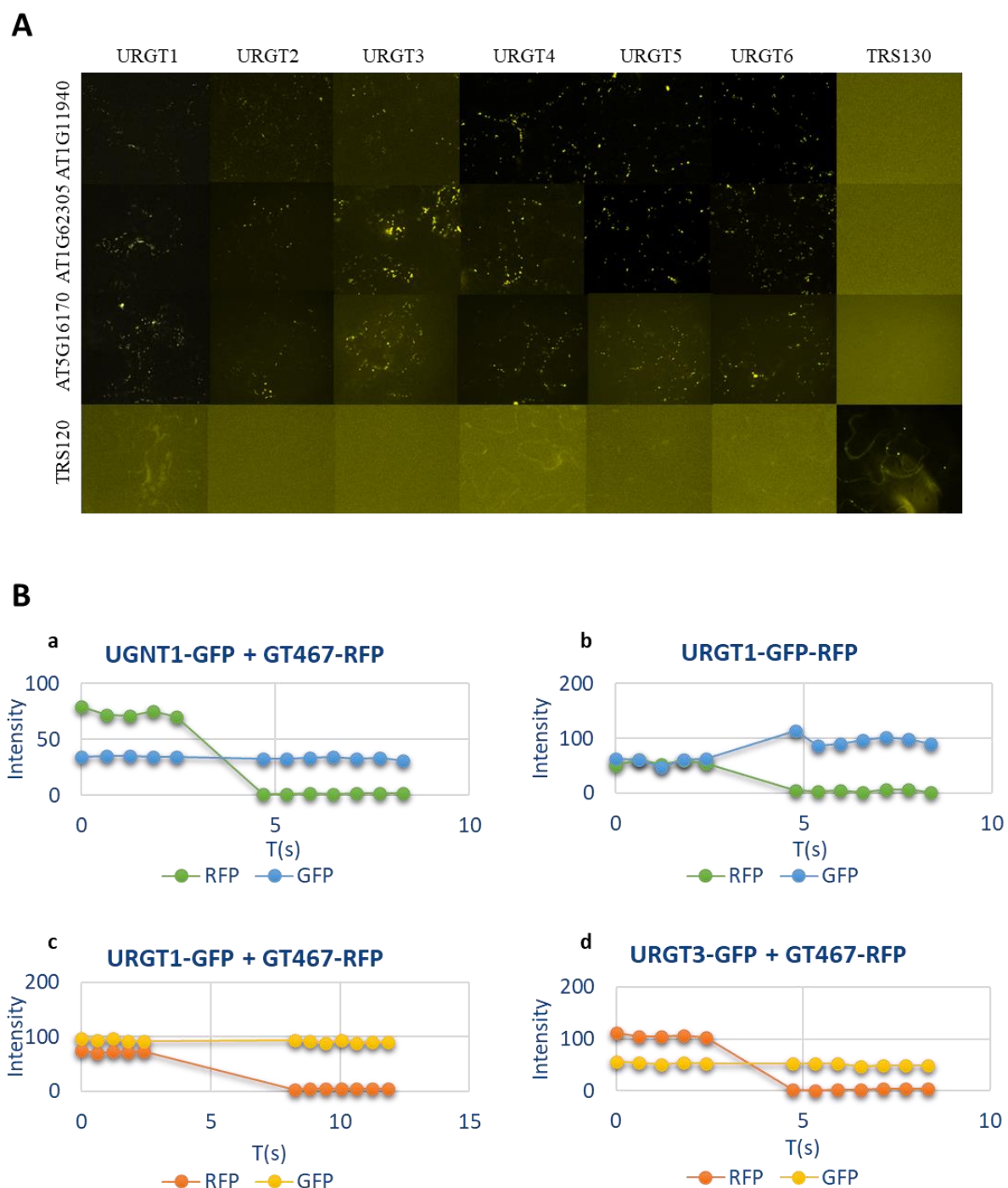
Used for construct:	Primer	Sequence (5'-3')
pUBC-GFP-Dest	GFP FW + SpeI	CCTCTAGAACTAGTATGGTGAGC
	GFP RV + PstI	GGCCGCTTATAACTTGTACAGCTCGTCCATG
	GFP FW + AvrII	ATGCATCCTAGGTGAGCAAGGGCGAGG
	GFP RV + XbaI	CTAGATCTAGACTTGTACAGCTCGTCCATGC
pUBC-RFP-HDEL	RFP FW + AvrII Mk2	GTGCCTAGGGCCTCCTCCGAGGACG
	RFP RV + XbaI Mk2	GTCTTAATTAATTGGCGCCGGTGGAG
	P1 XhoI FW	GGGGCYCGAGAYGAAGGGYCYCAYCYCCACCYC
	P2 SpeI RV	GGGGACTAGTTTCATCAGCCAGAGCTCCTCC
	P3 PstI+STOP+HDEL RV	GGGGTTATAATCATAAATCGTCGTGGGCGCCGGTGGAGTG
pUBC-NAG-RFP	NAG FW	ATGGCGAGGATCTCGTGTGAC
	NAG RV	AAAGTTCTTCGTCTGGCGGTTC
pUBC-GALT1-RFP	GALT1 FW XhoI	AACACTCGAGATGAAGAGATTTTATGGAGGGCTTC
	GALT1 RV SpeI	GCCATACTAGTAAAGTCAAGGGAGAGTAATCCG
pUBC-MUR3-RFP	MUR3 FW XhoI	AACACTCGAGATGTTTCAAGGGTTTCTATGAGG
	MUR3 RV SpeI	GCCATACTAGTAGCATAAATAGGTTGCAATTGGAG
pABindGFP - N-terminal version	AttR2 + SpeI RV	GATCGACTAGTACCACTTTGTACAAGAAAG
	AscI+GFP FW	CCTCGAGGCGCGCCATGGTGAGCAAGGG
	AttR2 + PacI RV	GCGTTAATTAACCACTTTGTACAAGAAAG

Supplemental Table S6. Oligonucleotides used for genotyping of plant lines

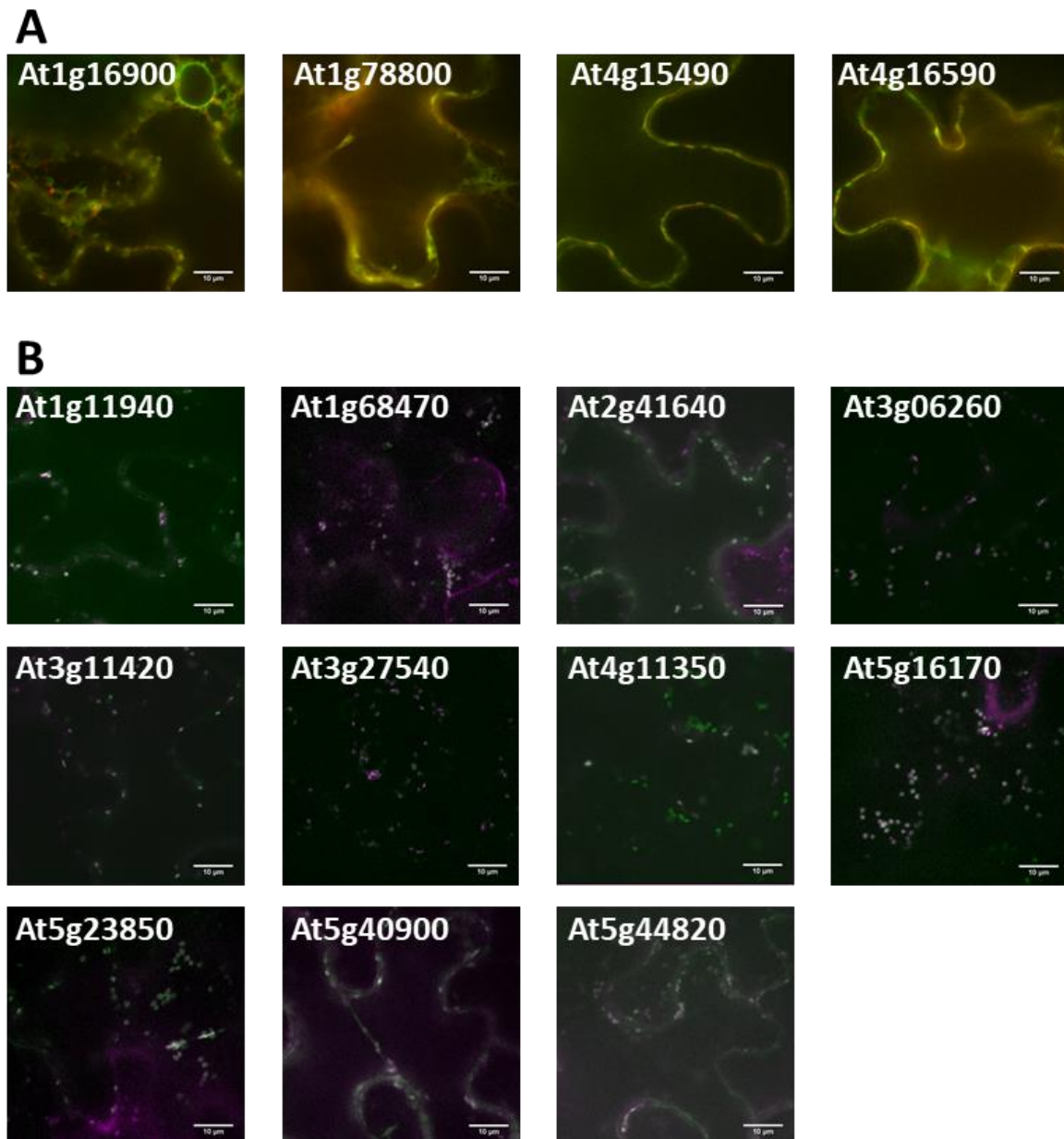
Gene	Primer	Sequence (5'-3')
SALK lines	SALK insertion	ATTTTGCCGATTTTCGGAAC
WISC lines	WISC insertion	AACGTCCGCAATGTGTTATTA
GABI lines	LBG_GABI	CCCATTGGACGTGAATGTAG
SAIL lines	SAIL_LB3	CATAACCAATCTCGATACAC
At3g11420	URGTi1.1-3 LP	CGAAAGTCGTACATCGTTG
	URGTi1.1-3 RP	CCCTTTAGACACTACACGTGATC
At2g37730	URGTih2.1 LP	CCCGGTTCAAGTTACACTTGC
	URGTih2.1 RP	GGCAGCTAGTAAACCGTATGG
At5g23850	URGTi3.1-3 LP	GTTCTGCTAGAGAAAAAGCCG
	URGTi3.1-3 RP	AGTCTGATTGCTTGTAACTTCC
At1g78800	URGT4.1-2 LP	TCTCTGTGTTCTTTGGGTTGG
	URGT4.1-2 RP	CAGCGGTTGAACAAGATGTG
At3g27540	URGTi6.1 LP	TCTTTGTGTTGATGCCATTTGG
	URGTi6.1 RP	CTCTTATGCAATTACCAGGCAAC
At1g68470	URGTi7.1 LP	GACAGCTTGGGATTTTCATGC
	URGTi7.1 RP	AACCTTGCTAGTCTTGAGAGTG
At5g44820	URGTi9.1 LP	CCAGAAACGGTTAGCTCTCATC
	URGTi9.1 RP	CCTTGACGAACAATAAACTCTTC
At4g19970	URGTi10.3 LP	CACTGCACAATCGCTTAACG
	URGTi10.3 RP	TGTCAGAAAGTTCGGTATGAGC
At1g21070	URGT2.1-3 LP	GCATCAAGGAAAGCATCAATGG
	URGT2.1-3 RP	AAGGTTCTTTGAGACAAAGAACG
At5g42420	URGT3.1 LP	TTCCGTTGGCTTCTACCAAG
	URGT3.1 RP	TGCGGCAAGTTCAACATC

Supplemental Table S7. Oligonucleotides used for amplification of full sequence CDS of selected genes

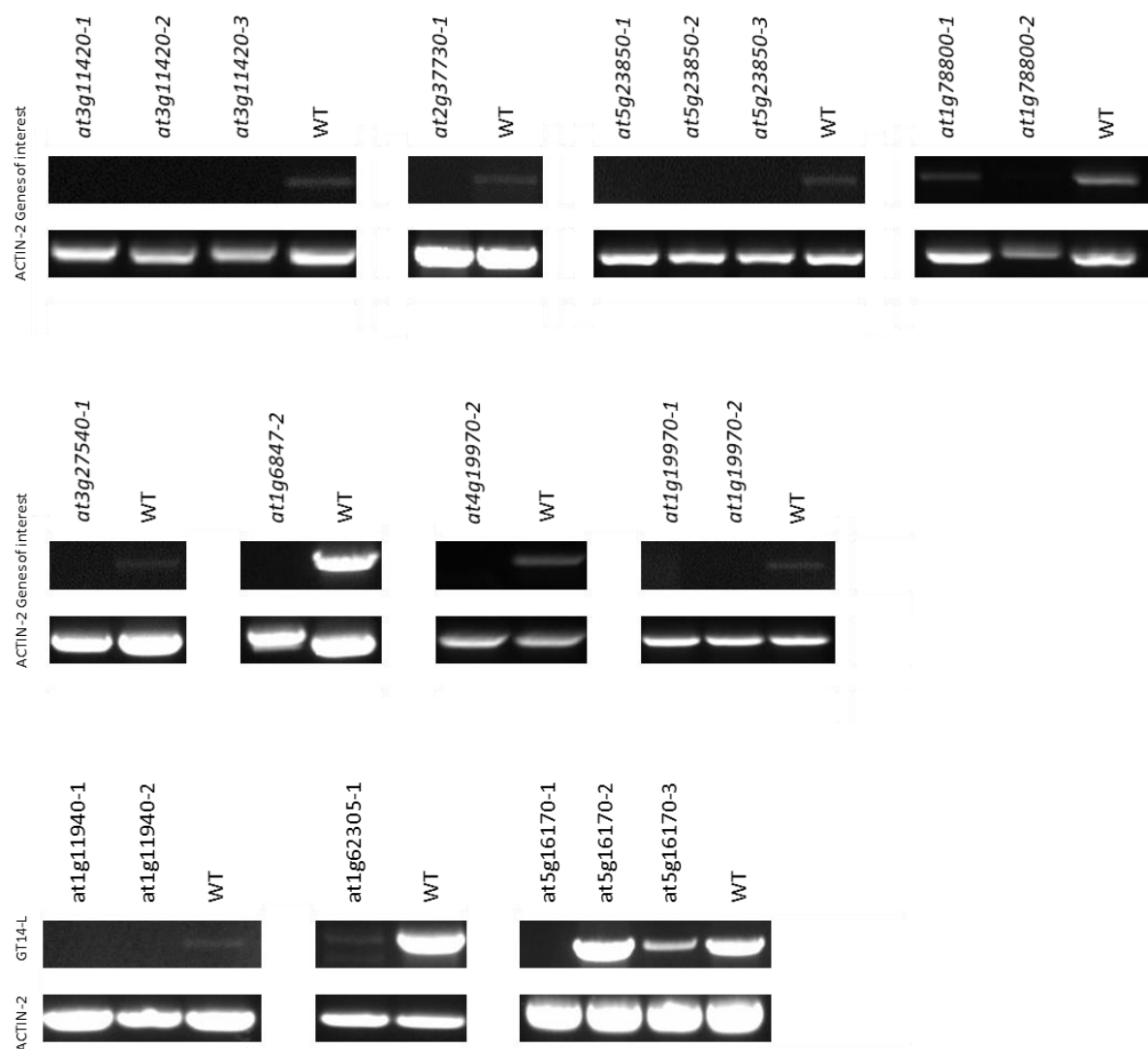
Gene	Primer	Sequence (5'-3')
At1g21070	FW CDS URG2	CACCATGGAGAAAGCAGAGAACGA
	RV CDS URG2	TGCTTTATTATTTCCAAGCTC
At3G18780	ACTIN2LP	GCCATCCAAGCTGTTCTCTC
	ACTIN2RP	GGGACTAAAAACGCAAAACGA
At3g11420	URGTi 1 LP	TGATCCAAAAGATGTGAGTTGA
	URGTi 1 RP	TGAAATCTTTTTGGTTCGGTTT
	At3g11420_fl_FW	ATGAAAGCAAATCGCGGCGG
	At3g11420_fl_RV	TTATATCTTCTCCAAAGATCTGC
AT2G37730	URGTih 2 LP	TCGATCCGTTTAAGCTCCAC
	URGTih 2 RP	TTTTGATATTCGGTTTGCTGA
	AT2G37730_fl_FW	ATGGCCATCTTCGATCCGTT
	AT2G37730_fl_RV	CTATATAGTCTCTGTAGATAAAC
AT5G23850	URGTi 3 LP	TTTTTCTCCCGCATGTCATT
	URGTi 3 RP	AAGAAACGACATCAGCTCGTA
	AT5G23850_fl_FW	ATGAGAAACTCGCCGTCCA
	AT5G23850_fl_RV	TCAAGAGCCGGTTTGATTCT
AT1G78800	URGTi 4 LP	AAAAATTGTTACGGGGAAA
	URGTi 4 RP	GGCCCATATAGCACCATCTC
	AT1G78800_fl_FW	ATGGCGAAAAAAGAAGGTTT
	AT1G78800_fl_RV	TCAATCTTCTTAGGACTTGAT
AT3G27540	URGTi 6 LP	CCGACGCGTCAACTAACTT
	URGTi 6 RP	GCACATACCAATCCAGAACTG
	AT3G27540_fl_FW	CGATGGATACATCAATTCGAAGA
	AT3G27540_fl_RV	TCATATAAACCTGTCTGTTTTCTC
AT1G68470	URGTi 7 LP	TCAAATAAAAGTCAACAGAATTCATAA
	URGTi 7 RP	GACAAGCGTGTTTCTATAACTGTCA
	AT1G68470_fl_FW	ATGACTTTCAATAAGAGGCAAGTC
	AT1G68470_fl_RV	TTACAATGAAACCACAACATTGTCT
AT5G40900	URGTi8 fullseq mark3 LP	AGTGAAATCACCTTACCTTGTC
	URGTi8 fullseq mark3 RP	CATCTGCATAACTGGCTTGG
AT5G44820	URGTi 9 LP	AAAAACACTTCCAAACAAATAATGA
	URGTi 9 RP	CTAGCATTCGGCTCGGAAA
	AT5G44820_fl_FW	ATGAGTTCTTCGTCTTCTCATC
	AT5G44820_fl_RV	TCAATCTTCCAAACACTTCATAGG
AT4G19970	URGTih 10 LP	AGGGAAATAGGAAAACAAAATTCA
	URGTih 10 RP	AATTAAGAAACAATAACCGACTCCA
	AT4G19970_fl_FW	ATGAAGAGCACGAACGACGA
	AT4G19970_fl_RV	TTAATAACCACACCTAGGCGGAAT
AT1G21070	URGT2 LP	TGGTAAACGACAAGCGTCAA
	URGT2 RP	CAATTTTGATTTTATGATTTGTTT
AT5G42420	URGT3 LP	TGTCAGATATTTTTCTCAACGTGAA
	URGT3 RP	GCTACGTAACAAGAACATGTGG
AT1G11940	AT1G11940 LP	CGTCGAACTATCAACATCATTCT
	AT1G11940 RP	TCGAAGCCAGTTTCTTCAGAC
	AT1G11940_fl_FW	GAAGAAATCACAGCCGCAGA
	AT1G11940_fl_RV	TTATAAGGTTTTGTTTGCTGATGA



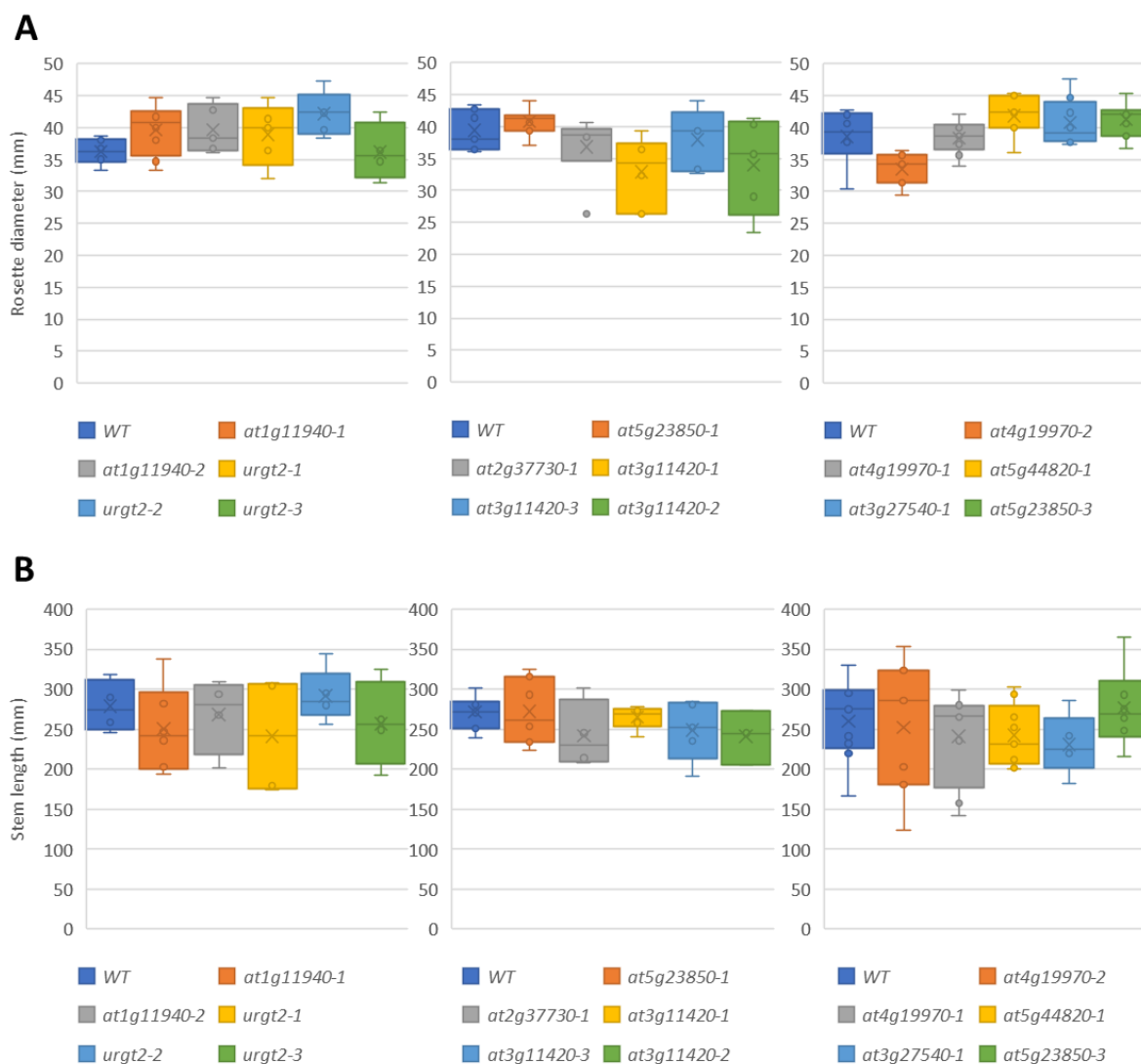
Supplemental Figure S1. Secondary methods used for validation of interactions – typical examples of observed results. **(A)** BiFC. Three GTs seemingly interact with six URGs in all of the combinations. TRS120 and TRS130 constitute negative controls. **(B)** APB-FRET. (a) negative control, (b) technical positive control, (c), (d) tested interactions. While covalently bound RFP and GFP engaged in FRET, this result could not be observed for any of the tested protein couples, including GAUT1-GAUT7 positive control (not shown in the figure).



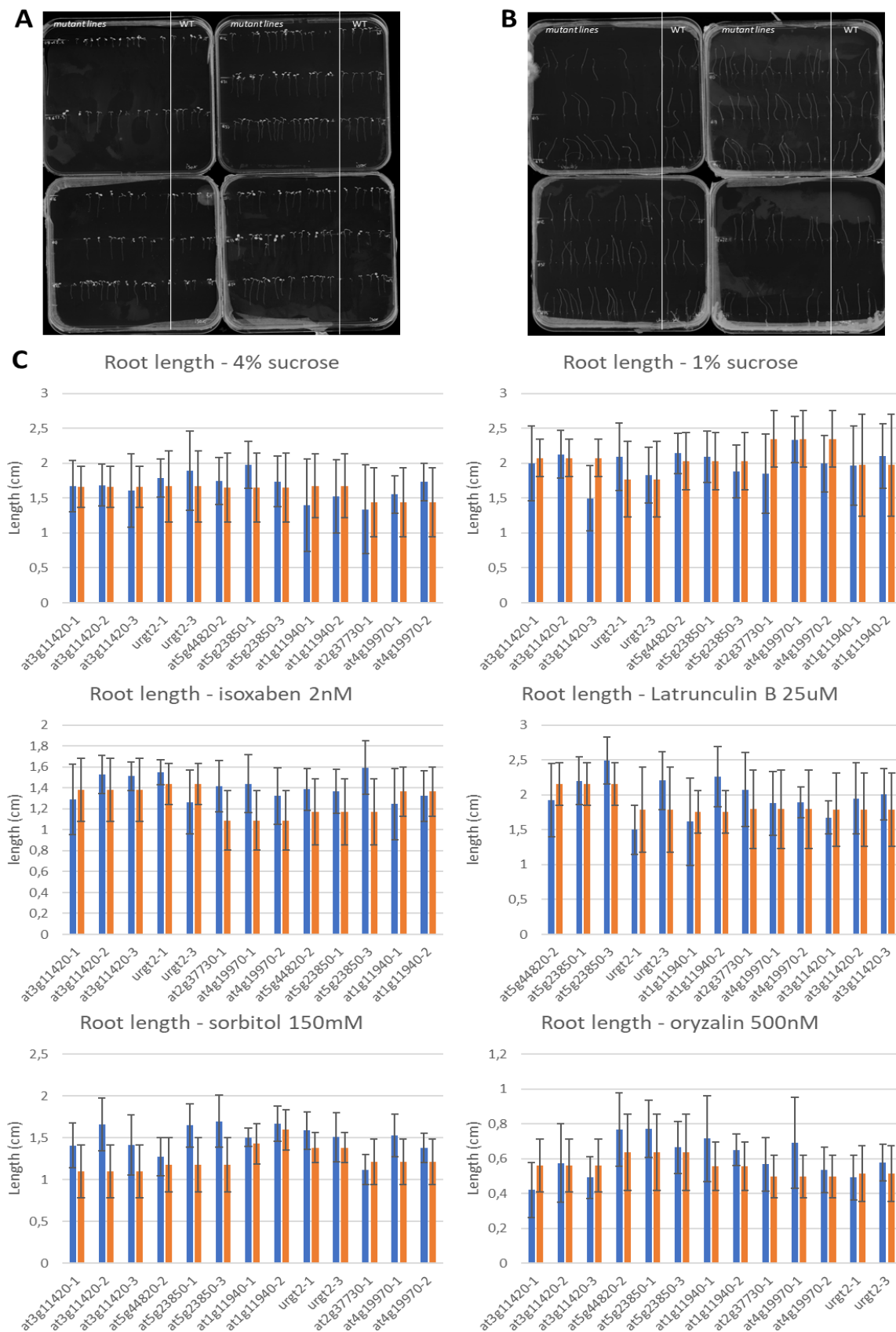
Supplemental Figure S2. Subcellular localisations of the GTs identified to be interacting with members of the URG1 family. Images of tobacco leaves expressing the *P_{35S}::GT-YFP* constructs colocalising with the ER (signal peptide from AtWAK2 fused to RFP and HDEL) and Golgi (transmembrane domain from soybean endo- β -mannanase 1 fused to RFP) markers (Nelson *et al.*, 2007). Red represents ER marker, green – GT, magenta – Golgi marker. **(A)** four of the GTs are being expressed in the ER (AT4G15490, AT1G78800, AT4G16590, AT1G16900). **(B)** The remaining 11 are localised exclusively to the Golgi apparatus.



Supplemental Figure S3. Validation of loss-of-function mutant lines. RT-PCR analysis to confirm the lack of the transcript in mutant plants. ACTIN-2 gene was used as a control. Lines *at1g78800-1* and *at5g16170-2* had residual transcript and were not further analysed.



Supplemental Figure S4. Analysis of the phenotype of loss-of-function GT mutant lines. **(A)** Rosette diameter measurements of GT mutant lines and *urgt* lines as control **(B)** Stem length measurements of GT mutant lines and *urgt* lines as control.



Supplemental Figure S5. Analysis of the phenotype of loss-of-function mutant lines in stress conditions. Seedlings grown on MS plates; selected lines and treatments are presented. **(A)** Exemplary root growth plates. **(B)** Exemplary DGH plates. **(C)** Measurements of root length of 7-day-old seedlings.