

# **Identification and characterization of plant lipid droplet-associated proteins**

Dissertation

for the award of the degree  
“Doctor rerum naturalium”  
of the Georg-August-Universität Göttingen

within the doctoral program  
“Plant Responses To Eliminate Critical Threats”  
of the Georg-August University School of Science (GAUSS)

submitted by  
**Philipp William Niemeyer**  
From Gehrden, Germany

Göttingen 2023

---

### **Thesis Committee**

Prof. Dr. Till Ischebeck, Green Biotechnology, Institute of Plant Biology and Biotechnology (IBBP), University of Münster

Prof. Dr. Andrea Polle, Department of Forest Botany and Tree Physiology, Büsgen-Institute, University of Göttingen

Prof. Harry Brumer, PhD, Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, Canada

### **Members of the Examination Board**

Referee: Prof. Dr. Till Ischebeck, Green Biotechnology, Institute of Plant Biology and Biotechnology (IBBP), University of Münster

2nd Referee: Prof. Dr. Andrea Polle, Department of Forest Botany and Tree Physiology, Büsgen-Institute, University of Göttingen

### **Further members of the Examination Board**

Prof. Dr. Kai Heimel, Department for Microbial Cell Biology, Institute for Microbiology and Genetics, University of Göttingen

Prof. Dr. Jan de Vries, Department of Applied Bioinformatics, Institute for Microbiology and Genetics, University of Göttingen

Prof. Dr. Ivo Feußner, Department for Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences, University of Göttingen

Prof. Dr. Gerhard Braus, Department of Molecular Microbiology and Genetics, Institute for Microbiology and Genetics, University of Göttingen

Date of oral examination: 29<sup>th</sup> of June, 2023

---

## Table of contents

<b>List of abbreviations</b>	<b>i</b>
<b>Abstract</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The lipid droplet, a unique organelle . . . . .	1
1.2 Synthesis of lipid droplets . . . . .	2
1.2.1 <i>De novo</i> synthesis of acyl chains . . . . .	2
1.2.2 Synthesis of neutral lipids . . . . .	4
1.2.3 Formation of a lipid droplet at the ER membrane . . . . .	5
1.2.4 Lipolysis/lipophagy marks the end of lipid droplets . . . . .	9
1.3 The proteome of lipid droplets . . . . .	10
1.3.1 Targeting of proteins to the LD surface . . . . .	10
1.3.2 Important plant LD proteins . . . . .	12
1.3.3 Identification of new lipid droplet proteins . . . . .	15
1.3.4 Lipid droplet-associated methyltransferases . . . . .	20
1.4 Functions of lipid droplets . . . . .	22
1.4.1 Lipid droplets as an energy source . . . . .	22
1.4.2 The involvement of LDs in protein degradation . . . . .	23
1.4.3 The role of LD in oxylipin synthesis and pathogen defense . . . . .	24
1.5 The role of LDs in desiccation tolerance and other abiotic stresses . . . . .	25
1.6 Aims of this thesis . . . . .	30
<b>2 Article I: A seed-like proteome in oil-rich tubers</b>	<b>32</b>
<b>3 Manuscript I: Core desiccation response strategies of plant development revealed by the dynamic proteome of <i>Physcomitrium patens</i> during spore germination.</b>	<b>50</b>

---

---

<b>4 Unpublished results I: The LD proteome of <i>Arabidopsis thaliana</i> roots</b>	<b>98</b>
<b>5 Unpublished results II: Characterization of lipid droplet-associated methyltransferases</b>	<b>105</b>
5.1 CNMTs possibly diverged from LIME-like proteins . . . . .	106
5.2 The LD targeting of LIME proteins . . . . .	108
5.2.1 The LD targeting of LIME relies on multiple protein segments . . . . .	109
5.2.2 LD targeting of LIME homologs is conserved . . . . .	114
5.2.3 Localization of LIME proteins expressed under the endogenous promoter	116
5.3 ColP supports dimeric conformation of LIME proteins . . . . .	117
5.4 Mutant lines were obtained for phenotype determination experiments . . . . .	119
5.5 LIME proteins do not affect stomata closure . . . . .	121
5.6 LIME proteins do not affect mildly dehydrated or UV-stressed plants . . . . .	122
5.7 LIME proteins seem to play no role in root development . . . . .	122
5.8 LIME is not important for seed germination under salt and heat stress . . . . .	124
5.9 LIME proteins do not affect sterol composition . . . . .	125
5.10 Non-targeted <i>ex vivo</i> approach does not reveal putative LIME1 substrates . . . . .	127
<b>6 Experimental procedures of unpublished results</b>	<b>129</b>
6.1 Generation of <i>Arabidopsis lime</i> mutant lines . . . . .	129
6.2 Analysis of potential phenotypes . . . . .	130
6.3 Plasmid construction . . . . .	131
6.4 Transient expression in tobacco pollen tubes . . . . .	133
6.5 <i>N. benthamiana</i> agrobacterium infiltration . . . . .	133
6.6 ColP for identification of potential LIME interactors . . . . .	134
6.7 ColP to immobilize proteins for <i>ex vivo</i> assay . . . . .	134
6.8 Axenic root cultivation . . . . .	135
6.9 Isolation of root LD-enriched and total protein fractions . . . . .	136

---

---

Table of contents

---

6.10 Protein processing and LC-MS/MS measurements . . . . .	136
6.11 Phylogenetic analysis of LIME proteins . . . . .	137
6.12 Sterol extraction and GC-MS . . . . .	139
6.13 <i>Ex vivo</i> assay . . . . .	139
6.14 Additional bioinformatics . . . . .	140
6.15 Materials . . . . .	141
6.15.1 Software . . . . .	141
6.15.2 Chemicals . . . . .	142
6.15.3 Media . . . . .	142
6.15.4 Lists of primers and microscope settings . . . . .	143
<b>7 Discussion</b>	<b>147</b>
7.1 Unmasking LD proteome's role in desiccation tolerance . . . . .	147
7.2 Abscisic acid and transcriptional regulators in desiccation tolerance . . . . .	152
7.3 The shift from dormant to vegetative LD proteome . . . . .	154
7.4 Shedding light on the LD proteome of Arabidopsis roots . . . . .	157
7.5 Proteomic data reveal diversity in LD proteomes across species and tissues . . . . .	163
7.6 New insights to LIME proteins, a highly conserved protein family . . . . .	168
7.7 Putative targeting of LIMEs to the LD in the context of established mechanisms . . . . .	172
<b>8 Concluding remarks</b>	<b>176</b>
<b>References</b>	<b>177</b>
<b>Supplementary materials</b>	<b>218</b>
<b>Acknowledgements</b>	<b>221</b>

---

## List of abbreviations

aa	amino acid
ACP	acyl carrier protein
ALDH	aldehyde dehydrogenase
ALDI	associated with lipid droplets I
ATG	autophagy-related
BIA	benzylisoquinoline alkaloid
bp	base pair
CAS	cycloartenol synthase
CB5-E	cytochrome B5 isoform E
CLO	caleosin
CNMT	<i>S</i> -coclaurine- <i>N</i> -methyltransferase
CoA	coenzyme A
CoIP	co-immunoprecipitation
COPI	coat protein complex I
CRISPR	clustered regularly interspaced short palindromic repeats
DAG	diacylglycerol
dai	days after imbibition
DGAT	diacylglycerol acyltransferase
DTT	dithiothreitol
ER	endoplasmic reticulum
ERAD	endoplasmic reticulum-associated protein degradation
ERD	early responsive to dehydration
FA	fatty acids
FAD	fatty acid desaturase
FAT	fatty acyl-ACP thioesterase
FAX	fatty acyl export
FDR	false discovery rate
FEY	forever young
FIT	fat storage-inducing transmembrane protein
G3P	glycerol-3-phosphate
GC-MS	gas chromatography coupled mass spectrometry

## Abbreviations

---

GFP	green fluorescent protein
GPAT	acyl-CoA:glycerol-3-phosphate-acyltransferase
HODE	hydroxy-octadecadienoic acid
HOT	hydroxyoctadecatrienoic acid
HPLC-MS/MS	high-performance liquid chromatography tandem mass spectroscopy
HPLF	hydroperoxide lyase
HPODE	hydroperoxide-octadecadienoic acid
HPOT	hydroperoxy-octadecatrienoic acid
HSD	hydroxysteroid dehydrogenase, steroleosin
iBAQ	intensity-based absolute quantification
LACS	long-chain acyl-CoA synthase
LB	lysogeny broth
LD	lipid droplets
LDAD	LD-associated protein degradation
LDAH	LD-associated hydrolase
LDAP	LD-associated protein
LDDH	LD dehydrogenase
LDIP	LDAP-interacting protein
LDNP	LD-localized NTF2 family protein
LDPS	LD protein of seeds
LDS	LDs and stomata
LDSP	LD surface protein
LEA	late embryogenesis abundant
LEW	leaf wilting
LFQ	label-free quantitation
LIDL	LD lipase
LIME	LD-associated methyltransferase
LIPA	LD plasma membrane adaptor
LOX	lipoxygenase
LPA	lysophosphatidic acid
LPAAT	acyl-CoA:LPA acyltransferase
LPCAT	lysophosphatidylcholine acyltransferase
LPEAT	acyl-CoA:lysophosphatidylethanolamine acyltransferase
MAGL	monoacyl glycerol lipase

## Abbreviations

---

MLDP	major LD protein
MS	Murashige and Skoog
MSTFA	trimethylsilyl-N-methyl trifluoroacetamide
MTBE	methyl-tert-butylether
NCBI	National Center for Biotechnology Information
NCS	norcochlorine synthase
NMT	nitrogen-methyltransferases
OBL	oil body lipase
OLE	oleosin
PA	phosphatidic acid
PAD	phytoalexin deficient
PALD	protein associated with LDs
PAP	phosphatidic acid phosphatase
PCME	prenylcysteine methylesterase
PDAT	phospholipid:diacylglycerol acyl transferase
PE	phosphatidylethanolamine
PMSF	phenylmethylsulfonyl fluoride
PSAT	phospholipid:sterol acyltransferase
PTLD	pollen tube LD protein
PUX	plant ubiquitin X regulatory domain-containing protein
PXA	peroxisomal ABC-transporter
RD	response to desiccation
rpm	revolutions per minute
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
SDP	sugar dependent
SE	sterol esters
sgRNA	single-guide ribonucleic acid
SLDP	seed LD protein
SMT	sterol methyltransferase
TAG	triacylglycerols
TBS	tris-buffered saline
THAD	thalian-diol desaturase
THAH	cytochrome P450s thalianol hydroxylase



## Abbreviations

---

THAS	thalianol synthase
TNMT	tetrahydroprotoberberine- <i>N</i> -methyltransferase
UV	ultra violet
VAP	vesicle associated membrane protein (VAMP) – associated protein
wt	wild type
$\alpha$ -DOX	$\alpha$ -dioxygenase

## Abstract

The molecular study of plant lipid droplets, first described over a century ago, has garnered attention since the identification of oleosins in the 1980s'. In recent years, considerable progress has been achieved in identifying and characterizing plant lipid droplet proteome constituents. However, most of our understanding originates from investigations conducted on oil seeds, seedlings, mesocarp, pollen, and leaves, with a disproportionate focus on angiosperms. Lipid droplets are present across all domains of life, although the proteins and their composition identified thus far exhibit significant diversity, even within the same species' tissues. Therefore, a high degree of specialization and functional versatility can be expected. Indeed, lipid droplets have been associated with numerous biological processes, encompassing lipid metabolism and remodeling, energy storage, signaling, and the production of defense compounds. They may also enhance resilience to abiotic stress. Lately, lipid droplets and a subset of their proteins have been proposed to be a remnant of an ancient desiccation tolerance mechanism that persists in angiosperm seeds and pollen. The advantage obtained by early land plants through this mechanism may have been pivotal for the great success of plant terrestrialization.

In this thesis, we applied a bottom-up label-free proteomic approach on desiccation tolerant and oil-rich tubers of yellow nutsedge and compared it against the same species' leaves and roots, as well as desiccation sensitive and oil-depleted tubers of the closest related species, purple nutsedge. The acquired data unveiled a proteomic fingerprint consisting of antioxidants, heat shock proteins, and late embryogenesis abundant proteins resembling desiccation-tolerant seeds' composition. Furthermore, our findings strongly indicate that seed-type lipid droplet proteins oleosins, caleosins, steroleosins, and a seed lipid droplet protein play central roles in the seed-like desiccation tolerance program. The putative co-option of those features in tubers may be achieved through transcription factors ABSCISIC ACID INSENSITIVE3, WRINKLED1, and LEAFY COTYLEDON1 regulation. Applying a similar approach to germinating spores of *Physcomitrium patens* and comparing them to vegetative gametophytes, most of these desiccation tolerance-mediating traits reappear, further supporting the hypothesis that they existed prior to the emergence of seeds. It further sheds light on the LD proteome composition of a bryophyte spore.

In another proteomic study, we described lipid droplet protein composition for the first time in the subterranean tissue of *Arabidopsis thaliana* roots. As a result, we were able to identify

12 novel proteins localized to the lipid droplets, in addition to 31 previously described. Their assigned functions indicate diverse metabolic roles, involvement in developmental processes, and responses to stress, which highlights the nature of lipid droplets as active organelles in vegetative tissues.

Protein characterization is a time-consuming process, and for several identified LD proteins, characterization has not yet been conducted. To initially characterize the lipid droplet-associated methyltransferase (LIME) protein family, a range of techniques was employed, including *in silico* phylogeny, cell biological approaches, phenotype evaluation, and biochemical analyses. This approach provided initial insights into a highly conserved protein family with an unique and complex lipid droplet targeting mechanism. It is putatively the evolutionary origin of complex alkaloid-producing methyltransferases found in opium poppy and other species, but likely has a distinct function.

Overall, this work identifies new plant lipid droplet proteins and contributes to our understanding of the roles of these in desiccation tolerance. It further gives first insights into the LIME protein family. Therefore, it will serve as a foundation for further research.

# 1 Introduction

## 1.1 The lipid droplet, a unique organelle

One of the most critical structural elements of the eukaryotic cell is the membrane-enclosed subdivision into cellular compartments, the organelles. Those organelles enable the specialization and spatial segregation of processes within the cell, typically separated by a single or a double phospholipid bilayer. This highly organized system eventually led to even more complex multicellular life forms, which in turn are equipped with specialized cell types for specific tasks in their organs.

Among the organelles, lipid droplets (LDs) are distinct due to several unique features. Their basic structure is a spherically shaped phospholipid monolayer, which separates a hydrophobic core of neutral lipids from the aqueous environment. Commonly, this hydrophobic core consists mainly of neutral lipids such as triacylglycerols (TAGs) and, to varying extent, sterol esters (SEs) (Guzha et al., 2023). However, also other hydrophobic compounds can be stored in lipid droplets, for instance, wax esters in the desert shrub Jojoba (*Simmondsia chinensis*), carotenoids in the green microalgae *Dunaliella bardawil* or poly-terpenes in the LD homologous structures of the rubber tree (*Hevea brasiliensis*) and dandelion (*Taraxacum brevicorniculatum*) rubber particles (Davidi et al., 2014b; Sturtevant et al., 2020; Berthelot et al., 2014; Laibach et al., 2015).

The main component of the surrounding monolayer in plant oilseeds, mammalian cells, and yeast is predominantly phosphatidylcholine, with the latter two having an over-representatively high content of lysophospholipids (Tzen et al., 1993; Bartz et al., 2007; Grillitsch et al., 2011), possibly favoring positive membrane curvature (Thiam et al., 2013b). Several proteins with various enzymatic or structural functions are attached to the surface of the monolayer (Guzha et al., 2023). Most of those proteins were discovered recently thanks to the rapid progress in affordable, highly sensitive HPLC-MS/MS mass spectrometry-based proteomic techniques (Kretzschmar et al., 2020, 2018; Doner et al., 2021; Horn et al., 2013; Fernández-Santos et al., 2020). Therefore, numerous proteins have not been characterized yet, and their function remains unknown. Depending on the organism and cell type, the size of LDs varies between 0.1 – 100 µm in diameter (Thiam et al., 2013b). In plant oil seeds, LD diameters of approximately 0.5 – 2 µm have been described (Huang, 2018; Tzen et al., 1993).

From a biophysical standpoint, LDs are an example of a direct emulsion. Thus, they are oil droplets in an aqueous environment, stabilized by the phospholipid-monolayer as a surfactant between the oil and water phase (Thiam et al., 2013b). By its very nature, LDs could form spontaneously in the presence of neutral and membrane lipids within the cytosol. This is a potential reason for another unique feature of LDs, their presence as the only organelle in all domains of life, thus in Eukarya, Bacteria, and even Archaea (Murphy, 2012).

Due to this simple structure, LDs were originally considered inert storage compartments (Coleman, 2020). However, a rising number of dynamic cellular functions besides neutral lipid storage have been unveiled in recent years, such as roles in membrane remodeling, metabolism, stress response, as well as hormone and lipid signaling (Ischebeck et al., 2020). Nevertheless, a more detailed description of the intracellular role and functions of LDs remains pending.

### 1.2 Synthesis of lipid droplets

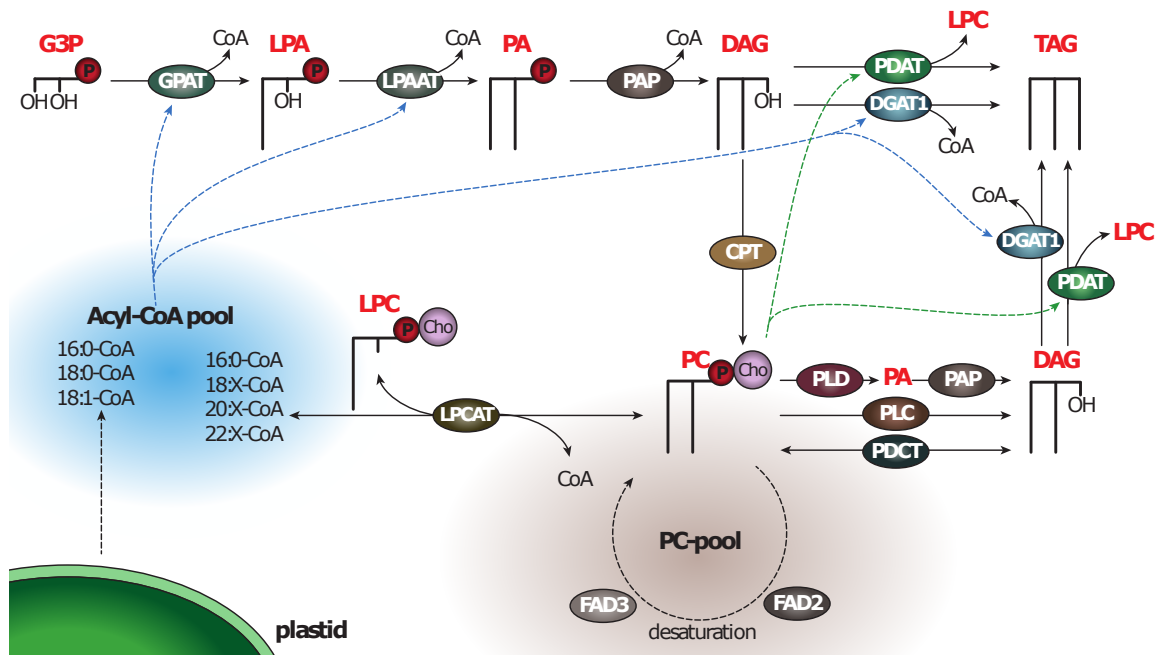
It is well established that cytosolic LDs derive in Eukarya from the endoplasmic reticulum (ER)(reviewed in Kang et al. (2022)) and are, to a certain degree, conserved between mammals, yeast, insects, and plants. Most of our knowledge of this process in plants derives from the model plant *Arabidopsis thaliana* and is outlined below.

#### 1.2.1 *De novo* synthesis of acyl chains

The formation of LDs starts with the synthesis of neutral lipids, eventually accumulating between the leaflets of the ER membrane bilayer (Chapman et al., 2019). The main components of these neutral lipids are fatty acids (FAs). They serve as a basic building block for TAGs, three FAs esterified to a glycerol backbone, wax esters, an FA esterified with a fatty alcohol, and sterol esters, a sterol backbone esterified with an FA. Unlike in mammals and fungi, where FA synthesis occurs in the cytosol, plants synthesize FAs *de novo* in the plastid stroma (Walker and Harwood, 1985). The fatty acid biosynthesis is reviewed by Hölzl and Dörmann (2019) and briefly summarized herein. The process is initiated by the acetyl-CoA carboxylase performed carboxylation of acetyl-CoA to malonyl-CoA, which is further used as a building block for fatty acids synthesis by a type II fatty acid synthase (FAS) multi-enzyme complex (Brown et al., 2009). The first step catalyzed by this complex is the transfer of malonyl linked to coenzyme A onto the acyl carrier protein (ACP) by malonyl-CoA:acyl carrier protein (ACP) transacylase.

This step is followed by a Claisen condensation of malonyl-ACP with acetyl-CoA by  $\beta$ -ketoacyl-ACP synthase III to form 3-ketobutyryl-ACP, the base C4 units to which additional C2 units are subsequently added. However, further enzymatic reactions must occur before extending the carbon chain. The process includes the reduction of the 3-keto group by 3-ketoacyl-ACP reductase, dehydration of the resulting hydroxy group by 3-hydroxyacyl-ACP dehydratase, and finally, the reduction of the trans- $\Delta^2$  double bonds by enoyl-ACP reductase. This is followed by the condensation of further C2 units derived from malonyl-CoA by  $\beta$ -ketoacyl-ACP synthase I followed by the described reduction and dehydration cycle, until a chain length of 16 carbons is reached. A potential further elongation to C18:0-stearoyl-ACP is carried out through  $\beta$ -ketoacyl-ACP synthase II. Subsequently, a subset of stearoyl-ACP can get desaturated at the  $\Delta^9$ -position to oleoyl-ACP by a stearoyl-ACP desaturase. For further desaturations in the plastid, acyl-ACPs need to be incorporated into plastidial lipids, a process referred to as the prokaryotic lipid pathway. Nevertheless, *de novo* synthesized FAs are either monounsaturated or saturated acyl-ACPs, which get also exported to the cytosol to fill the cell's accessible acyl-CoA pool. The export is initiated by acyl-ACP thioesterases (FAT) that hydrolyze the acyl-ACPs. One such isoform, FATA, predominantly acts on oleoyl-ACP, while FATB predominantly hydrolyses saturated acyl-ACPs (Salas and Ohlrogge, 2002). This is followed by the export of the resulting free FA out of the plastid into the cytosol, possibly mediated by FATTY ACYL EXPORT 1 (FAX1) and subsequently condensed with CoA by the long-chain acyl-CoA synthase (LACS) (Li et al., 2015). Those newly assembled acyl-CoAs can be further modified in the ER. For instance, they can be elongated or further desaturated and afterward redistributed into the acyl-CoA pool again (Bates, 2016). In the latter case, oleoyl-CoA gets incorporated into lysophosphatidylcholine at the *sn*-2 position by the lysophosphatidylcholine acyltransferase (LPCAT) to form phosphatidylcholine. Then, desaturation of  $\Delta^{12}$  to 18:2 linoleic acid and  $\Delta^{15}$  to 18:3 linolenic acid is performed by fatty acyl desaturases FAD2 and FAD3, respectively (Okuley et al., 1994; Arondel et al., 1992). The edited acyl group can get released either by phospholipase A2 or reverse enzymatic activity of LPCAT, eventually starting the cycle again. This cycle is referred to as the Lands cycle, one of the primary sources of the acyl-CoA pool for reusing (edited) acyl groups from membrane lipids (Bates, 2016). The diversity of the acyl-CoA pool is further enhanced in some species and tissues by unusual fatty such as cyclic, conjugated, and hydroxylated fatty acids (Shockey et al., 2018; Crombie and Holloway, 1984; da Silva Ramos et al., 1984).

## 1.2.2 Synthesis of neutral lipids



**Figure 1: Multiple pathways contribute to the synthesis of triacylglycerols.** Plastid-derived *de novo* synthesized acyl-CoAs are transported into the cytosol, providing a source for the cytosolic acyl-CoA pool. These acyl groups can undergo further modifications in the phosphatidylcholine (PC) pool by being incorporated at the *sn*-2 position of lysophosphatidylcholine (LPC) via the enzyme acyl-CoA:lysophosphatidylcholine acyltransferase (LPCAT). The PC pool enables additional modifications, such as the introduction of double bonds by fatty acid desaturases (FADs) or elongation. The modified acyl chains can be redistributed to the cytosolic acyl-CoA pool via LPCAT. Diacylglycerols (DAGs) get sequentially synthesized by esterifying an acyl-CoA with the *sn*-1 position of glyceraldehyde-3-phosphate (G3P) through acyl-CoA:G3P acyltransferase (GPAT). The resulting lysophosphatidic acid (LPA) is then esterified at the *sn*-2 position with another acyl-CoA via acyl-CoA:LPA acyltransferase (LPAAT), forming phosphatidic acid (PA). PA's phosphate headgroup is subsequently removed by PA phosphatase (PAP), generating 1,2-diacylglycerol (DAG). An alternative route of DAG synthesis is either by the action of phospholipase C (PLC) or through the coordinated activities of phospholipase D (PLD) and PAP, which cleaves the phosphate group of PC. For esterification at the *sn*-3 position of DAG, the acyl group can be obtained from either the cytosolic acyl-CoA pool catalyzed by diacylglycerol acyltransferases (DGAT) or from the *sn*-2 position of PC catalyzed by phospholipid:diacylglycerol acyltransferases (PDAT), yielding in triacylglycerol (TAG) respectively. Figure originates from Ischebeck et al. (2020).

The most prominent neutral lipids stored in LDs are TAGs. TAG biogenesis is best studied in *Arabidopsis thaliana* (Figure 1); however, it should also be comparable in most terrestrial plants. TAG can be synthesized starting with esterification of the *sn*-1 position of glyceraldehyde-3-phosphate (G3P) with an acyl-CoA catalyzed by acyl-CoA:G3P acyltransferase (GPAT) (reviewed in Bates (2016)). The resulting lysophosphatidic acid (LPA) is, in turn, esterified at the *sn*-2 position with another acyl-CoA, catalyzed by the acyl-CoA:LPA acyltransferase (LPAAT), generating phosphatidic acid (PA). In the next step, the phosphate headgroup of PA gets removed by PA phosphatase (PAP), resulting in 1,2-diacylglycerol (DAG). Alternatively, DAG can be generated using preexisting membrane phospholipids by cleaving the headgroup entirely by phospholipase C or cleaving the headgroup excluding the phosphate to generate PA by phospholipase D with subsequent cleavage of the phosphate group by PAP. The last step is the esterification of a third acyl chain to DAG on the *sn*-3 position resulting in

the formation of TAG. This is either performed by diacylglycerol acyl transferase (DGAT) or phospholipid:diacylglycerol acyl transferase (PDAT). PDAT transfers the acyl group attached to the *sn*-2 position of PC, whereas DGAT accesses the acyl-CoA pool. In Arabidopsis, three DGAT family proteins with distinct sequences and non-redundant functions are found. DGAT1 appears to be the most important protein associated with seed oil synthesis, as *dgat1-1* knockout lines have significantly lower seed oil content (Zhang et al., 2009). This content is further decreased when PDAT1 is silenced in seeds, demonstrating that PDAT can at least partially compensate for DGAT1 activity there. A double knockout of DGAT1 and PDAT1 turns the plant pollen lethal, further demonstrating that those other DGAT family proteins do not restore the roles of DGAT1 (Zhang et al., 2009; Bates et al., 2013). In leaves, DGAT1 appears to be the primary acyl transferase involved in TAG synthesis as well Slocombe et al. (2009). DGAT2, on the other hand, contributes to TAG synthesis in plants, but it has been reported to preferentially incorporate unusual fatty acids (Li et al., 2010). Finally, unlike DGAT1 and DGAT2, DGAT3 is a metal enzyme that is not ER membrane-bound (Hernández et al., 2012), but somewhat soluble in the cytosol (Aymé et al., 2018). It is widely expressed in Arabidopsis, incorporates linolenic and linoleic acid preferentially, and may execute a role in the regulation of acyl exchanges that take place between TAGs and the cytosolic acyl-CoA pool (Hernández et al., 2012; Aymé et al., 2018).

Sterol-esters are also synthesized at the ER (Bouvier-Nave et al., 2010). The sterol backbone derives from sequentially added isoprenoid building blocks originating from the cytosolic mevalonate pathway (Aharoni et al., 2005). A phospholipid:sterol acyltransferase is primarily responsible for the esterification of sterol backbone and acyl group, favoring to utilize acyl-groups from the *sn*-2 position of phospholipids and is most active on phosphatidylethanolamine (Banaś et al., 2005). Alternatively, synthesis is facilitated by an acyl-CoA sterol acyltransferase, which predominantly esterifies saturated acyl-CoAs with cycloartenol (Chen et al., 2007). Sterol-esters are considered as temporary storage compounds for excessive sterols (Bouvier-Nave et al., 2010).

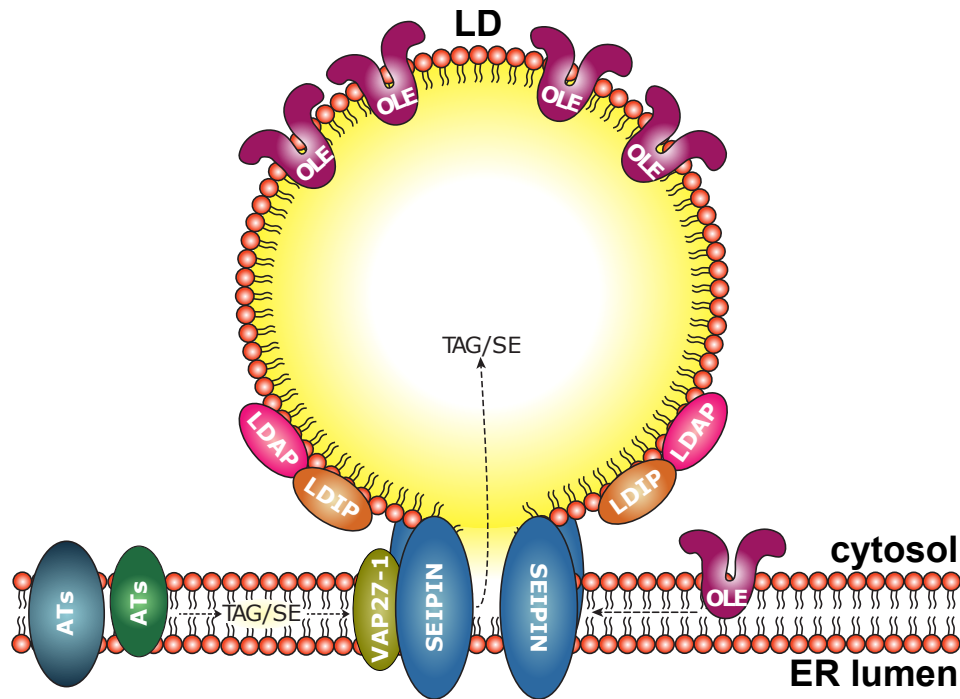
### **1.2.3 Formation of a lipid droplet at the ER membrane**

The deposition of synthesized neutral lipids between the ER bilayer leaflets initiates the onset of lipid droplet biogenesis. Driven by biophysical effects, single neutral lipid molecules start to coalesce and demix from other molecules, to nucleate to a lipid lens-like structure (Walther et al., 2017; Nettebrock and Bohnert, 2020; Renne et al., 2020; Thiam and Ikonen, 2021).



It is unknown at this time whether the neutral lipids stay close to their site of synthesis or diffuse freely before eventually condensing with one another, using heterogeneities in the membrane composition as nucleation sites (Thiam and Forêt, 2016; Ischebeck et al., 2020). However, ER subdomains for nascent LD assembly have been proposed in yeast (Joshi et al., 2018; Nettebrock and Bohnert, 2020; Pyc et al., 2017a). Proteins may also significantly affect the formation of LD by potentially lowering the critical nucleation concentration barrier (Thiam and Ikonen, 2021). A vital example of that are the seipin proteins, named after the Berardinelli-Seip congenital lipodystrophy medical disorder (Cartwright and Goodman, 2012). Seipins are functionally conserved proteins encoded as a single-copy gene in humans, insects, and yeast, or as a multi-copy gene in plants, forming barrel-shaped oligomeric complexes comprising 10 - 12 subunits, either determined via cryo-electron microscopy or in silico prediction (Chapman et al., 2019; Cai et al., 2015; Sui et al., 2018; Arlt et al., 2022; Yan et al., 2018). The fundamental structural unit of seipin comprises a conserved luminal spanning domain. This domain mediates protein-protein interactions with the luminal domains of other subunits, thereby facilitating the organization of the oligomer. The oligomer is embedded between transmembrane helices, which position the N- and C-termini towards the cytosolic site (Chapman et al., 2019). Those ring-like structures can trap small lipid lenses, potentially recognizing them by packing defects, hydrophobic patches of neutral lipids not shielded from the cytosol by membrane lipids. The seipin ring might then stabilize them during ripening, and might support supplementary oil accumulation by acting as a lipid vent (Zoni et al., 2021; Prasanna et al., 2021; Bohnert, 2018; Salo et al., 2019; Sui et al., 2018; Pyc et al., 2021). Hence, seipins mark the LD-ER junction sites and settle the area of an LD formation (Scholz et al., 2022).

Due to the increase of neutral lipids, the ER membrane starts to bulge and bud towards the cytosol (Walther et al., 2017). However, it is puzzling how the budding acquires directionality. In mammals and yeast, the ER-LD junction site located FAT STORAGE-INDUCED TRANSMEMBRANE 2 (FIT2) has been associated with regulating directionality due to changing membrane properties by removing DAG (Choudhary et al., 2015). Although plants lack a homolog of FIT2, mouse FIT2 has been expressed ectopically in *Nicotiana benthamiana* leaves or stably transformed in *Arabidopsis* lines, which results in an increase in LD size and quantity as well as an overall rise in the oil content of the leaf or seed tissue, respectively (Cai et al., 2017). This additionally indicates the conservation of targeting and functional complementarity across species of at least a subspecies of proteins. Reduced membrane surface tension of



**Figure 2: Simplified model of LD biogenesis.** Triacylglycerols and sterol esters synthesized by several acyltransferases (ATs) at the ER are channeled to the maturing LD through a ring-shaped oligomeric complex of seipin proteins tethered to the ER membrane by vesicle-associated membrane protein (VAMP)-associated protein 27-1 (VAP27-1). The LD-associated protein (LDAP) associates from the cytosol to the LD and stabilizes the LDAP-interacting protein (LDIP), which dissociates from the seipin complex when a nascent LD forms, on the surface of the maturing LD. LD class I proteins, such as oleosins (OLE), are synthesized as ER membrane proteins and relocated to the LD surface via the ER-LD tether complex by a yet unidentified mechanism. Those proteins may serve as surfactants, thus protecting the LD membrane surface and potentially contributing to its bulging. Figure was modified from Ischebeck et al. (2020); Scholz et al. (2022).

the cytosolic facing leaflet caused by membrane modification of the phospholipid composition and its accompanying intrinsic membrane curvature together with asymmetric protein insertion is biophysically essential for correct budding directionality (Choudhary et al., 2018; Chorlay et al., 2019). This process seems to be coordinated by proteins, as either lack of FIT2 or seipin in mammalian and yeast cells leads to a loss of budding directionality (Choudhary et al., 2015). The lack of FIT2 homologs in plants suggests that seipins may be solely responsible for cytosolic budding in plants.

In *Arabidopsis*, three partially redundant paralogs of seipin are known, possibly evolutionary adapted to serve plant-specific specialized tasks (Cai et al., 2015; Ischebeck et al., 2020). While SEIPIN1 is predominantly expressed in embryos, SEIPIN2 and SEIPIN3 are commonly found in all tissues (Taurino et al., 2018). Loss of budding polarity and abnormally enlarged LDs with associated pollen fertility and seed germination reduction is observed with the loss of function mutation in all three paralogs (Taurino et al., 2018). Additionally, depending on the seipin isoform, ectopic over-expression in *Nicotiana benthamiana* leaves results in larger or

smaller LDs with a higher overall TAG concentration (Cai et al., 2015). Consequently, seipins are essential for optimally regulated LD proliferation and size determination.

Another protein involved in the budding process of seed LDs might be plant-specific oleosin (Miquel et al., 2014). Oleosins are thought to migrate toward the LD and stabilize the nascent LD with their hydrophobic integral proline-knot motive after being co-translationally integrated into the ER membrane (Loer and Herman, 1993; Huang and Huang, 2017). Additionally, the budding direction was affected by oleosin expression in tobacco BY2 cells, demonstrating a cytosolic direction in the native targeting and a luminal direction in cases where oleosin had luminal targeting (Huang and Huang, 2017). Hence, oleosin directly influences the budding of nascent LDs. However, most plant tissues lack oleosins but harbor lipid droplets, indicating that oleosins are not essential for LD formation.

Recently, the model of the LD formation machinery has been extended (Pyc et al., 2021; Guzha et al., 2023). It was observed that the VESICLE ASSOCIATED MEMBRANE PROTEIN (VAMP) – ASSOCIATED PROTEIN 27-1 (VAP27-1) and LD ASSOCIATED PROTEIN (LDAP) – INTERACTING PROTEIN (LDIP) loss-of-function mutants in *Arabidopsis* had similar enormously enlarged LD phenotype as *seipin2 seipin3* mutants (Greer et al., 2020; Pyc et al., 2021; Taurino et al., 2018). LDAPs could also be involved in the underlying mechanism due to their interaction with LDIP (Pyc et al., 2017a). Based on all known protein-protein interactions between LDAPs, LDIP, VAP27-1, and all seipin paralogs, Guzha et al. (2023) refined the previously proposed model of Pyc et al. (2021) of the LD formation machinery. First, LDIP interacts with all seipin paralogs when initiating the oil-accumulation process within the seipin-oligomers (Pyc et al., 2021). In parallel, VAP27-1 binds to the N-terminal FFAT motifs of the SEIPIN2/SEIPIN3 oligomeric complex of the emerging LD, possibly stabilizing that structure at the ER-LD bordering region (Greer et al., 2020). The expansion of LD size due to the accumulation of neutral lipids facilitates the association of LDAPs with the surface of the LD monolayer. Then, the LDIP-seipin interaction dissociates, and LDIP gets stabilized at the LD surface through its interaction with LDAP instead (Figure 2; Coulon et al. (2020); Pyc et al. (2021)).

As the LD develops into its mature form, additional proteins either target LDs from the ER (class I LD proteins) or the cytosol (class II proteins) (Kory et al., 2016). Although the exact mechanism by which the LD and ER are eventually separated from one another is unknown, it is plausible that some LDs may stick to the ER while others diffuse into the cytosol. This has

been shown in yeast, where it has been noted that LDs may maintain their connection to the ER (Jacquier et al., 2011).

#### **1.2.4 Lipolysis/lipophagy marks the end of lipid droplets**

The neutral lipids stored in the LD are valuable components as carbon and energy source. At some point, the plant uses those resources through lipolysis, which also degrades the LD (Ischebeck et al., 2020). The TAG breakdown is carried out by the lipases SUGAR DEPENDENT 1 (SDP1) and SDP1-LIKE, as single and, more strikingly, double knockout Arabidopsis lines of these enzymes display diminishing TAG breakdown in seedlings and leaves, consequently causing oil accumulation there (Kelly et al., 2011; Eastmond, 2006; Kelly et al., 2013). In vitro studies demonstrated the ability of SDP1 to hydrolyze two fatty acids from the TAG's glycerol backbone, while LD-associated MONOACYL GLYCEROL LIPASE 8 (MAGL8) might cleave of the residual fatty acid (Eastmond, 2006; Kim et al., 2016b). Apparently, SDP1 initially localizes on peroxisomes and reaches the LD through peroxules, an extension of the peroxisomal membrane that creates a peroxisome-LD contact site, where SDP1 may then associate with the LD surface (Thazar-Poulot et al., 2015; Eastmond, 2006). Before released free FAs are transported into the peroxisome, they presumably need to be activated by LACS to become their acyl-CoA esters (Fulda et al., 2004). The added CoA group possibly gets cleaved (De Marcos Lousa et al., 2013) when the acyl-CoAs are transported by the PEROXISOMAL ABC-TRANSPORTER 1 (PXA1) into the peroxisomal lumen (Hayashi et al., 2002; Baker et al., 2015; Zolman et al., 2001). Subsequently, the free FAs get reactivated to their acyl-CoA esters again by LACS6 or LACS7 in the peroxisome matrix, where they are finally degraded by  $\beta$ -oxidation (Fulda et al., 2004).

Autophagy, also known as lipophagy in this context, may be an alternative option for the turnover of LDs (Zienkiewicz and Zienkiewicz, 2020). Lipophagy is established best in mammals (Petan et al., 2018) and yeast (van Zutphen et al., 2014; Schepers and Behl, 2021); however, its roles and mechanisms in plants are elusive. Two autophagy routes have been proposed in plants (Zienkiewicz and Zienkiewicz, 2020; Fan et al., 2019). On the one hand, macroautophagy is the process of encapsulating entire macro structures, such as LDs, within a double-membrane structure known as an autophagosome, which is mediated by autophagy-related (ATG) family proteins (Yoshimoto and Ohsumi, 2018). Subsequently, these autophagosomes fuse with the tonoplasts and enter the vacuole lumen as autophagic bodies, where they undergo degradation (Soto-Burgos et al., 2018). On the other hand, microautophagy is a process

in which cytoplasmic components are directly ingested via the invagination of the tonoplast into the vacuolar lumen for degradation (Fan et al., 2019). There is increasing evidence to suggest that autophagy plays a role in the degradation of LDs in plants (Fan et al., 2019; Kurusu et al., 2014) with research indicating that microautophagy may be involved in this process in *Arabidopsis* leaves (Fan et al., 2019). Despite the progress that has been made in understanding the role of autophagy in lipid degradation, further investigation is required to fully characterize the extent and mechanism by which autophagy contributes to the degradation of TAGs and LDs in plants.

### 1.3 The proteome of lipid droplets

The functions of LDs are carried out by various proteins associated with the LDs' surface. In addition to structural proteins that help maintain the integrity of LDs, several proteins with potential enzymatic roles in (sterol) metabolism, response to abiotic and biotic stress, and proteins with unknown functions have been described to localize at the LD monolayer. A selection of *bona fide* proteins identified in embryophytes can be found in Table 1. In seeds, about 1 – 4 % (w/w) of the LD weight can be assigned to proteins, approximately in the same range as the membrane phospholipids (Huang, 1992). Nevertheless, the mechanisms by which proteins are targeted to the LD monolayer are poorly understood (Song et al., 2022).

#### 1.3.1 Targeting of proteins to the LD surface

According to animal and yeast LD research, there are two primary pathways for targeting proteins on the LD surface. As previously introduced, class I proteins may initially target the ER membrane, commonly using a hydrophobic hairpin motive, and are then transferred to the LD membrane. Class II proteins, on the other hand, possibly target the LD membrane directly from the cytosol, typically via amphipathic helices that may detect packing defects in the membrane due to membrane curvature and the unusual hydrophobic matrix underlying the membrane phospholipids (Kory et al., 2016; Prévost et al., 2018; Olzmann and Carvalho, 2019; Song et al., 2022). In addition to amphipathic helices, class II proteins could also anchor to the LD surface via lipid modifications, like an added fatty acid moiety as a lipid-anchor or by protein-protein interaction with other LD proteins acting as a protein-anchor (Kory et al., 2016). However, this classification is a simplified concept to enable a starting point for discussing LD membrane targeting and does not fully cover the complexity of the

potential targeting mechanisms (Bersuker and Olzmann, 2017). Furthermore, there is no known targeting consensus sequence specific to LDs, preventing a convenient bioinformatical identification of LD proteins (Guzha et al., 2023; Olzmann and Carvalho, 2019).

For class II proteins, extensive molecular dynamics and *in vitro* studies led to a monolayer targeting model with amphipathic helices (Prévost et al., 2018; Chorlay and Thiam, 2020). Accordingly, the amphipathic helix of the protein that eventually targets the monolayer initially is present in an unfolded state in the cytosol (Prévost et al., 2018). Several factors contribute to the regulation of the targeting in this process, such as membrane electrostatics, sequence of the amphipathic helix, membrane curvature, oil composition of the LD core, membrane packing density, and membrane packing defects (Čopič et al., 2018; Chorlay and Thiam, 2020; Bigay and Antonny, 2012; Prévost et al., 2018; Thiam et al., 2013b). However, amphipathic helices can also target bilayer membranes (Bigay and Antonny, 2012); therefore, regulation strategies must be implemented to favor targeting the LD surface (Prévost et al., 2018). By nature of the lipid droplet, the packing density of the membrane molecules in the monolayer is lower than in comparable-sized bilayered membranes (Chorlay and Thiam, 2020). Reduced membrane packing density and the underlying hydrophobic phase may be causing packing defects that exceed the critical life span for protein-membrane interaction of 5 ns, determined by molecular dynamics (Vanni et al., 2013; Prévost et al., 2018).

The unfolded amphipathic helix of a targeting protein may initially interact weakly with the LD membrane, possibly stabilized by positively charged amino acid residues interacting with negatively charged headgroups of the phospholipids (Prévost et al., 2018). Subsequently, the amphipathic helix folds into the gap caused by the packing defect (Prévost et al., 2018), technically acting as a surfactant to reduce the interfacial energy between the inner hydrophobic oily phase and the outer aqueous phase (Thiam and Dugail, 2019). In this process, the amphipathic helix interacts with the neutral lipids of the hydrophobic core, resulting in the oil composition of the core possibly influencing targeting (Chorlay and Thiam, 2020). Furthermore, the sequence defining the amphipathic helix seems to be balanced between the parameters of its overall length, the size, and the hydrophobicity of its amino acid residues (Thiam and Dugail, 2019; Čopič et al., 2018). Hence, frequently occurring persistent packing defects (Prévost et al., 2018), in combination with interactions with neutral lipids may contribute to the specificity of amphipathic helices for LD targeting (Chorlay and Thiam, 2020; Thiam and Dugail, 2019; Prévost et al., 2018).

Conversely, the pathway of class I proteins targeting the LD coming from ER is more enigmatic. Two major LD targeting pathways are postulated (Song et al., 2022; Bersuker and Olzmann, 2017). The first pathway represents the early targeting of ER proteins to nascent LDs. Therefore, proteins may access the LD by diffusion through the LD assembly complex, with the seipin oligomer possibly serving as a gatekeeping diffusion barrier (Song et al., 2022). Furthermore, LD proteins might favor binding to the LD surface due to packing defects and the underlying hydrophobic phase, similar to class II proteins (Bersuker and Olzmann, 2017). In contrast, proteins with topologies unfavorable for binding at a monolayer, such as transmembrane domains, are excluded due to a hydrophobic mismatch (Bersuker and Olzmann, 2017).

However, a separate pathway is necessary for LD proteins unable to bypass the potential gatekeeping role of the LD assembly complex. In the so-called late-targeting pathway, LD protein targeting is possibly mediated by a membrane-bridge connection between the ER and the mature LD (Wilfling et al., 2013). Consequently, a recent study in *Drosophila melanogaster* S2R<sup>+</sup> cells proposed a model for establishing a membrane-bridge formation mediated by a complex membrane-fusion machinery at the ER exit site comprising SNAREs, tethers, and a Rab protein, enabling late-targeting proteins to diffuse toward the LD membrane (Song et al., 2022). Nevertheless, additional research is needed to determine if plant LD proteins utilize a similar targeting strategy.

### **1.3.2 Important plant LD proteins**

The prime example of class I proteins in plants are oleosins (OLE), which harbor a well-defined hydrophobic proline-knot motive necessary for LD targeting (Abell et al., 1997; Jolivet et al., 2017; Abell et al., 2002). According to the model, expression experiments in yeast revealed that oleosins seem to insert first co-translationally into the ER membrane via the signal recognition particle pathway, which also requires the Sec61 translocon before they eventually target the LD (Beaudoin et al., 2000). Oleosins were the first plant LD proteins identified in 1986 in the maize scutella (Qu et al., 1986) and have been predominantly found as the main structural LD proteins in angiosperm seeds (Huang, 1992; Wu et al., 1999), pollen (Jiang et al., 2007; Kim et al., 2002), and the tapetum (Ross and Murphy, 1996) since then. However, they are usually not expressed in vegetative tissues (Guzha et al., 2023; Klepikova et al., 2016). Oleosins have a putative enzymatic activity as a MAG acyltransferase and phospholipase (Parthibane et al., 2012); however, their primary task is a structural function as stabilizing

emulsifying surfactants that prevent coalescence and shielding from lipases (D'Andrea, 2016; Leprince et al., 1997). It is noteworthy that the specificity of oleosins for targeting the LD monolayer is striking, even in organisms (Beaudoin and Napier, 2002), tissues (Wahlroos et al., 2003), and cell compartments (Zhao et al., 2018) where oleosins are naturally absent. This feature renders oleosin an intriguing candidate for biotechnology applications by anchoring recombinant proteins to LDs (Huang, 2018; Bhatla et al., 2010).

**Table 1: Selection of *bona fide* embryophytic LD proteins.** The table presents proteins with high-confidence LD localization determined through microscopic imagery. It also indicates whether a homologous protein was found to be LD-associated in *Arabidopsis thaliana* (In *At*?), provides information on the tissue and condition (if available) where the protein was initially described as an LD protein, specifies the publication where the protein was first mentioned as LD-associated, and describes its (putative) function at the LD.

Name	In <i>At</i> ?	Firstly described in	Published in	Function in
Caleosin (CLO)	YES	Sesame seeds	Chen et al. (1999)	Surfactant, oxylipin metabolism
CYCLOARTHENOL SYNTHASE 1 (CAS1)	YES	Tobacco pollen tubes	Kretzschmar et al. (2018)	Sterol metabolism
$\alpha$ -DIOXYGENASE 1 ( $\alpha$ -DOX1)	YES	Arabidopsis leaves	Shimada et al. (2014)	Oxylipin metabolism
EARLY RESPONSE TO DEHYDRATION 7 (ERD7)	YES	Senescent Arabidopsis leaves	Doner et al. (2021)	Stress response
Glycerol-3-phosphate acyltransferase (GPAT)	YES	Bacterial infected Arabidopsis leaves	Fernández-Santos et al. (2020)	Cutin synthesis
9/13-hydroperoxide lyase (HPLF)	NO	Medicago truncatula	De Domenico et al. (2007)	Oxylipin metabolism
LD dehydrogenase (LDDH)	YES	Arabidopsis seeds and seedlings	Kretzschmar et al. (2020)	Metabolism
LD lipase (LIDL)	YES	Arabidopsis seeds and seedlings	Kretzschmar et al. (2020)	Metabolism
LD PROTEIN IN SEEDS (LDPS)	YES	Arabidopsis seeds and seedlings	Kretzschmar et al. (2020)	Unknown
LDAP-interacting protein (LDIP)	YES	Arabidopsis cDNA library	Pyc et al. (2017a)	LD formation
LD-associated hydrolase (LDAH)	YES	Arabidopsis seeds and seedlings	Kretzschmar et al. (2020)	Metabolism
LD-associated methyltransferase (LIME)	YES	Arabidopsis seeds and seedlings	Kretzschmar et al. (2020)	Metabolism
LD-associated protein (LDAP)	YES	Avocado mesocarp	Horn et al. (2013)	Surfactant, LD formation
LDS AND STOMATA 1 (LDS1)	YES	Arabidopsis leaves	Ge et al. (2022)	Stomata development, LD formation
Lipoxygenase (LOX)	NO	Cucumber and soybean cotyledons	Feußner and Kindl (1992)	Oxylipin metabolism
Monoacylglycerol lipase 8 (MAGL8)	YES	Arabidopsis germinating seeds and leaves	Kim et al. (2016b)	Lipid metabolism
Oil body lipase (OBL)	YES	Castor bean	Eastmond (2004)	Lipid metabolism
Oleosin (OLE)	YES	Maize scutella	Qu et al. (1986)	Surfactant, LD formation
PHOSPHOLIPID:STEROL ACYLTRANSFERASE 1 (PSAT1)	NO	Tomato	Lara et al. (2018)	Sterol metabolism
PHYTOALEXIN-DEFICIENT 3 (PAD3)	YES	Bacterial infected Arabidopsis leaves	Fernández-Santos et al. (2020)	Phytoalexin production



## 1. Introduction

Name	In <i>At</i> ?	Firstly described in	Published in	Function in
Plant Ubiquitin X regulatory domain-containing protein 10 (PUX10)	YES	Tobacco pollen tubes	Kretzschmar et al. (2018)	Protein degradation
POLLEN TUBE LD PROTEIN (PTLD)	NO	Tobacco pollen tubes	Kretzschmar et al. (2018)	Unknown
PROTEIN ASSOCIATED WITH LDS (PALD)	YES	Arabidopsis pollen	Li et al. (2022)	Pollen longevity
Seed lipid droplet protein (SLDP)	YES	Arabidopsis seeds and seedlings	Kretzschmar et al. (2020)	Unknown
STEROL METHYLTRANSFERASE 1 (SMT1)	YES	Tobacco pollen tubes	Kretzschmar et al. (2018)	Sterol metabolism
Steroleosin (HSD)	YES	Sesame seeds	Lin et al. (2002)	Brassinosteroids synthesis, sterol metabolism

Oleosins are feasibly the evolutionary heirs of the structurally related caleosins (CLO) (Jiang and Tzen, 2010), which bear many conserved protein phosphorylation sites and an N-terminal calcium-binding EF-hand motif (Chen et al., 1999; Næsted et al., 2000). Similarly, caleosin contains a proline knot motive that is suggested to stabilize the LD surface (Jiang and Tzen, 2010; Liu et al., 2009). Caleosins, unlike oleosins, are less abundant but more widely expressed, with the isoforms CLO3 and CLO4 even detected in vegetative tissues such as leaves (Shimada et al., 2018; Shen et al., 2014). They also appear to function in biotic and abiotic stress responses (Hanano et al., 2015; Aubert et al., 2010; Hanano et al., 2006).

Another classic example of a plant LD protein family are the predominantly seed-expressed steroleosins, membrane-anchored via a hydrophobic proline knob motif and possessing a C-terminal hydroxysteroid dehydrogenase domain (Lin et al., 2002). Moreover, steroleosins may play a vital role in the metabolism of brassinosteroids (Lin et al., 2002), a class of steroid hormones suggested to be involved in various stress responses, growth, and developmental processes in plants (Singh and Savaldi-Goldstein, 2015; Clouse and Sasse, 1998). Oleosin, caleosin, and steroleosin constitute the main angiosperm seed LD proteome (Chapman et al., 2012).

Especially the oleosins are established as the primary structural LD protein in angiosperm seed and pollen (Guzha et al., 2023; Ischebeck et al., 2020). Noteworthy, other species and tissues have independently evolved proteins with analogous structural functions to oleosins, but with possible distinct evolutionary origins. For instance, the oleaginous alga *Nannochloropsis oceanica* implements a unique LIPID DROPLET SURFACE PROTEIN (LDSP) (Vieler et al., 2012), although other green algae such as *Chlamydomonas reinhardtii* (Moellering and

Benning, 2010), *Chromochloris zofingiensis* (Wang et al., 2019), *Dunaliella salina* (Davidi et al., 2012), *Lobosphaera incisa* (Siegler et al., 2017), and *Scenedesmus quadricauda* (Javee et al., 2016) have a structurally unrelated MAJOR LIPID DROPLET PROTEIN (MLDP). On the other hand, MLDP is distantly related to human major structural LD protein perilipin (Pyc et al., 2021) and plant LDAPs (de Vries and Ischebeck, 2020), which serve as the primary structural LD protein in several vegetative tissues such as leaves (Gidda et al., 2016). Furthermore, LDAPs are more ubiquitously expressed than oleosins and can also be found in seeds and pollen (Chi et al., 2016). LDAPs were first identified in the oil-rich mesocarp of avocados, sharing striking homologies with small rubber particle proteins from rubber-producing plants, emphasizing the similarities of rubber particles and LD structures (Horn et al., 2013). They appear to have a role in drought tolerance (Kim et al., 2016a) and in correctly compartmentalizing and mobilizing TAG during early post-germinative growth (Gidda et al., 2016). As mentioned previously, LDAPs are also thought to be involved in LD biosynthesis in conjunction with LDIP (Pyc et al., 2021).

The generalized functionality of most LD structure proteins may be the protection of the LD integrity. Besides the primary structural proteins presented so far, the LD surface is decorated with a wide variety of proteins, including some with putative enzymatic activity (Table 1). Most of them have been identified in the last decade due to rapid advances in modern highly sensitive proteomic approaches.

### **1.3.3 Identification of new lipid droplet proteins**

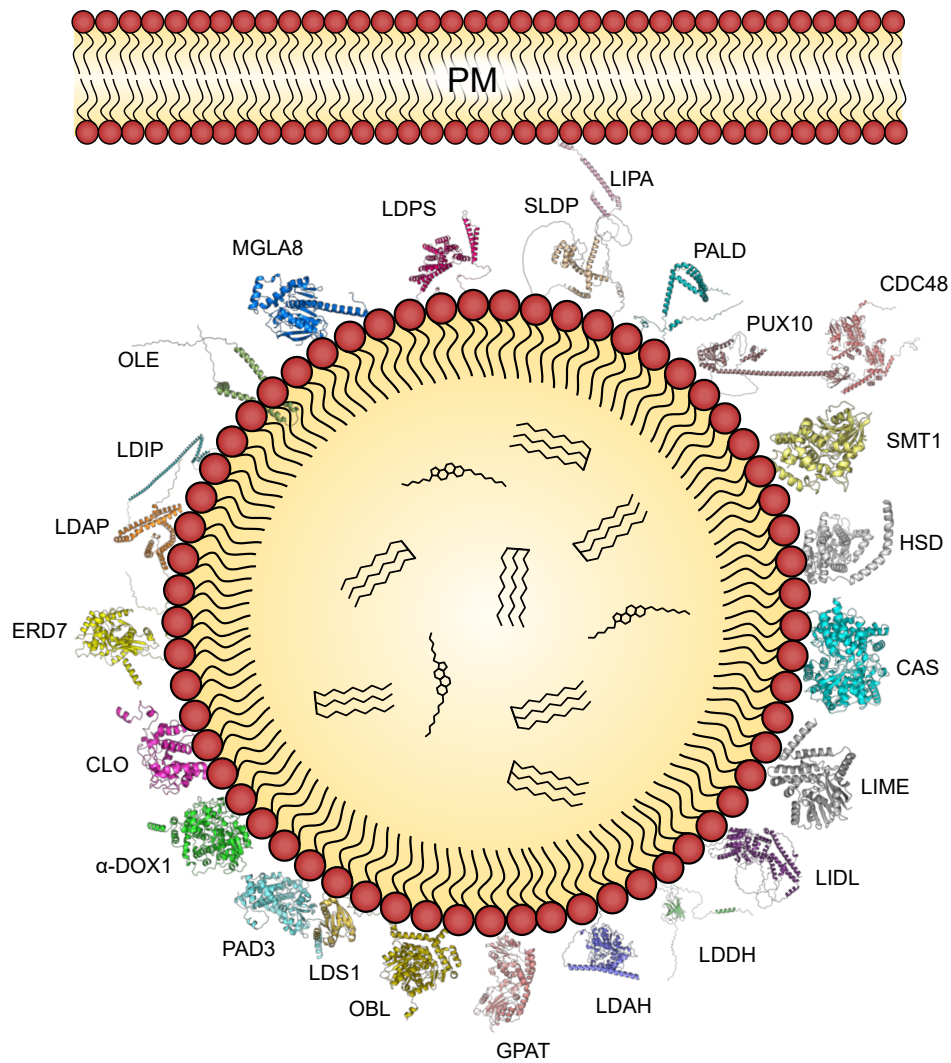
The identification of novel proteins associated with lipid droplets requires the surmounting of both technical and physiological obstacles. On the physiological side, the LD proteome is diverse across species, and some proteins are unique to specific organisms or phylogenetic subdivisions, as described in the previous section for MLDP and LDSP that are unique to algae (Vieler et al., 2012; Moellering and Benning, 2010). Moreover, even though a protein may be LD-associated with a particular organism, it should not be assumed that evolutionarily conserved orthologs will exhibit comparable targeting.

An example of this phenomenon is the *Arabidopsis* LIPOXYGENASE 1 (LOX1), an evolutionarily conserved ortholog of a highly abundant LOX in LDs of cucumber cotyledons (Feußner and Kindl, 1992). In contrast to the cucumber LOX, which displayed targeting to LDs in tobacco

seeds and leaves (Hause et al., 2000), Arabidopsis LOX1 did not co-localize with LDs in *Nicotiana benthamiana* or Arabidopsis leaves (Fernández-Santos et al., 2020).

Furthermore, LDs are dynamic organelles, and their proteome can drastically change during developmental progress (Kretzschmar et al., 2020; Jolivet et al., 2013). Consequently, the LD proteome composition is tissue and developmental stage-specific (Guzha et al., 2023; Shen et al., 2014; Kretzschmar et al., 2020). It is also conceivable that specific proteins only emerge under particular stress environments (Fernández-Santos et al., 2020). Thus, the LD proteome depends on species, tissue, physiological state, and potential environmental influences. Modern proteomic techniques have been used to collect the LD proteomes of several species and tissues from terrestrial plants (Table 2 and algae (Dadras et al., 2022; Vieler et al., 2012; Huang et al., 2013; Nguyen et al., 2011; Nojima et al., 2013; Siegler et al., 2017) in order to make progress of identifying novel LD-associated proteins. In this regard, Arabidopsis is the organism with the most variety of acquired proteomes and the most sophisticated representation of a land plant's LD proteome, as shown in Figure 3. Due to technical limitations, the known plant LD proteome might be less complex than the actual proteins that are stably or transiently associated with LDs. The following discourse highlights the technical limitations and challenges that are of paramount importance.

In order to isolate LDs from plant material, most studies utilized a protocol that involved homogenizing the sample and then separating the oil from the other crude cell components (Table 2). This is typically accomplished by subjecting the mixture to centrifugation, which forces the oil to float to the top of the aqueous buffer to form a so-called fat pad due to its lower density (Vermachova et al., 2014). Advanced centrifugation protocols can be utilized to separate LD subpopulations, which could exhibit distinct proteomic profiles, as demonstrated in cultured mammalian tissue cells (Zhang et al., 2016). Currently, most plant studies have been carried out on oil-rich tissues, with a particular emphasis on seeds, as they are known to contain high oil amounts (Vigeolas et al., 2007; Liu et al., 2015). On the other hand, tissues with a low oil content are underrepresented due to the challenging processing of a critical mass of material suitable for generating a high-quality floating fat pad. The majority of vegetative tissues fall under this category, with a few exceptions, such as the oil-rich mesocarp of avocado (Horn et al., 2013), olives (Bartolini et al., 2014), and oil palm (Tranbarger et al., 2011) or the underground tubers of yellow nutsedge (*Cyperus esculentus*) (Turesson et al., 2010). To date, the most extensively studied oil-deficient tissue are the senescent or drought-stressed leaves of Arabidopsis (Doner et al., 2021; Fernández-Santos et al., 2020; Brocard et al., 2017), owing



**Figure 3: Overview of the known LD proteome in *Arabidopsis thaliana*, the most advanced representation of a plant LD proteome.** All depicted proteins have been demonstrated to localize to LDs by microscopy, or to directly interact with a LD protein (CDC48, LIPA). The shown protein models are predictions acquired from the AlphaFold2 database (<https://alphafold.ebi.ac.uk/>) and visualized by PyMOL (open source version 2.5.0) software. Some predictions consist of large unfolded regions, indicating low reliability of those. The hydrophobic core of the LD is predominantly filled with TAGs and, to some extent, sterol esters, illustrated in a stylized form. The illustration has an exemplary visualization purpose only and does not accurately represent the color, scale, fold, interaction, and spatial orientation of the proteins in a living environment. Abbreviations: CLO, caleosin; CAS, CYCLOARTENOL SYNTHASE; CDC48, AAA ATPase Cell Division Cycle 48;  $\alpha$ -DOX1,  $\alpha$ -DIOXYGENASE 1; ERD7, EARLY RESPONSIVE TO DEHYDRATION 7; GPAT, glycerol-3-phosphate acyltransferase; HSD, steroleosin; LDAH, lipid droplet-associated hydrolase; LDAP, lipid droplet-associated protein; LDIP, LDAP-INTERACTING PROTEIN; LDDH, lipid droplet dehydrogenase; LDPS, LIPID DROPLET PROTEIN OF SEEDS; LDS1, LIPID DROPLETS AND STOMATA 1; LIDL, lipid droplet lipase; LIME, lipid droplet-associated methyltransferase; LIPA, LIPID DROPLET-PLASMA MEMBRANE ADAPTOR; MAGL8, MONOACYLGLYCEROL LIPASE 8; OBL, oil body lipase; OLE, oleosin; PAD3, PHYTOALEXIN-DEFICIENT 3; PALD, PROTEIN ASSOCIATED WITH LDS; PM, plasma membrane; PUX10, PLANT UBIQUITIN X REGULATORY DOMAIN-CONTAINING PROTEIN 10; SLDP, seed lipid droplet protein; SMT1, STEROL METHYLTRANSFERASE 1.

to its slightly elevated oil content in this stage (Troncoso-Ponce et al., 2013). Nevertheless, although LDs have been identified in various vegetative tissues (Pyc et al., 2017b), research on the LD proteome for many of them, such as roots, is lacking.

The purity of the isolated LD within the fat pad presents another difficulty. On the one hand, plants exhibit a second type of lipid droplets, plastoglobuli, located in their plastids with a distinct proteomic profile (Lundquist et al., 2012; Davidi et al., 2012). On the other hand, LDs

at least in mammalian and yeast cells have numerous membrane contact sites with various organelles such as mitochondria, Golgi vesicles, glyoxisomes, the cytoskeleton, the plasma membrane, and most prominently, the ER, their site of biosynthesis (Scholz et al., 2022; Thiam and Dugail, 2019; Bersuker and Olzmann, 2017; Jolivet et al., 2013). This association has the potential to result in contaminations when fragments of the foreign membrane, including its incorporated proteins, remain attached at the LD. Some protocols recommend extensive washing of the fat pad, utilizing detergents, in order to eliminate those contaminations (Jolivet et al., 2013). However, this approach may result in the loss of loosely attached proteins at the LD surface. Moreover, this methodology may inadvertently lead to the inclusion of extraneous proteins, making the issue even more challenging (Jolivet et al., 2013). More recent approaches employ minimal washing steps and instead opt to quantitatively compare total protein fractions against lipid droplet-enriched fractions to identify proteins enriched with lipid droplets (Horn et al., 2021; Kretzschmar et al., 2020; Dadras et al., 2022). This approach possesses the supplementary benefit of enabling the detection of low-abundant LD proteins (Horn et al., 2021). In Brassicaceae seeds, the proportion of oleosins within the total LD protein content is approximately 75 to 80 % (Jolivet et al., 2004, 2009, 2013) that can extend up to 85 % in *Arabidopsis* when caleosin and steroleosins are also included (Kretzschmar et al., 2020). Consequently, the remaining proteins present in seeds constitute a minimal fraction of the total proteome, making the use of this strategy advantageous for their identification. Nonetheless, the correct LD-targeting of each LD protein candidate must be individually confirmed by model systems localization studies (Müller et al., 2017; Kretzschmar et al., 2020; Doner et al., 2021). The application of ultra-sensitive high-performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS) is another factor that facilitates the identification of proteins present at low levels and is utilized by all presented studies Table 2.

**Table 2: Selection of plant LD proteomic studies.** An overview of proteomic studies acquired with HPLC-MS/MS ordered by their publication date. The table presents the studied organisms, the corresponding tissue(s) and extraction conditions for the proteins, as well as the publication source of the data.

Organism	Tissue(s)/condition	Publication
<i>Arabidopsis thaliana</i>	Seeds	Jolivet et al. (2004)
Castor ( <i>Ricinus communis</i> )	Endosperm	Eastmond (2004)
Rapeseed ( <i>Brassica napus</i> )	Seeds	Katavic et al. (2006)
<i>Arabidopsis thaliana</i>	Seeds	d'Andréa et al. (2007)
Rapeseed ( <i>Brassica napus</i> )	Seeds	Jolivet et al. (2009)
<i>Jatropha curcas</i>	Seeds	Popluechai et al. (2011)
Maize ( <i>Zea mays</i> )	Embryos	Tnani et al. (2011)
<i>Arabidopsis thaliana</i>	Seeds	Vermachova et al. (2011)

## 1. Introduction

Organism	Tissue(s)/condition	Publication
Rapeseed ( <i>Brassica napus</i> )	Developing seeds	Jolivet et al. (2011)
<i>Gevuina avellana</i>	Seeds	Acevedo et al. (2012)
<i>Madia sativa</i>	Seeds	Acevedo et al. (2012)
Sunflower ( <i>Helianthus annuus</i> )	Seeds	Furse et al. (2013)
Avocado ( <i>Persea americana</i> )	Mesocarp	Horn et al. (2013)
Hazelnut ( <i>Corylus avellana</i> )	Seeds	Zuidmeer-Jongejan et al. (2014)
<i>Jatropha curcas</i>	Seeds	Liu et al. (2015)
Sunflower ( <i>Helianthus annuus</i> )	Developing seeds	Thakur and Bhatla (2015)
<i>Arabidopsis thaliana</i>	Senescent leaves	Brocard et al. (2017)
Chinese tallow ( <i>Triadica sebifera</i> )	Mesocarp, seeds	Zhi et al. (2017)
Peanut ( <i>Arachis hypogaea</i> )	Seeds	Zaaboul et al. (2018)
Tobacco ( <i>Nicotiana tabacum</i> )	Pollen tubes	Kretzschmar et al. (2018)
<i>Arabidopsis thaliana</i>	Siliques, seeds, seedlings	Kretzschmar et al. (2020)
Sesame ( <i>Sesamum indicum</i> )	Seeds	Hamada et al. (2020)
Hazelnut ( <i>Corylus avellana</i> )	Seeds	Lamberti et al. (2020)
<i>Arabidopsis thaliana</i>	Bacterial-infected and senescent leaves	Fernández-Santos et al. (2020)
<i>Arabidopsis thaliana</i>	Senescent/ drought-stressed leaves	Doner et al. (2021)
Hybrid aspen ( <i>Populus tremula</i> × <i>Populus tremuloides</i> )	Shoot meristem	Veerabagu et al. (2021)

Prior to detection via HPLC-MS/MS, the proteins must undergo a series of preparatory steps. First, the fat pad is delipidated with an organic solvent, which also leads to the precipitation of the proteins (Horn et al., 2021; Vermachova et al., 2014). Henceforth, the proteins are redissolved and loaded onto an acrylamide gel and separated either through a two-dimensional polyacrylamide gel electrophoresis or a reducing sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (Jolivet et al., 2013). An alternative strategy is to load the proteins onto an SDS polyacrylamide gel and to briefly run the gel to minimize separation, thus immobilizing the protein in the gel matrix as a single band (Horn et al., 2021; Kretzschmar et al., 2020). Subsequently, bands are excised from the gel before tryptic digestion (Vermachova et al., 2014). Depending on whether a selection of individual bands or a band encompassing all proteins is chosen, the proteome is either partially or entirely covered (Jolivet et al., 2013). The resulting digested peptides are ultimately loaded onto the HPLC-MS/MS for analysis. The final obstacle is the need for a high-quality proteome reference database, usually in silico generated from transcriptomics via ORF prediction and translation tools (Entizne et al., 2020). Those reference databases are essential for software such as MaxQuant to infer the identity of proteins from measured peptide data (Tyanova et al., 2016). Therefore, an excellent transcriptomic acquisition is crucial for a complete proteomic dataset. Despite the rapid advances of such transcriptomic datasets, comprehensive data for all species remains limited.

There is still much potential for discovering novel plant LD proteins, as only a fraction of the possible organisms and physiological stages have been studied to date. The plant LD proteome may still be poorly understood compared to the yeast and mammalian ones, which each contain up to 40 or 150 possible LD proteins, respectively (Olzmann and Carvalho, 2019). Therefore, a significant amount of research is still required in this regard.

#### **1.3.4 Lipid droplet-associated methyltransferases**

Identifying novel LD proteins is only the first step in understanding the dynamic roles of the LD proteome, as determining the putative functions of identified proteins is even more challenging. While some functional insights have been gained on a number of LD-associated proteins (Guzha et al., 2023), most of them remain poorly understood. One example is the recently identified lipid droplet-associated methyltransferases (LIMEs) present as two paralogs (LIME1/LIME2) with 90 % sequence identity and 355 amino acid length in *Arabidopsis thaliana* (Kretzschmar et al., 2020). LIME1 is ubiquitously expressed at the transcriptional level in various plant tissues, with particularly high levels observed in germinating seeds and young leaf tissues and with the exception of developing and mature seeds. In contrast, LIME2 expression is primarily confined to the root apical meristem (Klepikova et al., 2016). Correspondingly, Kretzschmar et al. (2020) have shown that LIME proteins were not detectable in *Arabidopsis* non-germinated seeds and their levels increase over the detection limit after germination, with the highest levels observed in 48-hour-old seedlings. However, it should be noted that this study was not able to differentiate between the two paralogous isoforms of LIME1 and LIME2 (Kretzschmar et al., 2020). Moreover, LIME1 has been detected in the lipid droplet-enriched fraction of senescent leaves (Brocard et al., 2017), corresponding with transcriptomic data suggesting its presence in these tissues (Klepikova et al., 2016). Furthermore, orthologs of LIME1 have been identified in the lipid droplet-enriched fractions of the algae *Chlamydomonas reinhardtii* (Nguyen et al., 2011) and *Mesotaenium endlicherianum* (Dadras et al., 2022), suggesting that the targeting of this protein to lipid droplets is a conserved mechanism that predates the evolution of terrestrial plants.

According to the predicted functional annotation of LIME by the National Center for Biotechnology Information (NCBI), it belongs to the *S*-adenosyl-*L*-methionine-dependent (SAM) methyltransferases superfamily of proteins. Furthermore, the closest functionally characterized homolog in the UniProt database used for predicting the putative function of LIME is an (*S*)-COCLAURINE-*N*-METHYLTRANSFERASE (CNMT) from the opium poppy (*Papaver*

*somniferum*), with a 51 % sequence identity to Arabidopsis LIME. CNMTs belong to the sub-class of nitrogen-methyltransferases (NMTs) and catalyze the conversion of (*S*)-coclaurine to (*S*)-*N*-methylcoclaurine by utilizing SAM as a methyl donor in the benzyloisoquinoline alkaloid (BIA) biosynthesis pathway (Choi et al., 2001). This pathway leads to the production of complex alkaloids such as morphine, papaverine, noscapine, and sanguinarine (Beaudoin and Facchini, 2014). However, it is plausible that LIMEs have a different metabolic function in Arabidopsis.

One indication that LIMEs may have a different metabolic function in Arabidopsis is that despite the presence of several proteins involved in the BIAs pathway, no such alkaloids have yet been detected (Facchini et al., 2004). This suggests that the ancestor of Arabidopsis may have had the capability to produce these alkaloids but has lost this function, or, that the BIAs-producing species gained this as a new functionality due to evolution (Facchini et al., 2004). In addition, Arabidopsis lacks a functional homolog of the vital entry-point enzyme of the BIA pathway, the norcoclaurine synthase (NCS), which is essential for the pathway to proceed (Facchini et al., 2004; Vimolmangkang et al., 2016; Zhang et al., 2013). Most critical, recombinant Arabidopsis LIME2 did not show methylation activity on isoquinoline substrates, emphasizing the possibility of another function in Arabidopsis (Liscombe and Facchini, 2007). Moreover, it is worth noticing that the subcellular localization of the poppy CNMT and Arabidopsis LIME differs, as the CNMT is found in the cytosol (Hagel and Facchini, 2012), while LIME is associated with the membrane of lipid droplets (Kretzschmar et al., 2020).

The methylations of oxygen, nitrogen, carbon, and sulfur atoms of acceptor molecules via methyltransferases, utilizing *S*-adenosyl methionine (SAM) as the methyl donor, is a common reaction in all living organisms that may have been present since the last universal common ancestor (Abdelraheem et al., 2022). A total of five classes of protein folds of SAM-binding methyltransferases have been identified, whereas most belong to class I (Schubert et al., 2003). The prevalence of this class I methyltransferases within the eukaryotic genome has been estimated to range between 0.6 - 1.6 %, emphasizing the importance of this type of enzymes (Katz et al., 2003). The methylation occurs on various molecules, including proteins, lipids, and nucleic acids (Miller et al., 2003; Schubert et al., 2003). Therefore, methyltransferases are involved in crucial cellular processes such as epigenetic gene regulation (Siedlecki and Zielenkiewicz, 2006), signaling (Zubieta et al., 2003), and influencing primary and secondary metabolism (Liscombe et al., 2012). Moreover, methyltransferases are involved in the biosynthesis of various compounds, including antioxidants (Cheng et al., 2003) and defense compounds (Liu et al., 2006). This further underscores the challenge of determining



the function of LIME due to the multifaceted nature of methyltransferases, which are also known to act on multiple substrates (Choi et al., 2002). There remain many unresolved questions for LIME, including how the protein is targeted to the lipid droplet, whether this targeting is conserved across species, the potential substrates of the enzyme, differentiated function between LIME1 and LIME2, and its overall role in the broader metabolic context in relation to LDs.

### **1.4 Functions of lipid droplets**

The first description of lipid droplets dates back to the end of the 19th century; however, their recognition as a dynamic organelle did not occur until recent times (Farese and Walther, 2009). It has been approximately 100 years since their first description before they have been acknowledged as such, making them the most recently identified organelle (Coleman, 2020). Currently, LDs across eukaryotes are known by a plethora of synonymous names, including oleosomes, spherosomes, fat bodies, adiposomes, lipid bodies, and oil bodies (Ischebeck et al., 2020; Farese and Walther, 2009). However, in recent years, due to the similarities observed across a range of different species, the unified term "lipid droplets" has increasingly been adopted within the field of plant biology (Ischebeck et al., 2020). In plants, two subtypes of lipid droplets have been described with high confidence. On the one hand, there are cytosolic lipid droplets, the subtype that has been described so far. On the other hand, there are plastid-specific plastoglobules with a unique proteome, membrane composition, and functions (Lundquist et al., 2012; Bréhélin and Nacir, 2013; Hölzl and Dörmann, 2019). It is possible that a third type of lipid droplets, referred to as nuclear lipid droplets, exists at least in mammalian and yeast cells, but their characterization is limited (Layerenza et al., 2013; Uzbekov and Roingeard, 2013; Soltysik et al., 2019, 2021). Nonetheless, as this thesis focuses on cytosolic LDs, neither plastoglobules nor nuclear LDs will be addressed. The following section will provide a selective overview of various functions associated with lipid droplets, including but not limited to their role in energy storage, protein degradation, and abiotic stress response.

#### **1.4.1 Lipid droplets as an energy source**

LDs play a crucial role in a variety of biological processes and functions. The simplest is the storage of hydrophobic energy-dense compounds, predominantly but not exclusively TAG. Specifically, TAG accumulation occurs during the maturation process in plant seeds (Baud

and Lepiniec, 2010) and can be mobilized upon germination as a carbon and energy source for heterotrophic growth before the plants can energetically self-sustain by photosynthesis (Graham, 2008). The importance of TAG breakdown is demonstrated in the impaired or arrested post-germinative growth phenotypes of mutants of the major TAG lipases *sdp1* and *sdp1-like* or the peroxisomal acyl-transporter *pxa1* critical for  $\beta$ -oxidation (Kelly et al., 2011; Eastmond, 2006; Zolman et al., 2001). However, exogenous sugar can compensate for those phenotypes (Kelly et al., 2011; Eastmond, 2006; Zolman et al., 2001). It appears that these enzymes alone are not essential but beneficial for post-germinative growth. In addition, the breakdown of lipids has been proposed as the energy source for H<sup>+</sup>-ATPase-dependent acidification of the stomata based on the observation that *sdp1 pxa1* mutants exhibit a delay in stomata opening and a reduced response to blue light-induced acidification of the cell wall, a process considered to be the primary driving forces for stomata opening (McLachlan et al., 2016).

#### **1.4.2 The involvement of LDs in protein degradation**

The breakdown of lipids from the LD during post-germinative growth requires the removal of protective surface proteins by a machinery that resembles the endoplasmic reticulum-associated protein degradation (ERAD) pathway (Kretzschmar et al., 2018; Deruyffelaere et al., 2018). Consequently, oleosins, steroleosins, and caleosins may be ubiquitinated upon germination, indicating that they are a possible target for proteasome-dependent degradation (Deruyffelaere et al., 2015, 2018; Kretzschmar et al., 2018; Hsiao and Tzen, 2011). The PLANT UBX-DOMAIN CONTAINING PROTEIN 10 (PUX10) then associates with the LD surface of seedlings, interacting with the ubiquitinated proteins and bringing them close to the AAA ATPase CELL DIVISION CYCLE 48A (CDC48A), which may have been recruited by the ubiquitin regulator X (UBX) domain of PUX10 (Kretzschmar et al., 2018; Park et al., 2007). Subsequently, CDC48A may utilize an ATP-dependent unfolding process to remove ubiquitin-labeled proteins from the membrane, rendering them accessible for degradation by the proteasome (Baek et al., 2013). In comparison to the wild-type, *pux10* mutants displayed delayed degradation of seed-specific proteins and reduced lipid droplet (LD) size (Kretzschmar et al., 2018; Deruyffelaere et al., 2018). This may be attributed to the continued presence of oleosins, which are known to shield the LD and prevent coalescence (Deruyffelaere et al., 2018). Although LD protein degradation is impaired, it is not entirely depleted, indicating that PUX10 effectively participates in protein degradation but is not essential, inferring the presence

of alternative routes within the ubiquitin-dependent proteasomal pathway. Nonetheless, LDs may have the function of protein turnover via an LD-associated protein degradation pathway (LDAD) (Deruyffelaere et al., 2018).

#### **1.4.3 The role of LD in oxylipin synthesis and pathogen defense**

The abilities of LDs are not limited to carbon storage and mobilization. There is growing evidence that LDs are the synthesis site and resource hub for defense and abiotic stress-signaling compounds. A prominent example of that is the production of the antifungal oxylipin 2-hydroxy-octadecatrienoic acid (2-HOT) in *Colletotrichum higginsianum* infected leaves of *Arabidopsis* by a cooperative two-step process (Shimada et al., 2014; Shimada and Hara-Nishimura, 2015). Specifically, the enzyme  $\alpha$ -DOX1 catalyzes the oxygenation of  $\alpha$ -linolenic acid, which is hypothesized to originate from the lipid droplet inner core, into the unstable intermediate compound 2-hydroperoxy-octadecatrienoic acid (2-HPOT) (Shimada et al., 2014). Subsequently, CALEOSIN3 might catalyze the reduction from 2-HPOT to the more stable 2-HOT with phytoalexin function against *Colletotrichum higginsianum* by utilizing its peroxygenase activity, which may also result in the production of epoxy-FAs by oxidizing double bonds in unsaturated FAs (Shimada et al., 2014; Hanano et al., 2006; Blée et al., 2014).

It was also hypothesized that in a process similar to the aforementioned biosynthesis of 2-HOT, the antifungal oxylipin 13-hydroxyoctadecatrienoic acid (13-HOT) with phytoalexin function against *Botrytis cinerea*, *Cladosporium herbarum*, and *Phytophthora parasitica* might occur through a LOX dependent manner (Shimada and Hara-Nishimura, 2015; Prost et al., 2005). Accordingly, a 13-LOX would oxygenate  $\alpha$ -linolenic acid to 13-HPOT, which is then converted by CALEOSIN3 peroxygenase activity to 13-HOT (Shimada and Hara-Nishimura, 2015). This hypothesis is supported by the observation that overexpression lines of CALEOSIN3 in *Arabidopsis* exhibit an enhanced production of 13-HOT, whereas knockout lines display a decreased compound production (Blée et al., 2014). In previous studies, LOX localization to lipid droplets has been observed sporadically in different plant tissues, including cucumber cotyledons (Feußner and Kindl, 1992; Rudolph et al., 2011), sunflower seeds/cotyledons (Rodríguez-Rosales et al., 1998; Yadav and Bhatla, 2011), and olive pollen (Zienkiewicz et al., 2013). However, ectopic expressed recombinant *Arabidopsis* LOX1 did not show LD association in *N. benthamiana* leaves, indicating species specificity of LOX association to LDs (Fernández-Santos et al., 2020). Nonetheless, it cannot be excluded that other *Arabidopsis* LOX enzymes might associate with the LD surface, nor that cytosolic LOX can be supplied with

precursors from the LD without direct membrane-association. Additional enzymes involved in the metabolism of oxylipin-derived volatiles have been identified at the LD surface, including the hydroperoxide lyase from *Medicago truncatula* (De Domenico et al., 2007; Wasternack and Feussner, 2018) and the Arabidopsis OIL BODY LIPASE 1 (OBL1) homolog in tomato (Garbowicz et al., 2018; Ischebeck et al., 2020). These observations suggest a more significant role for lipid droplets in the metabolism of oxylipins. Since oxylipins are also engaged in stress signaling, signaling, and direct defense action are intertwined (Wasternack and Feussner, 2018).

Apart from oxylipin-producing enzymes, the PHYTOALEXIN DEFICIENT 3 (PAD3) enzyme involved in the production of the phytoalexin camalexin (Zhou et al., 1999) was identified as being partially located at LDs of *Pseudomonas syringae* pv. *tomato* infected Arabidopsis leaves (Fernández-Santos et al., 2020). The relevance of camalexin as an antifungal defense compound is demonstrated by the camalexin-deficient mutant *pad3*, which has an increased susceptibility to the necrotrophic ascomycetes *Alternaria brassicicola* (Nafisi et al., 2007), *Leptosphaeria maculans* (Bohman et al., 2004), and *Botrytis cinerea* (Glazebrook, 2005; Fernández-Santos et al., 2020). However, the resistance of *pad3* to the bacterial strain *Pseudomonas syringae* is comparable to the wild-type (Fernández-Santos et al., 2020). Additionally, in the liverwort *Marchantia polymorpha*, it has been observed that hydrophobic terpenoids are sequestered into lipid droplets, potentially serving as a defense compound against herbivorous organisms (Romani et al., 2020). LDs are thus woven into a complex network of abiotic resistance mechanisms of plants.

### 1.5 The role of LDs in desiccation tolerance and other abiotic stresses

One of humanity's most significant challenges of the 21st century are extreme weather events caused by climate change with substantial socio-economical and ecological impact (Lyon et al., 2022; Costa and Farrant, 2019). In many places of the world, rising global temperatures will cause an increase in extreme weather events resulting in increased aridification on the one hand and a higher risk of flooding on the other (Bailey-Serres et al., 2019; Hopkin, 2005). Extreme weather events have already significantly impacted crop yields (Lesk et al., 2016). Hence, there is a demand for crops that withstand drought and desiccation conditions to guarantee food supply in the future (Farrant and Hilhorst, 2022). There is a plausible role of LDs involved in the regulation and maintenance of abiotic stress mechanisms, including heat, cold, drought, and desiccation tolerance (de Vries and Ischebeck, 2020). Thus, a

comprehensive understanding of the involvement of LDs in these mechanisms may prove beneficial in achieving the goal of developing drought and desiccation-tolerant crops, making it important to understand drought resistance and desiccation tolerance and its evolutionary origins.

Upon conquering land about 590 million years ago, plants were faced with a plethora of abiotic stressors, with drought being a prevalent factor (Fürst-Jansen et al., 2020). As sessile organisms, they evolved diverse strategies to strengthen their resilience to evade water scarcity, subdivided into distinct paths of drought resistance and desiccation tolerance. Three major drought resistance strategies have evolved in response to low water content in the environment (Blum, 2011). The drought escape strategy is characterized by plants that thrive under favorable conditions and survive periods of drought through desiccation-tolerant reproductive structures (Kooyers, 2015). Vegetative tissues directly faced with low environmental water either employ dehydration avoidance to limit water loss or dehydration tolerance to sustain metabolic activity at up to 60 – 70 % relative water content loss (Blum, 2011; Zhang and Bartels, 2018). In contrast, desiccation tolerance refers to the ability to endure extreme long-duration intracellular water loss of over 90 % while maintaining a state of metabolic inactivity (Alpert, 2005). Interestingly, for some niche plants, complete desiccation seems to be favored over drought tolerance to avoid metabolic stress (Tuba et al., 1998; Lyall and Gechev, 2020). However, this resilience comes at the cost of reduced metabolic efficiency and limited growth during habitable growth conditions and is, therefore, mainly an advantage for organisms in harsh environments to fill their ecological niche (Oliver et al., 2000).

It has been proposed that LDs are a component of the apparatus maintaining drought resistance and desiccation tolerance, controlled by an ancient genetic framework (de Vries and Ischebeck, 2020). Correspondingly, most reproductive tissues, such as seeds, pollen, and spores, are desiccation tolerant and store high amounts of oil in LDs (Gaff and Oliver, 2013), even though the relative oil amount can be significantly different (Ischebeck, 2016; Huang et al., 2009; Mello et al., 2010). LDs possibly fill up the cell volume to prevent it from collapsing, contributing to the resilience of those tissues by reducing mechanical shrinking stress besides being a carbon source for post-germinative growth (Lyall and Gechev, 2020). Thus, LDs line up with cell protective components such as late embryogenesis abundance (LEA) proteins, heat-shock proteins (HSPs), antioxidants, and sugars, commonly found as the footprint of desiccation-tolerant tissues (Oliver et al., 2020; Berjak, 2006; Hilhorst et al., 2018). LEA proteins and sugars may additionally stabilize intracellular macromolecules by vitrification of

the cytosol, providing a ridged hydrophilic shell that replaces the water molecules, while HSPs may stabilize protein folds as chaperons and antioxidants protect against reactive oxygen species (ROS)(Hoekstra et al., 2001; Kranner and Birtić, 2005; Al-Whaibi, 2011). In the insect *Drosophila melanogaster*, it has been observed that LDs provide protection against reactive oxygen species (ROS), suggesting that LDs might also act as intrinsic oxidation protection in desiccated plant cells (Bailey et al., 2015). The maintenance of the plant LD membrane integrity may be facilitated by the presence of oleosins, which prevent the coalescence of LDs, possibly via electrostatic repulsion when the intracellular space is reduced due to water loss (Siloto et al., 2006; Huang, 1992). Double knockout lines of oleosins in *Arabidopsis* seeds exhibited aberrantly increased LDs, reduced germination rate, and susceptibility to freezing, emphasizing the vital role of oleosins in the desiccation tolerance of seeds (Shimada et al., 2008). The occurrence of oleosins paired with a high abundance of LDs has also been observed in desiccation-tolerant vegetative tissues (Hilhorst et al., 2018). For instance, transcript levels of oleosins and abundance of lipid droplets increased upon the desiccation of leaves of the resurrection grass *Oropetium thomaeum* (VanBuren et al., 2017). Similarly, oleosin transcripts along with LD-accumulation have been observed in the desiccation-tolerant gametophytes of the moss *Physcomitrium patens* (Huang et al., 2009), the resurrection clubmoss *Selaginella lepidophylla* leaves (Brighigna et al., 2002; VanBuren et al., 2018b), and in the desiccation-tolerant oil-rich tubers of the monocot *Cyperus esculentus* (Yang et al., 2016). In addition, oleosin transcripts have been observed as a commonly noted characteristic for vegetative desiccation tolerant plants when subjected to dehydration (Costa et al., 2017a; VanBuren et al., 2018a; Xu et al., 2018). *Oleosin* expression is regulated by the stress hormone abscisic acid (ABA) induced transcription factors LEAFY COTYLEDON 2 (LEC2) (Che et al., 2009) and ABSCISIC ACID INSENSITIVE 3 (ABI3) (Mönke et al., 2012), both key regulators of a network establishing seed maturation (Santos-Mendoza et al., 2008). However, while *lec2* mutants remain desiccation tolerant (Meinke et al., 1994), *abi3* mutants in *Arabidopsis* (Nambara et al., 1995), and *Medicago truncatula* (Terrasson et al., 2013) seeds, as well as in *Physcomitrium patens* protonema (Khandelwal et al., 2010), turned desiccation sensitive. In contrast, the upregulation of *ABI3* was able to reactivate desiccation tolerance in germinating *Arabidopsis* and *Medicago* seeds (Dekkers et al., 2015). Moreover, *ABI3* is suggested to be a central regulator of mediating vegetative desiccation tolerance in *Xerophyta viscosa* (Costa et al., 2017a; VanBuren et al., 2017). Thus, it can be inferred that *ABI3* plays a crucial role in establishing desiccation tolerance, thereby modulating oleosin occurrence. Nevertheless, *LEC2* may be involved in the transcriptional regulation of other seed-specific

but likely desiccation-irrelevant LD proteins, such as HSD1 (Baud et al., 2009). Despite the likely interplay between LDs, oleosins, and desiccation tolerance, the underlying mechanisms remain to be elucidated.

Nevertheless, the accumulation of TAG and LDs is a phenomenon also occurring with other abiotic stressors, including drought, heat, extended darkness (Kunz et al., 2009), and cold stress, indicating a broader spectrum of LD-mediated stress resilience. An overarching theme may be the maintenance of membrane integrity and fluidity by acting as a sink and source for acyl groups. The membrane's fluidity is determined by the ratio of (poly-)unsaturated to saturated acyl groups in the membrane that needs to be adjusted upon temperature changes to preserve membrane integrity (Murata and Los, 1997). When heat stress is applied to *Arabidopsis* seedlings and leaves, a putative membrane remodeling mechanism stabilizes the membrane by rapidly channeling polyunsaturated acyl groups into TAG, likely causing an increase in LD abundance (Mueller et al., 2015; Higashi et al., 2015). This process is fully reversible when stress is removed; thus, plants have a cache of membrane acyl groups for dynamic temperature membrane adjustments that may be required during daily temperature shifts (Mueller et al., 2015). Moreover, the reduction of sterol esters under heat stress suggests that plants may regulate membrane rigidity by releasing free sterols derived from LD-stored sterol esters (Krawczyk et al., 2022a; Dufourc, 2008).

Additionally, LDs may act as a direct sink for free fatty acids, which are particularly detrimental to the membrane integrity and which are released when membrane phospholipids break down (Yang and Benning, 2018; Listenberger et al., 2018; Fan et al., 2013). Prolonged periods of darkness inhibit photosynthetic activity in plants, thereby necessitating the utilization of alternative energy and carbon sources as a result of the depletion of stored starch (Kunz et al., 2009). This may be accomplished through the process of  $\beta$ -oxidation of fatty acids (Flügge et al., 2011). LDs may serve as an acyl-chain cache of decomposed membrane components prior to their degradation via  $\beta$ -oxidation in the peroxisome. The peroxisomal fatty acid transporter *pxa1* mutants exhibit an accumulation of free FAs (Kunz et al., 2009), a phenotype that is abrogated in the absence of the main TAG lipase, SDP1 (Fan et al., 2017). Thus, the hypothesized pathway is that free fatty acids derived from the breakdown of membrane lipids are directed towards endoplasmic reticulum-associated triacylglycerol synthesis, where they are initially incorporated into LDs before being hydrolyzed by SDP1 and channeled through PXA1 to the peroxisome for degradation (Fan et al., 2017). Another possible function of LDs in algae and diatoms is to act as a metabolic sink for electrons

and carbon in the presence of nitrogen deficiency, which otherwise impairs photosynthesis (Le Moigne et al., 2022; Leyland et al., 2020; Mock and Kroon, 2002).

Beyond the sink function of potentially harmful components, the abiotic stress response may also be mediated by some members of the LD proteome. For instance, LDAP's presence influences plants' susceptibility to drought (Gidda et al., 2016). On the one hand, expression levels of *LDAPs* increase while *Arabidopsis* is experiencing drought stress. Furthermore, drought resistance was enhanced by ectopic overexpression of *Arabidopsis LDAP1-3* coupled with an increased number of LDs (Kim et al., 2016a; Gidda et al., 2016). Equivalent effects were observed for dandelion (*T. brevicorniculatum*) *LDAP* ectopically overexpressed in *Arabidopsis*, indicating a conservation of the mechanism (Laibach et al., 2018). Consequently, a double knockout *Arabidopsis* mutant line *ldap1/ldap3* showed the opposite effect of a higher susceptibility to drought (Gidda et al., 2016). Interestingly, the transcription levels of an *LDAP* homolog of the resurrection grass *Oropetium thomaeum* increased in response to desiccation alongside oleosin, implying a similar role in maintaining LD membrane integrity due to its surfactant properties (VanBuren et al., 2017). These results suggest that LDAP, in addition to its function in compartmentalization and biosynthesis of LDs, is also an important contributor to plant drought tolerance.

However, the *Arabidopsis* LD protein CALEOSIN3, alternatively called RESPONSE TO DESICCATION 20 (RD20), contributes to plants' drought evasion strategy by influencing stomata opening and, thus, the transpiration rate (Aubert et al., 2010). In mutant *Arabidopsis* lines lacking RD20, the transpiration rates significantly but mildly increased due to reduced stomata aperture-closing (Aubert et al., 2010). In contrast, overexpression of *RD20* did not result in improved drought resistance; thus, its influence is moderate (Aubert et al., 2010). As transcript levels of *RD20* strongly depend on ABA concentrations, Aubert et al. (2010) hypothesized a function as a stress-related signaling hub. Moreover, ectopic overexpression of the rice (*Oryza sativa*) *CALEOSIN5* gene suggests a potential role in oxylipin signaling as it disrupts jasmonate-mediated cold tolerance through a potential negative regulation of jasmonates (Zeng et al., 2022; Sharma and Laxmi, 2016). Caleosins appear to play a role in response to various abiotic stressors, including drought, heat, salt, wounding, and cold, potentially by modulating hormone-mediated stress signaling pathways as they are differentially expressed in plants subjected to these stressors (Zeng et al., 2022; Aubert et al., 2010). In conclusion, LDs are a vital element in the abiotic stress response with significant importance in desiccation tolerance. Those traits possibly root back before terrestrialization, as homologs



of several remote homologs of stress-related LD proteins can already be found in green algae (de Vries and Ischebeck, 2020; Lyall and Gechev, 2020; Oliver et al., 2000; Huang et al., 2013). Nevertheless, the involved regulations, mechanistic, and key elements remain poorly understood and require further investigation.

## 1.6 Aims of this thesis

The field of plant lipid droplet biology is relatively new. In the last decades, lipid droplets have been recognized as being more than storage compartments; they are now understood to be dynamic organelles. However, the knowledge of processes is still limited, including, among other things, fundamental questions regarding formation, origin, turnover, diverged functions in subpopulations, degradation of lipid droplet proteins, protein targeting, and targeting sequences. Generally, this thesis aims to contribute to our understanding of plant lipid droplets, their role in plant physiology, and biochemistry from three distinct standpoints; (i) understanding the connection between desiccation tolerance and LDs; (ii) identification of novel LD-associated proteins from roots; (iii) characterization of a recently identified LD-associated methyltransferase.

- i. Lipid droplets are linked to desiccation tolerance. However, there is a significant knowledge gap in their function in vegetative desiccation-tolerant plant tissues. This project aimed to gain insights via comprehensive developmental proteomic datasets mining for vegetative and reproductive partially desiccation tolerant tissues of yellow nutsedge and *Physcomitrium patens* to identify the key proteins involved in their resilience.
- ii. Lipid droplet proteomes have been acquired chiefly for oil-rich tissues, such as seeds, seedlings, pollen, pollen tubes, and oil-rich mesocarp. In contrast, few studies have been conducted on vegetative tissues with low oil content. Nevertheless, LDs have been described to be present also in roots, even though the absolute amount of neutral lipids is below 1 % of the root dry weight. Nevertheless, the lipid droplet proteome of roots is understudied so far. The tissue-specificity of LD proteomes suggests that a deeper understanding of LDs in roots could reveal new candidates among the growing list of plant LD proteins. The known plant LD proteome has almost tripled in size in the past five years, but is still outnumbered by the mammalian LD proteome comprising 100-150 proteins, leading to the assumption that further plant LD proteins are not unveiled yet.

- iii. An increasing number of proteins associating with the LDs monolayer have been discovered recently with the rise of HPLC-MS/MS analysis techniques and computational power. However, the discrete function of most of these newly discovered proteins concerning lipid droplets remains mysterious due to a lack of their characterization. As yet uncharacterized, the lipid droplet-associated methyltransferase has been chosen for further analysis comprising its targeting sequence, developing phenotypes, and potential substrates for methylation.

## **2 Article I: A seed-like proteome in oil-rich tubers**




The article was published online in The Plant Journal in September 2022. The online version, including supplementary figures and files, can be found under the following DOI.

<https://doi.org/10.1111/tpj.15964>

Author contribution:

Philipp William Niemeyer planned and performed all experiments except LC-MS/MS proteomic measurements and LC-MS phytohormone measurements. He did most of the initial proteomic data processing and analysis, wrote the manuscript's first draft, and drafted the raw figures, except for phylogenetic trees and clusters. For the revision process, he re-edited figures 1, 3, 4, S1, S2, S3, S4, newly created figure 5, re-analyzed proteomic data, including additional yam, potato, and purple nutsedge data, and contributed to the revision of the manuscript text.

# A seed-like proteome in oil-rich tubers

Philipp William Niemeyer<sup>1</sup> , Iker Irisarri<sup>2</sup> , Patricia Scholz<sup>1</sup> , Kerstin Schmitt<sup>3</sup> , Oliver Valerius<sup>3</sup> ,  
Gerhard H. Braus<sup>3</sup> , Cornelia Herrfurth<sup>1,4</sup> , Ivo Feussner<sup>1,4</sup> , Shrikant Sharma<sup>5</sup> , Anders S. Carlsson<sup>5</sup> ,  
Jan de Vries<sup>2</sup> , Per Hofvander<sup>5,\*</sup>  and Till Ischebeck<sup>1,6,\*</sup> 

<sup>1</sup>Department of Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences and Göttingen Center for Molecular Biosciences (GZMB), University of Göttingen, 37077, Göttingen, Germany,

<sup>2</sup>Department of Applied Bioinformatics, Göttingen Center for Molecular Biosciences (GZMB) and Campus Institute Data Science (CIDAS), Institute for Microbiology and Genetics, University of Göttingen, 37077, Göttingen, Germany,

<sup>3</sup>Department for Molecular Microbiology and Genetics, Genetics and Göttingen Center for Molecular Biosciences (GZMB) and Service Unit LCMS Protein Analytics, Institute for Microbiology, University of Göttingen, 37077, Göttingen, Germany,

<sup>4</sup>Department of Plant Biochemistry, Service Unit for Metabolomics and Lipidomics, Göttingen Center for Molecular Biosciences (GZMB), University of Göttingen, 37077, Göttingen, Germany,

<sup>5</sup>Department of Plant Breeding, SLU Alnarp, Swedish University of Agricultural Sciences, Box 190, SE-234 22, Lomma, Sweden, and

<sup>6</sup>Green Biotechnology, Institute of Plant Biology and Biotechnology (IBBP), University of Münster, 48143, Münster, Germany

Received 12 March 2021; revised 9 August 2022; accepted 26 August 2022; published online 1 September 2022.

\*For correspondence (e-mail till.ischebeck@uni-muenster.de and per.hofvander@slu.se).

## SUMMARY

There are numerous examples of plant organs or developmental stages that are desiccation-tolerant and can withstand extended periods of severe water loss. One prime example are seeds and pollen of many spermatophytes. However, in some plants, also vegetative organs can be desiccation-tolerant. One example are the tubers of yellow nutsedge (*Cyperus esculentus*), which also store large amounts of lipids similar to seeds. Interestingly, the closest known relative, purple nutsedge (*Cyperus rotundus*), generates tubers that do not accumulate oil and are not desiccation-tolerant. We generated nanoLC-MS/MS-based proteomes of yellow nutsedge in five replicates of four stages of tuber development and compared them to the proteomes of roots and leaves, yielding 2257 distinct protein groups. Our data reveal a striking upregulation of hallmark proteins of seeds in the tubers. A deeper comparison to the tuber proteome of the close relative purple nutsedge (*C. rotundus*) and a previously published proteome of *Arabidopsis* seeds and seedlings indicates that indeed a seed-like proteome was found in yellow but not purple nutsedge. This was further supported by an analysis of the proteome of a lipid droplet-enriched fraction of yellow nutsedge, which also displayed seed-like characteristics. One reason for the differences between the two nutsedge species might be the expression of certain transcription factors homologous to ABSCISIC ACID INSENSITIVE3, WRINKLED1, and LEAFY COTYLEDON1 that drive gene expression in *Arabidopsis* seed embryos.

**Keywords:** lipid droplets, yellow nutsedge, *Cyperus esculentus*, tubers, seeds, proteome, *Cyperus rotundus*, *Arabidopsis thaliana*.

## INTRODUCTION

Drought is the prime stressor that land plants have to overcome since their dawn (Fürst-Jansen et al., 2020). One strategy that land plants employ is to limit water loss and to transport water via a vascular system (Harrison & Morris, 2018; Lu et al., 2020). Another is to produce desiccation-tolerant cells and tissues (Oliver et al., 2020). Common to drought and desiccation tolerance is the accumulation of osmolytes and proteins with protective functions (Oliver et al., 2020). Also common to both strategies

is the accumulation of neutral lipids, foremost triacylglycerol (TAG), in cytosolic lipid droplets (LDs), with especially high levels being reached in embryonic tissues (de Vries & Ischebeck, 2020). In most flowering plants, desiccation tolerance is limited to seeds and to some extent pollen. However, there are also plants that have desiccation-tolerant vegetative organs such as *Sporobolus stapfianus* (Neale et al., 2000), *Craterostigma plantagineum* (Bartels, 2005), and *Eragrostis nindensis* (Pardo et al., 2020); Desiccation is furthermore found in non-flowering plants such as

bryophytes (Alejo-Jacuinde et al., 2020) and clubmosses (Gao et al., 2017). Therefore, it is likely that desiccation tolerance evolved several times independently (Le & McQueen-Mason, 2006; Oliver et al., 2000; Pardo et al., 2020). It is however also likely that these independent origins of desiccation tolerance are underpinned by the co-option of existing regulatory programs for resilience (VanBuren et al., 2017).

Here, we investigated yellow nutsedge (*Cyperus esculentus*), a monocot, perennial C4 plant (Defelice, 2002). This species produces stolon-derived underground tubers that can fully desiccate and remain viable (Stoller & Sweet, 1987). Furthermore, they can accumulate 25–30% of their dry mass in lipids, especially TAG (Turesson et al., 2010; Yang et al., 2016). In this regard, they are unique as even the tubers of a close relative, purple nutsedge (*Cyperus rotundus*), are neither desiccation-tolerant nor oil accumulating (Iqbal et al., 2012; Wills, 1987). We studied the proteomes of yellow and purple nutsedge tubers and compared them to previously published proteomes of Arabidopsis seeds and seedlings. This comparison revealed striking similarities between seeds and yellow but not purple nutsedge tubers, and between seedling establishment and tuber sprouting of yellow nutsedge. Our results indicate that yellow nutsedge evolved the desiccation tolerance and oil-richness in their tubers probably by co-opting a protein framework normally present predominantly in seeds by the expression of master regulators.

## RESULTS

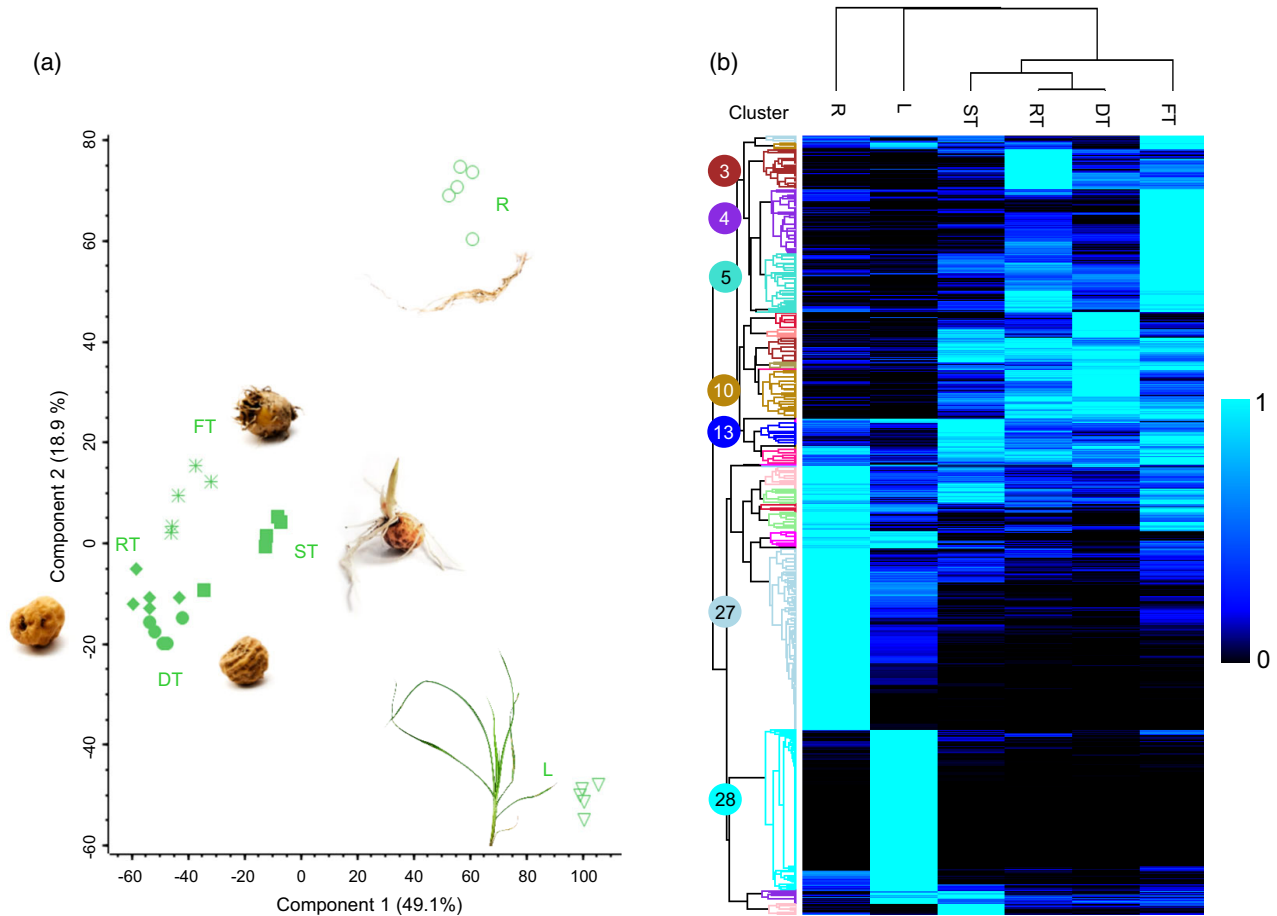
### Tubers of yellow nutsedge are enriched in hallmark proteins of seeds

To understand the molecular basis of tuber resilience in yellow nutsedge, we studied the proteome of four stages of tuber development (freshly harvested, dried, rehydrated for 48 h, and sprouted) and compared it to roots and leaves (Figure S1). The tubers contained 6% water before and 35% water after rehydration. From the first three tuber stages, we additionally isolated LD-enriched fractions. After a tryptic digestion step, all peptide samples were analyzed by LC-MS/MS in five biological replicates and all protein groups were quantified using a label-free MS1-based quantification algorithm resulting in relative intensity-based absolute quantification (riBAQ) values (Data S1). For functional annotation of these proteins, we used all nutsedge library entries as queries for a BLASTp against the Arabidopsis TAIR10 primary transcript protein release library (Lamesch et al., 2012) (Data S2). We also searched for post-translational modifications and found 23 proteins with phosphorylation sites (Ser, Thr, Tyr), 14 with ubiquitination sites (Lys), and 23 with acetylation sites (Lys, see Data S1).

Then, we studied first the total proteomes. A principal component analysis (PCA) revealed that the tuber proteomes were mostly separated by developmental stage but much more distinct from roots and leaves (Figure 1a). In order to identify tuber-specific proteins, the data were hierarchically clustered (Figure 1b). Nine of the resulting 30 clusters (clusters 3–6, 8–11, and 13, see Data S3) contained 477 protein groups highly enriched in at least one of the tuber stages and less abundant in leaves and roots. These protein groups corresponded to 433 Arabidopsis homologs that were subjected to a gene ontology (GO) search (Mi et al., 2019). This search indicates that proteins assigned to certain GO terms are overrepresented in number in a group of proteins in comparison to their number in the whole proteome. The search showed that 54 of these enriched proteins were associated with carbohydrate metabolism and 41 with cellular responses to stress (Data S3) but similar numbers were found in the other clusters combined (Data S3). Overrepresented were also starch metabolism and several seed-related processes, including seed maturation, dormancy, germination, and seedling development (Data S3). None of these seed-related terms were overrepresented in the other combined clusters (Data S3).

Seven of the 10 most abundant proteins that were at least 20-fold enriched in one of the tuber stages in comparison to leaves and roots (Figure 2 and Data S4 and S5) have Arabidopsis homologs that are almost exclusively expressed in seeds based on transcript data (Klepikova et al., 2016). Furthermore, these homologs are strongly enriched in proteomes of Arabidopsis seeds compared to 60-h-old seedlings (Figure 2; Data S5, Kretschmar et al., 2020) and are also found in the seed proteomes of *Nicotiana tabacum* (Kretschmar et al., 2018) and *Brassica napus* (Han et al., 2013) seeds (Data S5). For comparison, we also isolated proteins from purple nutsedge and analyzed them in the same manner as the yellow nutsedge proteins. Here, only fresh tubers (Figure S1) were analyzed as the tubers of purple nutsedge cannot undergo a desiccation phase. In total, 1174 proteins could be identified (Data S6, see Data S7 for BLASTp results against the Arabidopsis and yellow nutsedge proteomes). A comparison to yellow nutsedge in regard to the tuber-enriched proteins revealed that purple nutsedge homologs of these proteins were either not found or of lower abundance (Figure 2a). Further, the proteome of the purple nutsedge tubers was distinct from all yellow nutsedge stages (Figure S2).

We additionally compared the expression levels of the genes coding for the tuber-specific proteins by investigating the recently published transcriptomes of yellow nutsedge and the respective homologs in purple nutsedge (Ji et al., 2021; Figure 2b). In this study, tubers were harvested 20, 50, and 90 days after tuber formation. All of the seven 'seed' protein-coding genes had at least 10 times higher transcript levels in yellow nutsedge than in purple



**Figure 1.** Developmental tuber stages have a distinct proteome. Comparative proteomics of total protein extracts of tubers (fresh [FT], dry [DT], rehydrated for 48 h [RT], and sprouted [ST]), leaves (L), and roots (R) (Data S1). (a) Principal component analysis plot. (b) The values were normalized setting the average of the stage with the highest abundance to 1 and these values were hierarchically clustered. Clusters are labeled with numbers in circles. Proteins sorted by cluster can be found in Data S3. *n* = 5 biological replicates per stage.

nutsedge when averaging the three tuber stages per species analyzed in the study (taking also into account transcripts with very high similarity; Data S5).

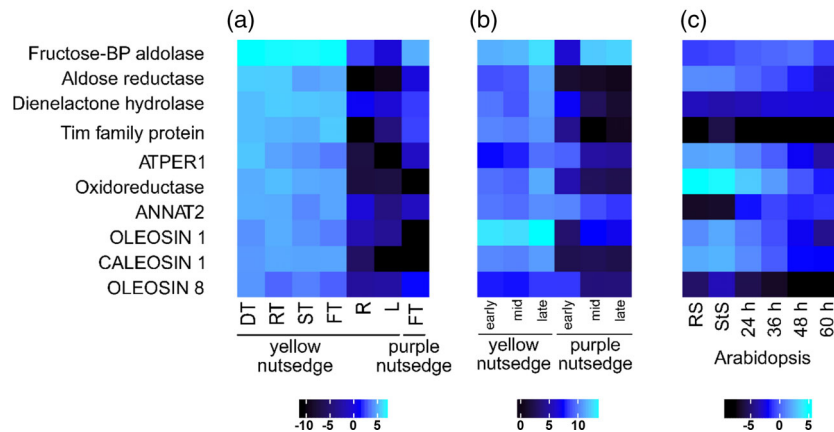
**Proteomes of tubers of yellow nutsedge and Arabidopsis seeds correlate**

To further explore the similarity of tubers of yellow nutsedge and seeds of Arabidopsis, and also to see differences to purple nutsedge, we compared the developmental patterns of the six yellow nutsedge stages/tissues investigated in this study and six stages of seedling establishment from Arabidopsis (Kretschmar et al., 2020). In the Arabidopsis study, seeds rehydrated for 30 min and seeds stratified for 74 h at 4°C were compared to seedlings grown for up to 72 h after stratification. Nutsedge proteins were then assigned to their closest Arabidopsis homolog based on our BLAST analysis (Data S2 and S7). Abundances of proteins with the highest homology to the same Arabidopsis protein were added. In total, 1067 protein homologs were present in the proteomes of both Arabidopsis and at least

one of the nutsedge species and were found in at least four replicates of at least one stage. Furthermore, to allow for meaningful comparisons, we defined, separately for the nutsedge and the Arabidopsis samples and for each protein, the highest average abundance in a given stage as 1. This dataset was then analyzed in three ways.

First, a PCA revealed that the yellow nutsedge tuber and Arabidopsis seed samples clustered closely together with the exception of fresh tubers that were more closely related to seedlings (Figure 3a). The purple nutsedge tubers, on the other hand, did not cluster with any Arabidopsis samples but appeared more similar to the root samples of yellow nutsedge.

Second, hierarchical clustering of the stages showed again a similarity of Arabidopsis seed and yellow but not purple nutsedge tuber stages. Furthermore, separating the proteins into 30 clusters (Figure 3b and Data S8) yielded two clusters (3 and 4), comprising 78 proteins with a marked elevation in seeds rehydrated for 30 min and seeds stratified for 74 h at 4°C, and in dry and rehydrated tubers.



**Figure 2.** Highly abundant and tuber-specific proteins display higher expression in yellow than in purple nutsedge tubers and have seed-enriched homologs in Arabidopsis.

(a) The protein compositions of yellow nutsedge tubers (fresh, dry, rehydrated, and sprouted) were compared to the ones of leaves and roots and to the fresh tubers of purple nutsedge. Displayed are the 10 most abundant proteins found in the yellow nutsedge tuber samples that were at least 20-fold enriched in comparison to leaves and roots (by highest average). Given are  $\log_2$ -transformed % riBAQ values. (b) Expression of homologous proteins identified in transcriptomes of yellow and purple nutsedge based on previously published data (Ji et al., 2021; closely related homologs were accounted for). Given are  $\log_2$ -transformed TMM values. (c) Protein abundance of the closest homologs in Arabidopsis rehydrated seeds (RS), stratified seeds (StS), and seedlings 24–60 h after stratification as previously published (Kretschmar et al., 2020). Given are  $\log_2$ -transformed % riBAQ values. See Data S4 for a full list of enriched proteins and Data S5 for numerical values of the three heatmaps. Arabidopsis homologs were determined by BLASTp analysis (Data S2).

These proteins included known LD-associated proteins such as oleosins and caleosins, late embryogenesis abundant (LEA) proteins, and class I small heat shock proteins. All these protein families are known to be highly abundant in seeds (Battaglia et al., 2008; Huang, 2018; Wehmeyer et al., 1996) but likewise important in stress responses (Aubert et al., 2010; Guo et al., 2020; Hanin et al., 2011; Shimada et al., 2008). Cluster 17 noteworthy contains 66 proteins that are highly expressed in sprouting tubers and seedlings (Figure 3b and Data S8). Included in this cluster are proteins involved in  $\beta$ -oxidation and the glyoxylate cycle, needed for the conversion of lipids into carbohydrates. Further, tubers and seeds again clustered together.

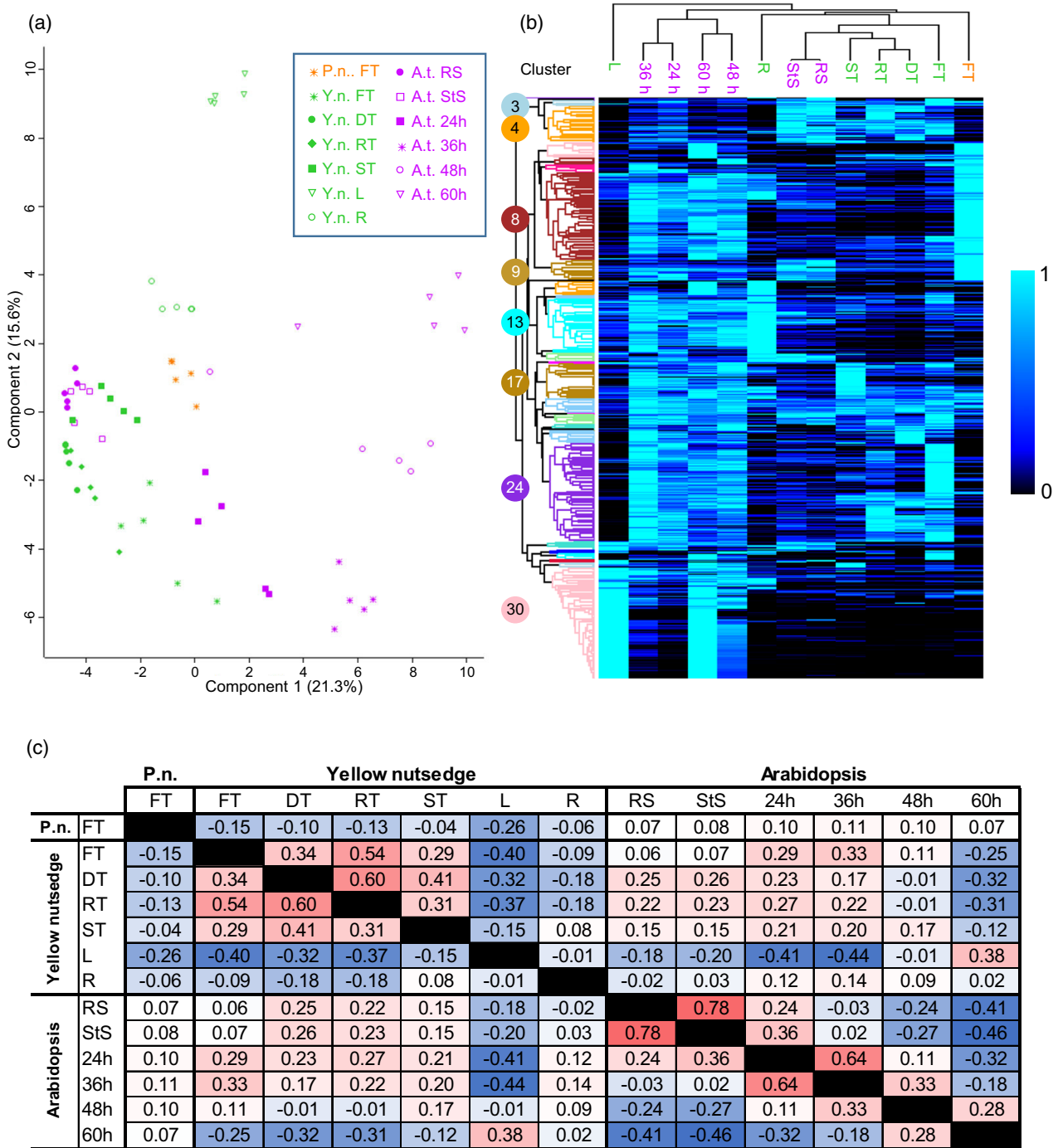
Third, we calculated correlation coefficients between the averages of the different developmental stages to estimate similarities (Figure 3c). Here it was evident that, while tubers were most similar to each other, they also showed a positive correlation with the Arabidopsis seed and young seedling stages. Reversely, Arabidopsis seeds had higher correlation coefficients to nutsedge tubers than to 36-h-old Arabidopsis seedlings. The purple nutsedge stage analyzed showed a negative correlation to all of the yellow nutsedge stages and a weak positive correlation to all of the Arabidopsis stages.

#### Homologs to proteins with important functions in seeds and seedlings are enriched in tubers

GO term analyses based on the total abundance of proteins (refer to Data S9 for yellow nutsedge, Data S10 for purple nutsedge, and Data S11 for a comparison of both species to Arabidopsis) bolstered the similarities of yellow

nutsedge tubers and Arabidopsis seeds (Figure 4). For instance, the terms ‘embryo development’, ‘lipid storage’, ‘maintenance of seed dormancy’, and ‘seed germination’ were especially high in seeds and tubers but reduced in older stages. The term ‘maintenance of seed dormancy’ only comprises one protein, a homolog to the Arabidopsis protein 1-CYSTEINE PEROXIREDOXIN 1 (PER1), which enhances primary seed dormancy by suppressing abscisic acid (ABA) catabolism and gibberellic acid biosynthesis (Chen, Ruan, et al., 2020). Another protein, homolog to GEM-RELATED 5 (GER5, AT5G13200), displays a similar pattern (Data S3). *GER5* is an ABA-responsive gene that also regulates germination in Arabidopsis seeds (Baron et al., 2014). Based on these results, we tested if ABA levels were increased in nutsedge tubers. Indeed, ABA levels ( $51 \text{ pmol g}^{-1}$  fresh weight) in dry tubers decreased to 7% (Figure 5; Data S12) of the dry tuber levels during sprouting. Also in roots and leaves of yellow nutsedge the levels were much lower. In the tubers of purple nutsedge, no ABA could be detected. The tubers were not enriched in any other hormones we could detect in comparison to other tissues, with jasmonoyl-isoleucine (JA-Ile) and salicylic acid (SA) levels being especially high in roots and leaves, respectively.

The GO term analysis (Figure 4) also revealed differences between tubers and seeds. Tubers had higher levels of proteins associated with starch biosynthesis and glycolytic processes, likely reflecting that the tubers store starch and free sugars in higher amounts (Linssen et al., 1989). Another difference is the amount of storage proteins present. While members of cupin storage protein



**Figure 3.** Comparison of protein expression patterns between nutsedge tissues and Arabidopsis seeds and seedlings. Total protein extracts of yellow (Y.n.) and purple (P.n.) nutsedge tubers (fresh [FT], dry [DT], rehydrated for 48 h [RT], and sprouted [ST]), leaves (L), and roots (R) were compared to Arabidopsis (A.t.) seeds (rehydrated [RS] and stratified [StS]) and seedlings (24–60 h in the light after stratification). (a) Principal component analysis comparing the individual samples. (b, c) For the hierarchical clustering (b, see Data S8 for clusters) and the Pearson correlation analysis (c), samples were averaged for each stage. Clusters in (b) are labeled with numbers in circles. Arabidopsis protein data were previously published in Kretzschmar et al. (2020).

families are enriched in tubers and decrease during sprouting (Data S9, GO term 45735), Arabidopsis seeds contain a far greater amount of storage proteins amounting to more

than half of the total protein based on the proteomic data; this amount is far lower in tubers (<0.5%). Tubers on the other hand contain many proteins associated with protein



GO Term	Purple nutsedge				Yellow nutsedge				Arabidopsis					
	FT	FT	DT	RT	ST	R	L	RS	StS	24h	36h	48h	60h	
9793 Embryo development	20	42	33	47	36	18	10	377	347	271	195	128	75	
1990137 Plant seed peroxidase activity	n.f.	19	15	20	19	0.26	0.01	2.0	2.4	1.4	0.66	0.35	0.37	
10231 Maintenance of seed dormancy	0.28	8.1	35	18	13	0	0	2.2	2.4	1.3	0.62	0.19	0.03	
9737 Response to abscisic acid	28	96	144	126	113	21	25	472	444	339	261	239	198	
9269 Response to desiccation	7.6	11	37	23	15	26	18	2.8	3.0	2.1	1.4	1.8	2.5	
9845 Seed germination	7.8	29	28	43	30	9.5	2.6	102	107	74	42	27	17	
6633 Fatty acid biosynthetic process	2.4	11	6.4	7.4	9.9	3.9	3.8	2.2	1.9	1.9	2.5	2.7	2.8	
10344 Seed oilbody biogenesis	n.f.	16	22	33	21	0.16	0.07	10.6	12.5	7.3	4.6	2.5	1.2	
19915 Lipid storage	0.2	22	35	38	34	2.7	0.7	11	13	7.3	4.6	2.5	1.2	
12511 Monolayer-surrounded lipid storage body	n.f.	20	30	34	28	0.35	0.18	6.6	9.1	5.2	3.0	1.6	0.8	
8106 Alcohol dehydrogenase (NADP+) activity	0.13	19	22	31	20	0.47	0.46	19	17	7.1	3	1	0.56	
6099 Tricarboxylic acid cycle	9.5	20	17	20	36	26	8.8	7.5	11	47	63	44	32	
6635 Fatty acid beta-oxidation	0.3	1.5	1.9	2	4.4	0.37	0.04	2.3	2.1	7.8	8.9	8.2	6.2	
6097 Glyoxylate cycle	0.52	3.2	2.8	3.7	11	0.32	0.09	2.9	5.9	37	50	32	20	
6096 Glycolytic process	133	279	332	291	310	45	97	20	21	29	39	55	61	
19252 Starch biosynthetic process	17	21	14	17	6.8	5.5	6.7	0.23	0.42	0.75	1.3	1.8	2.7	
45735 Nutrient reservoir act.	10	3.1	4.1	3.8	1.7	38	9.9	520	503	370	234	127	66	
6457 Protein folding	71	26	46	45	27	9.2	6.7	14	17	40	61	71	60	
2000280 Regulation of root development	4.81	0.02	0.00	0.01	0.00	30.88	2.05	0.00	0.00	0.00	0.00	0.00	0.00	
9809 Lignin biosynthetic process	35	2.7	1.0	1.4	1.3	20	3.0	1.3	1.1	2.8	4.8	4.8	4.5	

**Figure 4.** GO term comparison.

Total protein extracts of yellow and purple nutsedge tubers (fresh [FT], dry [DT], rehydrated for 48 h [RT], and sprouted [ST]), leaves (L), and roots (R) were compared to Arabidopsis seeds (rehydrated [RS] and stratified [StS]) and seedlings (24–60 h in the light after stratification). Several GO terms are displayed that show similarities but also differences between tubers and seeds and between yellow and purple nutsedge. Proteins were assigned to each GO terms and their per mille rBAQ values of the total proteome were added up. A list of all GO terms can be found in Data S9–S11. *n* = 5 for all stages. Arabidopsis protein data were previously published in Kretzschmar et al. (2020).

folding, including members of various heat shock protein families that might protect proteins during desiccation. In contrast, these types of proteins are less abundant in Arabidopsis seeds and accumulate only during seedling establishment, indicating that here other strategies are employed to preserve protein integrity.

In purple nutsedge, the GO terms ‘glycolytic process’, ‘starch biosynthetic process’, and ‘protein folding’ are similarly high as in yellow nutsedge. However, lipid- and LD-related terms were much lower, as were the terms ‘response to abscisic acid’ and ‘response to desiccation’. Much higher on the other hand was the term ‘lignin biosynthetic process’. A proteomic analysis of purple nutsedge tubers that were dried prior to protein extraction did not reveal any strong enrichment of typical seed proteins (Data S13) or proteins belonging to GO terms that are high in seeds and yellow nutsedge tubers (Data S9).

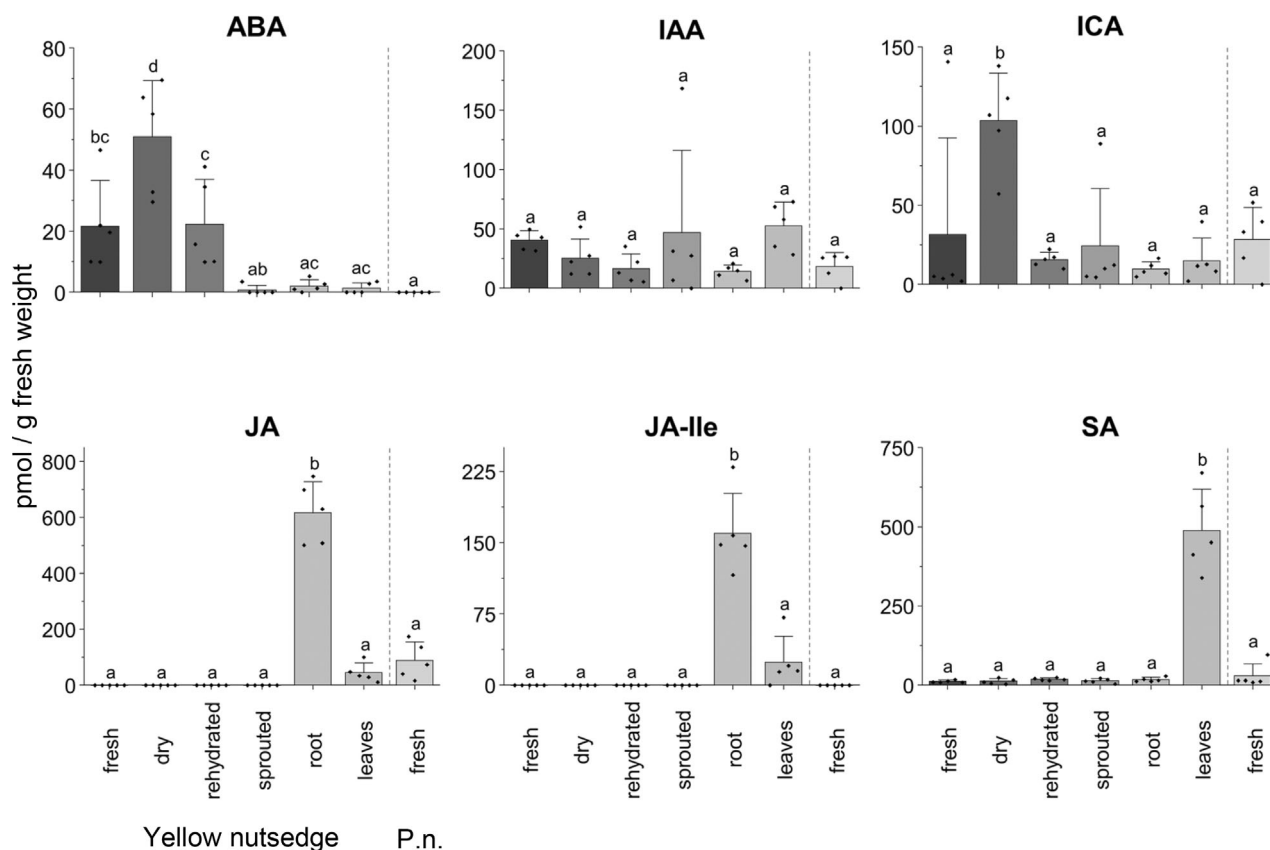
We also compared the proteome of nutsedge tubers to the previously published proteomes of the starch-rich tubers of potato (*Solanum tuberosum*; Lebecka et al., 2019) and yam (*Dioscorea alata*; Sharma & Deswal, 2021) (Figure S3 and Data S11, S14, and S15). As expected, these proteomes did not harbor a seed-like signature as apparent for example by the proteins grouped in the GO terms ‘maintenance of seed dormancy’, ‘response to desiccation’, and ‘seed oilbody biogenesis’, which were either not found or only summed up to a comparable low abundance. The term ‘response to abscisic acid’ was however relatively high in potato, as was the term ‘nutrient reservoir activity’. Proteins involved in glycolytic processes

were higher in both yam and potato than in Arabidopsis seeds, but not as high as in yellow and purple nutsedge.

#### Lipid droplet proteins known from seeds abundant in tubers of yellow nutsedge

Most plant cells store TAG predominantly in cytosolic LDs (Ischebeck et al., 2020). LDs are presumably present in all cell types but can strongly differ in their protein composition (Brocard et al., 2017; Fernández-Santos et al., 2020; Horn et al., 2013; Kretzschmar et al., 2018, 2020). Oil-rich but non-desiccation-tolerant tissues such as avocado (*Persea americana*) mesocarp lack most LD-associated proteins, which are most abundant in seeds (Horn et al., 2013). The described similarities between tubers and seeds and the notion that both accumulate oil raise the question if the tuber proteins that are associated with LDs are also similar to seed proteins. Therefore, we analyzed the proteome of LD-enriched fractions of three tuber stages and could see that this proteome was clearly distinct from that of the total cellular fraction (Figure S4; Data S1).

Homologs to known LD proteins made up approximately 60–70% of the proteins in LD-enriched fractions (Data S16), with only approximately 5% being found in total extracts (purple nutsedge tubers for comparison only contained 0.2% LD proteins in the total extracts). With one exception, all these homologs were significantly enriched in the LD fraction (Figure 6a), indicating that these homologs are also LD-associated. The most abundant among these proteins were the oleosins (Figure 6b), especially the isoforms with higher sequence similarity to the



**Figure 5.** Hormone analysis.

Hormone levels were determined by UPLC-nanoESI-MS/MS in different tissues of yellow nutsedge and fresh tubers of purple nutsedge (P.n.).  $n = 5$  for each stage. Error bars indicate standard deviation. ABA, abscisic acid; IAA, indole-3-acetic acid; ICA, indole carboxylic acid; JA, jasmonic acid; JA-Ile, jasmonoyl-isoleucine; SA, salicylic acid.

Arabidopsis main seed oleosins OLE1, 2, 4, and 5 (Klepikova et al., 2016; Kretschmar et al., 2020; Figure S5a). Oleosins have been described to have a structural role as surfactants and to be seed- and pollen-specific proteins (Chapman et al., 2012).

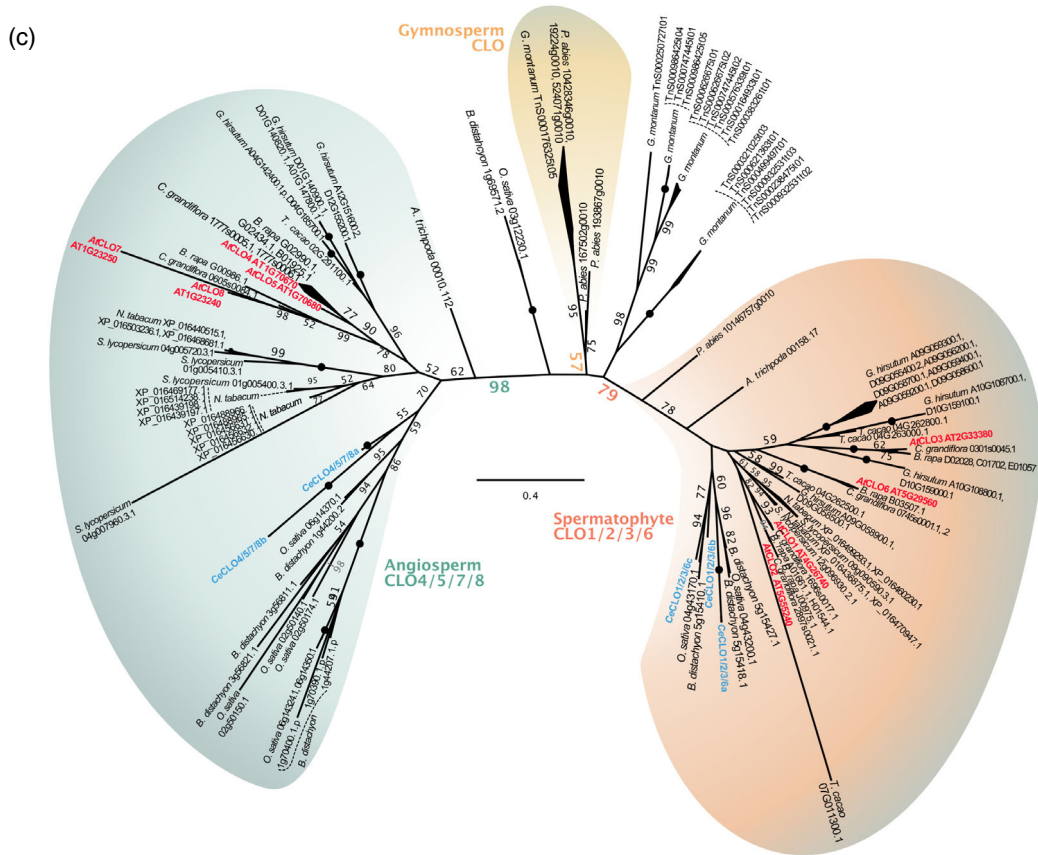
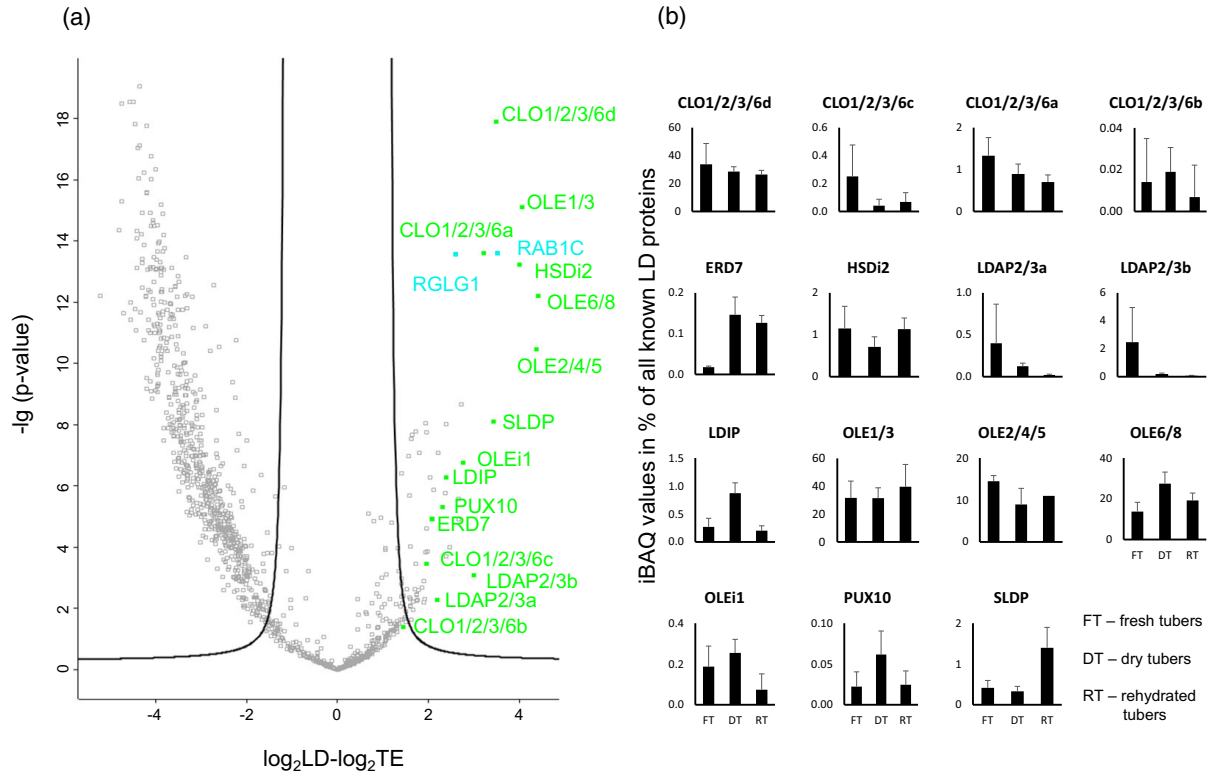
Steroleosins are likewise predominantly found in seeds and not in vegetative tissues (Shimada et al., 2018). They are speculated to play a role in the metabolism of brassinosteroids (Baud et al., 2009; Li et al., 2007), hormones involved in development, and seed dormancy (Peres et al., 2019). Two steroleosin isoforms were

detected in tubers (Figure S5b). Interestingly, seed lipid droplet protein (SLDP), another LD-associated protein (Kretschmar et al., 2020) that is seed- and seedling-specific in Arabidopsis (Klepikova et al., 2016) and is part of a membrane contact site between LDs and the plasma membrane (Krawczyk et al., 2022; Scholz et al., 2022), was also detected in tubers.

Caleosins (CLOs), in comparison, have been found in seeds, pollen, and leaves (Chen et al., 1999; Jiang et al., 2008; Shimada & Hara-Nishimura, 2015). In Arabidopsis, CLO1 and CLO2 are the dominant caleosins of

**Figure 6.** LD proteomes of yellow nutsedge tubers resemble those of seeds and seedlings.

(a) A volcano plot was constructed to visualize proteins, which are significantly and consistently LD-enriched in yellow nutsedge. All 15 LD-enriched samples were compared to all total cellular fractions of fresh, dry, and rehydrated tubers (Data S1). The  $\log_2$ -transformed values and  $P$ -values were calculated. Known LD proteins are indicated in green, while interesting candidates are shown in blue (see Data S16 for a list of all candidates). Black lines indicate a false discovery rate of 0.01. (b) The riBAQ values of LD-associated proteins in the LD-enriched fraction were calculated as a percentage of the riBAQ of all known LD-associated proteins (see Data S9 for data). (c) Unrooted maximum likelihood phylogeny of caleosin homologs detected in the predicted proteomes of spermatophyte genomes and the yellow nutsedge transcriptomes. Yellow nutsedge and Arabidopsis caleosins are depicted in blue and red, respectively. The angiosperm and gymnosperm depicted protein clades contain only proteins from species of the respective plant clades while the spermatophyte depicted protein clade contains proteins from both plant clades. Abbreviations: CLO, caleosin; ERD, early responsive to dehydration; HSD, steroleosin; LDAP, LD-associated protein; LDIP, LDAP-interacting protein; OLE, oleosin; PUX, plant UBX domain-containing protein; RAB, Rab GTPase; RGLG, RING domain ligase, SLDP, seed LD protein.



seeds (Kretschmar et al., 2020); CLO3 is likely the most abundant leaf LD protein (Brocard et al., 2017; Fernández-Santos et al., 2020) and is upregulated under stress (Aubert et al., 2010). We explored whether the four tuber caleosins of yellow nutsedge showed a specific phylogenetic relationship to the seed or leaf isoforms of Arabidopsis. We however found that CLO1, 2, and 3 from Arabidopsis and the four caleosins detected in tubers belong to a larger clade within the CLO gene family (here coined 'CLO1/2/3/6') within which lineage-specific duplications occurred; the most recent common ancestor of monocots and dicots likely had a single CLO1/2/3/6 homolog (Figure 6c).

Two further ubiquitously present LD proteins involved in the proper formation of LDs are lipid droplet-associated protein (LDAP) (Gidda et al., 2016) and its interaction partner LDAP-interacting protein (LDIP; Pyc et al., 2017, 2021). The abundance of LDAPs is low in both desiccated tubers (Figure 6b) and seeds (Kretschmar et al., 2020). Their interaction partner LDIP is abundant in the desiccated structures of both species. The found LDAPs were classified in the same clade as LDAP2 and 3 from Arabidopsis (Figure S6), which are the main LDAPs found in seeds and seedlings.

ERD7, a recently identified LD-associated protein that was found in the proteome of drought-stressed Arabidopsis leaves (Doner et al., 2021) and was shown to play a role in membrane remodeling under cold stress (Barajas-Lopez et al., 2021), and a homolog were also detected in all tuber stages of yellow nutsedge.

Apart from known LD proteins, further proteins were found enriched in the LD fraction (Data S16). Especially strongly enriched was a small G protein, RAB1C. Its mammalian homolog RAB18 is known to localize to LDs in mammals (Jayson et al., 2018) but LD-binding G proteins have so far not been described in plants. It appears possible though that they are involved in LD formation or trafficking.

Another LD-enriched protein, the E3 ubiquitin ligase RING domain ligase 1 (RGLG1; Cheng et al., 2012), would be interesting to study in the future considering that several LD proteins are ubiquitinated prior to degradation (Deruyffelaere et al., 2015, 2018; Hsiao & Tzen, 2011; Kretschmar et al., 2018). RGLG1 from Arabidopsis is predominantly expressed in seeds (Klepikova et al., 2016) and was enriched in the LD fractions of seedlings (Kretschmar et al., 2020). It also plays a role in drought responses (Cheng et al., 2012).

### The seed-like proteome evolved in a relatively short amount of time

There are striking similarities between Arabidopsis seeds and tubers of yellow but not purple nutsedge. It is therefore conceivable that oil accumulation and desiccation

tolerance of both organs are based on a similar proteome and the underlying gene expression pattern that could be regulated by a similar network of hormones and transcription factors, as for example the transcription factors ABSCISIC ACID INSENSITIVE3 (ABI3), WRINKLED1 (WRI1), and LEAFY COTYLEDON1 (LEC1), which show increased expression levels in the tubers of yellow nutsedge in comparison to purple nutsedge (Figure S7; Ji et al., 2021).

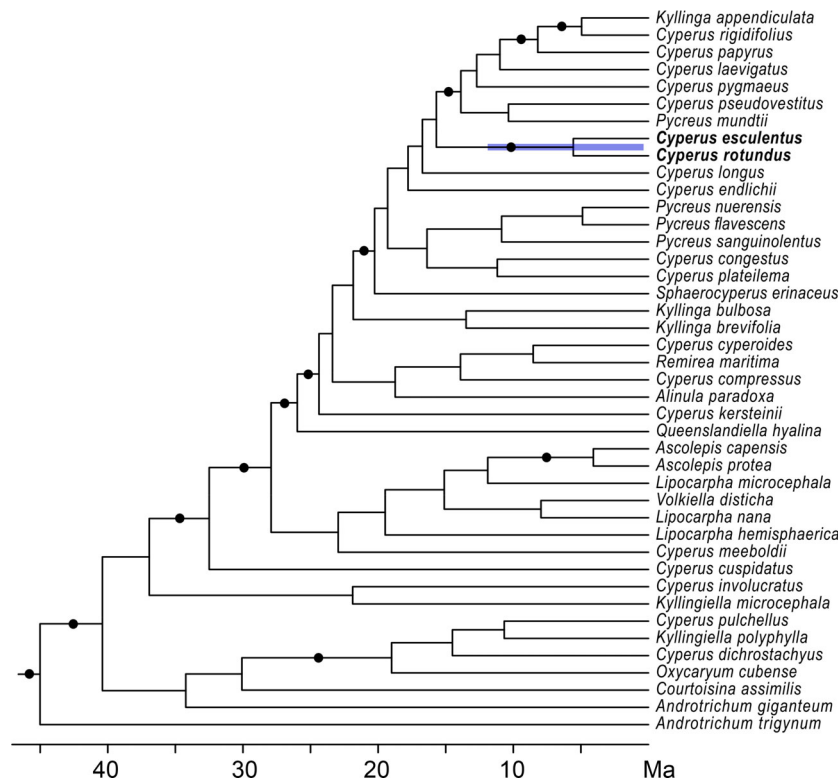
In order to estimate the time frame in which the difference in oil-richness between yellow and purple nutsedge evolved, we performed a Bayesian molecular clock analysis (Figure 7, Figure S8, and Data S17 and S18) using three plastid regions (*rbcL*, *trnL* intron, and *trnL-trnF* spacer; 2460 aligned nucleotides) for 267 species of Cyperaceae. Our results indicate that purple and yellow nutsedge diverged only around 5.6 million years ago (95% highest posterior density interval, 0.3–12 million years ago), a similar time frame to that estimated for the divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata* (Maurus & Quesneville, 2014).

## DISCUSSION

### Yellow nutsedge tubers contain a seed-like proteome

The tubers of yellow nutsedge have a similar function as seeds in other plant species. While they represent a means of vegetative, not sexual propagation, the tubers can, like seeds, endure desiccation for several years before sprouting under favorable conditions (Stoller & Sweet, 1987), giving rise to photosynthetically active offspring within days. This rapid sprouting is supported by storage compounds in the form of TAG and starch (Turesson et al., 2010; Yang et al., 2016).

These comparable functions are mirrored by striking similarities to seed (Data S5) but not tuber proteomes (Figure S3) of other dicotyledons on several levels. The tubers of yellow nutsedge contain LEA proteins that are known to protect the cells during desiccation (Hanin et al., 2011) and homologs to other seed-specific proteins in Arabidopsis like PER1 and a plastidial aldehyde reductase, whose function in seeds has so far not been elucidated (Data S5). The tubers harbor LD-associated proteins (Figure 6) like oleosins highly prevalent in seeds (Huang, 2018) that have to the best of our knowledge so far not been found in the proteomes of vegetative tissues from flowering plants (Brocard et al., 2017; Horn et al., 2013). We also found a homolog to SLDP1 and steroleosins that have so far only been found in seeds and seedlings at the proteome level (Ischebeck et al., 2020). As steroleosins are implicated in brassinosteroid metabolism, this warrants future work on their role in development, dormancy, and/or sprouting of yellow nutsedge. Furthermore, similar to seeds (Chen, Li, et al., 2020), tuber dormancy and desiccation tolerance might be regulated by the phytohormone ABA. ABA levels



**Figure 7.** Yellow and purple nutsedge diverged approximately 5.6 million years ago according to a Bayesian molecular clock analysis.

Excerpt of the time-calibrated phylogeny of the Cyperaceae inferred with a Bayesian relaxed molecular clock on three plastid markers (*rbcl*, *trnL* intron, and *trnL-trnF* spacer; 2460 aligned nucleotides) and 267 species. The divergence between yellow and purple nutsedge was inferred to be 5.6 million years ago (95% highest posterior density interval, 0.3–12 million years ago). Dots represent branches with >90% ultrafast bootstrap support. The full tree is depicted in Figure S8 and Data S18.

are very high in tubers (Figure 5) and decrease during rehydration and sprouting; this notion is corroborated by abounding ABA-related proteins like PER1 and GER5 (Data S3).

The described similarities also extend to seedling establishment and tuber sprouting, as can be seen for example for  $\beta$ -oxidation and the glyoxylate cycle, which play key roles during *Arabidopsis* seedling establishment (Ischebeck et al., 2020). Similar to seedling establishment, the respective GO terms also increase in sprouting tubers (Figure 4). Yellow nutsedge thus seems to utilize its oil to synthesize carbohydrates, despite storing additionally large amounts of sugars and starch that are also degraded during sprouting (Stoller et al., 1972).

#### A short evolutionary timeframe for the emergence of yellow nutsedge's unique features

Based on our molecular clock analysis (Figure 7 and Figure S8), purple and yellow nutsedge diverged a comparably short time ago. This means that a whole molecular program for seed-like characteristics only found in yellow nutsedge must have emerged rapidly and might have evolved as an adaptation strategy to changes in its environment.

One possibility for swift alterations in a whole set of genes and proteins is a shift in their regulatory regime, foremost master regulators in the form of transcription factors (Das Gupta & Tsiantis, 2018). Indeed, homologs of the transcription factors WRI1, LEC1, and ABI3, which play important roles in seed development and maturation (Chen, Li, et al., 2020), are expressed more highly in yellow than in purple nutsedge (Figure S7). It is therefore conceivable that these transcription factors are part of a regulatory network that simply shifted its tissue specificity, resulting in the co-option of an established molecular program by a different organ. This underlines the potential for a complete reprogramming of vegetative tissues in crop plants by genetic engineering (Hofvander et al., 2016; Vanhercke et al., 2017; Vanhercke et al., 2019). Strategies observed in tubers of yellow nutsedge have the potential to guide such biotechnological approaches.

#### A recurrent framework of proteins is used in desiccation tolerance and oil accumulation

While desiccation tolerance and oil-richness in yellow nutsedge tubers might have evolved by co-opting a 'seed-like' gene network, it is likely that such a network already existed prior to the evolution of seeds. Indeed, comparable

expression patterns occur across non-seed land plants and even in streptophyte algae to sustain a variety of oil-rich desiccation-tolerant structures (de Vries & Ischebeck, 2020). For example, the oil-rich and desiccation-tolerant spores of *Physcomitrium patens* also abound in transcripts coding for typical LD proteins such as oleosins and LEA proteins (Fernandez-Pozo et al., 2019; Huang et al., 2009); similar transcripts also accumulate in drought-stressed algae (de Vries & Ischebeck, 2020).

In seed plants, pollen is another tissue that stores large amounts of TAG and is desiccation-tolerant in many plants (Ischebeck, 2016). Also here, oleosins are expressed and LEA proteins accumulate; LEA protein abundance rises during pollen maturation, followed by degradation during pollen tube growth (Ischebeck et al., 2014).

However, like yellow nutsedge, some plants also accumulate oil in vegetative tissues. One striking example is the resurrection grass *Oropetium thomaeum* that shows an increase in LDs after desiccation (VanBuren et al., 2017). This increase is accompanied by expression of oleosin and LEA protein-coding genes. Homologs to the transcription factor ABI3 might act as master regulators to boost the expression of a 'seed-like' gene pattern in this species as well.

In conclusion, TAG accumulation and expression of a 'seed-like' gene network often come along in desiccation-tolerant structures. A better understanding of the core components of this network could help to improve not only seed and pollen longevity, but also drought resistance of vegetative tissues.

## EXPERIMENTAL PROCEDURES

### Plant material

Yellow nutsedge (*C. esculentus* L. var. *sativus*) tubers were either taken dry and rehydrated for 2 days in water under continuous airflow or allowed to sprout in soil and harvested 1–2 days after sprouting. Yellow nutsedge plants were harvested from plants grown in a greenhouse under 14–16 h artificial light of mercury vapor lamps ( $150 \mu\text{mol m}^{-2} \text{sec}^{-1}$ ) complemented with sunlight at a temperature of 20–22°C at daytime or 16–18°C at nighttime. Roots and leaves were harvested from 33-day-old plants and fresh tubers were harvested from 4-month-old plants. Purple nutsedge was grown under the same conditions and tubers were harvested from 2-month-old plants. Tubers were either used fresh or dried for 7 days at 35°C at 20% relative humidity.

### Isolation of total and LD-enriched protein fractions

For each condition, five biological replicates were processed. Biological replicates used for LD isolation comprise five individual tubers, tissues of germinated tubers, leaves, and roots derived from single individuals. The tubers were ground in a pre-cooled mortar with 1–2 g sea sand and 10 ml grinding buffer (50 mM Tris, pH 7.4, 10 mM KCl, 200 mM sucrose, 200  $\mu\text{M}$  phenylmethylsulfonyl fluoride). Subsequently, 2 ml of the suspension was transferred into a 2-ml reaction tube and spun for

10 s at 1000 *g*. A 50- $\mu\text{l}$  aliquot, referred to as total extract, was taken from the supernatant and put into 1 ml ethanol. The remaining material was centrifuged for 20 min at 20 000 *g* at 4°C. The floating fat pad was mechanically picked with a spatula and washed three times in 1.7 ml fresh grinding buffer by centrifuging for 20 min at 20 000 *g* at 4°C. The resulting crude fat pad was resuspended in 1 ml ethanol and referred to as the LD-enriched fraction. Total extract and the LD-enriched fraction were stored at –20°C for at least 1 day to enhance protein precipitation.

Total extracts of purple nutsedge were obtained from deep frozen tubers ground in a pre-cooled mortar with 1–2 g sea sand and 5 ml grinding buffer (50 mM Tris, pH 7.4, 10 mM KCl, 6 M urea, 200  $\mu\text{M}$  phenylmethylsulfonyl fluoride). Afterwards, 1.5 ml of the suspension was transferred to a 2-ml reaction cup and solid sodium dodecyl sulfate (SDS) was added to a final concentration of 5% (w/v). Subsequently, the sample was incubated for 1 h at 4°C under constant agitation, followed by centrifugation for 15 min at 19 000 *g* at 4°C. A 200- $\mu\text{l}$  fraction of the supernatant was recovered. For protein precipitation, 50  $\mu\text{l}$  50% (v/v) trichloroacetic acid was added and the samples were incubated for 30 min on ice. Then, 1 ml acetone was added to the suspension and the fraction was stored for at least 24 h at –20°C for protein precipitation.

### Peptide sample preparation

The protein pellets were dissolved in 6 M urea and 5% (w/v) SDS. Protein concentrations were determined with a Pierce BCA protein assay kit (Thermo Fisher Scientific, Waltham, MA, USA). Next, 10  $\mu\text{g}$  of protein for yellow nutsedge or 40  $\mu\text{g}$  of protein for purple nutsedge was run on an SDS-PAGE gel until they entered the separation gel. A single gel piece per sample containing all proteins was excised, tryptically digested, and derivatized (Shevchenko et al., 2006). Peptides were desalted over Empore™ Octadecyl C18 47 mm extraction disks 2215 (Supelco, St. Paul, MN, USA) as previously described (Rappsilber et al., 2007).

### LC-MS/MS analysis

Dried peptide samples were reconstituted in 20  $\mu\text{l}$  LC-MS sample buffer (2% acetonitrile, 0.1% formic acid). Next, 3  $\mu\text{l}$  of each sample was subjected to reverse phase LC for peptide separation using an RSLCnano Ultimate 3000 system (Thermo Fisher Scientific). Peptides were loaded on an Acclaim PepMap 100 pre-column (100  $\mu\text{m} \times 2 \text{ cm}$ , C18, 5  $\mu\text{m}$ , 100 Å; Thermo Fisher Scientific) with 0.07% trifluoroacetic acid at a flow rate of 20  $\mu\text{l min}^{-1}$  for 3 min. Analytical separation of peptides was done on an Acclaim PepMap RSLC column (75  $\mu\text{m} \times 50 \text{ cm}$ , C18, 2  $\mu\text{m}$ , 100 Å; Thermo Fisher Scientific) at a flow rate of 300  $\text{nl min}^{-1}$ . The solvent composition was gradually changed within 94 min from 96% solvent A (0.1% formic acid) and 4% solvent B (80% acetonitrile, 0.1% formic acid) to 10% solvent B within 2 min, to 30% solvent B within the next 58 min, to 45% solvent B within the following 22 min, and to 90% solvent B within the last 12 min of the gradient. All solvents and acids for LC-MS were Optima grade (Thermo Fisher Scientific). Eluting peptides were on-line ionized by nano-electrospray (nESI) using a Nanospray Flex Ion Source (Thermo Fisher Scientific) at 1.5 kV (liquid junction) and transferred into a Q Exactive HF mass spectrometer (Thermo Fisher Scientific). Full scans in a mass range of 300–1650 *m/z* were recorded at a resolution of 30 000 followed by data-dependent top 10 HCD fragmentation at a resolution of 15 000 (dynamic exclusion enabled). LC-MS method programming and data acquisition were performed with XCalibur 4.0 software (Thermo Fisher Scientific).

### Generation of a protein library of yellow nutsedge for peptide identification

Total RNA was extracted from frozen *C. esculentus* mixed tissue samples and eight different stages of tuber development by homogenizing material in Plant RNA Reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). RNA integrity and concentration were determined using the Experion RNA StdSens analysis kit (BioRad, Hercules, CA, USA). Total RNA was treated with DNase (TurboDNase; Ambion, Carlsbad, CA, USA) before sequencing. Libraries for RNA sequencing (RNA-seq) were prepared at BGI (Shenzhen, China). Libraries were sequenced using an Illumina HiSeq 2000 sequencing platform as unpaired-end reads. The sequenced libraries yielded a total of 156 853 170 50-bp RNA-seq reads which were processed using CLC Genomics Workbench 7.0.4 removing duplicated reads. This resulted in 76 819 654 remaining reads, which were used for assembly into contigs with a minimum length of 200 bp using default settings. In total, 63 620 119 short-read sequences were mapped in the assembly, resulting in 38 909 contigs with an average contig length of 636 bp. Two different software programs were used to predict open reading frames (ORFs) or proteins potentially encoded by the produced contigs (putative transcripts). TransDecoder (<https://github.com/TransDecoder>) Release v5.0.1 (Haas et al., 2013) was applied using default settings. GeneMark.hmm eukaryotic (<http://exon.gatech.edu/GeneMark/gmhmm.cgi>) was applied (Lomsadze et al., 2005) using Arabidopsis as species model and with protein sequence as selected output for processing of the contigs. An additional protein reference was based on a previously published yellow nutsedge transcriptome of 99 558 assembled transcripts with an average length of 787 bp (Yang et al., 2016). ORFs were extracted and translated using Geneious 8.1.8 (Biomatters Ltd., Auckland, New Zealand). The deduced protein libraries are available on ProteomeXchange/PRIDE (Vizcaino et al., 2014) under identifier PXD021894.

### Generation of a protein library of purple nutsedge for peptide identification

Total RNA was extracted from three to five tubers of purple nutsedge at five developmental stages in triplicate using Purelink® Plant RNA Reagent (Thermo Fisher Scientific) following the manufacturer's instructions. The extracted RNA was treated with DNase using the TURBO DNA-free™ Kit (Thermo Fisher Scientific) and analyzed with an Agilent 2100 Bioanalyzer to confirm RNA quality and integrity. Strand-specific paired-end libraries were generated and sequenced using an Illumina HiSeq (NovaSeq 6000 S2 PE150 XP) platform at Eurofins Genomics Europe Sequencing GmbH (Germany).

The RNA-seq reads from all 15 samples were filtered for quality, base contents, and adapters and subsequently trimmed for polyG/polyA using fastp (v0.20.1). The resulting 1023 million reads were assembled into a *de novo* transcriptome assembly for purple nutsedge using TRINITY (v2.11.0). All filtered reads were mapped back to the transcriptome assembly and the back-mapping rate was calculated using Bowtie2 (v2.4.2) and RSEM (v1.2.28). The completeness of the transcriptome was assessed using BUSCO (v5.0.0) with the embryophyta\_odb10 (2019/11/27) dataset in transcriptome mode.

The transcriptome assembly was annotated using the Trino-tate pipeline by the following steps.

First, the longest ORFs in transcriptome assembly were identified using Transdecoder (v5.5.0). The resulting candidate coding

sequences were further refined based on sequence and domain homology to known protein sequences in the Swiss-Prot (access date: 22-02-21) and PFAM (v34.0) databases, using NCBI BLAST+ (v2.11.0) and HMMER (v3.3.2), respectively.

In addition, RNA-seq data of purple nutsedge tubers (Ji et al., 2021) were downloaded from NCBI (SRA) and coding sequences were predicted using TransDecoder v5.5.0 with default settings.

The combined protein libraries are available on ProteomeXchange/PRIDE (Vizcaino et al., 2014) under identifier PXD031123.

### Calculation of riBAQ values

MS data were processed with MaxQuant software version 1.6.2.10 or 2.0.3.1 (Cox et al., 2014; Cox & Mann, 2008) with standard settings except that: 'Match between runs' was turned on; 'iBAQ' was selected for label-free quantification; and FTMS recalibration was turned on. The three abovementioned nutsedge protein libraries were used. The libraries, the metadata file, raw data files, MaxQuant search files, and ProteinGroup and Peptide search results created by MaxQuant are available on ProteomeXchange/PRIDE (Vizcaino et al., 2014) under identifiers PXD021894, PXD031123, and PXD035931. iBAQ values were determined with MaxQuant 1.6.2.10. The data were further processed with Perseus software version 1.6.2.2 (Tyanova et al., 2016). Reverse hits, contaminants, and proteins only identified by a modified peptide were removed from the data matrix. Then, all iBAQ values were divided by the total iBAQ value per sample and multiplied by 1000, yielding riBAQ values in per mille. These values were the basis for all further data analysis.

### Analysis of yam and potato data

Proteomic data of yam (Sharma & Deswal, 2021) and potato tubers (Lebecka et al., 2019) were analyzed using the respective protein databases (Cormier et al., 2019; Pham et al., 2020) by MaxQuant and Perseus software as outlined above.

### BLAST

Homologs of nutsedge, yam, and potato proteins in *A. thaliana* were detected by using a BLASTp approach (Altschul et al., 1990; Gish & States, 1993). For this, protein data predicted from nutsedge transcriptomes were used as queries against the *A. thaliana* TAIR10 primary transcript protein release using BLAST 2.5.0+ or 2.11.0+.

### PCA

PCA of the nutsedge data was conducted on riBAQ values. Only proteins were considered that were identified by a least two peptides and that were found in at least three replicates of one of the stages. For the PCA comparing yellow nutsedge and Arabidopsis data, only proteins were considered that had a homolog in both species and were found in at least four out of five replicates of at least one stage. The value for each protein was normalized within the individual species by setting the average of the highest stage to 1. PCA plots were created with Perseus 1.6.2.2 (Tyanova et al., 2016) using standard settings (Category enrichment in components was turned off).

### Hierarchical clustering

The clustering was performed on average values for each stage/tissue. Furthermore, the highest average of each protein within each species was set to 1. Only proteins were considered that were

found in all samples of at least one stage. In the comparison of yellow nutsedge and *Arabidopsis* only proteins were incorporated that had a homolog in both species. Plots were created with Perseus 1.6.2.2 (Tyanova et al., 2016) using Euclidian distance, average linkage, and no constraints. The data were pre-processed with k-means. The number of clusters was set to 300, the maximum number of iterations was set to 100, and the number of restarts was set to 100. In the end, the proteins were grouped in 30 clusters.

### Analysis of gene expression

RNA-seq data of purple and yellow nutsedge tubers (Ji et al., 2021) were downloaded from NCBI (SRA). Transcriptomes were assembled *de novo* using Trinity v2.11.0. Gene expression was estimated with RSEM after read mapping with BOWTIE2, followed by cross-sample normalization (trimmed mean of M [TMM]) and retaining the most abundant isoform per gene, following established protocols (Haas et al., 2013). Homologs for the 10 most abundant proteins identified by proteomics were searched on the two transcriptome assemblies (containing only the most highly expressed isoforms) using tBLASTn and an e-value threshold of  $1e^{-6}$ . Homology was confirmed upon multiple sequence alignment (mafft; Katoh & Standley, 2013) and when necessary confirmed by maximum likelihood phylogenetic inference (IQ-TREE; Nguyen et al., 2015). For each species, TMM expression values were added for protein fragments, isoforms, and close homologs, averaged across time points and replicates, and summarized on a heatmap after  $\log_2$  transformation.

### Pearson correlation

Pearson correlation was performed on average values for each stage/tissue. Furthermore, the highest average of each protein within each species was set to 1. Only proteins were considered that were found in all samples of at least one stage and that had a homolog in both species. The correlation coefficients were calculated with Perseus 1.6.2.2 (Tyanova et al., 2016).

### GO term enrichment

GO term enrichment analysis of proteins contained in certain clusters was conducted using an online tool based on the PANTHER classification system (<http://geneontology.org/>; <http://pantherdb.org/webservices/go/overrep.jsp>; Mi et al., 2019).

### Quantitative GO term analysis

The quantitative GO term analysis was based on riBAQ values. First, nutsedge, yam, or potato proteins were assigned *Arabidopsis* identifiers if the e-value of the BLASTp result was lower than  $10^{-5}$ . If several proteins were assigned the same *Arabidopsis* identifier, the values were added. Then, all proteins were assigned one or more GO terms and the values were added for each term.

### Volcano plots

Only proteins were considered that were found in at least four out of five replicates of at least one stage/subcellular fraction. Missing values were imputed with Perseus 1.6.2.2 (Tyanova et al., 2016) using a width of 0.3 and a downshift of 1.8 separately for each column. The volcano plot was then created using a *t*-test, a number of 250 randomizations, a false discovery rate of 0.01, and an  $S_0$  value of 2.

### Analysis of LD proteins

All proteins that were homologous to a known LD protein from *Arabidopsis* were considered as LD proteins. The abundance of

each LD protein was divided by the total abundance of all LD proteins in the sample.

### Phytohormone measurements

Phytohormones were extracted with methyl-tert-butyl ether, reversed phase-separated using an ACQUITY UPLC system (Waters Corp., Milford, MA, USA), and analyzed by nano-electrospray ionization (nanoESI) (TriVersa Nanomate; Advion Biosciences, Ithaca, NY, USA) coupled with an AB Sciex 4000 QTRAP tandem mass spectrometer (AB Sciex) employed in scheduled multiple reaction monitoring modes (Herrfurth & Feussner, 2020) with the following modifications. For quantification, 10 ng  $D_4$ -SA, 10 ng  $D_6$ -ABA, 10 ng  $D_5$ -jasmonic acid (JA; all three from C/D/N Isotopes Inc., Pointe-Claire, Canada), and 20 ng  $D_5$ -IAA (Eurisotop, Freising, Germany) were added at the beginning of the extraction procedure. For ABA, indole-3-acetic acid (IAA), indole carboxylic acid (ICA), and SA analysis, the following mass transitions were included: 137/93 (declustering potential [DP]  $-25$  V, entrance potential [EP]  $-6$  V, collision energy [CE]  $-20$  V) for SA, 141/97 (DP  $-25$  V, EP  $-6$  V, CE  $-22$  V) for  $D_4$ -SA, 160/116 (DP  $-40$  V, EP  $-6.5$  V, CE  $-22$  V) for ICA, 174/130 (DP  $-35$  V, EP  $-9$  V, CE  $-14$  V) for IAA, 179/135 (DP  $-35$  V, EP  $-9$  V, CE  $-14$  V) for  $D_5$ -IAA, 263/153 (DP  $-35$  V, EP  $-4$  V, CE  $-14$  V) for ABA, and 269/159 (DP  $-30$  V, EP  $-5$  V, CE  $-16$  V) for  $D_6$ -ABA. JA-Ile was quantified with  $D_5$ -JA as internal standard.

### Phylogenetic analyses

In order to construct phylogenies, the datasets of the proteins of interest were supplemented with additional homologs from other seed plants and, in case of LDAPs, land plants and green algae. For this, well-described *A. thaliana* proteins were used as query sequences in a BLASTp search against protein data from genome releases. In each case, the detected protein homologs in the yellow nutsedge transcriptomes were added.

For the phylogenetic analysis of caleosins and steroleosins, homologs were mined from genome data of *Theobroma cacao* (Argout et al., 2011), *Picea abies* (Nystedt et al., 2013), *Oryza sativa* (Ouyang et al., 2007), *N. tabacum* (Sierro et al., 2014), *Gnetum montanum* (Wan et al., 2018), *Gossypium hirsutum* (Li et al., 2015), *Capsella grandiflora* (Slotte et al., 2013), *Brassica rapa* (Wang et al., 2011), *Brachypodium distachyon* (The International Brachypodium Initiative, 2010), *Amborella trichopoda* (Amborella Genome, 2013), *A. thaliana* (Lamesch et al., 2012), and *Solanum lycopersicum* (The Tomato Genome Consortium, 2012).

For the phylogenetic analysis of oleosins, homologs were mined from genome data of *O. sativa* (Ouyang et al., 2007), *A. thaliana* (Lamesch et al., 2012), and *S. lycopersicum* (The Tomato Genome Consortium, 2012).

For the phylogenetic analysis of LDAPs, the same selection of homologs as in de Vries and Ischebeck (2020) was chosen, i.e., *Theobroma cacao* (Argout et al., 2011), *Triticum aestivum* (The International Wheat Genome Sequencing Consortium, 2018), *Selaginella moellendorffii* (Banks et al., 2011), *S. lycopersicum* (The Tomato Genome Consortium, 2012), *Sphagnum fallax* (v.0.5, DOE-JGI, <http://phytozome.jgi.doe.gov/>), *P. patens* (Lang et al., 2018), *P. abies* (Nystedt et al., 2013), *O. sativa* (Ouyang et al., 2007), *N. tabacum* (Sierro et al., 2014), *Marchantia polymorpha* (Bowman et al., 2017), *Gnetum montanum* (Wan et al., 2018), *G. hirsutum* (Li et al., 2015), *Carica papaya* (Ming et al., 2008), *Capsella grandiflora* (Slotte et al., 2013), *B. rapa* (Wang et al., 2011), *Brassica oleracea* (Liu et al., 2014), *B. distachyon* (The International Brachypodium Initiative, 2010), *A. trichopoda* (Amborella Genome, 2013), *A. thaliana* (Lamesch et al., 2012), *A.*



*lyrata* (Hu et al., 2011), the hornworts *Anthoceros agrestis* and *Anthoceros punctatus* (Li et al., 2020), the ferns *Azolla filiculoides* and *Salvinia cucullata* (Li et al., 2018), and the Charophyceae *Chara braunii* (Nishiyama et al., 2018), the genomes of the Zygnematophyceae *Mesotaenium endlicherianum* and *Spirogloea muscicola* (Cheng et al., 2019), and the transcriptomes of the filamentous Zygnematophyceae *Zygnema circumcarinatum* SAG2419 (Rippin et al., 2017) and SAG698-1a (de Vries et al., 2018), as well as *Spirogyra pratensis* MZCH10213 and *Mougeotia* sp. MZCH240 (de Vries et al., 2020) and genomes of the early-diverging streptophyte algae *Mesostigma viride* and *Chlorokybus atmophyticus* (Wang et al., 2020).

All protein sequences were aligned using MAFFT v7.453 L-INS-I (Katoh & Standley, 2013). Sequences were cropped to retain the conserved region/protein features and re-aligned; sequences that did not cover at least 50% of the conserved region were removed. Maximum likelihood phylogenies were computed using IQ-TREE (Nguyen et al., 2015) multicore version 1.5.5 for Linux 64-bit. In total 500 bootstrap replicates were computed. Each run of IQ-TREE entailed determining the best model for protein evolution according to the Bayesian information criterion (BIC) score via ModelFinder (Kalyaanamoorthy et al., 2017). The best models were LG+G4 for CLO, JTT+G4 for OLE and HSD, and JTT+G4 for LDAP.

### Estimation of divergence times

The dataset of Escudero and Hipp (2013) was enriched with homologous sequences for yellow and purple nutsedge (identified via BLASTn against the new transcriptome assemblies). This dataset consists of three plastid markers (*rbcl*, *trnL* intron, and *trnL-trnF* spacer) for a large diversity of Cyperaceae and outgroups. We inferred a partitioned maximum likelihood phylogeny using IQ-TREE (Nguyen et al., 2015) v1.6.12 with BIC-selected best-fit models and partitions. Divergence times were estimated using a Bayesian molecular clock using MCMCTREE (PAML v.4.9; Yang, 2007) with the seven calibrations as in Escudero and Hipp under an uncorrelated molecular clock model (CorrTest  $P > 0.5$ ). Mimicking Escudero and Hipp, a bell-shaped gamma distribution with mean at 88 million years ago was used for the secondary calibration of the Cyperaceae–Juncaeae split ('G(100,113.6)') and truncated Cauchy distributions were used for the following six fossil calibrations: *Carex* and *Scleria* ('L(0.372,0.1,0.1,0.025)'), *Scirpus* ('L(0.284,0.1,0.1,0.025)'), *Cladium* and *Fimbristylis* ('L(0.257, 0.1,0.1,0.025)'), and *Juncus* ('L(0.339,0.1,0.1,0.025)'). Calculations used approximate likelihood calculations under a HKY85+? model and the time unit was set to 100 million years ago. The prior for the mean rate ('rgene\_gamma') was set as a diffuse gamma prior with the mean approximated reflecting the average root-to-tip paths in the inferred maximum likelihood tree ('G(2,5.68)') and the prior on the rate drift parameter was set to reflect severe violation of constant rates ('G(2,2)'). A uniform birth–death process was assumed for the tree prior. Two independent Markov chain Monte Carlo chains were run for 20 000 cycles (sampled every 10) after a burnin period of 2000. Convergence was assessed *a posteriori* with Tracer v1.7.1 and all parameters had ESS values of  $>100$ .

### AUTHOR CONTRIBUTIONS

PWN, IF, GHB, JdV, PH, and TI designed the work; PWN, CH, KS, OV, JdV, SS, ASC, PH, and TI performed the research; PWN, II, KS, GHB, CH, IF, JdV, ASC, SS, PH, and TI analyzed the data; and PWN, II, KS, JdV, PH, and TI wrote the manuscript. All authors critically read and revised the manuscript and approved the final version.

### ACKNOWLEDGMENTS

PWN, GB, JdV, and TI thank the German research foundation (DFG, Grants IS 273/7-1, IRTG 2172 PRoTECT, BR1502-15-1, SPP 2237 MAdLand, VR 132/4-1). JdV and II thank the European Research Council for funding (grant agreement no. 852725; ERC-StG 'TerreStrIAL') under the European Union's Horizon 2020 research and innovation program. OV and KS on behalf of the Service Unit LCMS Protein Analytics of the Göttingen Center for Molecular Biosciences (GZMB) thank the DFG for funding (INST 186/1230-1 FUGG to Stefanie Pöggeler). PH and SS were financed by grants from the Swedish Foundation for Strategic Research (SSF) and Trees and Crops for the Future (TC4F), a Strategic Research Area at SLU, supported by the Swedish Government. We are grateful for pre-publication access to the genome data of *Sphagnum fallax* (v0.5, DOE-JGI, <http://phytozome.jgi.doe.gov/>). Open Access funding enabled and organized by Projekt DEAL.

### CONFLICT OF INTEREST

The authors declare no conflict of interest.

### DATA AVAILABILITY STATEMENT

All relevant data can be found within the manuscript and its supporting materials. Proteomic raw data can be found in the PRIDE database (Vizcaino et al., 2014) under identifiers PXD021894, PXD031123, and PXD035931 (<https://www.ebi.ac.uk/pride/>).

### SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

**Figure S1.** Purple and yellow nutsedge developmental stages investigated in this study.

**Figure S2.** PCA plot of yellow and purple nutsedge.

**Figure S3.** GO term comparison – tubers of different species.

**Figure S4.** LD-enriched fractions could be reproducibly isolated.

**Figure S5.** Oleosin (OLE) and steroleosin (HSD) phylogeny.

**Figure S6.** Phylogeny of lipid droplet-associated proteins (LDAPs).

**Figure S7.** Expression levels of transcription factors.

**Figure S8.** Time-calibrated phylogeny of the Cyperaceae based on a Bayesian relaxed molecular clock analysis.

**Data S1** Normalized iBAQ values of yellow nutsedge.

**Data S2** Results of a BLASTp query of the yellow nutsedge proteins against the *Arabidopsis thaliana* TAIR10 primary transcript protein release.

**Data S3** Result of hierarchical clustering of the yellow nutsedge total proteome.

**Data S4** List of tuber-enriched proteins.

**Data S5** The 10 most abundant tuber-specific proteins and expression and protein abundance levels of homologs.

**Data S6** Normalized iBAQ values of purple nutsedge.

**Data S7** Results of a BLASTp query of the purple nutsedge data against the *Arabidopsis thaliana* TAIR10 primary transcript protein release and against the compiled yellow nutsedge database.

**Data S8** Result of hierarchical clustering of the yellow and purple nutsedge proteomes and the *Arabidopsis* proteome.

**Data S9** GO term analysis – individual proteins of yellow nutsedge.

**Data S10** GO term analysis – individual proteins of purple nutsedge.

**Data S11** GO term analysis – comparison to Arabidopsis seedling establishment.

**Data S12** Hormone analysis.

**Data S13** Analysis of dried purple nutsedge tubers.

**Data S14** Analysis of potato data.

**Data S15** Analysis of yam data.

**Data S16** Analysis of homologs to known LD proteins and further LD-enriched proteins.

**Data S17** Multiple sequence alignment of Cyperaceae used for phylogenetic and relaxed molecular clock analyses.

**Data S18** Time-calibrated phylogeny of the Cyperaceae based on a Bayesian relaxed molecular clock analysis (Newick format).

## REFERENCES

- Alejo-Jacuinde, G., Gonzalez-Morales, S.I., Oropeza-Aburto, A., Simpson, J. & Herrera-Estrella, L. (2020) Comparative transcriptome analysis suggests convergent evolution of desiccation tolerance in *Selaginella* species. *BMC Plant Biology*, **20**, 468.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.
- Amborella Genome, P. (2013) The Amborella genome and the evolution of flowering plants. *Science*, **342**, 1241089.
- Argout, X., Salse, J., Aury, J.M., Guittinan, M.J., Droc, G., Gouzy, J. et al. (2011) The genome of *Theobroma cacao*. *Nature Genetics*, **43**, 101–108.
- Aubert, Y., Vile, D., Pervent, M., Aldon, D., Ranty, B., Simonneau, T. et al. (2010) RD20, a stress-inducible caleosin, participates in stomatal control, transpiration and drought tolerance in *Arabidopsis thaliana*. *Plant and Cell Physiology*, **51**, 1975–1987.
- Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., DePamphilis, C. et al. (2011) The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science*, **332**, 960–963.
- Barajas-Lopez, J.D., Tiwari, A., Zarza, X., Shaw, M.W., Pascual, J.S., Punkkinen, M. et al. (2021) EARLY RESPONSE TO DEHYDRATION 7 remodels cell membrane lipid composition during cold stress in *Arabidopsis*. *Plant & Cell Physiology*, **62**, 80–91.
- Baron, K.N., Schroeder, D.F. & Stasolla, C. (2014) GEM-related 5 (GER5), an ABA and stress-responsive GRAM domain protein regulating seed development and inflorescence architecture. *Plant Science*, **223**, 153–166.
- Bartels, D. (2005) Desiccation tolerance studied in the resurrection plant *Croton stigmata plantagineum*. *Integrative and Comparative Biology*, **45**, 696–701.
- Battaglia, M., Olvera-Carrillo, Y., Garcarrubio, A., Campos, F. & Covarrubias, A.A. (2008) The enigmatic LEA proteins and other hydrophilins. *Plant Physiology*, **148**, 6–24.
- Baud, S., Dichow, N.R., Kelemen, Z., d'Andrea, S., To, A., Berger, N. et al. (2009) Regulation of HSD1 in seeds of *Arabidopsis thaliana*. *Plant & Cell Physiology*, **50**, 1463–1478.
- Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K. et al. (2017) Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell*, **171**, 287–304.
- Brocard, L., Immel, F., Coulon, D., Esnay, N., Tophile, K., Pascal, S. et al. (2017) Proteomic analysis of lipid droplets from *Arabidopsis* aging leaves brings new insight into their biogenesis and functions. *Frontiers in Plant Science*, **8**, 894.
- Chapman, K.D., Dyer, J.M. & Mullen, R.T. (2012) Biogenesis and functions of lipid droplets in plants. *Journal of Lipid Research*, **53**, 215–226.
- Chen, H., Ruan, J., Chu, P., Fu, W., Liang, Z., Li, Y. et al. (2020) AtPER1 enhances primary seed dormancy and reduces seed germination by suppressing the ABA catabolism and GA biosynthesis in *Arabidopsis* seeds. *The Plant Journal*, **101**, 310–323.
- Chen, J.C.F., Tsai, C.C.Y. & Tzen, J.T.C. (1999) Cloning and secondary structure analysis of caleosin, a unique calcium-binding protein in oil bodies of plant seeds. *Plant and Cell Physiology*, **40**, 1079–1086.
- Chen, K., Li, G.J., Bressan, R.A., Song, C.P., Zhu, J.K. & Zhao, Y. (2020) Abscisic acid dynamics, signaling, and functions in plants. *Journal of Integrative Plant Biology*, **62**, 25–54.
- Cheng, M.C., Hsieh, E.J., Chen, J.H., Chen, H.Y. & Lin, T.P. (2012) Arabidopsis RGLG2, functioning as a RING E3 ligase, interacts with AtERF53 and negatively regulates the plant drought stress response. *Plant Physiology*, **158**, 363–375.
- Cheng, S., Xian, W., Fu, Y., Marin, B., Keller, J., Wu, T. et al. (2019) Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell*, **179**, 1057–1067.
- Cormier, F., Lawac, F., Maledon, E., Gravillon, M.C., Nudol, E., Mournet, P. et al. (2019) A reference high-density genetic map of greater yam (*Dioscorea alata* L.). *Theoretical and Applied Genetics*, **132**, 1733–1744.
- Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N. & Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular Proteomics*, **13**, 2513–2526.
- Cox, J. & Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, **26**, 1367–1372.
- Das Gupta, M. & Tsiantis, M. (2018) Gene networks and the evolution of plant morphology. *Current Opinion in Plant Biology*, **45**, 82–87.
- de Vries, J., Curtis, B.A., Gould, S.B. & Archibald, J.M. (2018) Embryophyte stress signaling evolved in the algal progenitors of land plants. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, E3471–E3480.
- de Vries, J., de Vries, S., Curtis, B.A., Zhou, H., Penny, S., Feussner, K. et al. (2020) Heat stress response in the closest algal relatives of land plants reveals conserved stress signalling circuits. *Plant Journal*, **103**, 1025–1048.
- de Vries, J. & Ischebeck, T. (2020) Ties between stress and lipid droplets pre-date seeds. *Trends in Plant Science*, **25**, 1203–1214.
- Defelice, M.S. (2002) Yellow nutsedge *Cyperus esculentus* L.—snack food of the gods. *Weed Technology*, **16**, 901–907.
- Deruyffelaere, C., Bouchez, I., Morin, H., Guillot, A., Miquel, M., Froissard, M. et al. (2015) Ubiquitin-mediated proteasomal degradation of oleosins is involved in oil body mobilization during post-germinative seedling growth in *Arabidopsis*. *Plant and Cell Physiology*, **56**, 1374–1387.
- Deruyffelaere, C., Purkrtova, Z., Bouchez, I., Collet, B., Cacas, J.L., Chardot, T. et al. (2018) PUX10 associates with CDC48A and regulates the dislocation of ubiquitinated oleosins from seed lipid droplets. *Plant Cell*, **30**, 2116–2136.
- Doner, N.M., Seay, D., Mehling, M., Sun, S., Gidda, S.K., Schmitt, K. et al. (2021) *Arabidopsis thaliana* EARLY RESPONSIVE TO DEHYDRATION 7 localizes to lipid droplets via its senescence domain. *Frontiers in Plant Science*, **12**, 658961.
- Escudero, M. & Hipp, A. (2013) Shifts in diversification rates and clade ages explain species richness in higher-level sedge taxa (Cyperaceae). *American Journal of Botany*, **100**, 2403–2411.
- Fernandez-Pozo, N., Haas, F.B., Meyberg, R., Ullrich, K.K., Hiss, M., Perroud, P.F. et al. (2019) PEATmoss (Physcomitrella expression atlas tool): a unified gene expression atlas for the model plant *Physcomitrella patens*. *The Plant Journal*, **102**, 165–177.
- Fernández-Santos, R., Izquierdo, Y., López, A., Muñoz, L., Martínez, M., Cascón, T. et al. (2020) Protein profiles of lipid droplets during the hypersensitive defense response of *Arabidopsis* against *Pseudomonas* infection. *Plant & Cell Physiology*, **61**, 1144–1157.
- Fürst-Jansen, J.M., de Vries, S. & de Vries, J. (2020) Evo-physio: on stress responses and the earliest land plants. *Journal of Experimental Botany*, **71**, 3254–3269.
- Gao, B., Li, X., Zhang, D., Liang, Y., Yang, H., Chen, M. et al. (2017) Desiccation tolerance in bryophytes: the dehydration and rehydration transcriptomes in the desiccation-tolerant bryophyte *Bryum argenteum*. *Scientific Reports*, **7**, 7571.
- Gidda, S.K., Park, S., Pyc, M., Yurchenko, O., Cai, Y., Wu, P. et al. (2016) Lipid droplet-associated proteins (LDAPs) are required for the dynamic regulation of neutral lipid compartmentation in plant cells. *Plant Physiology*, **170**, 2052–2071.
- Gish, W. & States, D.J. (1993) Identification of protein coding regions by database similarity search. *Nature Genetics*, **3**, 266–272.

- Guo, L.-M., Li, J., He, J., Liu, H. & Zhang, H.-M. (2020) A class I cytosolic HSP20 of rice enhances heat and salt tolerance in different organisms. *Scientific Reports*, **10**, 1383.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J. *et al.* (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, **8**, 1494–1512.
- Han, C., Yin, X., He, D. & Yang, P. (2013) Analysis of proteome profile in germinating soybean seed, and its comparison with rice showing the styles of reserves mobilization in different crops. *PLoS One*, **8**, e56947.
- Hanin, M., Brini, F., Ebel, C., Toda, Y., Takeda, S. & Masmoudi, K. (2011) Plant dehydrins and stress tolerance: versatile proteins for complex mechanisms. *Plant Signaling & Behavior*, **6**, 1503–1509.
- Harrison, C.J. & Morris, J.L. (2018) The origin and early evolution of vascular plant shoots and leaves. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **373**, 20160496.
- Herrfurth, C. & Feussner, I. (2020) Quantitative Jasmonate profiling using a high-throughput UPLC-NanoESI-MS/MS method. *Methods in Molecular Biology*, **2085**, 169–187.
- Hofvander, P., Ischebeck, T., Turesson, H., Kushwaha, S.K., Feussner, I., Carlsson, A.S. *et al.* (2016) Potato tuber expression of Arabidopsis WRINKLED1 increase triacylglycerol and membrane lipids while affecting central carbohydrate metabolism. *Plant Biotechnology Journal*, **14**, 1883–1898.
- Horn, P.J., James, C.N., Gidda, S.K., Kilaru, A., Dyer, J.M., Mullen, R.T. *et al.* (2013) Identification of a new class of lipid droplet-associated proteins in plants. *Plant Physiology*, **162**, 1926–1936.
- Hsiao, E.S.L. & Tzen, J.T.C. (2011) Ubiquitination of oleosin-H and caleosin in sesame oil bodies after seed germination. *Plant Physiology and Biochemistry*, **49**, 77–81.
- Hu, T.T., Pattyn, P., Bakker, E.G., Cao, J., Cheng, J.F., Clark, R.M. *et al.* (2011) The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nature Genetics*, **43**, 476–481.
- Huang, A.H.C. (2018) Plant lipid droplets and their associated proteins: potential for rapid advances. *Plant Physiology*, **176**, 1894–1918.
- Huang, C.-Y., Chung, C.-I., Lin, Y.-C., Hsing, Y.-I.C. & Huang, A.H. (2009) Oil bodies and oleosins in *Physcomitrella* possess characteristics representative of early trends in evolution. *Plant Physiology*, **150**, 1192–1203.
- Iqbal, J., Hussain, S., Ali, A. & Javaid, A. (2012) Biology and management of purple nutsedge (*Cyperus rotundus* L.). *The Journal of Animal & Plant Sciences*, **22**, 384–389.
- Ischebeck, T. (2016) Lipids in pollen — they are different. *Biochimica et Biophysica Acta*, **1861**, 1315–1328.
- Ischebeck, T., Krawczyk, H.E., Mullen, R.T., Dyer, J.M. & Chapman, K.D. (2020) Lipid droplets in plants and algae: distribution, formation, turnover and function. *Seminars in Cell & Developmental Biology*, **108**, 82–93.
- Ischebeck, T., Valledor, L., Lyon, D., Gingl, S., Nagler, M., Meijon, M. *et al.* (2014) Comprehensive cell-specific protein analysis in early and late pollen development from diploid microsporocytes to pollen tube growth. *Molecular & Cellular Proteomics*, **13**, 295–310.
- Jayson, C.B.K., Arlt, H., Fischer, A.W., Lai, Z.W., Farese, R.V., Jr. & Walther, T.C. (2018) Rab18 is not necessary for lipid droplet biogenesis or turnover in human mammary carcinoma cells. *Molecular Biology of the Cell*, **29**, 2045–2054.
- Ji, H., Liu, D. & Yang, Z. (2021) High oil accumulation in tuber of yellow nutsedge compared to purple nutsedge is associated with more abundant expression of genes involved in fatty acid synthesis and triacylglycerol storage. *Biotechnology for Biofuels*, **14**, 54.
- Jiang, P.-L., Jauh, G.-Y., Wang, C.-S. & Tzen, J.T.C. (2008) A unique caleosin in oil bodies of lily pollen. *Plant & Cell Physiology*, **49**, 1390–1395.
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. & Jermini, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**, 587–589.
- Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.
- 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. *The Plant Journal*, **88**, 1058–1070.
- Krawczyk, H.E., Sun, S., Doner, N.M., Yan, Q., Lim, M.S.S., Scholz, P. *et al.* (2022) SEED LIPID DROPLET PROTEIN1, SEED LIPID DROPLET PROTEIN2 and LIPID DROPLET PLASMA MEMBRANE ADAPTOR mediate lipid droplet-plasma membrane tethering. *The Plant Cell*, **34**, 2424–2448.
- Kretzschmar, F.K., Doner, N., Krawczyk, H.E., Scholz, P., Schmitt, K., Valerius, O. *et al.* (2020) Identification of low-abundance lipid droplet proteins in seeds and seedlings. *Plant Physiology*, **182**, 1236–1245.
- Kretzschmar, F.K., Mengel, L.F., Müller, A., Schmitt, K., Biersch, K.F., Valerius, O. *et al.* (2018) PUX10 is a lipid droplet-localized scaffold protein that anchors CDC48 and is involved in the degradation of lipid droplet proteins. *The Plant Cell*, **30**, 2137–2160.
- Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R. *et al.* (2012) The Arabidopsis information resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, **40**, D1202–D1210.
- Lang, D., Ullrich, K.K., Murat, F., Fuchs, J., Jenkins, J., Haas, F.B. *et al.* (2018) The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal*, **93**, 515–533.
- Le, T.N. & McQueen-Mason, S.J. (2006) Desiccation-tolerant plants in dry environments. *Reviews in Environmental Science and Bio/Technology*, **5**, 269–279.
- Lebecka, R., Kistowski, M., Dębski, J., Szajko, K., Murawska, Z. & Marczewski, W. (2019) Quantitative proteomic analysis of differentially expressed proteins in tubers of potato plants differing in resistance to *Dickeya solani*. *Plant and Soil*, **441**, 317–329.
- Li, F., Asami, T., Wu, X., Tsang, E.W. & Cutler, A.J. (2007) A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. *Plant Physiology*, **145**, 87–97.
- Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., Kohel, R.J. *et al.* (2015) Genome sequence of cultivated upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nature Biotechnology*, **33**, 524–530.
- Li, F.W., Brouwer, P., Carretero-Paulet, L., Cheng, S., de Vries, J., Delaux, P.M. *et al.* (2018) Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature Plants*, **4**, 460–472.
- Li, F.W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z. *et al.* (2020) Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. *Nat Plants*, **6**, 259–272.
- Linssen, J.P., Cozijnsen, J.L. & Pilnik, W. (1989) Chufa (*Cyperus esculentus*): a new source of dietary fibre. *Journal of the Science of Food and Agriculture*, **49**, 291–296.
- Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I.A. *et al.* (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications*, **5**, 3930.
- Lomsadze, A., Ter-Hovhannisyann, V., Chernoff, Y.O. & Borodovsky, M. (2005) Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research*, **33**, 6494–6506.
- Lu, K.J., van 't Wout Hofland, N., Mor, E., Mutte, S., Abrahams, P., Kato, H. *et al.* (2020) Evolution of vascular plants through redeployment of ancient developmental regulators. *Proceedings of the National Academy of Sciences of the United States of America*, **117**, 733–740.
- Maumus, F. & Quesneville, H. (2014) Ancestral repeats have shaped epigenome and genome composition for millions of years in *Arabidopsis thaliana*. *Nature Communications*, **5**, 4104.
- Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X. *et al.* (2019) Protocol update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nature Protocols*, **14**, 703–721.
- Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J.H. *et al.* (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, **452**, 991–996.
- Neale, A.D., Blomstedt, C.K., Bronson, P., Le, T.N., Guthridge, K., Evans, J. *et al.* (2000) The isolation of genes from the resurrection grass *Sporobolus stapfianus* which are induced during severe drought stress. *Plant, Cell & Environment*, **23**, 265–277.
- Nguyen, L.-T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**, 268–274.
- Nishiyama, T., Sakayama, H., de Vries, J., Buschmann, H., Saint-Marcoux, D., Ullrich, K.K. *et al.* (2018) The Chara genome: secondary complexity and implications for plant terrestrialization. *Cell*, **174**, 448–464.e424.

- Nystedt, B., Street, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.C., Scofield, D.G. et al. (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, **497**, 579–584.
- Oliver, M.J., Farrant, J.M., Hilhorst, H.W.M., Mundree, S., Williams, B. & Bewley, J.D. (2020) Desiccation tolerance: avoiding cellular damage during drying and rehydration. *Annual Review of Plant Biology*, **71**, 435–460.
- Oliver, M.J., Tuba, Z. & Mishler, B.D. (2000) The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology*, **151**, 85–100.
- Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K. et al. (2007) The TIGR Rice Genome annotation resource: improvements and new features. *Nucleic Acids Research*, **35**, D883–D887.
- Pardo, J., Man Wai, C., Chay, H., Madden, C.F., Hilhorst, H.W.M., Farrant, J.M. et al. (2020) Intertwined signatures of desiccation and drought tolerance in grasses. *Proceedings of the National Academy of Sciences*, **117**, 10079–10088.
- Peres, A.L.G., Soares, J.S., Tavares, R.G., Righetto, G., Zullo, M.A., Mandava, N.B. et al. (2019) Brassinosteroids, the sixth class of phytohormones: a molecular view from the discovery to hormonal interactions in plant development and stress adaptation. *International Journal of Molecular Sciences*, **20**, 331.
- Pham, G.M., Hamilton, J.P., Wood, J.C., Burke, J.T., Zhao, H., Vaillancourt, B. et al. (2020) Construction of a chromosome-scale long-read reference genome assembly for potato. *Gigascience*, **9**, 1–11.
- Pyc, M., Cai, Y., Gidda, S.K., Yurchenko, O., Park, S., Kretschmar, F.K. et al. (2017) Arabidopsis lipid drop-associated protein (LDAP) - interacting protein (LDIP) influences lipid droplet size and neutral lipid homeostasis in both leaves and seeds. *The Plant Journal*, **92**, 1182–1201.
- Pyc, M., Gidda, S.K., Seay, D., Esnay, N., Kretschmar, F.K., Cai, Y. et al. (2021) LDIP cooperates with SEIPIN and LDAP to facilitate lipid droplet biogenesis in Arabidopsis. *Plant Cell*, **33**, 3076–3103.
- Rappilber, J., Mann, M. & Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, **2**, 1896–1906.
- Rippin, M., Becker, B. & Holzinger, A. (2017) Enhanced desiccation tolerance in mature cultures of the streptophytic green alga *Zygnema circumcarinatum* revealed by transcriptomics. *Plant and Cell Physiology*, **58**, 2067–2084.
- Scholz, P., Chapman, K.D., Mullen, R.T. & Ischebeck, T. (2022) Finding new friends and revisiting old ones - how plant lipid droplets connect with other subcellular structures. *The New Phytologist*. <https://doi.org/10.1111/nph.18390>
- Sharma, S. & Deswal, R. (2021) Dioscorea Alata tuber proteome analysis uncovers differentially regulated growth-associated pathways of tuber development. *Plant & Cell Physiology*, **62**, 191–204.
- Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, **1**, 2856–2860.
- Shimada, T.L. & Hara-Nishimura, I. (2015) Leaf oil bodies are subcellular factories producing antifungal oxylipins. *Current Opinion in Plant Biology*, **25**, 145–150.
- Shimada, T.L., Hayashi, M. & Hara-Nishimura, I. (2018) Membrane dynamics and multiple functions of oil bodies in seeds and leaves. *Plant Physiology*, **176**, 199–207.
- Shimada, T.L., Shimada, T., Takahashi, H., Fukao, Y. & Hara-Nishimura, I. (2008) A novel role for oleosins in freezing tolerance of oilseeds in *Arabidopsis thaliana*. *The Plant Journal*, **55**, 798–809.
- Sierro, N., Battey, J.N., Ouadi, S., Bakaher, N., Bovet, L., Willig, A. et al. (2014) The tobacco genome sequence and its comparison with those of tomato and potato. *Nature Communications*, **5**, 3833.
- Slotte, T., Hazzouri, K.M., Agren, J.A., Koenig, D., Maumus, F., Guo, Y.L. et al. (2013) The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. *Nature Genetics*, **45**, 831–835.
- Stoller, E.W., Nema, D.P. & Bhan, V.M. (1972) Yellow nutsedge tuber germination and seedling development. *Weed Science*, **20**, 93–97.
- Stoller, E.W. & Sweet, R.D. (1987) Biology and life cycle of purple and yellow nutsedges (*Cyperus rotundus* and *C. esculentus*). *Weed Technology*, **1**, 66–73.
- The International Brachypodium Initiative. (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*, **463**, 763–768.
- The International Wheat Genome Sequencing Consortium. (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, **361**, eaar7191.
- The Tomato Genome Consortium. (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, **485**, 635–641.
- Turesson, H., Marttila, S., Gustavsson, K.-E., Hofvander, P., Olsson, M.E., Bülow, L. et al. (2010) Characterization of oil and starch accumulation in tubers of *Cyperus esculentus* var. sativus (Cyperaceae): a novel model system to study oil reserves in nonseed tissues. *American Journal of Botany*, **97**, 1884–1893.
- Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T. et al. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, **13**, 731–740.
- VanBuren, R., Wai Ching, M., Zhang, Q., Song, X., Edger, P.P., Bryant, D. et al. (2017) Seed desiccation mechanisms co-opted for vegetative desiccation in the resurrection grass *Oropetium thomaeum*. *Plant, Cell & Environment*, **40**, 2292–2306.
- Vanhercke, T., Divi, U.K., El Tahchy, A., Liu, Q., Mitchell, M., Taylor, M.C. et al. (2017) Step changes in leaf oil accumulation via iterative metabolic engineering. *Metabolic Engineering*, **39**, 237–246.
- Vanhercke, T., Dyer, J.M., Mülle, R.T., Kilaru, A., Rahman, M.M., Petrie, J.R. et al. (2019) Metabolic engineering for enhanced oil in biomass. *Progress in Lipid Research*, **74**, 103–129.
- Vizcaino, J.A., Deutsch, E.W., Wang, R., Csordas, A., Reisinger, F., Rios, D. et al. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature Biotechnology*, **32**, 223–226.
- Wan, T., Liu, Z.M., Li, L.F., Leitch, A.R., Leitch, I.J., Lohaus, R. et al. (2018) A genome for gnetophytes and early evolution of seed plants. *Nat Plants*, **4**, 82–89.
- Wang, S., Li, L., Li, H., Sahu, S.K., Wang, H., Xu, Y. et al. (2020) Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. *Nat Plants*, **6**, 95–106.
- Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S. et al. (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, **43**, 1035–1039.
- Wehmeyer, N., Hernandez, L.D., Finkelstein, R.R. & Vierling, E. (1996) Synthesis of small heat-shock proteins is part of the developmental program of late seed maturation. *Plant Physiology*, **112**, 747–757.
- Wills, G.D. (1987) Description of purple and yellow nutsedge (*Cyperus rotundus* and *C. esculentus*). *Weed Technology*, **1**, 2–9.
- Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586–1591.
- Yang, Z., Ji, H. & Liu, D. (2016) Oil biosynthesis in underground oil-rich storage vegetative tissue: comparison of *Cyperus esculentus* tuber with oil seeds and fruits. *Plant & Cell Physiology*, **57**, 2519–2540.

### **3 Manuscript I: Core desiccation response strategies of plant development revealed by the dynamic proteome of *Physcomitrium patens* during spore germination.**

The final peer-reviewed version of this preliminary manuscript was published online on the 7th of December 2023 in The Plant Journal under the following DOI, including supplementary files: [doi.org/10.1111/tpj.16574](https://doi.org/10.1111/tpj.16574).

Author contribution:

Philipp William Niemeyer optimized, planned, and performed protein isolation and processing, except for LC-MS/MS measurements. Additionally, he made the image acquisition of all examined sample materials. He did the raw data processing and initial data analysis. Philipp William Niemeyer wrote the draft of the manuscript, except for parts of the cluster analysis, the phylogeny of LEA proteins in the results, and LOX proteins in the discussion. He was responsible for figures 1a,d, 2, and 5b and did the final edit for figures 1, 3, 4, 5, 8, and supplemental figure 1.

#### **Supplementary materials:**

- Data S1. Results of a BLASTp query of the *P. patens* proteins against the *Arabidopsis thaliana* TAIR10 primary transcript protein release.
- Data S2. Normalized relative iBAQ values of *P. patens* proteins.
- Data S3. Result of hierarchical clustering of *P. patens* proteins.
- Data S4. GO term analysis.
- Data S5. Normalized relative iBAQ values of *P. patens* and Arabidopsis proteins.
- Data S6. Hierarchical clustering of the *P. patens* and Arabidopsis proteins.
- Data S7. Log2 transformed and imputed data used for volcano plots.
- Data S8. Overview of LEA proteins.
- Data S9. Overview of LD proteins.
- Data S10. Overview of LOX and AOS.
- Data S11. Sequence alignment of LEA proteins.
- Data S12. Lea tree file.
- Data S13. Lea tree design Figure 6.
- Data S14. Lea tree design Figure S2.

1 **Core desiccation response strategies of plant development**  
2 **revealed by the dynamic proteome of *Physcomitrium patens***  
3 **during spore germination**

4  
5 **Lea Hembach<sup>1#</sup> Philipp William Niemeyer<sup>2#</sup>, Kerstin Schmitt<sup>3</sup>, Jaccoline Zegers<sup>4</sup>**  
6 **Dennis Brandt<sup>5</sup>, Janis Dabisch<sup>1</sup>, Oliver Valerius<sup>3</sup>, Gerhard H. Braus<sup>3</sup>, Markus**  
7 **Schwarzländer<sup>5</sup>, Jan de Vries<sup>4</sup>, Stefan Rensing<sup>6,7,8</sup>, Till Ischebeck<sup>1,2\*</sup>**

8  
9 <sup>1</sup>University of Münster, Institute of Plant Biology and Biotechnology (IBBP), Green  
10 Biotechnology, 48143 Münster, Germany

11 <sup>2</sup>University of Göttingen, Albrecht-von-Haller-Institute for Plant Sciences and Göttingen  
12 Center for Molecular Biosciences (GZMB), Department of Plant Biochemistry, 37077  
13 Göttingen, Germany

14 <sup>3</sup>University of Göttingen, Institute for Microbiology, Genetics and Göttingen Center for  
15 Molecular Biosciences (GZMB) and Service Unit LCMS Protein Analytics, Department  
16 for Molecular Microbiology and Genetics, 37077 Göttingen, Germany

17 <sup>4</sup>University of Göttingen, Institute for Microbiology and Genetics, Göttingen Center for  
18 Molecular Biosciences (GZMB) and Campus Institute Data Science (CIDAS),  
19 Department of Applied Bioinformatics, 37077 Göttingen, Germany

20 <sup>5</sup>University of Münster, Institute of Plant Biology and Biotechnology (IBBP), Plant Energy  
21 Biology, 48143 Münster, Germany

22 <sup>6</sup>University of Marburg, Plant Cell Biology, Department of Biology, Marburg, Germany.

23 <sup>7</sup>University of Freiburg, BIOS Centre for Biological Signalling Studies, Freiburg,  
24 Germany.

25 <sup>8</sup>Present address: University of Freiburg, Faculty of Chemistry and Pharmaceutical  
26 Sciences, Freiburg, Germany

27 #equal contribution

28  
29 **\*Correspondence:**

30 [till.ischebeck@uni-muenster.de](mailto:till.ischebeck@uni-muenster.de)

31  
32 **Running title:**

33 Proteome of *P. patens* spore germination

34 **Keywords:**

35 *Physcomitrium patens*, Spore, Spore germination, Seeds, Lipid droplets, Proteome, LEA  
36 proteins, *Arabidopsis thaliana*.

37

38

## 39 **ABSTRACT**

40 The establishment of seeds is considered one of the hallmarks of plant evolution.  
41 However, most of the underlying protein networks required for desiccation tolerance, the  
42 accumulation of storage compounds, and the regulation of dormancy have probably  
43 evolved much earlier. The same could be true for the molecular program that drives the  
44 transition from a heterotrophic offspring to an autotrophic plant. A comparison to the  
45 earliest land plants could reveal where some of the roots of the “seed program” could  
46 originate. Therefore, we investigated the proteome of five timepoints of *P. patens* spore  
47 germination as well as protonema and gametophores, and compared it to Arabidopsis  
48 data on seedling establishment. This quantitative comparison showed that not only  
49 spores are related to seeds but also germinating spores to young seedlings. We  
50 discovered many similarities in regard to desiccation tolerance, lipid droplet proteome  
51 composition, control of dormancy, and the pathways that transform fatty acids into  
52 sugars. However, there were also striking differences in whereby the spores of *P. patens*  
53 did not harbor any obvious storage proteins. Furthermore, we did not detect homologs  
54 to the main triacylglycerol lipase in Arabidopsis, SUGAR DEPENDENT1. Instead, we  
55 discovered a triacylglycerol lipase of the oil body lipase family and a lipoxygenase as  
56 being the overall most abundant proteins in spores. This finding indicates an alternative  
57 pathway for triacylglycerol degradation via oxylipin intermediates.

## 58 **INTRODUCTION**

59 Bryophytes are non-vascular land plants that form a taxonomic division comprising  
60 mosses, liverworts, and hornworts and have been used as a model for understanding  
61 the evolution of plant terrestrialization about 500 million years ago (Ligrone *et al.*, 2012;  
62 Kenrick, 2017; Rensing, 2018). This diverse group includes approximately 20,000 known  
63 species, and among them, the moss *Physcomitrium patens* has gained particular  
64 importance as a model for non-seed plants over the past 60 years (Shaw *et al.*, 2011;  
65 Cove, 2005; Rensing *et al.*, 2020). In 2008, the *P. patens* genome was the first of any  
66 bryophyte to be fully sequenced (Rensing *et al.*, 2008). *P. patens* exhibits a life cycle  
67 with a dominating gametophytic phase: Haploid spore germination leads first to the  
68 development of a two-dimensional filamentous network called protonema (Cove, 2005).  
69 Eventually, the protonema develops buds that differentiate into leafy gametophores  
70 (Cove, 2005). During the sexual reproductive phase, male antheridia and female  
71 archegonia sexual organs are formed on the tips of its gametophores (Hohe *et al.*, 2002),



72 and after fertilization, the zygote forms a sporophyte which produces cells that undergo  
73 meiosis, ultimately resulting in the formation of 3000-16000 haploid spores within a  
74 capsule (Nakosteen and Hughes, 1978; Engel, 1968).

75 Spores display functional and structural resemblances with the reproductive tissues of  
76 seeds and pollen, which have been present in the plant kingdom since the evolution of  
77 spermatophytes (Loconte and Stevenson, 1990; Huang *et al.*, 2009). Similar to pollen,  
78 spores are less-intricate structures, small-sized, haploid, and protected by a  
79 sporopollenin wall (Daku *et al.*, 2016; Wallace *et al.*, 2011). However, spores are the  
80 primary means of dispersal in the moss life cycle, and in this sense, they are more  
81 analogous to spermatophyte seeds (Vesty *et al.*, 2016). Spores, pollen, and seeds of  
82 most species exhibit desiccation-tolerant characteristics, underlain by a typical  
83 molecular and gene expression pattern (Oliver *et al.*, 2000; Gaff and Oliver, 2013;  
84 Matilla, 2022).

85 One important element of desiccation-tolerant tissues is the accumulation of late  
86 embryogenesis abundant (LEA) proteins as a protective measurement against  
87 dehydration and other abiotic stressors such as temperature and salt stress (Artur *et al.*,  
88 2019; Delahaie *et al.*, 2013; Cuming *et al.*, 2007; Saavedra *et al.*, 2006). These highly  
89 hydrophilic LEA proteins contain intrinsically disordered regions and are classified into  
90 eight multigene families in plants (Artur *et al.*, 2019). The protective functions facilitated  
91 by LEA proteins include the vitrification of the cytosol in cooperation with non-reducing  
92 sugars. This produces a glassy state when the organism is subjected to desiccation,  
93 thereby replacing water with a hydrophilic protein shell (Manfre *et al.*, 2008; Wolkers *et*  
94 *al.*, 2001). Furthermore, LEA proteins demonstrate anti-aggregation (Chakrabortee *et*  
95 *al.*, 2012) and chaperone-like properties (Kovacs *et al.*, 2008) and can also regulate  
96 membrane permeability (Liu *et al.*, 2009). The role of distinct LEA proteins is functionally  
97 multifaceted, and only a subpopulation of LEAs contribute to the establishment of  
98 desiccation tolerance (Artur *et al.*, 2019; Matilla, 2022).

99 Another feature of most spores, seeds, and pollen is the accumulation of neutral  
100 lipids such as triacylglycerol (TAG), which are stored in cytosolic lipid droplets (LDs)  
101 (Guzha *et al.*, 2023). Lipid droplets may enhance desiccation resilience by maintaining  
102 membrane integrity through filling the intracellular space, which reduces cell shrinkage  
103 and prevents membrane collapse (Lyll and Gechev, 2020). LD formation might derive  
104 from a generic, ancient, abiotic-stress-related machinery that originated even before  
105 embryophytes emerged on *terra firma* (de Vries and Ischebeck, 2020). In spores and

106 seeds, lipid droplets serve as an energy and carbon source during the post-germinative  
107 growth (Ischebeck *et al.*, 2020; Huang *et al.*, 2009), while in pollen, their might function  
108 might be to serve as an acyl-chain source during pollen tube formation and growth  
109 (Ischebeck, 2016; Müller and Ischebeck, 2018) as well as a carbon sink during pollen  
110 tube growth (Krawczyk *et al.*, 2022). A hallmark of lipid droplets embedded in  
111 desiccation-tolerant tissues is their outer coverage of oleosin (Murphy *et al.*, 1995). This  
112 integral surface protein likely shields the droplets and inhibits their coalescence during  
113 desiccation (Siloto *et al.*, 2006), potentially through electrostatic repulsion arising from  
114 their negative charge (Huang, 1992).

115 The accumulation of such proteins with protective functions requires an underlying  
116 regulatory network. In this context, the plant hormone abscisic acid (ABA) has a crucial  
117 role in the acquisition of desiccation tolerance (Cuming, 2019). ABSCISIC ACID  
118 INSENSITIVE 3 (ABI3) is a key transcription factor activated by ABA (Khandelwal *et al.*,  
119 2010). ABA3 serves as a trigger and component of a signaling pathway highly conserved  
120 in land plants (Oliver *et al.*, 2005). This includes inducing the expression of oleosins  
121 (Crowe *et al.*, 2000; Yang *et al.*, 2022) and several LEA proteins (Tian *et al.*, 2020),  
122 among other gene products related to dehydration. Moreover, ABI3 is essential for  
123 acquiring desiccation tolerance in seeds (Giraudat *et al.*, 1992) and vegetative  
124 desiccation tolerance in *P. patens* gametophores which can be induced via ABA  
125 treatment (Khandelwal *et al.*, 2010).

126 While it has been previously suggested that bryophytes such as *P. patens* adopt  
127 strategies to cope with spore desiccation in a similar way as angiosperms seeds (López-  
128 Pozo *et al.*, 2018; Huang *et al.*, 2009; Farrant *et al.*, 2009), we aimed at taking a deeper  
129 look at its proteome in spores, its changes during spore germination, and how it  
130 compares to vegetative tissues in moss and ultimately also to the seeds of *Arabidopsis*.  
131 Several proteomic datasets were acquired in the last two decades that give solid  
132 supportive evidence of abiotic stress and desiccation tolerance-associated proteins in  
133 gametophores and protonema (Cho *et al.*, 2006; Cui *et al.*, 2012; Luo *et al.*, 2020;  
134 Mamaeva *et al.*, 2022; Sarnighausen *et al.*, 2004; Skripnikov *et al.*, 2009; Toshima *et al.*  
135 *et al.*, 2014; Wang *et al.*, 2010; X., Wang *et al.*, 2009; X., Q., Wang *et al.*, 2009; Wang *et al.*  
136 *et al.*, 2008; Yotsui *et al.*, 2016). However, to our knowledge, no comprehensive proteomic  
137 datasets of bryophyte spores are available as of now. The available transcript data  
138 covers spores and one stage of germinating spores (Fernandez-Pozo *et al.*, 2020; Julca  
139 *et al.*, 2021; Perroud *et al.*, 2018; Hiss *et al.*, 2014) but especially in developmental

140 processes, there can be a substantial discrepancy between the transcriptome and  
141 proteome (Chaturvedi *et al.*, 2013).

142 To understand key strategies of developmental desiccation tolerance, we devise a  
143 comprehensive proteomic analysis of germinating spores, protonema, and  
144 gametophores of the *P. patens* ecotype Reute. We devise an in-depth analysis of the  
145 changes during spore germination and the transition to vegetative tissues. A  
146 comparison to Arabidopsis seed germination provides remarkable insight into the  
147 operation of shared protein networks required for desiccation tolerance.

148

## 149 **RESULTS**

### 150 **The *P. patens* spore proteome changes during germination and is distinct from** 151 **vegetative tissues**

152 Several studies indicated that the molecular programs and the protein networks first  
153 described as characteristics of seeds have evolved much earlier and also occur in non-  
154 seed plants (de Vries and Ischebeck, 2020; Oliver *et al.*, 2005; Oliver *et al.*, 2000). In  
155 order to get insight into conserved mechanisms and players we chose a system that (i)  
156 has emerged before (orthodox) seeds and (ii) shares key physiological characteristics,  
157 such as sudden activation after quiescence and desiccation. We chose *P. patens* spore  
158 germination as a tangible model and devised a proteomic approach of five stages of  
159 spore germination (0, 8, 24, 48, and 72 h) and two vegetative stages (protonema and  
160 gametophore) of the moss *P. patens* (Figure 1a) and compared it to the previously  
161 published proteome of Arabidopsis seeds and seedlings (Kretzschmar *et al.*, 2020).

162 LC-MS/MS analysis of five independent biological replicates of each *P. patens* stage  
163 was performed after a tryptic digest of the protein samples and all protein groups were  
164 quantified using the label-free MS1-based algorithm of MaxQuant (intensity-based  
165 absolute quantification, iBAQ). These values were then divided by the total iBAQ of the  
166 respective sample and multiplied by 1000, resulting in relative iBAQ (riBAQ) values. All  
167 *P. patens* library entries were then functionally annotated by performing a BLASTp  
168 search against the Arabidopsis TAIR10 primary transcript protein release database  
169 (Data S1) (Lamesch *et al.*, 2012). In total, 3447 protein groups could be identified and  
170 quantified (Data S2a). 2609 of these protein groups were identified based on two  
171 peptides and were found in at least three samples of one of the stages (Data S2b).  
172 These rigidly filtered protein groups were further analyzed.

173 For studying the proteome of all stages investigated of *P. patens*, a principal component  
174 analysis (PCA, Figure 1b) was generated. The PCA indicates that the proteomes  
175 between spores and vegetative stages differ the most, while the first three spore stages  
176 are closely related. The similarities between the individual biological replicates  
177 demonstrate the high reproducibility of the measurements.

178 Next, averages were calculated for each developmental stage and these were  
179 normalized setting the highest value for each protein to 1. These values were then  
180 hierarchically clustered (Figure 1c, Data S3). Several clusters contain proteins that are  
181 high in early spore stages (cluster 3-5 and 8), while others were found predominantly  
182 during late spore germination (cluster 1, 2, 6, and to some extent 7, 17, 18) or in the  
183 protonema and gametophore tissues (cluster 28 and 30).

184 Among the proteins enriched in spores that are quickly degraded during spore  
185 germination (cluster 4, Figure 1d) are a homolog to the phytochrome B type  
186 photoreceptor HY3 that plays a role in the regulation of de-etiolation of Arabidopsis  
187 seedlings (Wester *et al.*, 1994), and the RNA helicase LOS4 (LOW EXPRESSION OF  
188 OSMOTICALLY RESPONSIVE GENES 4) that is involved in mRNA exports and acts  
189 upstream of abiotic stress responses (Gong *et al.*, 2005). Another example is a putative  
190  $\beta$ -d-xylosidase. Homologs of this protein in Arabidopsis are involved in cell wall  
191 loosening by removing xylose and arabinose side chains from cell wall components such  
192 as rhamnogalacturonan I (Guzha *et al.*, 2022; Arsovski *et al.*, 2009).

193 Two small heat shock proteins contained in cluster 5 first rise in abundance and then  
194 sharply drop (Figure 1d). Homologs of these proteins are important for seed vigor and  
195 longevity in Arabidopsis and are ABA-regulated (Waters and Vierling, 2020). Further  
196 proteins in the clusters that increased during spore germination but were much lower in  
197 protonema and gametophores proteins involved in  $\beta$ -oxidation (for example,  
198 multifunctional protein in cluster 18) and the glyoxylate cycle (such as malate synthase  
199 in cluster 7) indicating that protein synthesis and conversion of TAG into carbohydrates  
200 are especially high in spores 48 h and 72 h old.

201

## 202 **GO term analysis highlights similarities between *P. patens* spore germination and** 203 **Arabidopsis seedling establishment**

204 In order to take a look at the involvement of the detected proteins in larger cellular  
205 processes, these proteins were assigned to their respective homologs of Arabidopsis  
206 (Data S1) before being categorized and combined to Gene Ontology (GO) Terms

207 published for Arabidopsis (Berardini *et al.*, 2004). The riBAQ values of all proteins of  
208 each GO Term were added and compared to previously published data of Arabidopsis  
209 seeds and seedlings (Data S4; Kretzschmar *et al.*, 2020).

210 Several distribution patterns of protein abundances (riBAQ) consolidated in GO terms  
211 are consistent between germinating *P. patens* spores and Arabidopsis seedling  
212 establishment (Figure 2). Examples are the GO terms 'Response to abscisic acid', 'Lipid  
213 droplet', 'Oxylipin biosynthetic process', and 'Maintenance of seed dormancy', which  
214 display their highest abundance spores and seeds and then decline. An increase in  
215 protein abundance during spore germination and seedling establishment, respectively,  
216 was observed for the GO terms 'Response to hydrogen-peroxide', 'Protein folding',  
217 'Glyoxylate cycle', and 'Fatty acid beta-oxidation'. The values for these terms then  
218 dropped in protonema and gametophores and 60 h old seedlings, respectively. High in  
219 these later stages were proteins assigned to the GO terms to 'Reductive pentose  
220 phosphate activity' and 'Photosynthesis'.

221 Conversely, certain ontogenetic patterns related to GO terms of seeds and spores show  
222 apparent dissimilarities. For example, the terms 'Pollen development' and 'Seed  
223 development' manifest a gradual decline in germinating spores. In contrast, values for  
224 these terms in *Arabidopsis* nearly double from rehydrated seeds to the 36 h seedling  
225 stage, and subsequently decrease again after 60 h to rehydrated seed baseline levels.  
226 Moreover, the GO term 'Cellular response to hypoxia' shows a relatively stable  
227 abundance in germinating spores, while being almost entirely abolished in seedlings.  
228 Nevertheless, despite the notable differences in the ontogenetic patterns between  
229 spores and seeds, certain critical developmental steps are shared by both types of  
230 structures. These similarities may be linked to the maintenance of desiccation tolerance,  
231 oil degradation, and the transition into the vegetative photosynthetically active tissue  
232 important in both seeds and spores.

233

### 234 **A quantitative comparison supports the similarities between the seed and spore** 235 **proteomes**

236 In order to quantitatively compare spore germination and seedling establishment, the  
237 two proteomic datasets of *P. patens* (excluding protonema and gametophores) and  
238 Arabidopsis (Kretzschmar *et al.*, 2020), the *P. patens* protein groups were assigned to  
239 the closest homolog present in the Arabidopsis dataset (E-value  $\leq 10^{-5}$ , Data S1b, Data  
240 S5a). If several *P. patens* protein groups were assigned to the same Arabidopsis

241 accession, the values were added (Data S5b). Then, the averages for each stage were  
242 calculated and the highest average of each of the two datasets was set to 1. Proteins  
243 only found in one dataset were removed (Data S5c). All samples of *P. patens* were  
244 analyzed in direct comparison to samples of *Arabidopsis* of rehydrated (RS), stratified  
245 (StS), and 24, 36, 48, or 60 h germinated seeds (Kretzschmar *et al.*, 2020) normalized  
246 in the same manner.

247 A PCA plot of this analysis (Figure 3a) displayed a separation of the developmental  
248 stages by component 1 that was similar in both species. However, component 2 strongly  
249 separated 0 h spores and 0 h seeds, indicating that they are not very closely related.  
250 However, hierarchical clustering (Figure 3b, Data S6) revealed that in this analysis  
251 spores of 0 h to 24 h are most closely related to seeds, 48 h spores to 24 and 36 h old  
252 seedlings, and 72 h spores to seedlings at 48 and 60 h. Also, the clustering of the  
253 proteins showed that spores and seeds contain very similar protein sets in part. Clusters  
254 4-7, for example, harboring proteins high in both seeds and spores that decrease over  
255 time, contain a total of 105 proteins that might be evolutionarily conserved in protecting  
256 desiccated stages or facilitating the early phase of the transition from a heterotrophic  
257 offspring dispersal stage to an autotrophic plant. Proteins in these clusters might be  
258 involved in cell-wall-related metabolism, such as a putative xylosidase/arabinosidase  
259 (cluster 4, Data S6) and several putative galactose oxidases (cluster 7). Further  
260 examples are heat shock proteins of the 17.6 kd family and so-called universal stress  
261 proteins (USPs, all cluster 6). These USPs are associated with diverse biotic and abiotic  
262 stress responses (Chi *et al.*, 2019) and, based on their high abundance in spores and  
263 seeds might play a role in protection there as well. Our analysis of this cluster further  
264 highlighted proteins that are so far of unknown function but might have a conserved  
265 function in seeds and spores.

266 We also took a closer look at proteins in cluster 27 that were high in 36-48 h old  
267 *Arabidopsis* seedlings and 36 h old germinated spores, as these proteins might be  
268 involved in the late transition phase to autotrophy. This cluster contains numerous  
269 proteins involved in protein synthesis, such as t-RNA synthases, translation initiation  
270 factors, and ribosomal proteins, indicating an increased rate during this time window in  
271 both seedlings and germinating spores.

272

273

274 **Spores are enriched in LD and LEA proteins and proteins involved in oxylipin**  
275 **metabolism**

276 A one-on-one comparison of all developmental stages (Figure S1) using imputed values  
277 (Data S7a) revealed significantly changed protein abundances between all stages  
278 except for the comparison of 0 h and 8 h. To get further insight into the specific proteome  
279 of 0 h spores, we compared these to 72 h old germinated spores and gametophores  
280 (Figure 4, Data S7b, c).

281 Two examples much higher in 0 h versus 72 h spores might play a role in detoxifying  
282 oxidized lipids and their downstream products (Data S7b). They are homologs to the  
283 proteins CeQORH (CHLOROPLAST ENVELOPE QUINONE OXIDOREDUCTASE  
284 HOMOLOG, At4g13010) and CHLADR (CHLOROPLAST ALDEHYDE REDUCTASE,  
285 At1g54870) that have been implemented with such functions in Arabidopsis plastids  
286 (Curien *et al.*, 2016; Yamauchi *et al.*, 2011). Both these proteins are found in dry seeds  
287 and are degraded during seedling establishment (Kretzschmar *et al.*, 2020).

288 A similar pattern in the seedling establishment is observed for an AWPM-19-like  
289 membrane family protein (AT1G04560). A *P. patens* homolog was strongly enriched in  
290 spores in comparison to gametophores (Data S7c), while a homolog in rice has been  
291 shown to be involved in the ABA transport (Yao *et al.*, 2018), indicating a possible  
292 function for the regulation of seed and spore dormancy in Arabidopsis and *P. patens*,  
293 respectively.

294 The comparisons furthermore displayed that the spores at 0 h were chiefly enriched in  
295 proteins homolog to (i) LEA (late embryogenesis abundant) proteins, (ii) known LD  
296 proteins (Guzha *et al.*, 2023), and (iii) proteins involved in oxylipin metabolism. These  
297 proteins will be discussed below.

298

299 **Late embryogenesis proteins show distinct expression patterns that might be**  
300 **evolutionarily conserved**

301 Two of the 14 LEA proteins stayed high during spore germination and were almost  
302 absent in vegetative tissues, while two further isoforms were highest in gametophores.  
303 However, most of the detected LEAs in this study were highest in spores, degraded  
304 within 72 h, and largely absent in protonema and gametophores. A similar decrease was  
305 observed for Arabidopsis LEAs during seedling establishment (Kretzschmar *et al.*,  
306 2020).

307 As LEA proteins appear to be important in both seeds and spores, we investigated the  
308 overall evolutionary history of LEAs across 24 species in the green line to get a deeper  
309 understanding of this class of proteins. Here, we put a focus on four species – *A.*  
310 *thaliana*, *O. sativa*, *C. esculentus*, *P. patens* – for which we also analyzed experimental  
311 protein and transcript data from different datasets (see Methods section for details).

312 In a phylogenetic tree of all LEA families of the 24 species, the individual eight multigene  
313 LEA families are grouped together in separate clades; we could not pinpoint the common  
314 ancestral proteins prior to diversification into all of the diverse LEA families (Figure S2).  
315 That said, both LEA\_5 and LEA\_4 groups are conserved across all analyzed species,  
316 suggesting that these two groups may be the evolutionary oldest (Figure S2). LEA\_4  
317 presents the most diverse group with the most different sub-cellular localizations (Figure  
318 6). More specifically, we identified two sub-clades of mitochondrial- and/or plastidal-  
319 targeted LEAs, and a distinct clade that was more associated with secretory pathway  
320 targeting (Figure 6). As two of these clades contain proteins from both Arabidopsis and  
321 *P. patens*, it can be speculated that the subcellular targeting has been evolutionarily  
322 conserved for a long time period.

323 The other LEA families appeared at different times during the evolution of the green  
324 lineage, with LEA\_2 and seed maturation protein (SMP) having emerged in the last  
325 common ancestor (LCA) of streptophytes and LEA\_1 and dehydrins in the LCA of  
326 embryophytes (Figure S2).

327 LEA\_6 and LEA\_3 families are exclusive to spermatophytes (Figure S2); both families  
328 are mostly associated with seed- and/or pollen-stages (Figure 6). Yet, seed- and pollen-  
329 associated LEA proteins are also found across other LEA families, often sharing  
330 homologues with the bryophyte *P. patens* (Figure 6) Moreover, the angiosperm-specific  
331 LEA\_3 family stands out for having members from rice and Arabidopsis, which were all  
332 predicted and partially confirmed to be targeted to mitochondria (Figure S2). Lastly, the  
333 LEA\_2 can be split into two subclades, with one of them being the only clear sub-clade  
334 of LEA proteins, which is predominantly associated with the vegetative stages (Figure  
335 6).

### 336 **Lipid Droplet proteins and their role during ripening and germination of spores** 337 **and seeds**

338 One hallmark of seeds and other desiccated structures is the occurrence of a large  
339 number of lipid droplets that store neutral lipids, such as triacylglycerol (Guzha *et al.*,



340 2023). However, these structures also occur in non-desiccated tissues and their function  
341 might be reflected by their proteomes (Ischebeck *et al.*, 2020). While we did not isolate  
342 LDs in this study, 15 homologs to known LD proteins were found in the dataset, allowing  
343 us to study their developmental pattern (Figure 7, Data S9). In total, LDs made up 11.9  
344 % of all proteins in the spores, which was much higher than in protonema (0.05 %) and  
345 gametophore (0.4 %) tissues and even higher than in Arabidopsis seeds (5.7 %;  
346 (Kretzschmar *et al.*, 2020).

347 The most abundant putative LD protein in *P. patens* spores is a homolog of OIL BODY  
348 LIPASE 1 (Figure 7, OBL1, Pp3c3\_4690V3.2). This protein is primarily found in  
349 developing and mature seeds of Arabidopsis (Klepikova *et al.*, 2016), as well as in  
350 emerging pollen tubes of Arabidopsis and *Nicotiana tabacum*, potentially supplying TAG-  
351 derived membrane lipids there (Müller and Ischebeck, 2018). During spore germination,  
352 its abundance decreases by approximately 88.7% within 72 hours and is nearly depleted  
353 in the protonema and gametophore stages. In addition to this dominant isoform, three  
354 additional homologs were identified. The second most abundant family of proteins in the  
355 spores were oleosins that are the main LD proteins in most seeds (Guzha *et al.*, 2023).  
356 Furthermore, homologs to the other two most abundant proteins in seeds, caleosins,  
357 and steroleosins, were found almost exclusively in spores. Steroleosins are thought to  
358 be important for the metabolism of brassinosteroids (Lin *et al.*, 2002), while caleosins  
359 might have a peroxygenase activity (Hanano *et al.*, 2023) and are considered to play a  
360 role in stress-related processes (Shimada *et al.*, 2014; Aubert *et al.*, 2010). The LIPID  
361 DROPLET ASSOCIATED PROTEIN 3 (LDAP3, Pp3c19\_21240V3.2) and its interaction  
362 partner LDAP INTERACTING PROTEIN (LDIP, Pp3c25\_6290V3.3) are also  
363 predominantly found in spores. However, LDAP remains at a constant level, whereas  
364 LDIP decreases during germination. While these proteins are needed for the proper  
365 formation of LDs (Pyc *et al.*, 2021), LDAP might have an additional coating function and  
366 is found also in vegetative tissues (Gidda *et al.*, 2016; Brocard *et al.*, 2017). Other  
367 proteins were not predominately found in spores. One example is PLANT UBX DOMAIN  
368 CONTAINING PROTEIN 10 (PUX10) that is involved in the degradation of ubiquitinated  
369 LD-associated proteins in Arabidopsis (Deruyffelaere *et al.*, 2018; Kretzschmar *et al.*,  
370 2018) and that we found in similar amounts in all analyzed stages. The two proteins  
371 involved in sterol metabolism a putative cycloartenolsynthase and sterol methyl  
372 transferase 1 showed the highest levels in late stages of spore germination and  
373 vegetative tissues, respectively.

## 374 **A bimodal expression pattern of lipoxygenases in *P. patens***

375 The share of lipoxygenases and allene oxide synthases was comparably high in 0 h  
376 spores, making up 9.6 % of all proteins and then dropping to 2.4% in 72 h spores. The  
377 levels in protonema (0.37 %) and gametophores (0.33 %) were much lower (Data S10),  
378 implying a pivotal role during spore germination. Consistently, *PpLOX4b*  
379 (*Pp3c14\_10640V3.1*) is the second most abundant protein in 0 h spores (Data S2). LOX  
380 proteins oxygenate polyunsaturated fatty acids, which can be metabolized to oxylipins,  
381 volatiles, or designate the oxygenated fatty acid for degradation (Andreou and Feussner,  
382 2009). Furthermore, specific isoforms were observed as LD-associated proteins in some  
383 species (Feußner and Kindl, 1992). The different types of plant LOXs can be classified  
384 based on their targeting signals for either plastidic or non-plastidic LOX, as well as their  
385 specificity for the oxygenation of  $\alpha$ -linoleic acid (Andreou and Feussner, 2009). Our  
386 proteomic dataset identified seven LOX proteins and two AOS proteins (Data S10). LOX  
387 proteins and their enzymatic activity were annotated according to (Anterola *et al.*, 2009).  
388 *Pp3c1\_29700V3.2.p* with the highest similarity to *PpLOX6* was named by us *PpLOX10*.  
389 Then, the consensus of putative subcellular targeting was determined by assessing  
390 three distinct localization prediction tools (Figure 8).

391 The protein abundance of LOX pathway-related genes shows a bimodal expression  
392 pattern. One pattern is a high abundance in the vegetative stages with much lower levels  
393 in germinating spores. This trend is particularly evident in those LOX and AOS enzymes,  
394 which are predicted to be localized within the chloroplasts and exhibit 12-S substrate  
395 specificity to arachidonic acid. The second pattern is observed in LOX and AOS  
396 enzymes, which are predicted to be non-chloroplastic in their localization and exhibit 13-  
397 S substrate specificity to  $\alpha$ -linolenic acid. These enzymes are primarily expressed in  
398 spores, and a general trend of decreasing relative abundance is observed during  
399 germination.

400

## 401 **DISCUSSION**

### 402 **Both spore and seed germination, and seedling establishment go through key** 403 **phases**

404 Desiccation-tolerant spores and seeds are the prime dispersal units of mosses and seed  
405 plants, respectively, and hallmarks of plant evolution. While they are analogous  
406 structures, they share many similarities not only in their function but also on the  
407 molecular level (Pacini, 2012). Furthermore, the similarities are not restricted to the

408 desiccated stage but extend to the establishment of a photoautotrophic plant from the  
409 heterotrophic energy-rich spore or seed. Both seeds and spores undergo four key  
410 phases that partially overlap. The first phase is defined by desiccation tolerance and  
411 already comprises the late phase of spore and embryo development. In the second  
412 phase, the cells are rehydrated and metabolism is activated. Furthermore, it is  
413 determined if the spore or seed germinates or remains dormant. The third phase is  
414 characterized by the degradation of TAG and the activation of the glyoxylate cycle to  
415 generate energy and carbohydrates for cellular growth. During the fourth phase, the  
416 photosynthetic apparatus is established, leading to full photoautotrophy. We will discuss  
417 the individual phases in the context of similarities and differences between *Arabidopsis*  
418 and *P. patens*.

419

### 420 **Phase 1 Several protein families mediate desiccation tolerance**

421 LEAs proteins play a key role in protecting proteins and membranes in desiccated  
422 structures (Amara *et al.*, 2014). In line with this function, most of these proteins are  
423 largely degraded within 72 h in *P. patens* spores and already after 36 h in seedlings  
424 (Kretzschmar *et al.*, 2020). An exception are two LEA\_2 and several LEA\_4 proteins  
425 that were found predominantly in vegetative tissues of *P. patens* based on proteomic  
426 and transcript data (Figures 5 and 6). Interestingly the LEA2 proteins are closely related  
427 to LEA proteins of seed plants that are also predominantly found in non-desiccated  
428 organs (Figure 6). Also, the organelle targeting of certain LEA proteins seems to be  
429 conserved between mosses and seed plants as members of a subclade of the LEA\_4  
430 family have been shown or are predicted to target plastids and mitochondria.

431 Throughout plant evolution, the number of LEA protein families has increased in plants,  
432 with LEA\_1, LEA\_3 and LEA\_6 not being present in *P. patens*. The members of the  
433 LEA\_1 and LEA\_6 families are all found in desiccated structures, while the LEA\_3 family  
434 have all been shown or been predicted to target mitochondria indicating a specialized  
435 function for these LEA clades. Interestingly dehydrin-type LEA proteins are found in  
436 *P. patens* but not in streptophyte algae that. They could have evolved in the early land  
437 plants and helped to cope with the specific stresses encountered on land.

438 Both spores of *P. patens* and *Arabidopsis* seeds contain a large number of LDs and also  
439 these structures have to be protected to prevent them from fusing. Oleosins that already  
440 coat the LDs of the streptophyte alga *Mesotaenium endlicherianum* (Dadras *et al.*, 2022)  
441 and are commonly found in desiccated tissues, including spores, pollen, and seeds

442 (Guzha *et al.*, 2023). Their importance during desiccation is corroborated by the  
443 diminished or deficient presence of oleosins in desiccation-sensitive (recalcitrant) seeds  
444 (Leprince *et al.*, 1997) and by the heightened susceptibility to freezing of oleosin-reduced  
445 *Arabidopsis* seeds (Shimada *et al.*, 2008). Conversely, augmented levels of oleosins  
446 were observed in desiccated vegetative tissues of resurrection plants (Costa *et al.*, 2017;  
447 VanBuren *et al.*, 2017; Xu *et al.*, 2018) and in desiccation-tolerant tubers of *Cyperus*  
448 *esculentus* (Niemeyer *et al.*, 2022; Yang *et al.*, 2016). Another prominent LD-associated  
449 protein family in both spores and seeds are the caleosins (Figure 7; Guzha *et al.*, 2023).  
450 In *Arabidopsis*, with CLO1 and 2 being highest in seeds (Kretzschmar *et al.*, 2020),  
451 CLO3 in leaves (Fernández-Santos *et al.*, 2020), and CLO4 in pollen (Ischebeck, 2016).  
452 Such an organ-specific expression also seems to be the case in *P. patens*, with one  
453 gene (Pp3c6\_2090V3) being highest expressed in spores, while the other two isoforms  
454 are strongly upregulated under ABA treatment and dehydration (Fernandez-Pozo *et al.*,  
455 2020).

456

#### 457 **Phase 2 Spore and seed germination are regulated by similar factors**

458 In *Arabidopsis* seeds, it was shown that metabolism is initiated very quickly during  
459 rehydration, with ATP being produced within minutes (Nietzel *et al.*, 2020). While a  
460 similar study has not been performed for *P. patens* spore germination, all the proteins  
461 required for energy generation are readily available, with the GO-term “glycolytic  
462 process” being highest in dry spores and then slowly decreasing, and subunits of the  
463 mitochondrial ATP synthase complex also being present (Data S2b).

464 By activating their metabolism, spores and seeds take the first steps required for  
465 germination. However, germination can also be suppressed by a complex regulatory  
466 network leading to spores and seeds staying dormant (Nonogaki, 2014; Vesty *et al.*,  
467 2016). Part of this regulatory network are hormones, including ABA and, at least in  
468 seeds, brassinosteroids.

469 While *P. patens* spores do not seem to have a primary dormancy, their germination can  
470 be suppressed, for example, by ABA similar to seeds (Vesty *et al.*, 2016; Sano and  
471 Marion-Poll, 2021). ABA might also be involved in the accumulation of certain proteins  
472 during spore maturation, as many of the proteins found in dry spores have homologs to  
473 ABA-responsive proteins in *Arabidopsis*. Furthermore, we detected a AWPM-19-like  
474 membrane family protein in both spores and *Arabidopsis* seeds that might be involved

475 in ABA transport and thereby signaling ABA signaling based on the function of a homolog  
476 in rice (Yao *et al.*, 2018).

477 Brassinosteroids are plant hormones involved in growth and developmental processes  
478 (Peres *et al.*, 2019; Clouse and Sasse, 1998) that also occur in moss with a proposed  
479 function in development (Yokota *et al.*, 2017; Morikawa *et al.*, 2009). Brassinosteroids  
480 impact spore germination in ferns (Gómez-Garay *et al.*, 2018) and in seeds of  
481 Arabidopsis (Leubner-Metzger, 2003) and rice (Xiong *et al.*, 2022). It remains uncertain  
482 whether this effect applies to bryophyte spores. Steroleosins are major LD proteins in  
483 Arabidopsis seeds (Kretschmar *et al.*, 2020), but are not found in drought-stressed  
484 leaves (Pyc *et al.*, 2021), indicating that they might not have a role in drought and  
485 desiccation responses but rather developmental processes. In line with this is a potential  
486 involvement in brassinosteroid metabolism by acting as dehydrogenases (Baud *et al.*,  
487 2009; Li *et al.*, 2007), although the exact nature of their involvement is currently  
488 uncertain.

489 Another factor that influences dormancy is light perceived by specific receptors. In seeds  
490 of Arabidopsis, red light induces seed germination, while far-red light inhibits it (Jiang *et al.*  
491 *et al.*, 2016). Involved in this process is the phytochrome B photoreceptor HYPOCOTYL3  
492 (HY3) that induces seed germination when perceiving red light but is inhibited by far-red  
493 light. A homolog of HY3 that we detected in dry spores was rapidly degraded within 48  
494 h. Based on this finding, it could play an analogous role, especially since it was shown  
495 that *P. patens* spore germination is inhibited by far-red light as well (Vesty *et al.*, 2016).  
496

### 497 **Phase 3 Heterotrophic degradation of oil**

498 With its high energy and carbon density, TAG is an optimal storage reserve. However,  
499 both spore germination and seedling establishment require mostly sugars for cell wall  
500 synthesis needed for cell expansion (Ischebeck *et al.*, 2020). As of this, the fatty acids  
501 derived from the TAG are degraded first by  $\beta$ -oxidation, and the resulting acetyl-CoA is  
502 converted to sugars by the glyoxylate cycle and gluconeogenesis. In Arabidopsis, there  
503 is a sharp increase in the abundance of proteins involved in  $\beta$ -oxidation and the  
504 glyoxylate cycle during the first 36 h and a decrease thereafter. In *P. patens* the levels  
505 of these enzymes are already much higher in dry spores, slightly increase during spore  
506 germination, and are still high after 72 h. In vegetative tissues, the levels are then much  
507 lower. The result indicates that also *P. patens* spores, likewise to Arabidopsis seedlings,  
508 utilize a large amount of TAG for sugar production but that this process takes longer.

509 Contrary to that, the way the fatty acids are released from TAG might be different in both  
510 species. In Arabidopsis, the majority of TAG is hydrolyzed by the lipase SUGAR-  
511 DEPENDENT 1 and its homolog (Kelly *et al.*, 2011), while the knockout of the major  
512 OBL-type TAG lipase does not affect TAG breakdown during seed germination (Müller  
513 and Ischebeck, 2018). The physiological function of this protein family is therefore mostly  
514 unclear in Arabidopsis, but a homolog in tomato (*Solanum lycopersicum*) is involved in  
515 the production of oxylipin-derived volatiles (Garbowicz *et al.*, 2018; Ischebeck *et al.*,  
516 2020).

517 The genome of *P. patens* codes for four homologs of SDP1, but none of these were  
518 found in our proteomic dataset and their expression levels are also more than 100-times  
519 lower than for the lipases of the OBL family. Furthermore, based on this study, the OBL  
520 Pp3c3\_4690V3 is the most abundant protein indicating a pivotal role in TAG  
521 degradation. Based on the findings of Garbowicz *et al.*, a plausible function of OBLs in  
522 germinating spores is the liberation of LOX-oxygenated fatty acids, especially  
523 considering that a putative cytosolic 13-S LOX was found to be the second most  
524 abundant protein. Interestingly, there is also evidence pointing towards a catabolic role  
525 of at least the LOX genes in cucumber seedlings and olive pollen tubes (Matsui *et al.*,  
526 1999; Feussner *et al.*, 1997; Feussner *et al.*, 2001; Zienkiewicz *et al.*, 2013). In the  
527 present case, an analogous process could apply, using a 13-S LOX enzyme to first  
528 generate a 13S-hydroperoxide-octadecadienoic acid (13-HPODE) from linoleic acid,  
529 which gets then reduced to 13S-hydroxy-octadecadienoic acid (13-HODE) and released  
530 to the cytosol and glyoxysomes for  $\beta$ -oxidation. The cucumber 13-LOX was described  
531 as LD associated in cotyledons (Feußner and Kindl, 1992; Hause *et al.*, 2000), which  
532 was also observed in cotyledons and pollen tubes of olive (Zienkiewicz *et al.*, 2013;  
533 Zienkiewicz *et al.*, 2014) and sunflower seeds (Yadav and Bhatla, 2011). Similarly, the  
534 main *P. patens* LOX could be associated with LDs and drive TAG breakdown in  
535 combination with the OBLs.

536

#### 537 **Phase 4 Establishment of the photosynthetic apparatus**

538 The constitution of the photosynthetic apparatus requires three main components: i) the  
539 proteins and enzymes involved in the light and the dark reactions, ii) the pigments of the  
540 light-harvesting complexes, and iii) the membrane lipids of the thylakoids. All these  
541 components are missing or strongly reduced in Arabidopsis seeds (Li *et al.*, 2017;  
542 Kretschmar *et al.*, 2020; Kehelpannala *et al.*, 2021), likely in part because desiccation

543 increases the production of reactive oxygen species during the light reaction (Oliver *et*  
544 *al.*, 2020). Proteins required for the light reaction and the Calvin cycle are almost absent  
545 in dry seeds and start to strongly increase 24 h after imbibition, with the increase maybe  
546 not being completed after 72 h. In *P. patens*, the levels in spores are already a little bit  
547 higher, and start to increase after 24 h, but are still much lower after 72 h than in  
548 protonema and gametophore tissues indicating that the build-up is far from being  
549 completed. The increase in Arabidopsis goes in hand with degradation of storage  
550 proteins that could deliver amino acids for this purpose. *P. patens* does not have obvious  
551 storage proteins, but the degradation of highly abundant enzymes such as OBLs and  
552 LOXs could provide amino acids for the synthesis of photosynthesis-related proteins.  
553 Proteins involved in chlorophyll synthesis start to increase after 24 h in both Arabidopsis  
554 and *P. patens* and continuously increase. In Arabidopsis, a similar trend is observed for  
555 proteins involved in fatty acid synthesis. Interestingly, this is not the case in *P. patens*,  
556 where fatty acid synthesis is highest in spores and then decreases with the levels. This  
557 difference indicates that Arabidopsis synthesizes most of the fatty acids required for  
558 membrane lipid synthesis *de novo*, while *P. patens* might make use of fatty acids derived  
559 from TAG breakdown

560

## 561 **EXPERIMENTAL PROCEDURES**

### 562 **Plant Material**

563 *Physcomitrium patens* of the ecotype Reute was grown as previously described  
564 (Gömann *et al.*, 2021). Protonema cultivated on BCDAT medium were harvested 8 days  
565 after propagation, and full-grown gametophores cultivated on BCD medium 35 days after  
566 propagation. Sporophytes were induced as previously described (Hiss *et al.*, 2017) and  
567 individually picked. Spores were placed in water to germinate and harvested after 15  
568 minutes (0 h), 8 h, 24 h, 48 h, and 72 h by spinning down the spores and removing the  
569 water. After harvesting, the material was flash-frozen in liquid nitrogen and stored at -  
570 80°C.

### 571 **Isolation of protein fractions from moss.**

572 Frozen spores were ground in a liquid nitrogen-cooled shaking mill (Retsch, Haan,  
573 Germany) with three 2.5 mm steel beads at 30 Hz for 2 minutes. Subsequently, 30-50  
574 µL protein solubilization buffer (6 M Urea, 5% SDS, 2 mM phenylmethylsulfonyl fluoride)  
575 was added, followed by flash freezing in liquid nitrogen. Protonema and gametophore

576 material was ground in a liquid nitrogen-cooled mortar. Following, 40 mg material was  
577 added into 100  $\mu$ L protein solubilization buffer and thawed on ice before flash freezing  
578 in liquid nitrogen again. All samples were then treated in an ice-cooled ultrasonic bath  
579 for 15 minutes and centrifuged at 4°C for 10 min at 19.000 g. The supernatant was  
580 collected and used for further processing.

581

### 582 **Preparation of peptide samples and LC-MS/MS.**

583 All proteomic metadata can be found in Table S1. For the determination of protein  
584 concentration, the Pierce BCA protein assay kit (Thermo Fisher Scientific, Waltham, MA,  
585 USA) was used according to the manufacturer's instructions. In the following, 20  $\mu$ g  
586 (germinating spores) or 30  $\mu$ g (protonema/gametophores) of protein per replicate were  
587 subjected to electrophoresis on an SDS-acrylamide gel, with the run being halted once  
588 the proteins had entered the separation gel by approximately 3-5 mm. Individually  
589 excised gel pieces were processed and tryptically digested as previously described  
590 (Shevchenko *et al.*, 2006). Subsequently, the resulting peptides were desalted via a  
591 custom Empore™ Octadecyl C18 47 mm extraction disks 2215 (Supelco, St. Paul, MN,  
592 USA) filled columns as described elsewhere (Rappsilber *et al.*, 2007). Purified peptides  
593 were suspended in 20  $\mu$ L LC-MS sample buffer (2% acetonitrile, 0.1% formic acid) prior  
594 to applying it to reverse phase LC-based peptide separation on an RSLCnano Ultimate  
595 3000 system (Thermo Fisher Scientific). Peptides were loaded with 0.07% trifluoroacetic  
596 acid onto a C18 Acclaim PepMap 100 pre-column (100  $\mu$ m x 2 cm, 5  $\mu$ m particle size,  
597 100 Å pore size; Thermo Fisher Scientific) set to a flow rate of 20  $\mu$ L/min for 3 min. The  
598 peptides were then separated analytically using a C18 Acclaim PepMap RSLC column  
599 (75  $\mu$ m x 50 cm, 2  $\mu$ m particle size, 100 Å pore size; Thermo Fisher Scientific) at a flow  
600 rate of 300 nL/min. The solvent composition underwent a gradient change, initiating at  
601 96% solvent A containing 0.1% formic acid and 4% solvent B consisting of 80%  
602 acetonitrile and 0.1% formic acid. Over a period of 94 min, the proportion of solvent B  
603 increased progressively to 10% within 2 min, followed by a further increase to 30% within  
604 the subsequent 58 min. This was followed by an additional rise to 45% solvent B within  
605 the following 22 minutes, culminating in a peak of 90% solvent B in the last 12 minutes  
606 of the gradient. Optima-grade (Thermo Fisher Scientific) solvents and acids were used  
607 for all LC-MS experiments. Eluting peptides were ionized on-line by nano-electrospray  
608 (nESI) using a Nanospray Flex Ion Source (Thermo Fisher Scientific) at 1.5 kV (liquid  
609 junction) and transferred to a Q Exactive HF mass spectrometer (Thermo Fisher



610 Scientific). Full scans were recorded in a mass range of 300-1650 m/z at a resolution of  
611 30,000, followed by data-dependent top 10 HCD fragmentation at a resolution of 15,000  
612 (dynamic exclusion enabled). XCalibur 4.0 software (Thermo Fisher Scientific) was used  
613 for the LC-MS method programming and data acquisition.

614

#### 615 **Proteomic RAW data processing**

616 Proteomic RAW data were processed in MaxQuant 2.0.3.1 software (Cox and Mann,  
617 2008; Tyanova, Temu and Cox, 2016) on default settings except enabling 'Label free  
618 quantification - LFQ', 'iBAQ', 'FTMS recalibration', 'Match between runs', and set  
619 'Intensity determination' to 'Total sum'. The *P. patens* primary transcript protein file (v3.3,  
620 DOE-JGI, <https://phytozome-next.jgi.doe.gov/>; (Goodstein *et al.*, 2012; Lang *et al.*,  
621 2018)) was used as a reference library. Filtering and relative iBAQ (riBAQ) normalization  
622 in per mill from the resulting MaxQuant proteinGroups file was performed in Perseus  
623 1.6.2.2 (Tyanova, Temu, Sinitcyn, *et al.*, 2016) as previously described (Horn *et al.*,  
624 2021).

625

#### 626 **BLAST**

627 In order to identify the closest homologs of *P. patens* proteins in Arabidopsis, a BLASTp  
628 approach was employed (Altschul *et al.*, 1990; Gish and States, 1993). Here, the  
629 *P. patens* primary transcript protein file (v3.3, DOE-JGI, [https://phytozome-](https://phytozome-next.jgi.doe.gov/)  
630 [next.jgi.doe.gov/](https://phytozome-next.jgi.doe.gov/) (Goodstein *et al.*, 2012; Lang *et al.*, 2018)) was used as a query against  
631 the Arabidopsis TAIR10 primary transcript protein release (Lamesch *et al.*, 2012).  
632 BLAST software release 2.11.0+ was used, and an E-value cutoff of  $< 10^{-5}$  was set.

633

#### 634 **Data analysis**

635 PCA plots, hierarchical clusters, and volcano plots were created with Perseus 1.6.2.2  
636 (Tyanova, Temu, Sinitcyn, *et al.*, 2016). The PCA was created using standard settings.  
637 The hierarchical clustering was performed based on euclidean distance for row clusters  
638 and euclidean (Figure 1) or Pearson distance (Figure 3) for column clusters. Data was  
639 preprocessed with k-means, and the number of iterations was set to 10 with 100 restarts.  
640 For the volcano plots, data were imputed with a width of 0.3 and a downshift of 1.8.  
641 Significance was assessed via *t*-test, setting the number of randomizations to 250, the  
642 false discovery rate to 0.05, and the  $S_0$  to 0.1.

643 For GO term analysis, all *P. patens* accessions were assigned an Arabidopsis accession  
644 based on homology determined by the BLAST search (lowest E-value). In case several  
645 *P. patens* were assigned to the same Arabidopsis protein, and their riBAQ values were  
646 added. Then, these values were assigned to each AGI-accession GO-term pair  
647 (retrieved from www.arabidopsis.org 1<sup>st</sup> of September 2022), and the protein abundance  
648 in each GO-term was determined by adding the riBAQ of all the proteins therein.

649

### 650 **Phylogenetic analysis of LEA proteins**

651 To identify homologs of LEA proteins, the previously identified LEA proteins from *A.*  
652 *thaliana* (Hundertmark and Hinch, 2008) were used as a query in a BLASTp against a  
653 protein database with an e value cutoff  $< 10^{-5}$ . The database of (predicted) proteins was  
654 assembled from the genomes of *Anthoceros agrestis* (Li *et al.*, 2020), *Amborella*  
655 *trichopoda* (AMBORELLA GENOME PROJECT *et al.*, 2013), *A. thaliana* (Lamesch *et*  
656 *al.*, 2012), *Azolla filiculoides* (Li *et al.*, 2018), *Brachypodium distachyon* (Vogel *et al.*,  
657 2010), *Chlorokybus melkonianii* (Wang *et al.*, 2020; Irisarri *et al.*, 2021), *Chara braunii*  
658 (Nishiyama *et al.*, 2018), *Chlamydomonas reinhardtii* (Merchant *et al.*, 2007),  
659 *Coccomyxa subellipsoidea* (Blanc *et al.*, 2012), *Ceratopteris richardii* (Marchant *et al.*,  
660 2022), *C. esculentus* (Yang *et al.*, 2016; Niemeyer *et al.*, 2022), *Gnetum montanum*  
661 (Wan *et al.*, 2018), *Isoetes taiwaniensis* (Wickell *et al.*, 2021), *Klebsormidium nitens*  
662 (Hori *et al.*, 2014), *Marchantia polymorpha* (Bowman *et al.*, 2017), *Mesotaenium*  
663 *endlicherianum* (Cheng *et al.*, 2019; Dadras *et al.*, 2022), *O. sativa* (Ouyang *et al.*,  
664 2007), *Picea abies* (Nystedt *et al.*, 2013), *Penium margaritaceum* (Jiao *et al.*, 2020),  
665 *P. patens* (Lang *et al.*, 2018), *Selaginella moellendorffii* (Banks *et al.*, 2011), *Spirogloea*  
666 *muscicola* (Cheng *et al.*, 2019), and *Ulva mutabilis* (De Clerck *et al.*, 2018). Sequences  
667 (Data S11) were aligned with MAFFT (Kato and Standley, 2013) using the alignment  
668 method L-INS-i. The the maximum-likelihood tree was constructed with IQ-Tree (Minh  
669 *et al.*, 2020) in which the WAG+I+G4 model was selected as most suitable by  
670 ModelFinder (Kalyaanamoorthy *et al.*, 2017) utilizing the option -madd LG4M, LG4X. The  
671 branch support was assessed with ultrafast bootstrap (Hoang *et al.*, 2018). All data on  
672 the tree construction can be found in Data S12-14. The visualization was done in iTOL  
673 version 2.0.3 (Letunic and Bork, 2007), and the given subcellular locations of the  
674 proteins were predicted with TargetP-2.0 (Armenteros *et al.*, 2019).

675

676

## 677 **Analysis of LEA expression data**

678 LEA mRNA transcripts from different RNAseq databases were retrieved for three  
679 species: TraVa for *Arabidopsis thaliana* ecotype Col-0 (Transcriptome Variation Analysis  
680 database, <http://travadb.org/>; (Klepikova *et al.*, 2016), RED for *Oryza sativa* cv.  
681 Nipponbare (Rice Expression Database, <http://expression.ic4r.org/>; (Xia *et al.*, 2017),  
682 and PEATmoss for *Physcomitrium patens* ecotypes Gransden and Reute  
683 (<https://peatmoss.plantcode.cup.uni-freiburg.de/>, (Fernandez-Pozo *et al.*, 2020)). LEA  
684 protein data were acquired from *A. thaliana* ecotype Col-0 (Kretzschmar *et al.*, 2020).  
685 2020), *Cyperus esculentus* L. var. *sativus* (Niemeyer *et al.*, 2022), and *P. patens* ecotype  
686 Reute (this publication).

687 Due to varying data availability, we chose different desiccation- and vegetative-  
688 associated stages for each species. For *A. thaliana* ecotype Col-0 transcripts, we  
689 compared seeds of the first yellowing silique, dry seeds, opened anthers, and anthers  
690 from mature unopened flowers with germinating seeds three days after imbibition,  
691 seedling cotyledons, seedling roots, and whole mature leaves; while for protein  
692 expression data, seeds 30 min after imbibition and seedlings 60 h after imbibition were  
693 chosen. For *P. patens* transcripts, both Gransden and Reute ecotypes were included;  
694 namely, we compared dry spores (Gransden) and brown sporophytes (Reute) against  
695 protonema (Gransden and Reute), gametophores (Gransden and Reute), adult  
696 gametophores (Reute) and leaflets (Gransden). For the protein analysis of *P. patens*  
697 ecotype Reute, dry spores were compared with gametophores and protonema. For *O.*  
698 *sativa* cv. Nipponbare transcripts, mature seeds and anthers from mature unopened  
699 flowers were contrasted to 14-day-old shoots, roots, and leaves. No protein expression  
700 data were analyzed for rice. Lastly, for *C. esculentum* L. var. *sativus* – for which only  
701 proteome data was used - dry tubers were compared to leaves and roots.

702 We determined the ratio between the highest raw value of the before-mentioned  
703 desiccation- and vegetative-associated stages. An LEA transcript was considered  
704 desiccation or vegetative stage-specific when the ratio of the raw normalized mRNA  
705 levels were at least 1.5-times higher in desiccation-associated tissues compared to  
706 vegetative-associated tissues and *vice versa*. For protein data, expression was  
707 considered vegetative or desiccated-specific, if the expression level ratio was at least 2-  
708 times or 4-times increased, respectively.

709

710

## 711 **Prediction of LOX subcellular targeting and substrate affinity**

712 The annotation of *P. patens* LOX and AOS proteins is based on their closest homologs  
713 in Arabidopsis and was determined by a BLASTp search. The substrate specificity for  
714 the LOX proteins was assigned based on a previous characterization study (Anterola et  
715 al., 2009). The subcellular localization of the identified LOX and AOS proteins was  
716 predicted using three different computational tools, namely LOCALIZER  
717 (<https://localizer.csiro.au/>; (Sperschneider *et al.*, 2017)), Green Targeting Predictor  
718 (<https://plantcode.cup.uni-freiburg.de/plantco/predloc/>; (Fuss *et al.*, 2013)), and TargetP  
719 2.0 (<https://services.healthtech.dtu.dk/services/TargetP-2.0/>; (Armenteros *et al.*, 2019)).  
720 The consensus prediction from two out of three tools was considered as an indication of  
721 the subcellular localization of the LOX proteins.

722

## 723 **ACKNOWLEDGMENTS:**

724 PWN, GB, JdV and TI thank the German research foundation (DFG, Grants IS 273/7-1,  
725 IS 273/9-1, IS 273/10-1, IRTG 2172 PRoTECT, BR1502-15-1, SPP 2237 MAdLand, VR  
726 132/4-1); JdV thanks the European Research Council for funding (Grant Agreement No.  
727 852725; ERC-StG 'TerreStriAL') under the European Union's Horizon 2020 research  
728 and innovation program. OV and KS on behalf of the Service Unit LCMS Protein  
729 Analytics of the Göttingen Center for Molecular Biosciences (GZMB) thank the DFG for  
730 funding (INST 186/1230-1 FUGG to Stefanie Pöggeler). We would like to thank Marco  
731 Göttig for the collection of spore capsules and Dr. Tegan Haslam for the maintenance  
732 of the moss cultivation and careful proofreading of the manuscript's introduction.

733

## 734 **AUTHOR CONTRIBUTIONS:**

735 L.H., P.W.N., G.H.B., M.S., J.d.V., S.R. and T.I. designed the work, L.H., P.W.N., K.S.,  
736 J. Z., D.B., O.V., J.D., J.d.V., and T.I. performed research, L.H., P.W.N., K.S., J.d.V.,  
737 and T.I. analyzed data, and L.H., P.W.N., J.Z., D.B., J.D. and T.I. wrote the manuscript.  
738 All authors critically read and revised the manuscript and approved the final version.

739

## 740 **CONFLICT-OF-INTEREST STATEMENT:**

741 The authors declare no conflict of interest.

742

743

744 **REFERENCES:**

- 745 **Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J.** (1990) Basic local  
746 alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.
- 747 **Amara, I., Zaidi, I., Masmoudi, K., Ludevid, M.D., Pagès, M., Goday, A. and Brini, F.** (2014)  
748 Insights into late embryogenesis abundant (LEA) proteins in plants: from structure to the  
749 functions. *American Journal of Plant Sciences*, **05**, 3440.
- 750 **AMBORELLA GENOME PROJECT, Albert, V.A., Barbazuk, W.B., et al.** (2013) The *Amborella*  
751 genome and the evolution of flowering plants. *Science*, **342**, 1241089.
- 752 **Andreou, A. and Feussner, I.** (2009) Lipoxygenases – Structure and reaction mechanism.  
753 *Phytochemistry*, **70**, 1504–1510.
- 754 **Anterola, A., Göbel, C., Hornung, E., Sellhorn, G., Feussner, I. and Grimes, H.** (2009)  
755 *Physcomitrella patens* has lipoxygenases for both eicosanoid and octadecanoid  
756 pathways. *Phytochemistry*, **70**, 40–52.
- 757 **Armenteros, J.J.A., Salvatore, M., Emanuelsson, O., Winther, O., Heijne, G. von, Elofsson,**  
758 **A. and Nielsen, H.** (2019) Detecting sequence signals in targeting peptides using deep  
759 learning. *Life Science Alliance*, **2**, e201900429.
- 760 **Arsovski, A.A., Popma, T.M., Haughn, G.W., Carpita, N.C., McCann, M.C. and Western, T.L.**  
761 (2009) AtBXL1 encodes a bifunctional  $\beta$ -d-xylosidase/ $\alpha$ -l-arabinofuranosidase required  
762 for pectic arabinan modification in *Arabidopsis* mucilage secretory cells. *Plant*  
763 *Physiology*, **150**, 1219–1234.
- 764 **Artur, M.A.S., Zhao, T., Ligterink, W., Schranz, E. and Hilhorst, H.W.M.** (2019) Dissecting  
765 the genomic diversification of late embryogenesis abundant (LEA) protein gene families  
766 in plants. *Genome Biology and Evolution*, **11**, 459–471.
- 767 **Aubert, Y., Vile, D., Pervent, M., Aldon, D., Ranty, B., Simonneau, T., Vavasseur, A. and**  
768 **Galaud, J.-P.** (2010) RD20, a stress-inducible caleosin, participates in stomatal control,  
769 transpiration and drought tolerance in *Arabidopsis thaliana*. *Plant and Cell Physiology*,  
770 **51**, 1975–1987.
- 771 **Banks, J.A., Nishiyama, T., Hasebe, M., et al.** (2011) The *Selaginella* genome identifies  
772 genetic changes associated with the evolution of vascular plants. *Science*, **332**, 960–  
773 963.
- 774 **Baud, S., Dichow, N.R., Kelemen, Z., et al.** (2009) Regulation of HSD1 in seeds of *Arabidopsis*  
775 *thaliana*. *Plant and Cell Physiology*, **50**, 1463–1478.
- 776 **Berardini, T.Z., Mundodi, S., Reiser, L., et al.** (2004) Functional annotation of the *Arabidopsis*  
777 genome using controlled vocabularies. *Plant Physiology*, **135**, 745–755.
- 778 **Blanc, G., Agarkova, I., Grimwood, J., et al.** (2012) The genome of the polar eukaryotic  
779 microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biology*,  
780 **13**, R39.
- 781 **Bowman, J.L., Kohchi, T., Yamato, K.T., et al.** (2017) Insights into land plant evolution  
782 garnered from the *Marchantia polymorpha* genome. *Cell*, **171**, 287–304.e15.

- 783 **Brocard, L., Immel, F., Coulon, D., et al.** (2017) Proteomic analysis of lipid droplets from  
784 *Arabidopsis* aging leaves brings new insight into their biogenesis and functions. *Frontiers*  
785 *in Plant Science*, **8**.
- 786 **Chakrabortee, S., Tripathi, R., Watson, M., Schierle, G.S.K., P. Kurniawan, D., F. Kaminski,**  
787 **C., J. Wise, M. and Tunnacliffe, A.** (2012) Intrinsically disordered proteins as molecular  
788 shields. *Molecular BioSystems*, **8**, 210–219.
- 789 **Chaturvedi, P., Ischebeck, T., Egelhofer, V., Lichtscheidl, I. and Weckwerth, W.** (2013) Cell-  
790 specific analysis of the tomato pollen proteome from pollen mother cell to mature pollen  
791 provides evidence for developmental priming. *Journal of Proteome Research*, **12**, 4892–  
792 4903.
- 793 **Cheng, S., Xian, W., Fu, Y., et al.** (2019) Genomes of subaerial *Zygnematophyceae* provide  
794 insights into land plant evolution. *Cell*, **179**, 1057-1067.e14.
- 795 **Chi, Y.H., Koo, S.S., Oh, H.T., et al.** (2019) The physiological functions of universal stress  
796 proteins and their molecular mechanism to protect plants from environmental stresses.  
797 *Frontiers in Plant Science*, **10**.
- 798 **Cho, S.H., Hoang, Q.T., Kim, Y.Y., Shin, H.Y., Ok, S.H., Bae, J.M. and Shin, J.S.** (2006)  
799 Proteome analysis of gametophores identified a metallothionein involved in various  
800 abiotic stress responses in *Physcomitrella patens*. *Plant Cell Reports*, **25**, 475–488.
- 801 **Clouse, S.D. and Sasse, J.M.** (1998) Brassinosteroids: Essential regulators of plant growth and  
802 development. *Annual Review of Plant Physiology and Plant Molecular Biology*, **49**, 427–  
803 451.
- 804 **Costa, M.-C.D., Artur, M.A.S., Maia, J., et al.** (2017) A footprint of desiccation tolerance in the  
805 genome of *Xerophyta viscosa*. *Nature Plants*, **3**, 1–10.
- 806 **Cove, D.** (2005) The moss *Physcomitrella patens*. *Annual Review of Genetics*, **39**, 339–358.
- 807 **Cox, J. and Mann, M.** (2008) MaxQuant enables high peptide identification rates, individualized  
808 p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature*  
809 *Biotechnology*, **26**, 1367–1372.
- 810 **Crowe, A.J., Abenes, M., Plant, A. and Moloney, M.M.** (2000) The seed-specific  
811 transactivator, ABI3, induces oleosin gene expression. *Plant Science*, **151**, 171–181.
- 812 **Cui, S., Hu, J., Guo, S., Wang, J., Cheng, Y., Dang, X., Wu, L. and He, Y.** (2012) Proteome  
813 analysis of *Physcomitrella patens* exposed to progressive dehydration and rehydration.  
814 *Journal of Experimental Botany*, **63**, 711–726.
- 815 **Cuming, A.C.** (2019) Evolution of ABA signaling pathways. In M. Seo and A. Marion-Poll, eds.  
816 *Advances in Botanical Research*. Academic Press, pp. 281–313.
- 817 **Cuming, A.C., Cho, S.H., Kamisugi, Y., Graham, H. and Quatrano, R.S.** (2007) Microarray  
818 analysis of transcriptional responses to abscisic acid and osmotic, salt, and drought  
819 stress in the moss, *Physcomitrella patens*. *New Phytologist*, **176**, 275–287.
- 820 **Curien, G., Giustini, C., Montillet, J.-L., Mas-y-Mas, S., Cobessi, D., Ferrer, J.-L., Matringe,**  
821 **M., Grechkin, A. and Rolland, N.** (2016) The chloroplast membrane associated  
822 ceQORH putative quinone oxidoreductase reduces long-chain, stress-related oxidized  
823 lipids. *Phytochemistry*, **122**, 45–55.

- 824 **Dadras, A., Fuerst-Jansen, J.M., Darienko, T., et al.** (2022) Environmental gradients reveal  
825 stress hubs predating plant terrestrialization. *bioRxiv*, 2022.10.17.512551.
- 826 **Daku, R.M., Rabbi, F., Buttigieg, J., Coulson, I.M., Horne, D., Martens, G., Ashton, N.W.**  
827 **and Suh, D.-Y.** (2016) PpASCL, the *Physcomitrella patens* anther-specific chalcone  
828 synthase-like enzyme implicated in sporopollenin biosynthesis, is needed for integrity of  
829 the moss spore wall and spore viability. *PLOS ONE*, **11**, e0146817.
- 830 **De Clerck, O., Kao, S.-M., Bogaert, K.A., et al.** (2018) insights into the evolution of  
831 multicellularity from the sea lettuce genome. *Current Biology*, **28**, 2921-2933.e5.
- 832 **Delahaie, J., Hundertmark, M., Bove, J., Leprince, O., Rogniaux, H. and Buitink, J.** (2013)  
833 LEA polypeptide profiling of recalcitrant and orthodox legume seeds reveals ABI3-  
834 regulated LEA protein abundance linked to desiccation tolerance. *Journal of*  
835 *Experimental Botany*, **64**, 4559–4573.
- 836 **Deruyffelaere, C., Purkrtova, Z., Bouchez, I., Collet, B., Cacas, J.-L., Chardot, T., Gallois,**  
837 **J.-L. and D'Andrea, S.** (2018) PUX10 is a CDC48A adaptor protein that regulates the  
838 extraction of ubiquitinated oleosins from seed lipid droplets in *Arabidopsis*. *Plant Cell*, **30**,  
839 2116–2136.
- 840 **Engel, P.P.** (1968) The induction of biochemical and morphological mutants in the moss  
841 *Physcomitrella patens*. *American Journal of Botany*, **55**, 438–446.
- 842 **Farrant, J.M., Lehner, A., Cooper, K. and Wiswedel, S.** (2009) Desiccation tolerance in the  
843 vegetative tissues of the fern *Mohria caffrorum* is seasonally regulated. *The Plant*  
844 *Journal*, **57**, 65–79.
- 845 **Fernandez-Pozo, N., Haas, F.B., Meyberg, R., et al.** (2020) PEATmoss (Physcomitrella  
846 Expression Atlas Tool): a unified gene expression atlas for the model plant  
847 *Physcomitrella patens*. *The Plant Journal*, **102**, 165–177.
- 848 **Fernández-Santos, R., Izquierdo, Y., López, A., Muñiz, L., Martínez, M., Cascón, T.,**  
849 **Hamberg, M. and Castresana, C.** (2020) Protein profiles of lipid droplets during the  
850 hypersensitive defense response of *Arabidopsis* against *Pseudomonas* infection. *Plant*  
851 *and Cell Physiology*, **61**, 1144–1157.
- 852 **Feussner, I., Balkenhohl, T.J., Porzel, A., Kühn, H. and Wasternack, C.** (1997) structural  
853 elucidation of oxygenated storage lipids in cucumber cotyledons: implication of lipid body  
854 lipoyxygenase in lipid mobilization during germination. *Journal of Biological Chemistry*,  
855 **272**, 21635–21641.
- 856 **Feußner, I. and Kindl, H.** (1992) A lipoyxygenase is the main lipid body protein in cucumber and  
857 soybean cotyledons during the stage of triglyceride mobilization. *FEBS Letters*, **298**,  
858 223–225.
- 859 **Feussner, I., Kühn, H. and Wasternack, C.** (2001) Lipoyxygenase-dependent degradation of  
860 storage lipids. *Trends in Plant Science*, **6**, 268–273.
- 861 **Fuss, J., Liegmann, O., Krause, K. and Rensing, S.A.** (2013) Green Targeting Predictor and  
862 Ambiguous Targeting Predictor 2: the pitfalls of plant protein targeting prediction and of  
863 transient protein expression in heterologous systems. *New Phytologist*, **200**, 1022–1033.
- 864 **Gaff, D.F. and Oliver, M.** (2013) The evolution of desiccation tolerance in angiosperm plants: a  
865 rare yet common phenomenon. *Functional Plant Biology*, **40**, 315–328.

- 866 **Garbowicz, K., Liu, Z., Alseekh, S., et al.** (2018) Quantitative trait loci analysis identifies a  
867 prominent gene involved in the production of fatty acid-derived flavor volatiles in tomato.  
868 *Molecular Plant*, **11**, 1147–1165.
- 869 **Gidda, S.K., Park, S., Pyc, M., et al.** (2016) Lipid droplet-associated proteins (LDAPs) are  
870 required for the dynamic regulation of neutral lipid compartmentation in plant cells. *Plant*  
871 *Physiology*, **170**, 2052–2071.
- 872 **Giraudat, J., Hauge, B.M., Valon, C., Smalle, J., Parcy, F. and Goodman, H.M.** (1992)  
873 Isolation of the Arabidopsis ABI3 gene by positional cloning. *The Plant Cell*, **4**, 1251–  
874 1261.
- 875 **Gish, W. and States, D.J.** (1993) Identification of protein coding regions by database similarity  
876 search. *Nature Genetics*, **3**, 266–272.
- 877 **Gömann, J., Herrfurth, C., Zienkiewicz, K., Haslam, T.M. and Feussner, I.** (2021)  
878 Sphingolipid  $\Delta 4$ -desaturation is an important metabolic step for glycosylceramide  
879 formation in *Physcomitrium patens*. *Journal of Experimental Botany*, **72**, 5569–5583.
- 880 **Gómez-Garay, A., Gabriel y Galán, J.M., Cabezuelo, A., Pintos, B., Prada, C. and Martín, L.**  
881 (2018) Ecological significance of brassinosteroids in three temperate ferns. In H.  
882 Fernández, ed. *Current Advances in Fern Research*. Cham: Springer International  
883 Publishing, pp. 453–466.
- 884 **Gong, Z., Dong, C.-H., Lee, H., Zhu, J., Xiong, L., Gong, D., Stevenson, B. and Zhu, J.-K.**  
885 (2005) A DEAD box RNA helicase is essential for mRNA export and important for  
886 development and stress responses in *Arabidopsis*. *The Plant Cell*, **17**, 256–267.
- 887 **Goodstein, D.M., Shu, S., Howson, R., et al.** (2012) Phytozome: a comparative platform for  
888 green plant genomics. *Nucleic Acids Research*, **40**, D1178–D1186.
- 889 **Guzha, A., McGee, R., Scholz, P., et al.** (2022) Cell wall-localized BETA-XYLOSIDASE4  
890 contributes to immunity of *Arabidopsis* against *Botrytis cinerea*. *Plant Physiology*, **189**,  
891 1794–1813.
- 892 **Guzha, A., Whitehead, P., Ischebeck, T. and Chapman, K.D.** (2023) Lipid droplets: packing  
893 hydrophobic molecules within the aqueous cytoplasm. *Annual Review of Plant Biology*,  
894 **74**, 195–223.
- 895 **Hanano, A., Blée, E. and Murphy, D.J.** (2023) Caleosin/Peroxygenases: multifunctional  
896 proteins in plants. *Annals of Botany*, **131**, 387–409.
- 897 **Hause, B., Weichert, H., Höhne, M., Kindl, H. and Feussner, I.** (2000) Expression of cucumber  
898 lipid-body lipooxygenase in transgenic tobacco: lipid-body lipooxygenase is correctly  
899 targeted to seed lipid bodies. *Planta*, **210**, 708–714.
- 900 **Hiss, M., Laule, O., Meskauskiene, R.M., et al.** (2014) Large-scale gene expression profiling  
901 data for the model moss *Physcomitrella patens* aid understanding of developmental  
902 progression, culture and stress conditions. *The Plant Journal*, **79**, 530–539.
- 903 **Hiss, M., Meyberg, R., Westermann, J., Haas, F.B., Schneider, L., Schallenberg-Rüdinger,**  
904 **M., Ullrich, K.K. and Rensing, S.A.** (2017) Sexual reproduction, sporophyte  
905 development and molecular variation in the model moss *Physcomitrella patens*:  
906 introducing the ecotype Reute. *The Plant Journal*, **90**, 606–620.



- 907 **Hoang, D.T., Chernomor, O., Haeseler, A. von, Minh, B.Q. and Vinh, L.S.** (2018) UFBoot2:  
908 improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, **35**,  
909 518–522.
- 910 **Hohe, A., Rensing, S.A., Mildner, M., Lang, D. and Reski, R.** (2002) Day length and  
911 temperature strongly influence sexual reproduction and expression of a novel MADS-box  
912 gene in the moss *Physcomitrella patens*. *Plant Biology*, **4**, 595–602.
- 913 **Hori, K., Maruyama, F., Fujisawa, T., et al.** (2014) *Klebsormidium flaccidum* genome reveals  
914 primary factors for plant terrestrial adaptation. *Nature Communications*, **5**, 3978.
- 915 **Horn, P.J., Chapman, K.D. and Ischebeck, T.** (2021) Isolation of Lipid Droplets for Protein and  
916 Lipid AnalysisLipid analysis. In D. Bartels and P. Dörmann, eds. *Plant Lipids: Methods  
917 and Protocols*. Methods in Molecular Biology. New York, NY: Springer US, pp. 295–320.
- 918 **Huang, A.H.C.** (1992) Oil bodies and oleosins in seeds. *Annual Review of Plant Biology*, **43**,  
919 177–200.
- 920 **Huang, C.-Y., Chung, C.-I., Lin, Y.-C., Hsing, Y.-I.C. and Huang, A.H.C.** (2009) Oil bodies and  
921 oleosins in *Physcomitrella* possess characteristics representative of early trends in  
922 evolution. *Plant Physiology*, **150**, 1192–1203.
- 923 **Hundertmark, M. and Hinch, D.K.** (2008) LEA (Late Embryogenesis Abundant) proteins and  
924 their encoding genes in *Arabidopsis thaliana*. *BMC Genomics*, **9**, 118.
- 925 **Irisarri, I., Darienko, T., Pröschold, T., Fürst-Jansen, J.M.R., Jamy, M. and Vries, J. de**  
926 (2021) Unexpected cryptic species among streptophyte algae most distant to land plants.  
927 *Proceedings of the Royal Society B: Biological Sciences*, **288**, 20212168.
- 928 **Ischebeck, T.** (2016) Lipids in pollen — They are different. *Biochimica et Biophysica Acta (BBA)*  
929 - *Molecular and Cell Biology of Lipids*, **1861**, 1315–1328.
- 930 **Ischebeck, T., Krawczyk, H.E., Mullen, R.T., Dyer, J.M. and Chapman, K.D.** (2020) Lipid  
931 droplets in plants and algae: Distribution, formation, turnover and function. *Seminars in  
932 Cell & Developmental Biology*, **108**, 82–93.
- 933 **Jiang, Z., Xu, G., Jing, Y., Tang, W. and Lin, R.** (2016) Phytochrome B and REVEILLE1/2-  
934 mediated signalling controls seed dormancy and germination in *Arabidopsis*. *Nat  
935 Commun*, **7**, 12377.
- 936 **Jiao, C., Sørensen, I., Sun, X., et al.** (2020) The *Penium margaritaceum* genome: hallmarks of  
937 the origins of land plants. *Cell*, **181**, 1097-1111.e12.
- 938 **Julca, I., Ferrari, C., Flores-Tornero, M., et al.** (2021) Comparative transcriptomic analysis  
939 reveals conserved programmes underpinning organogenesis and reproduction in land  
940 plants. *Nature Plants*, **7**, 1143–1159.
- 941 **Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Haeseler, A. von and Jermin, L.S.** (2017)  
942 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*,  
943 **14**, 587–589.
- 944 **Katoh, K. and Standley, D.M.** (2013) MAFFT Multiple Sequence Alignment Software version  
945 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–  
946 780.

- 947 **Kehelpannala, C., Rupasinghe, T., Pasha, A., Esteban, E., Hennessy, T., Bradley, D., Ebert,**  
948 **B., Provart, N.J. and Roessner, U.** (2021) An *Arabidopsis* lipid map reveals differences  
949 between tissues and dynamic changes throughout development. *The Plant Journal*, **107**,  
950 287–302.
- 951 **Kelly, A.A., Quettier, A.-L., Shaw, E. and Eastmond, P.J.** (2011) Seed storage oil mobilization  
952 is important but not essential for germination or seedling establishment in *Arabidopsis*.  
953 *Plant Physiology*, **157**, 866–875.
- 954 **Kenrick, P.** (2017) How land plant life cycles first evolved. *Science*, **358**, 1538–1539.
- 955 **Khandelwal, A., Cho, S.H., Marella, H., Sakata, Y., Perroud, P.-F., Pan, A. and Quatrano,**  
956 **R.S.** (2010) Role of ABA and ABI3 in desiccation tolerance. *Science*, **327**, 546–546.
- 957 **Klepikova, A.V., Kasianov, A.S., Gerasimov, E.S., Logacheva, M.D. and Penin, A.A.** (2016)  
958 A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based  
959 on RNA-seq profiling. *The Plant Journal*, **88**, 1058–1070.
- 960 **Kovacs, D., Kalmar, E., Torok, Z. and Tompa, P.** (2008) Chaperone activity of ERD10 and  
961 ERD14, two disordered stress-related plant proteins. *Plant Physiology*, **147**, 381–390.
- 962 **Krawczyk, H.E., Rotsch, A.H., Herrfurth, C., Scholz, P., Shomroni, O., Salinas-Riester, G.,**  
963 **Feussner, I. and Ischebeck, T.** (2022) Heat stress leads to rapid lipid remodeling and  
964 transcriptional adaptations in *Nicotiana tabacum* pollen tubes. *Plant Physiology*, **189**,  
965 490–515.
- 966 **Kretzschmar, F.K., Doner, N.M., Krawczyk, H.E., Scholz, P., Schmitt, K., Valerius, O.,**  
967 **Braus, G.H., Mullen, R.T. and Ischebeck, T.** (2020) Identification of low-abundance  
968 lipid droplet proteins in seeds and seedlings. *Plant Physiology*, **182**, 1326–1345.
- 969 **Kretzschmar, F.K., Mengel, L.A., Müller, A.O., Schmitt, K., Biersch, K.F., Valerius, O.,**  
970 **Braus, G.H. and Ischebeck, T.** (2018) PUX10 is a lipid droplet-localized scaffold protein  
971 that interacts with CELL DIVISION CYCLE48 and is involved in the degradation of lipid  
972 droplet proteins. *Plant Cell*, **30**, 2137–2160.
- 973 **Lamesch, P., Berardini, T.Z., Li, D., et al.** (2012) The Arabidopsis Information Resource  
974 (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, **40**, D1202–  
975 D1210.
- 976 **Lang, D., Ullrich, K.K., Murat, F., et al.** (2018) The *Physcomitrella patens* chromosome-scale  
977 assembly reveals moss genome structure and evolution. *The Plant Journal*, **93**, 515–  
978 533.
- 979 **Leprince, O., Aelst, A.C. van, Pritchard, H.W. and Murphy, D.J.** (1997) Oleosins prevent oil-  
980 body coalescence during seed imbibition as suggested by a low-temperature scanning  
981 electron microscope study of desiccation-tolerant and -sensitive oilseeds. *Planta*, **204**,  
982 109–119.
- 983 **Letunic, I. and Bork, P.** (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic  
984 tree display and annotation. *Bioinformatics*, **23**, 127–128.
- 985 **Leubner-Metzger, G.** (2003) Brassinosteroids promote seed germination. In S. Hayat and A.  
986 Ahmad, eds. *Brassinosteroids: Bioactivity and Crop Productivity*. Dordrecht: Springer  
987 Netherlands, pp. 119–128. Available at: [https://doi.org/10.1007/978-94-017-0948-4\\_5](https://doi.org/10.1007/978-94-017-0948-4_5)  
988 [Accessed May 30, 2023].

- 989 **Li, F., Asami, T., Wu, X., Tsang, E.W.T. and Cutler, A.J.** (2007) A putative hydroxysteroid  
990 dehydrogenase involved in regulating plant growth and development. *Plant Physiology*,  
991 **145**, 87–97.
- 992 **Li, F.-W., Brouwer, P., Carretero-Paulet, L., et al.** (2018) Fern genomes elucidate land plant  
993 evolution and cyanobacterial symbioses. *Nature Plants*, **4**, 460–472.
- 994 **Li, F.-W., Nishiyama, T., Waller, M., et al.** (2020) *Anthoceros* genomes illuminate the origin of  
995 land plants and the unique biology of hornworts. *Nature Plants*, **6**, 259–272.
- 996 **Li, Z., Wu, S., Chen, J., Wang, X., Gao, J., Ren, G. and Kuai, B.** (2017) NYEs/SGRs-mediated  
997 chlorophyll degradation is critical for detoxification during seed maturation in *Arabidopsis*.  
998 *The Plant Journal*, **92**, 650–661.
- 999 **Ligrone, R., Duckett, J.G. and Renzaglia, K.S.** (2012) Major transitions in the evolution of early  
1000 land plants: a bryological perspective. *Annals of Botany*, **109**, 851–871.
- 1001 **Lin, L.-J., Tai, S.S.K., Peng, C.-C. and Tzen, J.T.C.** (2002) Steroleosin, a sterol-binding  
1002 dehydrogenase in seed oil bodies. *Plant Physiology*, **128**, 1200–1211.
- 1003 **Liu, X., Wang, Z., Wang, L., Wu, R., Phillips, J. and Deng, X.** (2009) LEA 4 group genes from  
1004 the resurrection plant *Boea hygrometrica* confer dehydration tolerance in transgenic  
1005 tobacco. *Plant Science*, **176**, 90–98.
- 1006 **Loconte, H. and Stevenson, D.W.** (1990) Cladistics of the Spermatophyta. *Brittonia*, **42**, 197–  
1007 211.
- 1008 **López-Pozo, M., Fernández-Marín, B., García-Plazaola, J.I. and Ballesteros, D.** (2018)  
1009 Desiccation tolerance in ferns: from the unicellular spore to the multi-tissular sporophyte.  
1010 In H. Fernández, ed. *Current Advances in Fern Research*. Cham: Springer International  
1011 Publishing, pp. 401–426.
- 1012 **Luo, W., Komatsu, S., Abe, T., Matsuura, H. and Takahashi, K.** (2020) Comparative  
1013 proteomic analysis of wild-type *Physcomitrella patens* and an OPDA-deficient  
1014 *Physcomitrella patens* mutant with disrupted PpAOS1 and PpAOS2 genes after  
1015 wounding. *International Journal of Molecular Sciences*, **21**, 1417.
- 1016 **Lyall, R. and Gechev, T.** (2020) Multi-omics insights into the evolution of angiosperm  
1017 resurrection plants. In *Annual Plant Reviews online*. John Wiley & Sons, Ltd, pp. 77–110.
- 1018 **Mamaeva, A., Knyazev, A., Glushkevich, A. and Fesenko, I.** (2022) Quantitative proteomic  
1019 dataset of the moss *Physcomitrium patens* PSEP3 KO and OE mutant lines. *Data in*  
1020 *Brief*, **40**, 107715.
- 1021 **Manfre, A.J., LaHatte, G.A., Climer, C.R. and Marcotte, W.R.** (2008) Seed dehydration and  
1022 the establishment of desiccation tolerance during seed maturation is altered in the  
1023 *Arabidopsis thaliana* mutant *atem6-1*. *Plant and Cell Physiology*, **50**, 243–253.
- 1024 **Marchant, D.B., Chen, G., Cai, S., et al.** (2022) Dynamic genome evolution in a model fern.  
1025 *Nature Plants*, **8**, 1038–1051.
- 1026 **Matilla, A.J.** (2022) The orthodox dry seeds are alive: a clear example of desiccation tolerance.  
1027 *Plants*, **11**, 20.

- 1028 **Matsui, K., Hijiya, K., Tabuchi, Y. and Kajiwara, T.** (1999) Cucumber cotyledon lipoxygenase  
1029 during postgerminative growth. its expression and action on lipid bodies. *Plant*  
1030 *Physiology*, **119**, 1279–1288.
- 1031 **Merchant, S.S., Prochnik, S.E., Vallon, O., et al.** (2007) The *Chlamydomonas* genome reveals  
1032 the evolution of key animal and plant functions. *Science*, **318**, 245–250.
- 1033 **Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Haeseler, A.**  
1034 **von and Lanfear, R.** (2020) IQ-TREE 2: new models and efficient methods for  
1035 phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, **37**, 1530–  
1036 1534.
- 1037 **Morikawa, T., Saga, H., Hashizume, H. and Ohta, D.** (2009) CYP710A genes encoding sterol  
1038 C22-desaturase in *Physcomitrella patens* as molecular evidence for the evolutionary  
1039 conservation of a sterol biosynthetic pathway in plants. *Planta*, **229**, 1311–1322.
- 1040 **Müller, A.O. and Ischebeck, T.** (2018) Characterization of the enzymatic activity and  
1041 physiological function of the lipid droplet-associated triacylglycerol lipase AtOBL1. *New*  
1042 *Phytologist*, **217**, 1062–1076.
- 1043 **Murphy, D.J., Ross, J.H.E. and Pritchard, H.W.** (1995) Are oleosins only associated with oil  
1044 bodies from desiccation tolerant plant tissues? In J.-C. Kader and P. Mazliak, eds. *Plant*  
1045 *Lipid Metabolism*. Dordrecht: Springer Netherlands, pp. 558–560.
- 1046 **Nakosteen, P.C. and Hughes, K.W.** (1978) Sexual life cycle of three species of *Funariaceae* in  
1047 culture. *The Bryologist*, **81**, 307–314.
- 1048 **Niemeyer, P.W., Irisarri, I., Scholz, P., et al.** (2022) A seed-like proteome in oil-rich tubers. *The*  
1049 *Plant Journal*, **112**, 518–534.
- 1050 **Nietzel, T., Mostertz, J., Ruberti, C., et al.** (2020) Redox-mediated kick-start of mitochondrial  
1051 energy metabolism drives resource-efficient seed germination. *Proceedings of the*  
1052 *National Academy of Sciences*, **117**, 741–751.
- 1053 **Nishiyama, T., Sakayama, H., Vries, J. de, et al.** (2018) The *Chara* genome: secondary  
1054 complexity and implications for plant terrestrialization. *Cell*, **174**, 448-464.e24.
- 1055 **Nonogaki, H.** (2014) Seed dormancy and germination—emerging mechanisms and new  
1056 hypotheses. *Frontiers in Plant Science*, **5**.
- 1057 **Nystedt, B., Street, N.R., Wetterbom, A., et al.** (2013) The Norway spruce genome sequence  
1058 and conifer genome evolution. *Nature*, **497**, 579–584.
- 1059 **Oliver, M.J., Farrant, J.M., Hilhorst, H.W.M., Mundree, S., Williams, B. and Bewley, J.D.**  
1060 (2020) Desiccation tolerance: avoiding cellular damage during drying and rehydration.  
1061 *Annual Review of Plant Biology*, **71**, 435–460.
- 1062 **Oliver, M.J., Tuba, Z. and Mishler, B.D.** (2000) The evolution of vegetative desiccation  
1063 tolerance in land plants. *Plant Ecology*, **151**, 85–100.
- 1064 **Oliver, M.J., Velten, J. and Mishler, B.D.** (2005) Desiccation tolerance in bryophytes: a  
1065 reflection of the primitive strategy for plant survival in dehydrating habitats?. *Integrative*  
1066 *and Comparative Biology*, **45**, 788–799.
- 1067 **Ouyang, S., Zhu, W., Hamilton, J., et al.** (2007) The TIGR Rice Genome Annotation Resource:  
1068 improvements and new features. *Nucleic Acids Research*, **35**, D883–D887.

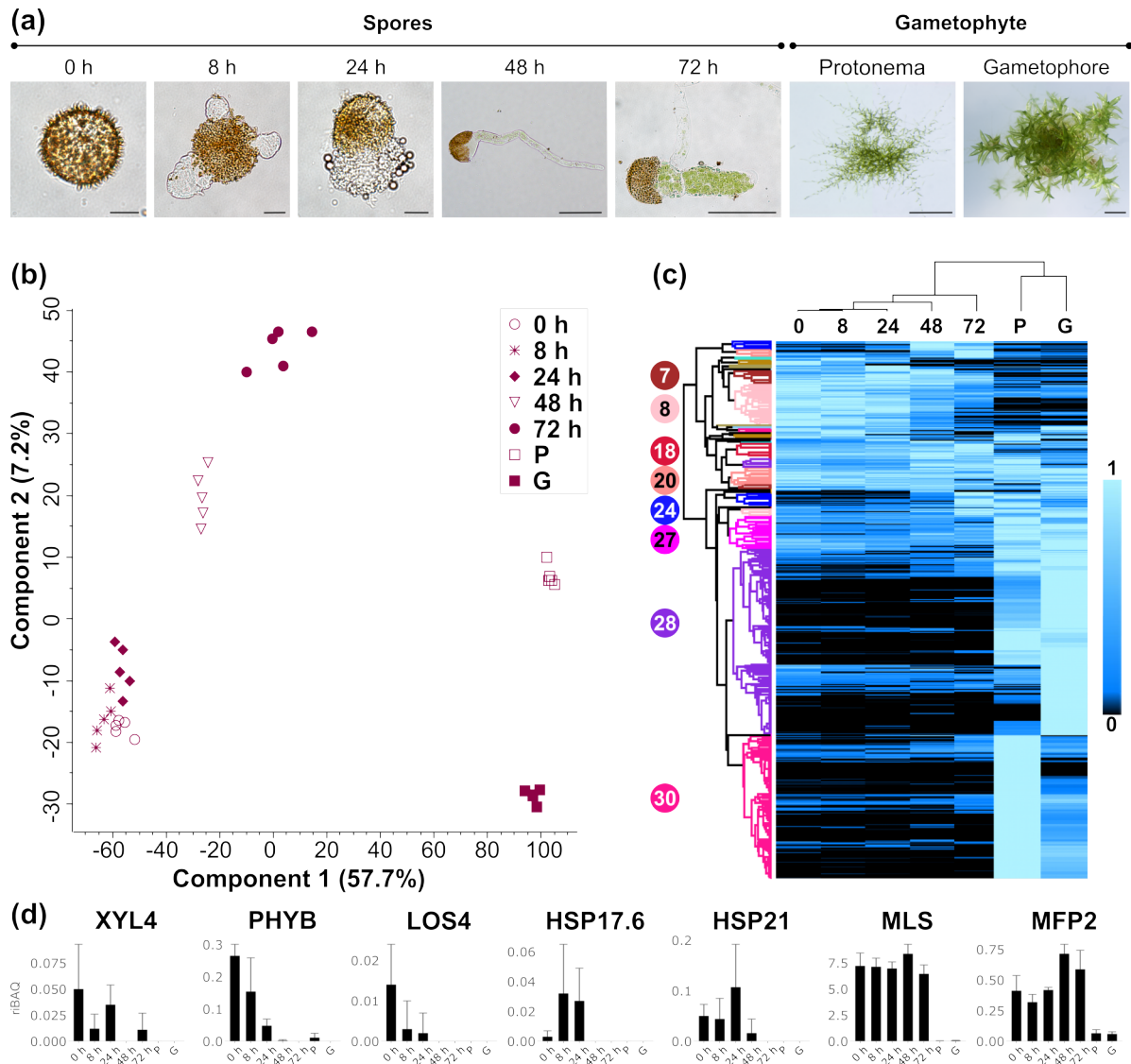
- 1069 **Pacini, E.** (2012) Pollen and seed analogies. *Plant Biosystems - An International Journal*  
1070 *Dealing with all Aspects of Plant Biology*, **146**, 738–748.
- 1071 **Peres, A.L.G.L., Soares, J.S., Tavares, R.G., Righetto, G., Zullo, M.A.T., Mandava, N.B. and**  
1072 **Menossi, M.** (2019) Brassinosteroids, the sixth class of phytohormones: a molecular  
1073 view from the discovery to hormonal interactions in plant development and stress  
1074 adaptation. *International Journal of Molecular Sciences*, **20**, 331.
- 1075 **Perroud, P.-F., Haas, F.B., Hiss, M., et al.** (2018) The *Physcomitrella patens* gene atlas project:  
1076 large-scale RNA-seq based expression data. *The Plant Journal*, **95**, 168–182.
- 1077 **Pyc, M., Gidda, S.K., Seay, D., et al.** (2021) LDIP cooperates with SEIPIN and LDAP to facilitate  
1078 lipid droplet biogenesis in *Arabidopsis*. *The Plant Cell*, **33**, 3076–3103.
- 1079 **Rappsilber, J., Mann, M. and Ishihama, Y.** (2007) Protocol for micro-purification, enrichment,  
1080 pre-fractionation and storage of peptides for proteomics using StageTips. *Nature*  
1081 *Protocols*, **2**, 1896–1906.
- 1082 **Rensing, S.A.** (2018) Great moments in evolution: the conquest of land by plants. *Current*  
1083 *Opinion in Plant Biology*, **42**, 49–54.
- 1084 **Rensing, S.A., Goffinet, B., Meyberg, R., Wu, S.-Z. and Bezanilla, M.** (2020) The moss  
1085 *Physcomitrium* (*Physcomitrella*) *patens*: a model organism for non-seed plants. *Plant*  
1086 *Cell*, **32**, 1361–1376.
- 1087 **Rensing, S.A., Lang, D., Zimmer, A.D., et al.** (2008) The *Physcomitrella* genome reveals  
1088 evolutionary insights into the conquest of land by plants. *Science*, **319**, 64–69.
- 1089 **Saavedra, L., Svensson, J., Carballo, V., Izmendi, D., Welin, B. and Vidal, S.** (2006) A  
1090 dehydrin gene in *Physcomitrella patens* is required for salt and osmotic stress tolerance.  
1091 *The Plant Journal*, **45**, 237–249.
- 1092 **Sano, N. and Marion-Poll, A.** (2021) ABA metabolism and homeostasis in seed dormancy and  
1093 germination. *International Journal of Molecular Sciences*, **22**, 5069.
- 1094 **Sarnighausen, E., Wurtz, V., Heintz, D., Van Dorsselaer, A. and Reski, R.** (2004) Mapping  
1095 of the *Physcomitrella patens* proteome. *Phytochemistry*, **65**, 1589–1607.
- 1096 **Shaw, A.J., Szövényi, P. and Shaw, B.** (2011) Bryophyte diversity and evolution: windows into  
1097 the early evolution of land plants. *American Journal of Botany*, **98**, 352–369.
- 1098 **Shevchenko, A., Tomas, H., Havli, J., Olsen, J.V. and Mann, M.** (2006) In-gel digestion for  
1099 mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, **1**,  
1100 2856–2860.
- 1101 **Shimada, T.L., Shimada, T., Takahashi, H., Fukao, Y. and Hara-Nishimura, I.** (2008) A novel  
1102 role for oleosins in freezing tolerance of oilseeds in *Arabidopsis thaliana*. *The Plant*  
1103 *Journal*, **55**, 798–809.
- 1104 **Shimada, T.L., Takano, Y., Shimada, T., et al.** (2014) Leaf oil body functions as a subcellular  
1105 factory for the production of a phytoalexin in *Arabidopsis*. *Plant Physiology*, **164**, 105–  
1106 118.
- 1107 **Siloto, R.M.P., Findlay, K., Lopez-Villalobos, A., Yeung, E.C., Nykiforuk, C.L. and Moloney,**  
1108 **M.M.** (2006) The accumulation of oleosins determines the size of seed oilbodies in  
1109 *Arabidopsis*. *The Plant Cell*, **18**, 1961–1974.

- 1110 **Skripnikov, A.Yu., Polyakov, N.B., Tolcheva, E.V., Velikodvorskaya, V.V., Dolgov, S.V.,**  
 1111 **Demina, I.A., Rogova, M.A. and Govorun, V.M.** (2009) Proteome analysis of the moss  
 1112 *Physcomitrella patens* (Hedw.) B.S.G. *Biochemistry (Moscow)*, **74**, 480–490.
- 1113 **Sperschneider, J., Catanzariti, A.-M., DeBoer, K., Petre, B., Gardiner, D.M., Singh, K.B.,**  
 1114 **Dodds, P.N. and Taylor, J.M.** (2017) LOCALIZER: subcellular localization prediction of  
 1115 both plant and effector proteins in the plant cell. *Scientific Reports*, **7**, 44598.
- 1116 **Tian, R., Wang, F., Zheng, Q., Niza, V.M.A.G.E., Downie, A.B. and Perry, S.E.** (2020) Direct  
 1117 and indirect targets of the *Arabidopsis* seed transcription factor ABSCISIC ACID  
 1118 INSENSITIVE3. *The Plant Journal*, **103**, 1679–1694.
- 1119 **Toshima, E., Nanjo, Y., Komatsu, S., Abe, T., Matsuura, H. and Takahashi, K.** (2014)  
 1120 Proteomic analysis of *Physcomitrella patens* treated with 12-oxo-phytodienoic acid, an  
 1121 important oxylipin in plants. *Bioscience, Biotechnology, and Biochemistry*, **78**, 946–953.
- 1122 **Tyanova, S., Temu, T. and Cox, J.** (2016) The MaxQuant computational platform for mass  
 1123 spectrometry-based shotgun proteomics. *Nature Protocols*, **11**, 2301–2319.
- 1124 **Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M. and Cox,**  
 1125 **J.** (2016) The Perseus computational platform for comprehensive analysis of  
 1126 (prote)omics data. *Nature Methods*, **13**, 731–740.
- 1127 **VanBuren, R., Wai, C.M., Zhang, Q., Song, X., Edger, P.P., Bryant, D., Michael, T.P.,**  
 1128 **Mockler, T.C. and Bartels, D.** (2017) Seed desiccation mechanisms co-opted for  
 1129 vegetative desiccation in the resurrection grass *Oropetium thomaeum*. *Plant, Cell &*  
 1130 *Environment*, **40**, 2292–2306.
- 1131 **Vesty, E.F., Saidi, Y., Moody, L.A., et al.** (2016) The decision to germinate is regulated by  
 1132 divergent molecular networks in spores and seeds. *New Phytologist*, **211**, 952–966.
- 1133 **Vogel, J.P., Garvin, D.F., Mockler, T.C., et al.** (2010) Genome sequencing and analysis of the  
 1134 model grass *Brachypodium distachyon*. *Nature*, **463**, 763–768.
- 1135 **Vries, J. de and Ischebeck, T.** (2020) Ties between stress and lipid droplets pre-date seeds.  
 1136 *Trends in Plant Science*, **25**, 1203–1214.
- 1137 **Wallace, S., Fleming, A., Wellman, C.H. and Beerling, D.J.** (2011) Evolutionary development  
 1138 of the plant spore and pollen wall. *AoB PLANTS*, **2011**, plr027.
- 1139 **Wan, T., Liu, Z.-M., Li, L.-F., et al.** (2018) A genome for gnetophytes and early evolution of seed  
 1140 plants. *Nature Plants*, **4**, 82–89.
- 1141 **Wang, S., Li, L., Li, H., et al.** (2020) Genomes of early-diverging streptophyte algae shed light  
 1142 on plant terrestrialization. *Nature Plants*, **6**, 95–106.
- 1143 **Wang, X., Kuang, T. and He, Y.** (2010) Conservation between higher plants and the moss  
 1144 *Physcomitrella patens* in response to the phytohormone abscisic acid: a proteomics  
 1145 analysis. *BMC Plant Biology*, **10**, 192.
- 1146 **Wang, X., Yang, P., Gao, Q., Liu, X., Kuang, T., Shen, S. and He, Y.** (2008) Proteomic analysis  
 1147 of the response to high-salinity stress in *Physcomitrella patens*. *Planta*, **228**, 167–177.
- 1148 **Wang, X., Yang, P., Zhang, X., Xu, Y., Kuang, T., Shen, S. and He, Y.** (2009) Proteomic  
 1149 analysis of the cold stress response in the moss, *Physcomitrella patens*. *PROTEOMICS*,  
 1150 **9**, 4529–4538.

- 1151 **Wang, X.Q., Yang, P.F., Liu, Z., et al.** (2009) Exploring the mechanism of *Physcomitrella patens*  
1152 desiccation tolerance through a proteomic strategy. *Plant Physiology*, **149**, 1739–1750.
- 1153 **Waters, E.R. and Vierling, E.** (2020) Plant small heat shock proteins – evolutionary and  
1154 functional diversity. *New Phytologist*, **227**, 24–37.
- 1155 **Wester, L., Somers, D.E., Clack, T. and Sharrock, R.A.** (1994) Transgenic complementation  
1156 of the hy3 phytochrome B mutation and response to PHYB gene copy number in  
1157 *Arabidopsis*. *The Plant Journal*, **5**, 261–272.
- 1158 **Wickell, D., Kuo, L.-Y., Yang, H.-P., et al.** (2021) Underwater CAM photosynthesis elucidated  
1159 by *Isoetes* genome. *Nat Commun*, **12**, 6348.
- 1160 **Wolkers, W.F., McCready, S., Brandt, W.F., Lindsey, G.G. and Hoekstra, F.A.** (2001)  
1161 Isolation and characterization of a D-7 LEA protein from pollen that stabilizes glasses in  
1162 vitro. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular*  
1163 *Enzymology*, **1544**, 196–206.
- 1164 **Xia, L., Zou, D., Sang, J., et al.** (2017) Rice Expression Database (RED): an integrated RNA-  
1165 Seq-derived gene expression database for rice. *Journal of Genetics and Genomics*, **44**,  
1166 235–241.
- 1167 **Xiong, M., Yu, J., Wang, J., et al.** (2022) Brassinosteroids regulate rice seed germination  
1168 through the BZR1-RAmy3D transcriptional module. *Plant Physiology*, **189**, 402–418.
- 1169 **Xu, Z., Xin, T., Bartels, D., et al.** (2018) Genome analysis of the ancient tracheophyte  
1170 *Selaginella tamariscina* reveals evolutionary features relevant to the acquisition of  
1171 desiccation tolerance. *Molecular Plant*, **11**, 983–994.
- 1172 **Yadav, M.K. and Bhatla, S.C.** (2011) Localization of lipoxygenase activity on the oil bodies and  
1173 in protoplasts using a novel fluorescence imaging method. *Plant Physiology and*  
1174 *Biochemistry*, **49**, 230–234.
- 1175 **Yamauchi, Y., Hasegawa, A., Taninaka, A., Mizutani, M. and Sugimoto, Y.** (2011) NADPH-  
1176 dependent reductases involved in the detoxification of reactive carbonyls in plants.  
1177 *Journal of Biological Chemistry*, **286**, 6999–7009.
- 1178 **Yang, Z., Ji, H. and Liu, D.** (2016) Oil biosynthesis in underground oil-rich storage vegetative  
1179 tissue: comparison of *Cyperus esculentus* tuber with oil seeds and fruits. *Plant Cell*  
1180 *Physiology*, **57**, 2519–2540.
- 1181 **Yang, Z., Liu, X., Wang, K., Li, Z., Jia, Q., Zhao, C. and Zhang, M.** (2022) ABA-INSENSITIVE  
1182 3 with or without FUSCA3 highly up-regulates lipid droplet proteins and activates oil  
1183 accumulation. *Journal of Experimental Botany*, **73**, 2077–2092.
- 1184 **Yao, L., Cheng, X., Gu, Z., et al.** (2018) The AWPM-19 family protein OsPM1 mediates abscisic  
1185 acid influx and drought response in rice. *The Plant Cell*, **30**, 1258–1276.
- 1186 **Yokota, T., Ohnishi, T., Shibata, K., Asahina, M., Nomura, T., Fujita, T., Ishizaki, K. and**  
1187 **Kohchi, T.** (2017) Occurrence of brassinosteroids in non-flowering land plants, liverwort,  
1188 moss, lycophyte and fern. *Phytochemistry*, **136**, 46–55.
- 1189 **Yotsui, I., Serada, S., Naka, T., Saruhashi, M., Taji, T., Hayashi, T., Quatrano, R.S. and**  
1190 **Sakata, Y.** (2016) Large-scale proteome analysis of abscisic acid and ABSCISIC ACID  
1191 INSENSITIVE3-dependent proteins related to desiccation tolerance in *Physcomitrella*  
1192 *patens*. *Biochemical and Biophysical Research Communications*, **471**, 589–595.

- 1193 **Zienkiewicz, A., Zienkiewicz, K., Rejón, J.D., Dios Alché, J. de Castro, A.J. and Rodríguez-**  
1194 **García, M.I.** (2014) Olive seed protein bodies store degrading enzymes involved in  
1195 mobilization of oil bodies. *Journal of Experimental Botany*, **65**, 103–115.
- 1196 **Zienkiewicz, A., Zienkiewicz, K., Rejón, J.D., Rodríguez-García, M.I. and Castro, A.J.** (2013)  
1197 New insights into the early steps of oil body mobilization during pollen germination.  
1198 *Journal of Experimental Botany*, **64**, 293–302.
- 1199

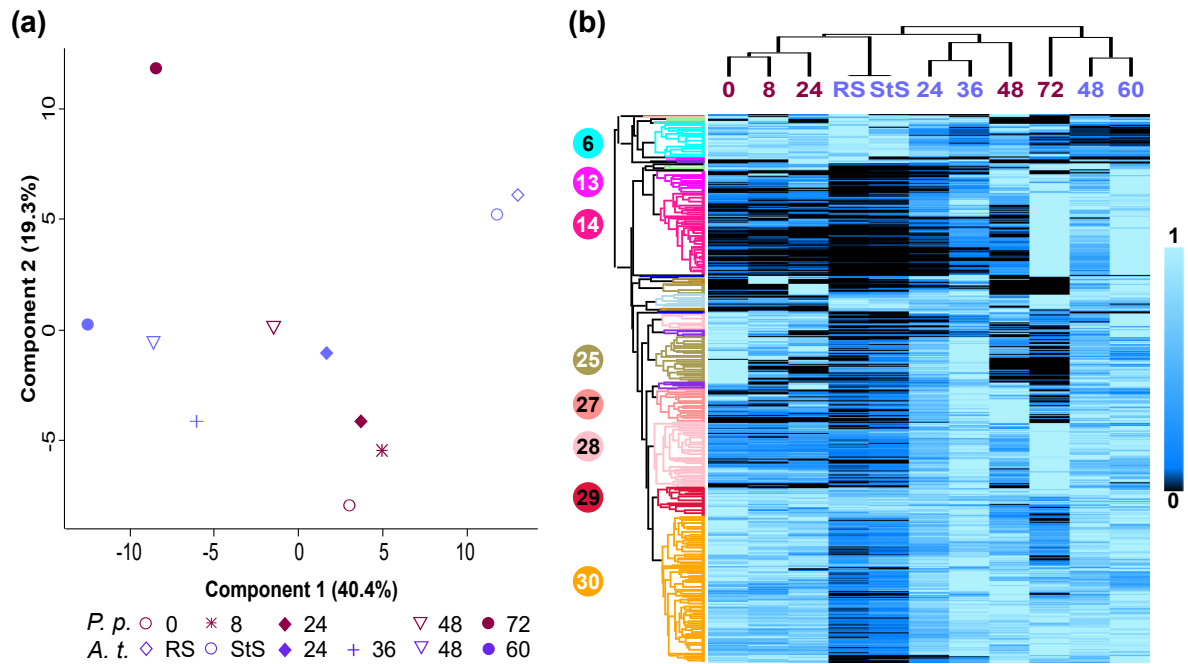




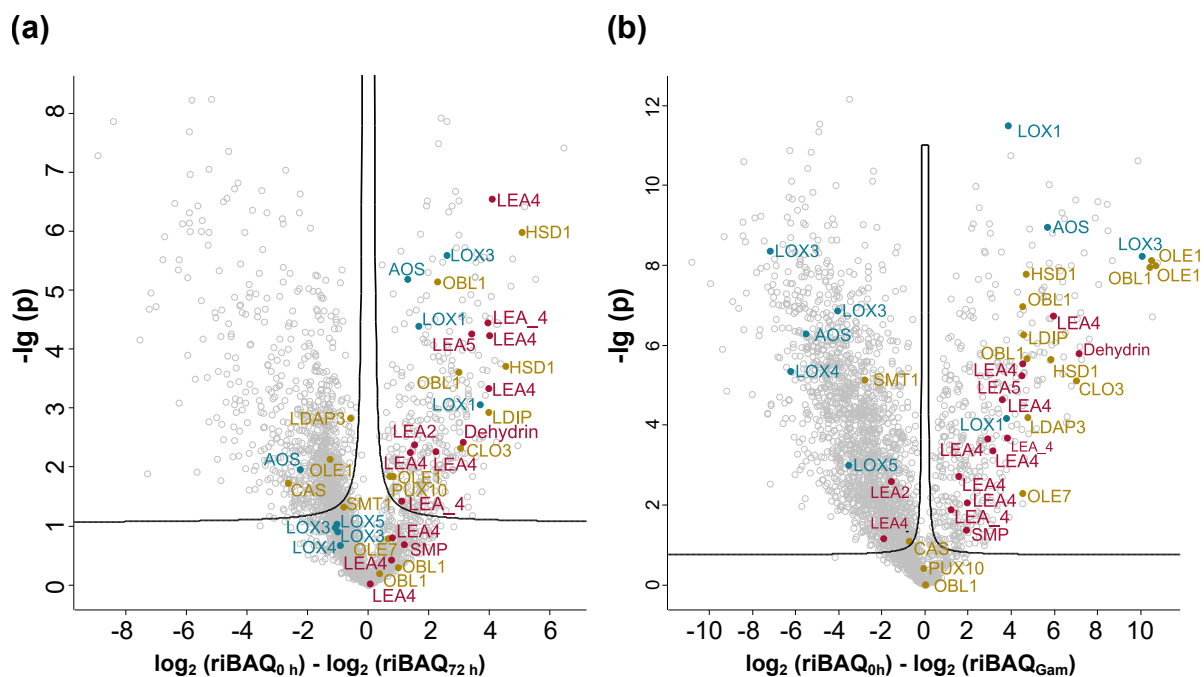
**Figure 1 Proteomic analysis of seven developmental stages of *P. patens*.** **(a)** Micrographs of developmental stages analyzed in this study: germinating spores of *P. patens* ecotype Reute 0 h - 72 h after rehydration, juvenile protonema and adult gametophores. First chloronema cells emerge after 8 h, continuously growing over the time course. Size bars, 0 h, 8 h, 24 h  $\cong$  10  $\mu$ m; 48 h, 72 h  $\cong$  50  $\mu$ m; protonema  $\cong$  1mm, gametophores  $\cong$  2mm. **(b)** Principal component analysis plot (n = 5 per stage). **(c)** Hierarchical clustering of the means (n = 5 per stage), with the highest mean of each protein group being set to 1. Proteins sorted by clusters can be found in Data S3. Clusters with more than 50 candidates are labeled with numbers. **(d)** Relative protein distribution of seven selected proteins from clusters 4, 5, 7, and 18 showing the relative iBAQ (riBAQ) in per mill. Error bars, SD. HSP17.6, 17.6 kDa small heat shock protein; HSP21, HEAT SHOCK PROTEIN 21; LOS4, LOW EXPRESSION OF OSMOTICALLY RESPONSIVE GENES 4; MFP2, MULTIFUNCTIONAL PROTEIN 2; MLS, MALATE SYNTHASE; PHYB, PHYTOCHROME B; XYL4, BETA-XYLOSIDASE 4.

GO ID	GO Term	No <i>Pp</i> IDs	<i>P. patens</i>						<i>A. thaliana</i>			
			0 h	8 h	24 h	48 h	72 h	P	G	RS	36 h	60 h
GO:0009737	Response to abscisic acid	73	113.6	104.1	97.3	76.0	38.2	14.3	15.8	615.6	336.9	189.3
GO:0005811	Lipid droplet	10	87.1	79.4	73.1	52.8	23.7	0.2	0.4	13.4	4.9	1.9
GO:0009555	Pollen development	47	70.4	63.8	57.2	38.4	17.2	11.6	10.4	4.6	7.4	4.0
GO:0048316	Seed development	61	102.2	93.2	87.3	64.3	34.0	9.8	8.9	1.3	2.0	1.7
GO:0031408	Oxylipin biosynthetic process	12	96.1	88.4	78.6	53.3	24.2	4.9	6.0	5.3	3.6	2.6
GO:0006096	Glycolytic process	49	26.4	25.2	22.6	23.3	14.9	16.6	20.8	19.0	32.8	45.7
GO:0009414	Response to water deprivation	80	106.8	98.7	92.9	70.5	36.0	15.2	27.9	5.4	12.1	10.4
GO:0010231	Maintenance of seed dormancy	3	19.3	21.2	18.9	15.8	8.4	1.0	4.2	3.3	0.8	0.1
GO:0071456	Cellular response to hypoxia	24	35.7	38.2	37.8	33.9	25.0	6.5	17.8	0.001	0.006	0.009
GO:0042542	Response to hydrogen peroxide	18	89.5	113.4	122.8	101.9	77.9	4.2	2.6	12.7	12.0	7.2
GO:0006457	Protein folding	89	115.3	135.7	145.7	132.1	105.7	33.7	28.7	18.5	41.5	36.9
GO:0006097	Glyoxylate cycle	2	7.7	7.7	8.2	12.9	9.6	0	0.1	1.7	31.7	11.8
GO:0006635	Fatty acid beta-oxidation	14	3.6	3.1	3.7	8.0	10.0	1.5	1.2	1.9	6.1	3.9
GO:0019253	Reductive pentose-phosphate cycle	26	3.1	2.6	3.4	5.7	12.6	41.4	37.3	4.6	54.4	179.7
GO:0015979	Photosynthesis	40	4.2	3.5	3.0	5.0	23.2	89.5	79.8	1.1	36.4	108.3

**Figure 2 Abundance of proteins in selected GO terms.** Depicted are averaged ( $n = 5$ ) and summed values in per mill relative iBAQ of all proteins assigned to the corresponding GO term for each individual tissue type, along with the number of proteins per GO term (No *Pp* IDs). For *P. patens*, the five stages of spore germination (0 h, 8 h, 24h, 48 h, 72 h), as well as protonema (P) and gametophores (G) are displayed, compared to three stages of Arabidopsis seedling establishment originally published in Kretzschmar et al., 2020. Given are rehydrated seeds (RS), and seedling grown for 36 h and 60 h.



**Figure 3 Comparison of *P. patens* spore germination and Arabidopsis seedling establishment.** For *P. patens*, the five stages of spore germination (0 h, 8 h, 24 h, 48 h, 72 h) and for Arabidopsis (rehydrated seeds) RS, stratified seeds (StS), 24 h, 36 h, 48 h, and 60 h of seed germination, originally published in Kretzschmar et al., 2020, are displayed. The highest mean of each protein (averaged values with  $n = 5$ ) was set to 1 individually for *P. patens* and Arabidopsis. **(a)** Principal component analysis plot. **(b)** Hierarchical cluster of which all proteins are sorted by clusters can be found in Data S6; clusters with more than 50 candidates are labeled with numbers.

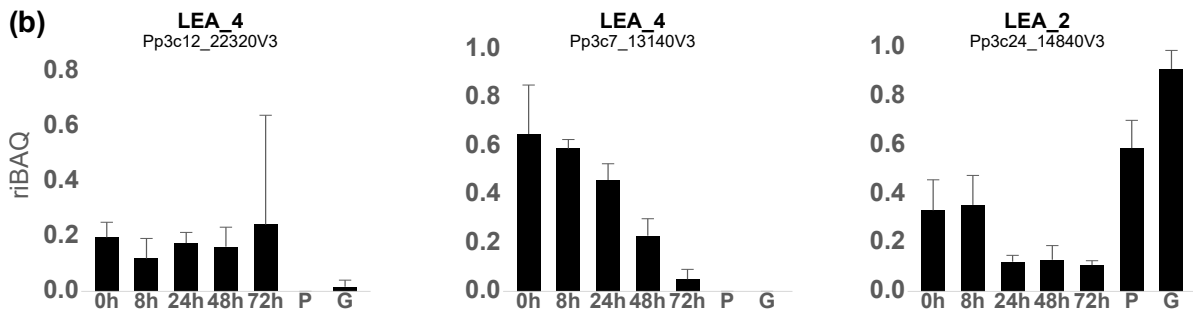


**Figure 4** Volcano plots visualize differences between 0 h spores and 72 h spores or gametophores. The  $\log_2$  transformed and imputed data were compared and  $p$ -values calculated. Yellow circles represent homologs of known LD-proteins, magenta circles LEA proteins, and blue circles oxylipin-related proteins. The black line indicates a false discovery rate of 0.05. Abbreviations: AOS, allene oxide synthase; CAS, cycloartenol synthase; CLO, caleosin; LEA, late embryogenesis abundant protein; LOX, lipoxygenase; HSD, steroleosin; LDAP, lipid droplet associated protein; LDIP, LDAP-interacting protein; OBL, oil body lipase; OLE, oleosin; PUX, plant UBX domain-containing protein; SMP, seed maturation protein; SMT, sterol methyltransferase.

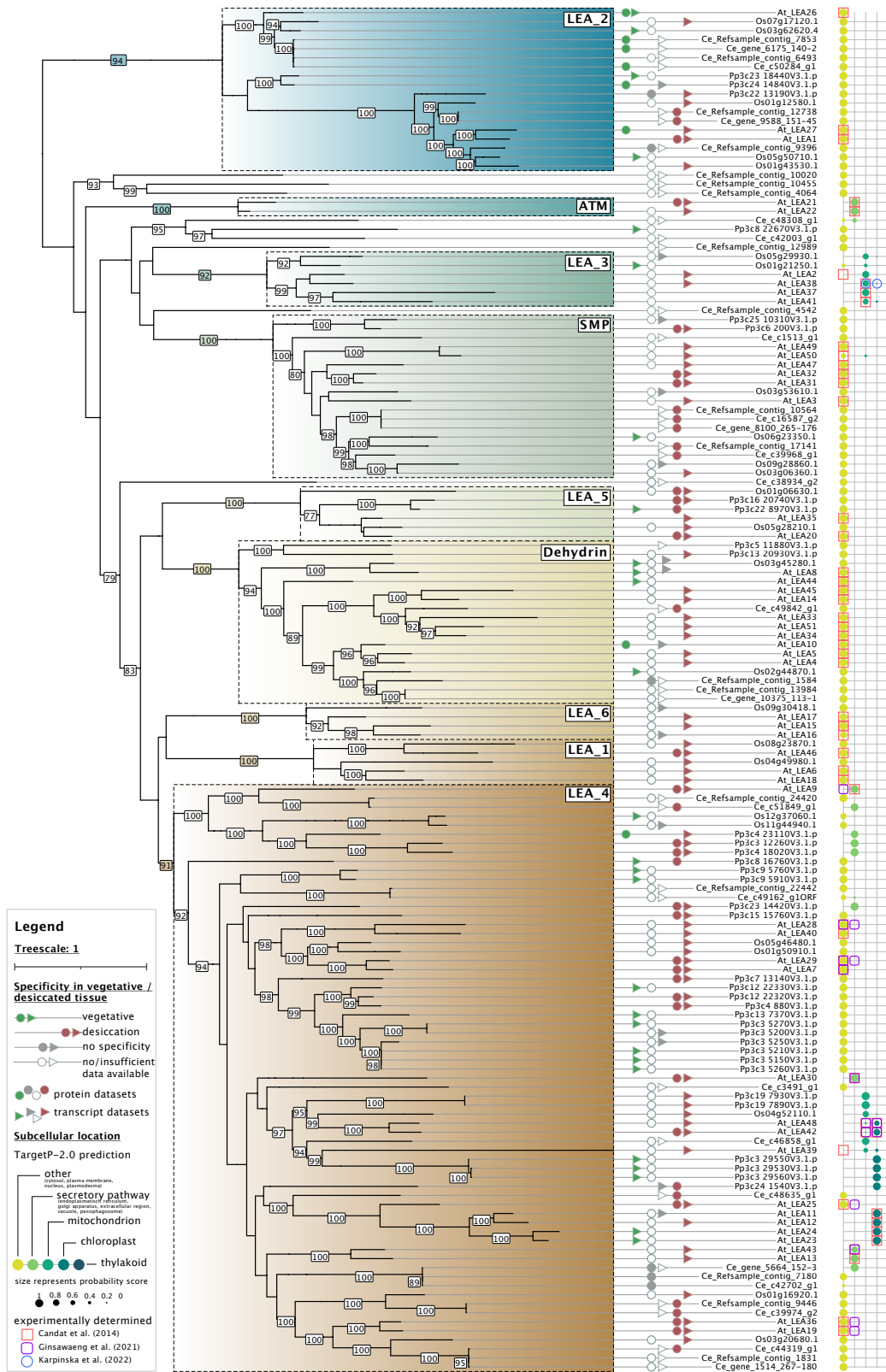
(a)

Cluster	<i>P. Patens</i> ID	<i>A. thaliana</i> ID	Family	0 h	8 h	24 h	48 h	72 h	P	G
5	Pp3c6_200V3	AT5G27980	SMP	0.06	0.06	0.09	0.01	0.03	0.00	0.01
7	Pp3c4_18020V3	AT1G72100	LEA_4	0.03	0.03	0.03	0.04	0.00	0.00	0.00
8	Pp3c5_11880V3	AT1G20440	Dehydrin	0.87	1.19	1.13	0.58	0.14	0.00	0.00
8	Pp3c3_12260V3	AT1G72100	LEA_4	0.22	0.15	0.08	0.05	0.00	0.02	0.01
8	Pp3c7_13140V3	AT2G18340	LEA_4	0.65	0.59	0.46	0.23	0.05	0.00	0.00
8	Pp3c16_20740V3	AT2G40170	LEA_5	0.22	0.27	0.30	0.21	0.00	0.00	0.00
8	Pp3c24_1540V3	AT3G53040	LEA_4	0.03	0.04	0.02	0.01	0.00	0.00	0.00
8	Pp3c23_14420V3	AT3G53040	LEA_4	0.30	0.35	0.27	0.20	0.01	0.05	0.01
8	Pp3c15_15760V3	AT3G53040	LEA_4	0.19	0.16	0.12	0.08	0.01	0.01	0.01
8	Pp3c8_16760V3	AT5G44310	LEA_4	0.09	0.07	0.05	0.02	0.01	0.02	0.00
18	Pp3c12_22320V3	AT1G52690	LEA_4	0.20	0.12	0.18	0.16	0.24	0.00	0.02
18	Pp3c4_880V3	AT3G15670	LEA_4	0.07	0.05	0.09	0.08	0.09	0.02	0.01
28	Pp3c24_14840V3	AT2G44060	LEA_2	0.33	0.35	0.12	0.13	0.11	0.59	0.91
28	Pp3c4_23110V3	AT3G53040	LEA_4	0.02	0.02	0.02	0.01	0.00	0.08	0.14

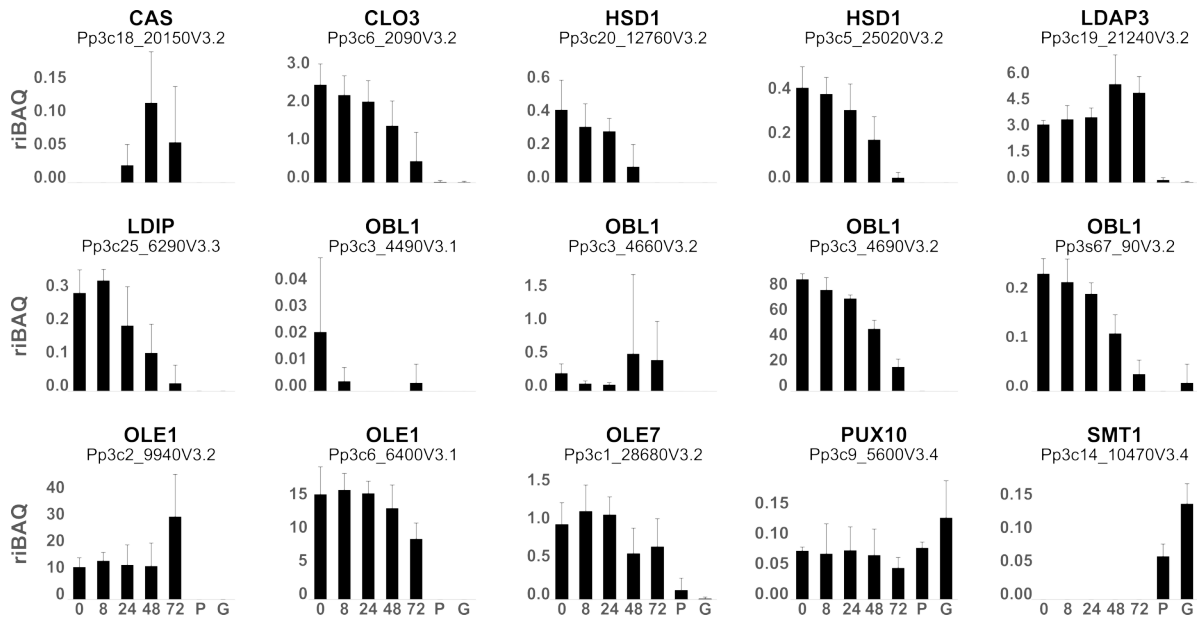
(b)



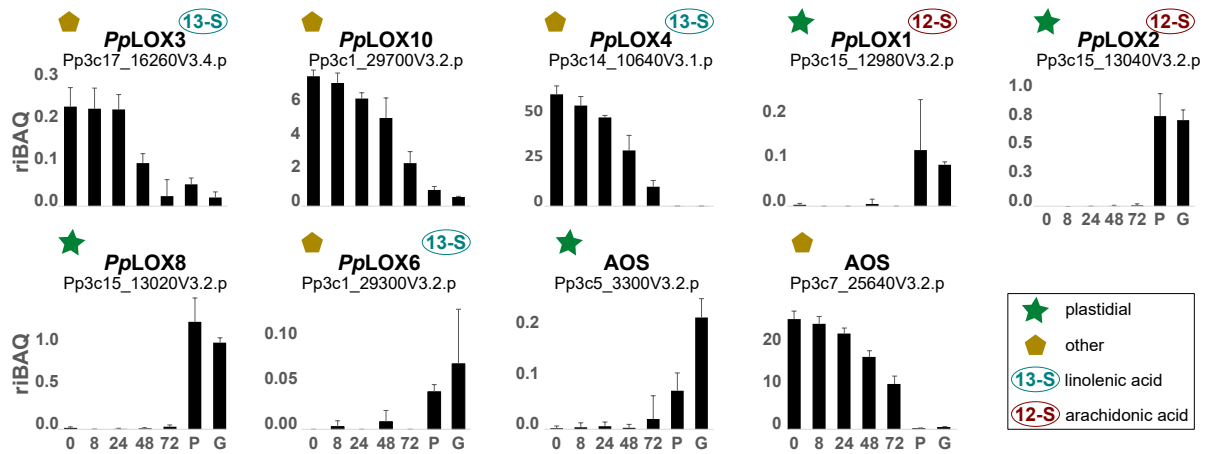
**Figure 5 Most LEA proteins are enriched in early stages of spore germination (a)** Late embryogenesis abundant (LEA) proteins identified in *Physcomitrium patens* clustered as visualized in Figure 1c. Given are averaged relative iBAQ values in per mill for all developmental stages (spores 0 h, 8 h, 24 h, 48 h, 72 h, protonema (P), gametophore (G); n = 5 for each stage), the cluster number, the identifier of *P. patens*, the identifier of the most similar *Arabidopsis* homolog as well as the protein family. **(b)** Relative protein distribution of three selected LEA-proteins showing the relative iBAQ (riBAQ) in per mill. Error bars, SD.



**Figure 6 Phylogenetic tree of LEA proteins of *P. patens* and three Angiosperms.** Phylogeny of LEA proteins including information on their tissue specificity and subcellular location. The tree was constructed with LEA proteins from 23 different Viridiplantae species and pruned for the LEAs of Arabidopsis (*At*) and their homologs found in *C. esculentus* (*Ce*), *O. sativa* (*Os*), and *P. patens* (*Pp*). Localizations were predicted with TargetP 2.0 or are based on experimental data (Candat et al., 2014, Karpinska et al., 2022, Ginsawaeng et al., 2021). Assignment as desiccation or vegetative specific is based on previously published data. See methods for details.

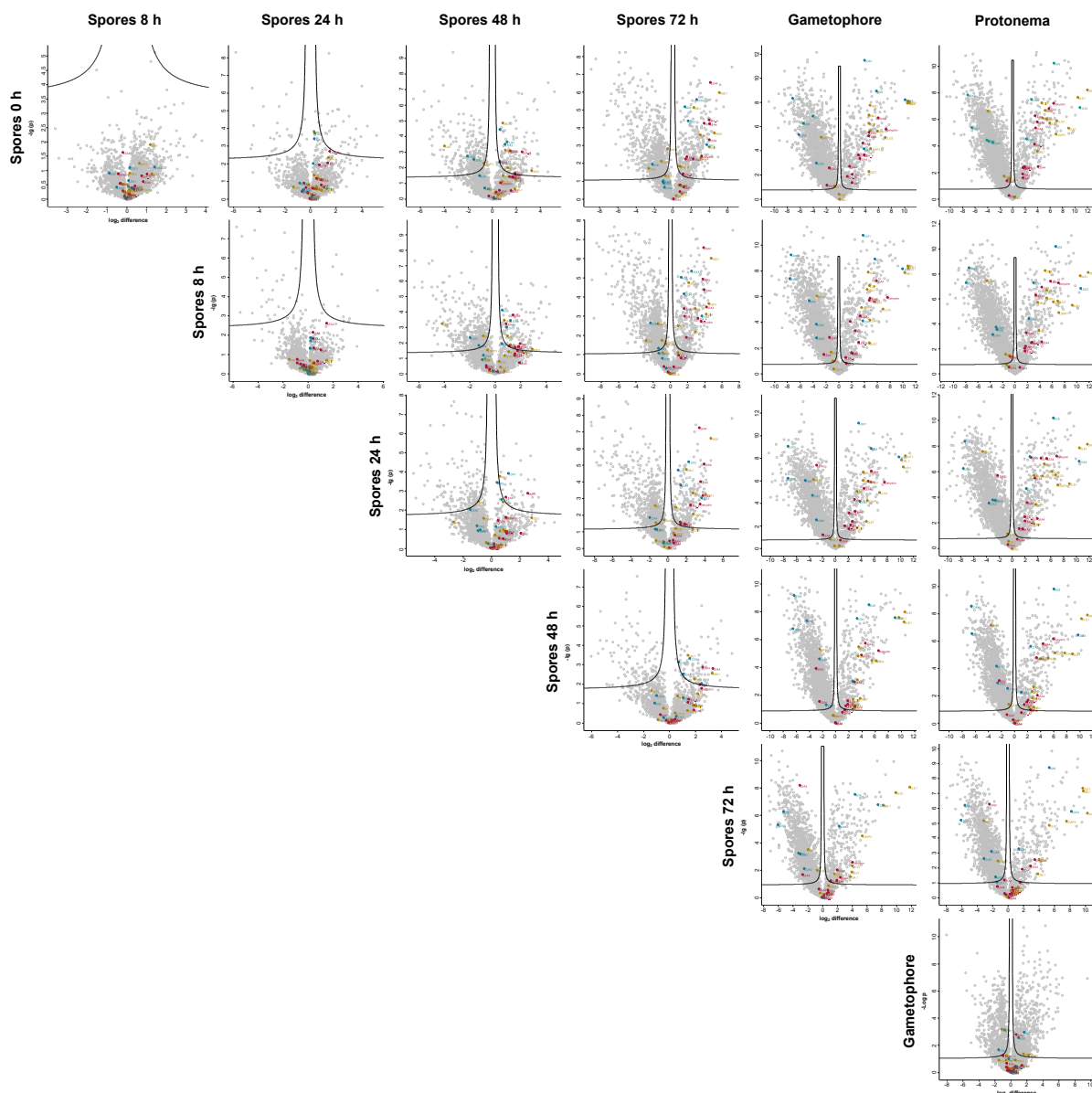


**Figure 7 Most LD-proteins are predominantly found in spores.** Abundance of protein homologs of known LD proteins (spores 0 h, 8 h, 24 h, 48 h, 72 h; P, protonema; G, gametophore).  $n = 5$  for each tissue. Error bars, SD. Abbreviations: CAS, cycloartenolsynthase; CLO, caleosin; HSD, hydroxysteroiddehydrogenase, steroleosin; LDAP, lipid droplet-associated proteins; LDIP, LDAP-interacting protein; OBL, oil body lipase; OLE, oleosin; PUX, plant UBX-domain containing protein; SMT, sterol methyl transferase.

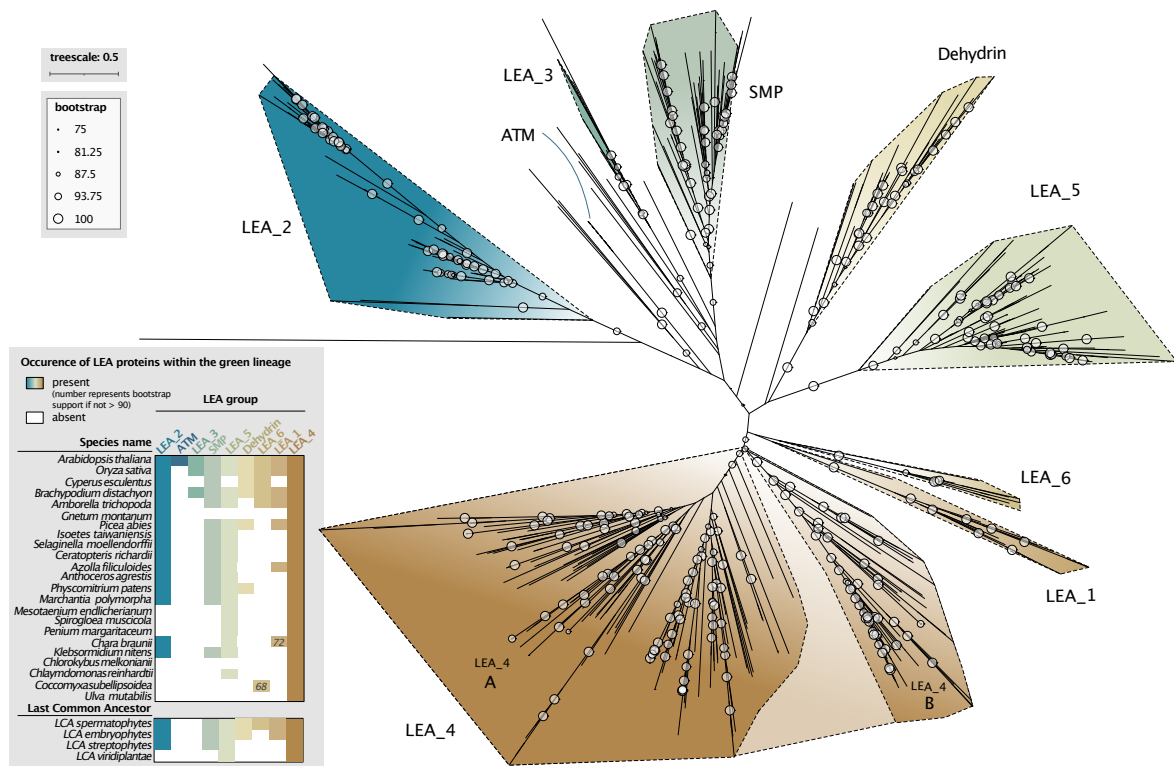


**Figure 8 LOX pathway-related proteins in *P. patens*.** Abundance of protein homologs of LOX pathway-related proteins (spores 0 h, 8 h, 24 h, 48 h, 72 h; P, protonema; G, gametophore).  $n = 5$  for each tissue. Error bars, SD. Subcellular localization was predicted using the software tools LOCALIZER, Green Targeting Predictor, and TargetP 2.0, and a consensus is given. The substrate specificity and the position of the oxygen addition was annotated according to Anterola et al., 2009.





**Figure S1. Pair-wise comparison of developmental stages through volcano plots.** Log<sub>2</sub> transformed and imputed data was compared between stages. The plots display the stage depicted on the left minus the stage depicted on the top. *p*-Values were calculated. Yellow circles represent homologs of known LD-proteins, magenta circles LEA proteins, and blue circles oxylipin-related proteins. The black line indicates a false discovery rate of 0.05. Abbreviations: AOS, allene oxide synthase; CAS, cycloartenol synthase; CLO, caleosin; LEA, late embryogenesis abundant protein; LOX, lipoxygenase; HSD, steroleosin; LDAP, lipid droplet associated protein; LDIP, LDAP-interacting protein; OBL, oil body lipase; OLE, oleosin; PUX, plant UBX domain-containing protein; SMP, seed maturation protein; SMT, sterol methyltransferase.



**Figure S2. Phylogenetic tree of LEA proteins in the green lineage.** Phylogenetic tree was built with the WAG+I+G4 substitution model. The occurrence of each LEA orthologues group in different species and the inferred presence of each LEA ortholog in several last common ancestors is given in the table on the bottom-left.

<b>Table S1: Metadata file for LC-MS/MS data processing with MaxQuant.</b>	
Proteome of <i>Physcomitrium patens</i> germinating spores, protonema, and gametophores	
<b>1. General features</b>	
Responsible persons	Prof. Till Ischebeck <sup>1</sup> , Dr. Oliver Valerius <sup>2</sup> , Dr. Kerstin Schmitt <sup>2</sup> , Prof. Gerhard H. Braus <sup>2</sup> 1 Institute of Plant Biology and Biotechnology (IBBP), University of Münster, Green Biotechnology, Münster 48143, Germany 2 Institute for Microbiology and Genetics and Service Unit LCMS Protein Analytics, Department for Molecular Microbiology and Genetics, University of Göttingen, Göttingen 37077, Germany
Instrument manufacturer, model	Thermo Fisher Scientific, Orbitrap Q Exactive HF
Experimental Design	Analysis of total protein fractions of germinating spores, protonema, and gametophores (phyllids) of <i>Physcomitrium patens</i> ecotype Reute
Group(s)	Total protein fraction
Biological and technical replicates	Biological replicates: 5 for different genotypes and developmental stages as listed below Technical replicates (independent LC-MS runs): PN_Q_15-39: 2 (Q/Q2), PN_Q_116-125: 1 (Q). PN_Q_15, 16, 17, 18, 19, PN_Q2_15, 16, 17, 18, 19: <i>P. patens</i> spores 0 h post rehydration PN_Q_20, 21, 22, 23, 24, PN_Q2_20, 21, 22, 23, 24: <i>P. patens</i> spores 8 h post rehydration PN_Q_25, 26, 27, 28, 29, PN_Q2_25, 26, 27, 28, 29: <i>P. patens</i> spores 24 h post rehydration PN_Q_30, 31, 32, 33, 34, PN_Q2_30, 31, 32, 33, 34: <i>P. patens</i> spores 48 h post rehydration PN_Q_35, 36, 37, 38, 39, PN_Q2_35, 36, 37, 38, 39: <i>P. patens</i> spores 72 h post rehydration PN_Q_116, 117, 118, 119, 120: <i>P. patens</i> protonema PN_Q_121, 122, 123, 124, 125: <i>P. patens</i> gametophore (phyllids)
Sample amount	35
<b>2. Electrospray Ionisation (ESI)</b>	
Supply type (static or fed)	fed
Interface manufacturer	Thermo Fisher Scientific
Sprayer type	Nanospray Flex Ion Source
<b>3.1 Post source component – Analyser</b>	
	Q Exactive HF: Orbitrap analyser
<b>3.2 Post source component – Activation/dissociation</b>	
Instrument component where the activation/dissociation occurs	Q Exactive HF: HCD cell
Gas type	Q Exactive HF: Nitrogen
Activation/dissociation type	Q Exactive HF: HCD

<b>4.1 Spectrum and peak list generation and annotation – Data acquisition</b>	
Software name and version	Xcalibur 4.0
Acquisition parameters	Data-dependent Top10
Software name and version	MaxQuant 2.0.3.1
<b>4.2 Spectrum and peak list generation and annotation – Resulting data</b>	
Location of source and processed files	The mass spectrometry proteomics data will be deposited to the ProteomeXchange Consortium via the PRIDE partner repository
<b>5. Description of the software and methods applied in the quantitative analysis</b>	
Quantification software	MaxQuant 2.0.3.1
Description of the selection and/or matching method of features, together with the description of the method of the primary extracted quantification values determination for each feature and/or peptide	<p>Upload of all .raw files into the software. Grouping of technical replicates as one Experiment (“set experiment”).</p> <p>Group-specific parameters:</p> <ol style="list-style-type: none"> <li>1) Type: default</li> <li>2) Digestion: default</li> <li>3) Modifications: default</li> <li>4) Cross-links: default</li> <li>5) Label-free quantification: LFQ, default</li> <li>6) Instrument: intensity determination: total sum, rest default</li> <li>7) First search: default</li> <li>8) Misc: default</li> </ol> <p>Global parameters</p> <ol style="list-style-type: none"> <li>1) Sequences: PhytozomeV13_20201118_Ppatens_318_v3.3.protein, rest default</li> <li>2) Identification: Match between runs ✓, rest default</li> <li>3) Protein quantification: default</li> <li>4) Label free quantification: iBAQ ✓, rest default</li> <li>5) Tables: default</li> <li>6) Folder locations: default</li> <li>7) MS/MS analyzer: FTMS recalibration ✓, rest default</li> <li>8) Advanced: default</li> <li>9) MS/MS fragmentation: default</li> </ol>
Confidence filter of features or peptides prior to quantification	Global parameters: identification: PSM FDR=0.01, protein FDR=0.01, rest default
Normalization	All values were divided by the total iBAQ intensities or total LFQ intensities in one sample and multiplied by 1000.

## **4 Unpublished results I: The LD proteome of *Arabidopsis thaliana* roots**

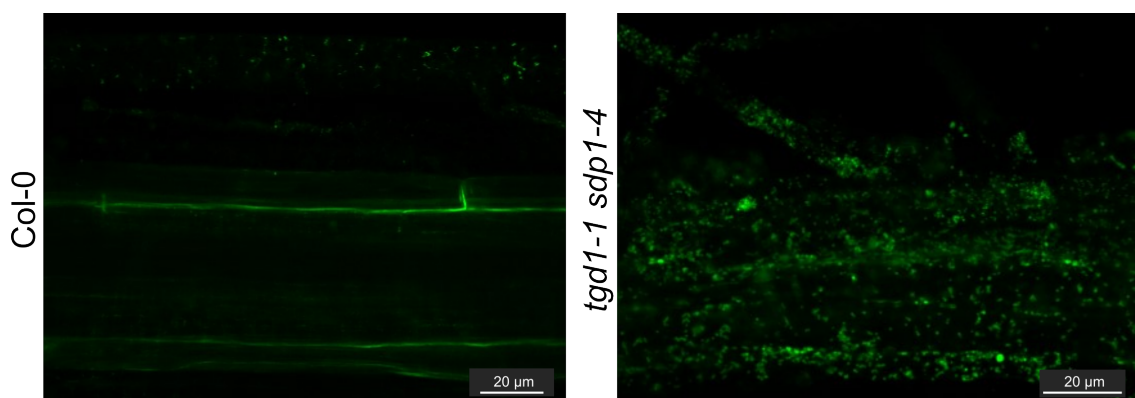
Author contribution:

For the LD proteome of roots, Philipp William Niemeyer designed, planned, and performed the experiments with Dr. Patricia Scholz in equal parts, except for LC-MS/MS measurements. He analyzed the resulting proteomic dataset. Displayed microscopy imagery was captured by Philipp William Niemeyer together with lab rotation student Fabienne Dreier. This research is intended for future publication.

## The LD proteome of *Arabidopsis thaliana* roots

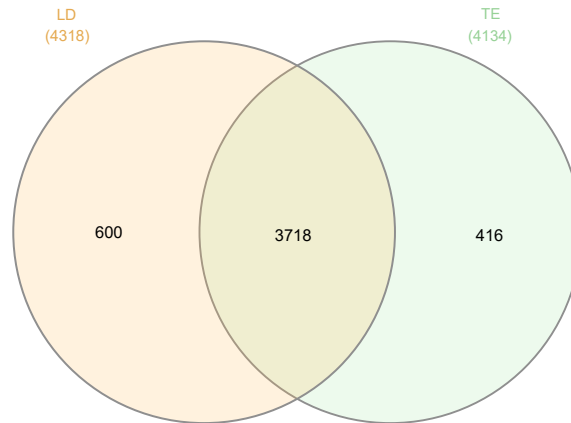
The current understanding of the plant lipid droplet proteome is primarily based on screenings conducted in seeds, seedlings, mesocarp, and pollen tubes, with more recent studies including leaves (introduction, Table 2). To our knowledge, the LD proteome of roots has not yet been described in any species. In order to isolate LD proteins from roots, two significant challenges had to be addressed. On the one hand, we had a demand for high yields of non-disruptively harvestable root biomass. On the other hand, the low natural triacylglycerol content of less than 1 % of dry weight (Kelly et al., 2013) in *Arabidopsis* roots needed to be elevated to obtain a floating LD-enriched phase suitable for LD protein isolation. In order to combine those two aspects, we employed an axenic root culture approach in high-sucrose media (Hétu et al., 2005) and cultivated the oil-rich mutant line *tgdl1-1 sdp1-4* (Fan et al., 2014). Accumulation of TAGs in roots was previously shown independently in *tgdl1* and *sdp1* mutants (Xu et al., 2005; Kelly et al., 2013). Especially the loss-of-function mutation *sdp1* and cultivation in high-concentrated sucrose media had beneficial effects on the oil content (Kelly et al., 2013). The combination of axenic root culture with high-concentrated sucrose media and a *tgdl1-1 sdp1-4* oil-rich mutant was necessary, as the employment of Col-0 wt lines in axenic culture showed a low oil content insufficiently for the enrichment of the LD phase, as microscopically confirmed (Figure 4). Thus, proteomic data were obtained from the LD-enriched phase of *tgdl1-1 sdp1-4* roots and total protein fractions were collected from five biological replicates for the analysis.

In total, 4734 different protein groups (unifications of indistinguishable proteins) were identified (Supplemental Dataset 1A). Thereof, 4318 protein groups were detected in the LD-enriched



**Figure 4: Confocal micrographs of *Arabidopsis* Col-0 and *tgdl1-1 sdp1-4* roots grown in high-sucrose media axenic root cultures.** Roots were stained with the lipophilic dye BODIPY 493/503. The *tgdl1-1 sdp1-4* roots exhibit a higher density of LDs in comparison to Col-0.

protein fraction and 4134 protein groups in the total protein fraction. Altogether, there are 600 protein groups exclusively in the LD-enriched protein fraction and 416 in the total protein fraction, respectively, while 3718 were found in both fractions (Figure 5). Those numbers indicate a successful LD enrichment and protein isolation.



**Figure 5: Venn diagram of proteomic dataset of *tgdl-1 sdp1-4* roots.** LD-enriched (LD) was compared with total protein (TE) fractions. All proteins with iBAQ value greater than 0 in any replicate were considered.

In the LD-enriched fraction, we detected 31 proteins that have been previously reported as LD-associated proteins, considering also the proteins cytochrome B5 isoform E (CB5-E) and LD-localized NTF2 family protein (LDNP) recently described by Scholz (2022) (Table 3). Of the 31 LD-associated proteins, 15 were found exclusively in the LD-enriched protein fraction. Furthermore, the LD fraction contained 222.7 ‰ described LD proteins in comparison to 3.3 ‰ in the total fraction, implying effective enrichment of LDs.

We mined our derived dataset for potential additional LD-associated proteins. To reduce noise and ensure data reliability, we implemented a stringent filter that only considered proteins meeting two criteria: first, they must have been identified with at least two unique peptides; and second, they must have been detected in all five replicates of either the LD or the total cellular protein fraction. By applying these strict criteria, we aimed to exclude false positives and increase the accuracy of our protein identification results.

Of the 31 known LD proteins, 23 satisfied the filter criteria (Table 3), suggesting high confidence in their presence in roots. Furthermore, it reduces the dataset to 2814 proteins. To ensure the accurate identification of potentially low-abundance LD proteins, the determination of  $\log_2$  differences between LD-enriched and total protein fractions to obtain enrichment and statistical significance calculation through Student's *t*-test was performed for each protein (Supplemental Dataset 1B). Ultimately, significance and difference were depicted as a volcano plot (Figure 6).

#### 4. Unpublished results I: The LD proteome of *Arabidopsis thaliana* roots

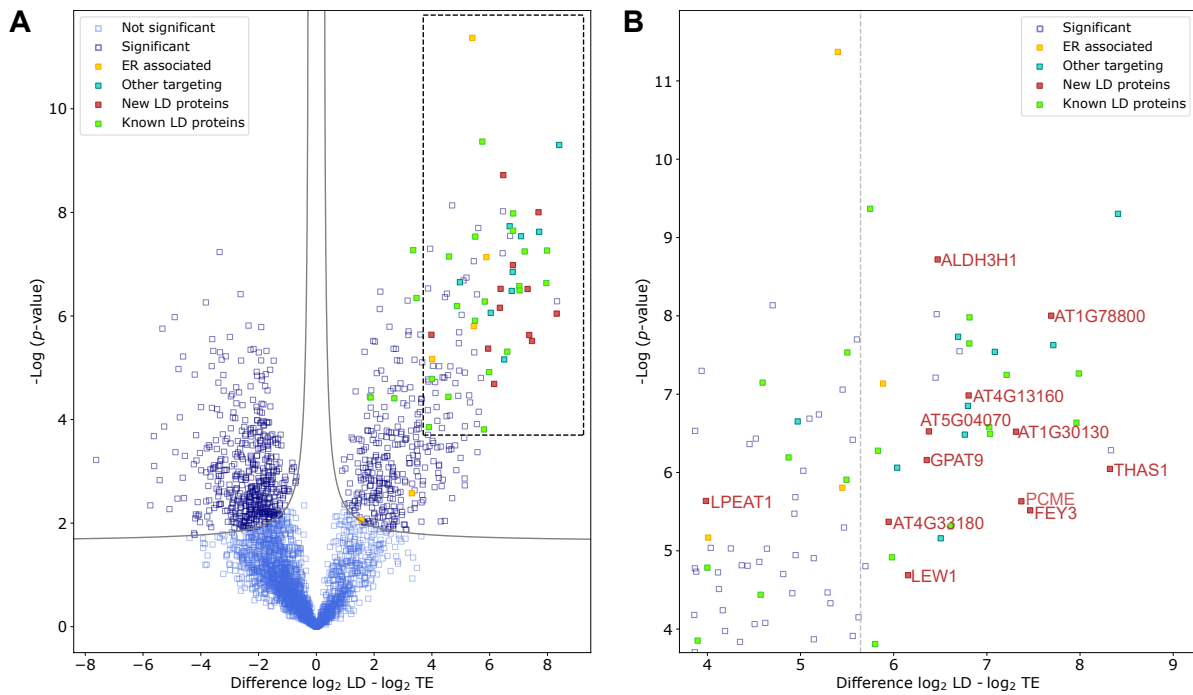
**Table 3: Previously described LD proteins identified in *Arabidopsis tgd1-1 sdp1-4* axenic root cultures.** The proteins were derived from LD-enriched (LD) and total (TE) protein fractions. The intensities of the proteomic dataset were computed using the iBAQ algorithm and normalized as per mille of all proteins in a sample (riBAQ). The values provided represent the average riBAQ (n=5) of the LD-enriched fractions. The enrichment of LD proteins was determined by dividing the average LD riBAQ by the average TE riBAQ (not detected in TE equals only detected in LD). Filter criteria are identified by minimum 2 peptides and 5 out of 5 replicates in either LD-enriched or TE fraction.

Protein name	AGI	riBAQ	LD enrichment	Filter criteria
CAS	AT2G07050	0.3	421.5	True
CB5-E	AT5G53560	1.0	3.5	True
CLO3	AT2G33380	0.2	n.d. in TE	True
CLO4	AT1G70670	2.1	23.4	True
ERD7	AT2G17840	3.2	51.2	True
GPAT4	AT1G01610	0.9	325.1	True
GPAT8	AT4G00400	0.02	n.d. in TE	False
HSD1	AT5G50700	0.10	n.d. in TE	False
HSD3	AT3G47360	0.09	n.d. in TE	False
HSD7	AT5G50690	0.5	n.d. in TE	True
LDAH1	AT1G10740	0.004	n.d. in TE	False
LDAH2	AT1G23330	0.05	n.d. in TE	True
LDAP1	AT1G67360	12.5	51.0	True
LDAP2	AT2G47780	0.02	n.d. in TE	False
LDAP3	AT3G05500	135.8	120.8	True
LDIP	AT5G16550	4.8	54.2	True
LDNP	AT5G04830	6.8	77.5	True
LDS1	AT1G43890	0.9	6.6	True
LIDL1	AT1G18460	0.08	n.d. in TE	True
LIDL2	AT1G73920	0.3	n.d. in TE	True
LIME1	AT4G33110	0.02	n.d. in TE	False
LIME2	AT4G33120	3.7	207.4	True
MAGL8	AT2G39420	0.3	15.7	True
OBL1	AT3G14360	0.2	n.d. in TE	True
OBL3	AT1G45201	0.5	11.3	True
OBL4	AT1G56630	0.4	n.d. in TE	True
OBL5	AT5G42930	0.04	n.d. in TE	False
OLE8	AT3G18570	0.05	n.d. in TE	False
PUX10	AT4G10790	0.2	10.3	True
SMT1	AT5G13710	1.9	29.9	True
$\alpha$ -DOX1	AT3G01420	45.8	46.3	True

Applying a false discovery rate of 0.01 resulted in a selection of 380 proteins significantly enriched in the LD-enriched protein fraction of roots, including all 23 known LD proteins. Their enrichment values ranged from 3.7 to 253.7, with a mean enrichment of 75.7 and a median enrichment of 54.7 (Figure 6A, Supplemental Dataset 1B).

For initial screening, seven protein candidates were selected from the 380 proteins significantly enriched in LDs, and their localization was examined by transient expression experiments in tobacco pollen tubes (Figure 7). Five selected proteins, namely THALIANOL SYNTHASE 1

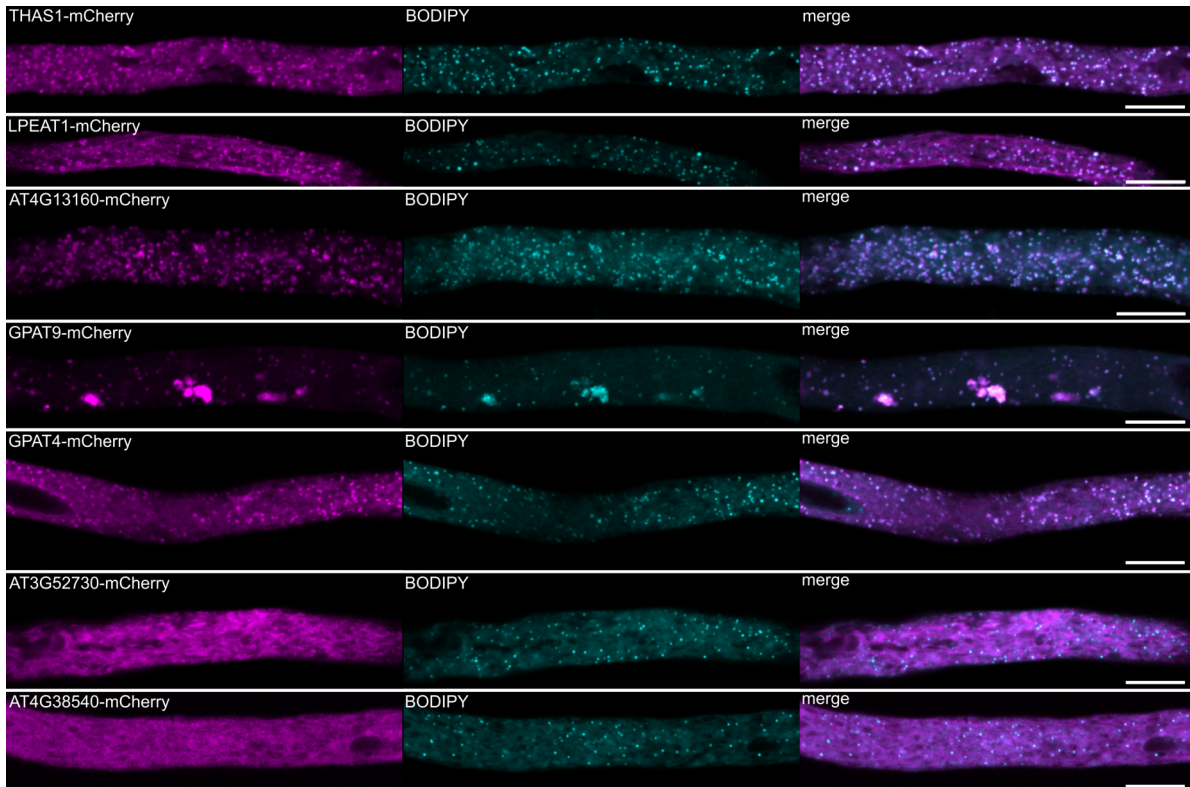




**Figure 6: Volcano plot of Arabidopsis root proteomic data.** The riBAQ dataset was imputed and the values were log<sub>2</sub> transformed. Then, the difference between the LD-enriched and total protein fractions was calculated for each protein. Additionally, the corresponding *p*-values (-log) were determined. A) a volcano plot was generated with both values. The dash-lined box indicates the enlargement shown in B). A false discovery rate (FDR) of 0.01 was used to distinguish between significant and non-significant differences. Furthermore, known LD proteins, as well as screened candidates, were highlighted on the plot, indicating their identified localization. The dashed line in B) marks 50-fold enrichment. ALDH3H1, ALDEHYDE DEHYDROGENASE 3H1; FEY3, FOREVER YOUNG 3; GPAT9, GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 9; LEW1, LEAF WILTING 1; LPEAT1, LYOPHOSPHATIDYLETHANOLAMINE ACYLTRANSFERASE1; PCME, PRENYLCYSTEINE METHYLESTERASE; THAS1, THALIANOL SYNTHASE 1.

(THAS1), a so far uncharacterized protein encoded by the gene AT4G13160, GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 4 (GPAT4), GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 9 (GPAT9), and LYOPHOSPHATIDYLETHANOLAMINE ACYLTRANSFERASE1 (LPEAT1) indeed exhibited at least partial co-localization with LDs in tobacco pollen tubes. Additionally, GPAT9 consistently showed putative ER aggregation artifacts. GPAT4 has recently been described as an LD-associated protein by Fernández-Santos et al. (2020); Thus, we could confirm their observed LD co-localization in tobacco pollen tubes. In contrast, the candidates encoded by the genes AT4G38540 and AT3G52730 did not exhibit LD-colocalization (Figure 7).

Except for LPEAT1, all newly discovered LD-localized proteins and GPAT4 exhibited enrichment levels higher than 50-fold (Table 4, Figure 6B), which represents approximately the above median range enrichment of known LD proteins. In order to screen for more potential LD-associated proteins, we applied a filter to identify proteins that exhibited an enrichment level of at least 50-fold in the LD-enriched fraction. This approach yielded a total of 37 proteins, including 12 known LD proteins and three newly discovered LD proteins, resulting in



**Figure 7: Potential LD protein candidates transiently expressed in tobacco pollen tubes.** The proteins of interest were cloned into pLAT52-mCC vectors (Müller et al., 2017), which were then coated onto gold particles and delivered into the pollen tubes via gene gun transformation. To visualize LDs, the pollen tubes were stained with the lipophilic dye BODIPY 493/503, and images were captured using confocal microscopy. Scale bar  $\cong$  10  $\mu$ m.

a remaining set of 22 high-potential LD protein candidates. Indeed, eight additional proteins were identified that are at least partially associated with LDs either when transiently expressed in tobacco pollen tubes (unpublished work of Dr. Patricia Scholz, Dr. Lea Hembach, Dr. Katharina Blersch, Janis Dabisch, Magdiel Lim Sheng Satha, and Siqi Sun) or *N. benthamiana* leaves (unpublished work of Alyssa Clews and Dr. You Wang, Department of Molecular and Cellular Biology, University of Guelph, Guelph, Ontario, Canada)(Table 4, Figure 6B). On the other hand, one candidate was ER-associated, eight had another targeting than ER or LD, and five candidates have not been screened yet. In addition to that, five candidates not fitting the 50-fold enrichment criteria were transiently expressed in tobacco pollen tubes (Supplementary Dataset 1C). However, neither showed LD-colocalization (unpublished work of Dr. Patricia Scholz, Dr. Lea Hembach, Dr. Katharina Blersch, Janis Dabisch, Magdiel Lim Sheng Satha, and Siqi Sun). Altogether, a total of 12 new proteins associated with LDs were identified by the usage of the root LD enrichment dataset (Table 4).

Interestingly, among the known LD proteins in root were the LD-associated methyltransferases (LIMEs, Table 3) that were previously found in LDs of seedlings (Kretschmar et al., 2020) and were investigated as part of this thesis (Section 5).

---

#### 4. Unpublished results I: The LD proteome of *Arabidopsis thaliana* roots

---

**Table 4: Protein candidates with confirmed LD localization.** The localization of these proteins was validated through transient expression in tobacco pollen tubes and/or *N. benthamiana* leaves. For certain proteins, their characterization remains incomplete, and they are referred to by their respective encoding genes. The table provides the average LD riBAQ values, LD enrichment ratios, and the corresponding  $-\log_{10}$  *p*-values for each candidate. ALDH3H1, ALDEHYDE DEHYDROGENASE 3H1; FEY3, FOREVER YOUNG 3; GPAT9, GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 9; LEW1, LEAF WILTING 1; LPEAT1, LYSOPHOSPHATIDYLETHANOLAMINE ACYLTRANSFERASE 1; PCME, PRENYLCYSTEINE METHYLESTERASE; THAS1, THALIANOL SYNTHASE 1.

AGI	Protein name	riBAQ	Enrichment	<i>p</i> -value
AT5G48010	THAS1	4.90	319.8	6.0
AT1G78800		1.00	206.6	8.0
AT4G27760	FEY3	0.91	176.7	5.5
AT5G15860	PCME	1.08	165.5	5.6
AT1G30130		0.51	159.0	6.5
AT4G13160		0.30	111.8	7.0
AT1G44170	ALDH3H1	6.55	88.7	8.7
AT5G04070		0.19	83.2	6.5
AT5G60620	GPAT9	0.54	81.8	6.2
AT1G11755	LEW1	0.26	71.2	4.7
AT4G33180		0.19	61.6	5.4
AT1G80950	LPEAT1	0.17	15.8	5.6

---

## **5 Unpublished results II: Characterization of lipid droplet-associated methyltransferases**

Author contribution:

Philipp William Niemeyer designed and performed all LIME-related except LC-MS/MS measurements for the ColP experiment and the LC-MS measurements for the *ex vivo* assay. He analyzed the data. Full-length LIME1-mVenus confocal microscopy image originated from Kretzschmar et al. (2020) as a reference for LIME1 truncation experiments. This research is intended for future publication.

## Characterization of lipid droplet-associated methyltransferases

The number of identified LD-associated proteins has shown a steady increase in recent years (Guzha et al. (2023); Kretzschmar et al. (2020), Section 4). Despite these advancements, the function of many of these proteins, particularly concerning their interaction with LDs, remains elusive. In an effort to illuminate one of these candidates, this chapter centers on the characterization of lipid droplet-associated methyltransferases (LIMEs), initially identified as proteins associated with LDs in *Arabidopsis* seedlings by Kretzschmar et al. (2020). For a detailed overview of the current knowledge of LIME proteins, refer to section Introduction 1.3.4.

This chapter comprises a multifaceted array of methodologies employed to characterize LIME proteins. Notably, a phylogenetic analysis of LIME proteins was conducted to discern potential homologs in various species and explore their association with characterized proteins implicated in BIAs biosynthesis. Additionally, investigations encompassed native tissue localization, potential targeting sites to LDs, and LD association of LIME homologs of a variety of plant species. Furthermore, an exploration of potential physiological functions through the generation of *Arabidopsis* mutant lines and the identification of candidate substrates has been performed. In essence, this work marks the pioneering effort to gain insights into the phylogeny, targeting, and functionality of LIME proteins unraveled in the following sections.

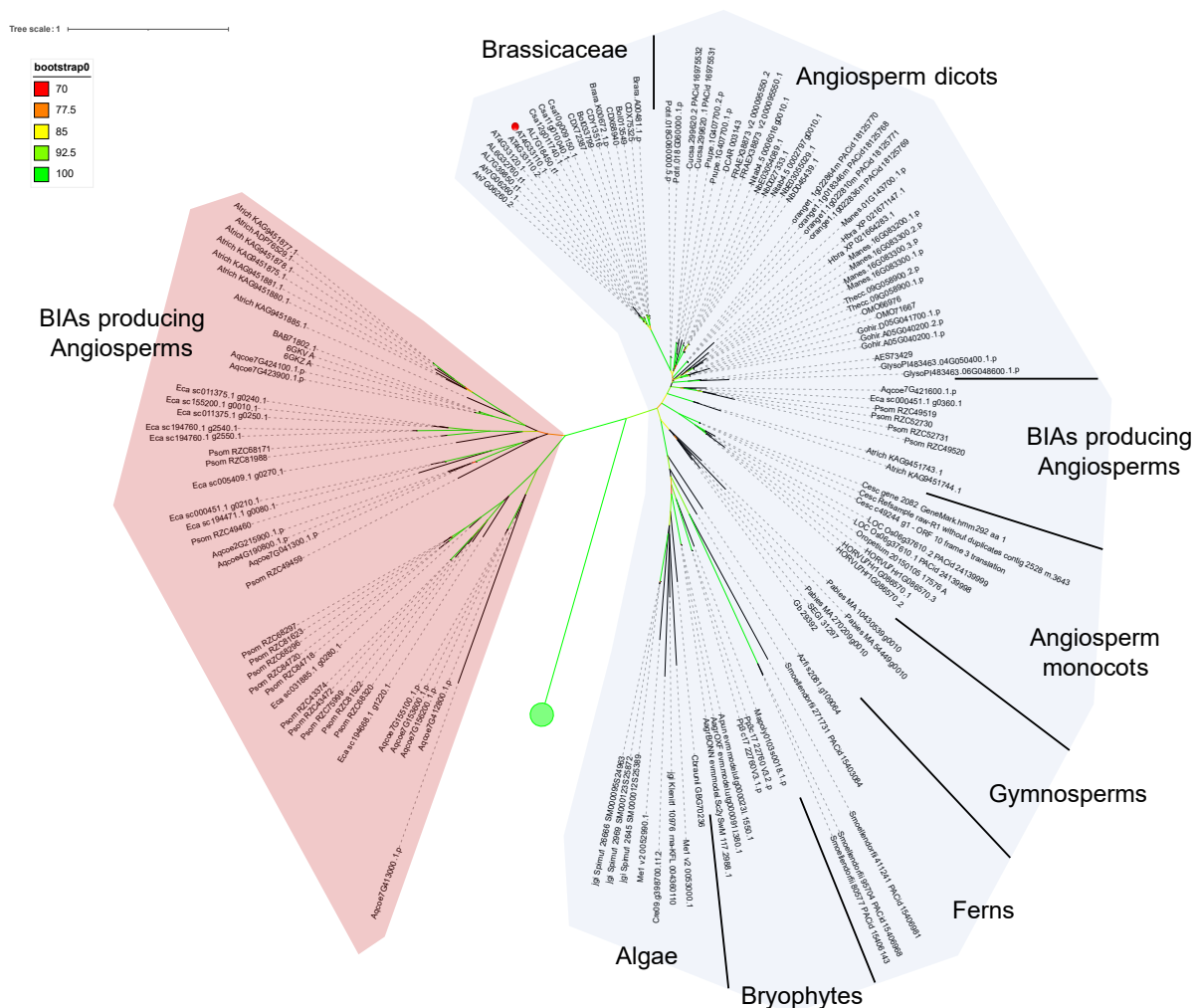
### 5.1 CNMTs possibly diverged from LIME-like proteins

The putative role of proteins is often inferred based on the characterized functions of closely related homologs from the same or other species. The closest characterized homologs of the LIME proteins are *S*-coclaurine-*N*-methyltransferases (CNMTs) and other methyltransferases associated with BIAs-producing species. However, it is uncertain whether LIMEs originated from the CNMT family or whether CNMTs and lime proteins have diverged from a common ancestor and evolved distinct functions. To address this question, a protein family analysis was conducted based on the primary sequences. In order to identify lime homologs, the protein sequence of *Arabidopsis* LIME1 was used as a query in a BLASTp approach against a protein database comprising a wide variety of representatives of the kingdom Viridiplantae with an *e* value cut off  $< 10^{-15}$ . In addition, the dataset contained the protein sequences of the five BIAs-producing species *Aquilegia coerulea*, *Aristolochia fimbriata*, *Coptis japonica*, *Eschscholzia*

## 5. Unpublished results II: Characterization of lipid droplet-associated methyltransferases

*californica*, and *Papaver somniferum*. An unrooted phylogenetic tree was generated with full-length retrieved sequences (Figure 8).

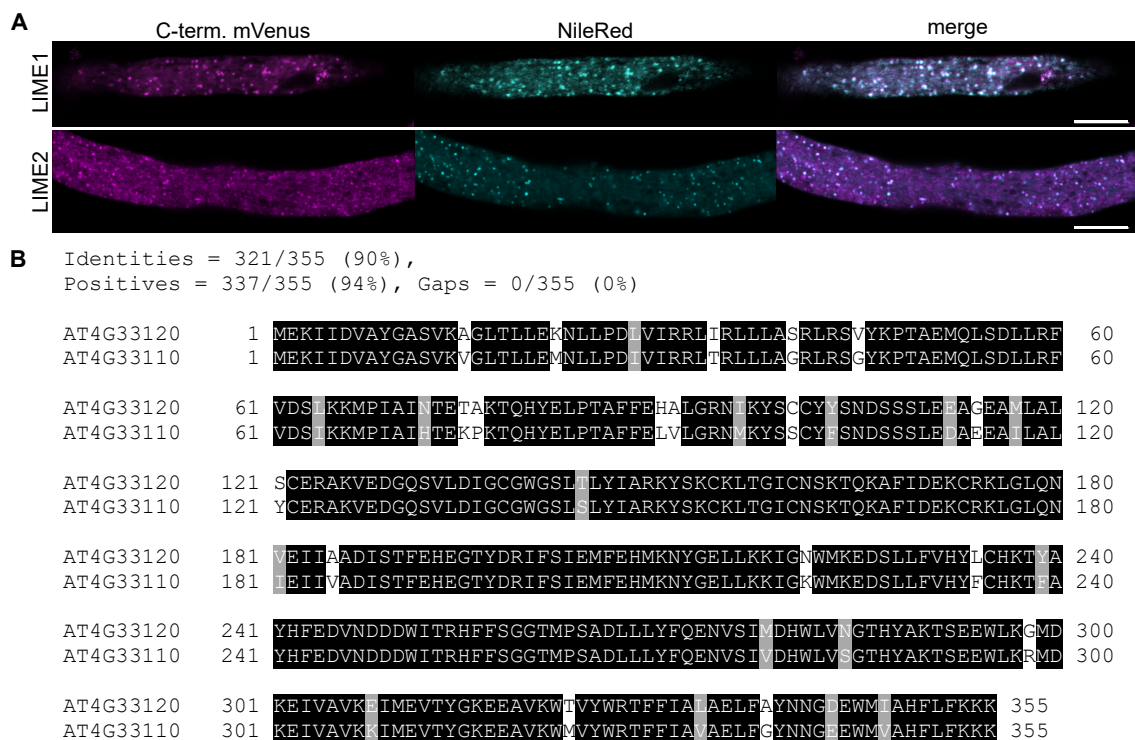
The phylogenetic tree supports the assumption that the CNMTs likely originate from gene duplication in a common Angiosperm ancestor of a LIME-like protein and diverged during the course of evolution. All five BIAs producing species form a distant clade (Figure 8, red background), harboring characterized CNMTs such as from *Coptis japonica* (6GKZ). Furthermore, this clade exclusively comprises BIAs-producing species. This implies that a specialization required for the unique needs of the complex BIA secondary metabolite pathway evolved. The branch of the clade is backed with ultrafast bootstrap support of 99.4, indicating the high reliability of the findings. The putative LIME clade (Figure 8, blue background) has a



**Figure 8: Unrooted maximum likelihood phylogenetic tree of *AtLIME1* homologs.** Homologs were found in selected transcriptome/genome-derived proteomes from Viridiplantae via BLASTp search ( $e$  value cut off  $10^{-15}$ ). BLASTp resulting proteins were aligned with MAFFT L-INS-i and the phylogenetic tree was calculated with W-IQ-TREE. Branch validation was performed by computing 1000 ultrafast bootstrap replicates. Only bootstraps in the range 70-100 are displayed. The putative CNMT clade from BIAs producing species was colored red, putative LIME clade was colored blue. The original *LIME1* sequence is highlighted with a red dot. Distinct groups were collapsed. Phylogenetic groups are annotated.

moderate ultrafast bootstrap support of 92.6. In comparison, a wide variety of species and phylogenetic groups, ranging from algae to Brassicaceae, suggest a broader occurrence of LIME homologs across the phylogenetic tree of plants and an ancient emergence in evolution. Interestingly, four out of five BIAs producing species also have a LIME-like homolog in the putative LIME clade, except for *Coptis japonica* (which had limited entries in the database). Thus, BIAs-producing species may rely on LIME-like functionality in addition to BIAs-related metabolism, implying a distinct molecular function of LIME-like proteins compared to CNMTs. The phylogenetic analysis also suggests that LIME1 and LIME2 in *Arabidopsis* derived from a recent duplication event. Suppose the *e*-value cut-off is set to  $< 10^{-100}$  to filter for the closest homologs. In that case, it can be observed that neither *Arabidopsis lyrata* nor *Arabidopsis halleri*, the species most closely related to *Arabidopsis thaliana*, possess two distinct isoforms of LIME-like homologs (Supplementary Figure 1). In addition, the branch distance of LIME2 outreaches LIME1, hinting at a functional specialization of LIME2.

## 5.2 The LD targeting of LIME proteins



**Figure 9: Visualization of LD localization and sequence alignment of full-length LIME proteins.** A) Transient expression of LIME1 and LIME2 in tobacco pollen tubes using the pLAT52-mVC vector (Müller et al., 2017), stained with Nile Red, and captured by confocal microscopy. LIME1 images were obtained from Kretzschmar et al. (2020) and serve as a reference. Scale bar  $\cong 10 \mu\text{m}$ . B) Sequence alignment of LIME1 (encoded by AT4G33110) and LIME2 (encoded by AT4G33120) highlighting the similarity between the two proteins.

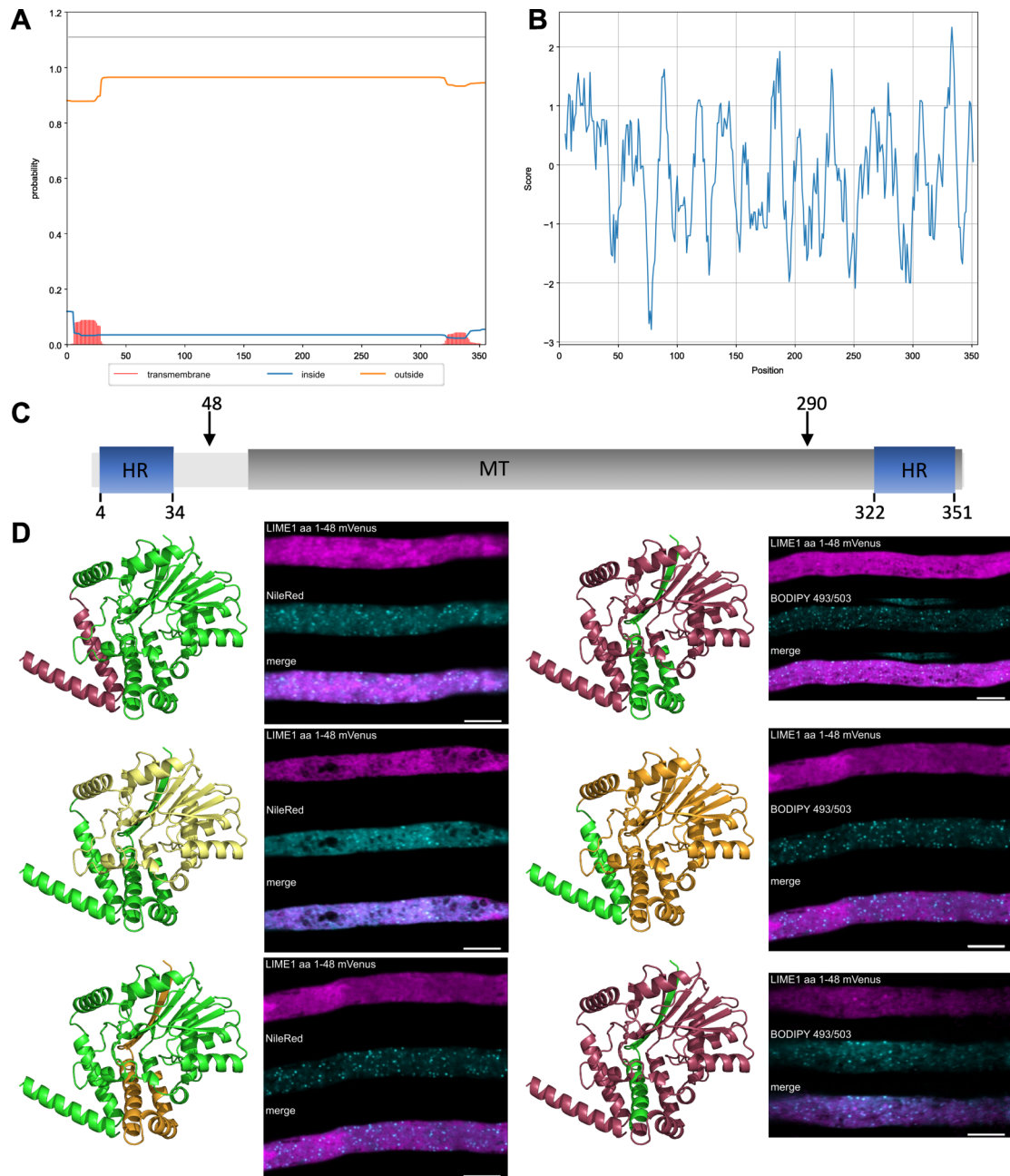
Arabidopsis LIME1 has been described as an LD-associated protein by Kretzschmar et al. (2020). In order to investigate whether LIME2 also targets the LD, recombinant LIME2-mVenus was transiently expressed in tobacco pollen tubes. Indeed, LD co-localization can be observed for LIME2 as well. However, targeting seems less specific towards LDs than LIME1, indicated by a variety of non-LD localized patterns in the microscopic image (Figure 9A). Interestingly, LIME1 and LIME2 have a sequence identity of 90 %, indicating a crucial role of a few amino acids in the sequence for targeting specificity (Figure 9B)

### 5.2.1 The LD targeting of LIME relies on multiple protein segments

An important aspect of the characterization of LD proteins is the identification of their targeting sequences anchoring the protein to the LD monolayer. This information can provide valuable insight into how LD proteins differ from other proteins regarding their membrane targeting mechanisms and ultimately might lead to the sufficient prediction of other targeting sequences by feeding deep learning algorithms. For the estimation of the putative LD binding regions of LIME1, the TMHMM 2.0 web tool was employed, which is specialized in predicting transmembrane helices by hidden Markov models (Krogh et al., 2001). Although classic transmembrane helices are not physically possible in the LD monolayer due to the hydrophobic inner matrix of the LD, predictions can still give valuable information on possible membrane insertions. Indeed, two regions with a small probability of transmembrane helices were predicted: one located close to the N-terminus between residues 4-34 and the other near the C-terminus spanning residues 322-351 (Figure 10A). Consistent with the TMHMM 2.0 prediction, a hydropathicity plot, calculated by the web tool ProtScale (Gasteiger et al., 2005), showed a hydrophobic region spanning from residue 0-41 and from residue 326-337 (Figure 10B). In contrast, a webtool for the prediction of organelle targeting sequences, TargetP 2.0 (Armenteros et al., 2019), predicted no target peptide, demonstrating the unknown nature of LD targeting sequences. To supplement *in silico* analysis, localization studies were carried out with truncated versions of the LIME proteins. A predicted domain structure containing the highly conserved InterPro SAM-dependent methyltransferase superfamily domain from residue 72-350 and the two putative hydrophobic regions binding to the LD monolayer guided the selection of LIME1 truncations used in localization studies (Figure 10C). In order to not cut on domain borders, truncation points were selected between amino acid residues 48-49 and 290-291, keeping a safe distance in case the putative hydrophobic domains span wider than anticipated. Based on this reasoning, truncations in the amino acid range 1-48, 49-290,



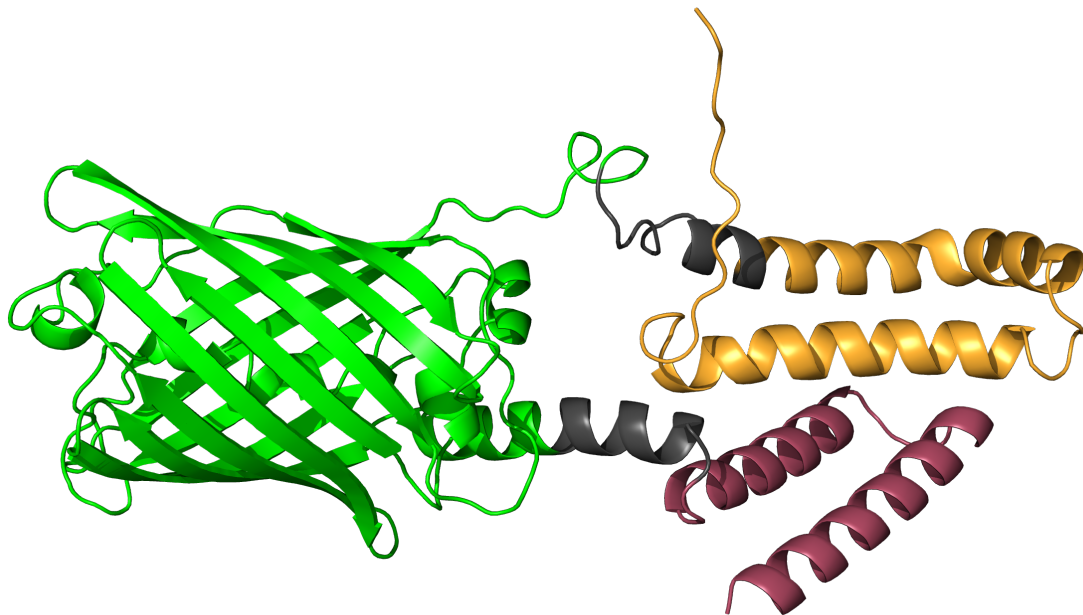
## 5. Unpublished results II: Characterization of lipid droplet-associated methyltransferases



**Figure 10: Putative domain structure of LIME1 and truncation analysis.** Membrane inserting protein sequences were *in silico* predicted using A) TMHMM 2.0 for potential transmembran helices (Krogh et al., 2001) and B) ProtScale hydropathicity plot (Gasteiger et al., 2005; Kyte and Doolittle, 1982). C) The InterPro SAM-dependent methyltransferase superfamily domain was assigned to residues 72-350. A domain model was constructed, highlighting two potential hydrophobic regions spanning residues 4-34 and 322-351. D) Truncations of the LIME protein spanning depicted amino acid ranges were generated and fused with a C-terminal fluorescent protein (mCherry or mVenus). Confocal microscopy was employed to visualize the truncations, while LDs were stained using BODIPY 493/503 or Nile Red. Scale bar  $\cong 10 \mu\text{m}$ . Protein model predictions using AlphaFold Colab (Jumper et al., 2021) depicted the regions of interest in the full-length protein, with red representing regions starting from the N-terminus, orange from the C-terminus, and yellow indicating regions starting from neither end.

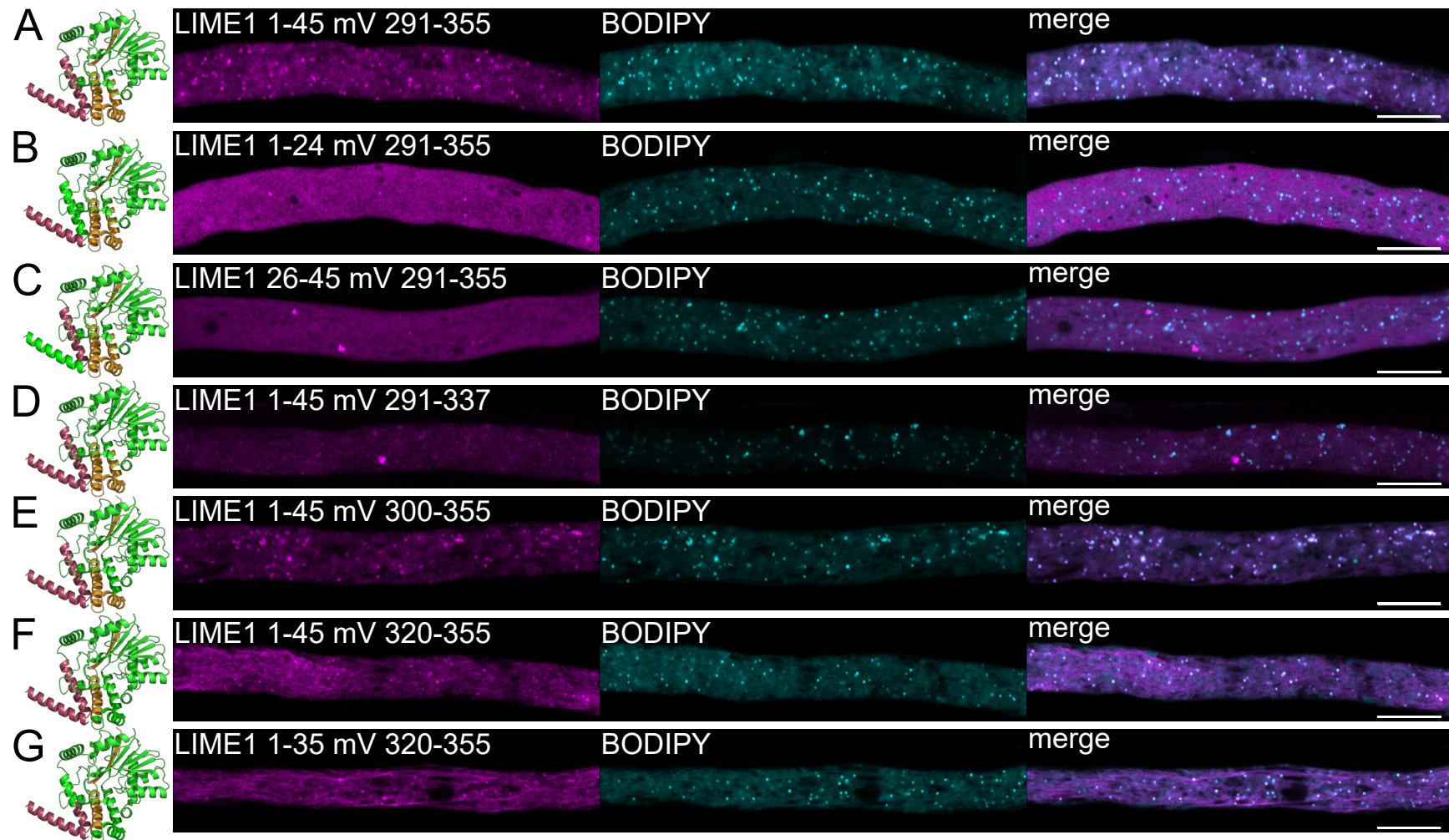
291-355, 1-290, and 49-355 were constructed. These truncations were then fused with a C-terminal fluorescence protein such as mVenus or mCherry. Strikingly, none of the truncated variants displayed LD localization when transiently expressed in tobacco pollen tubes (Figure 10D). Thus, neither the N-terminal nor the C-terminal putative hydrophobic region alone

were sufficient to mediate LD localization, nor was the region between them (Figure 10D). Furthermore, combining the protein termini with the interspacing region did not restore LD localization (Figure 10D). LD-localization could also not be restored when only the last 35 amino acids of the C-terminus were removed (amino acid range 1-320), indicating that the N-terminus is necessary for the localization but not sufficient (Figure 10D). Furthermore, it suggests several sequence motifs are necessary for appropriate LD localization.



**Figure 11: Model prediction of the recombinant LIME1 aa 1-45 mVenus LIME1 aa 291-355 fusion construct.** The protein is depicted with the N-terminal LIME1 aa 1-45 shown in red, mVenus in green, C-terminal LIME1 aa 291-355 in orange, and the linker region in grey. This configuration allows for a native-like proximity of the N- and C-termini. Model prediction was performed using the RoseTTAFold algorithm (Baek et al., 2021).

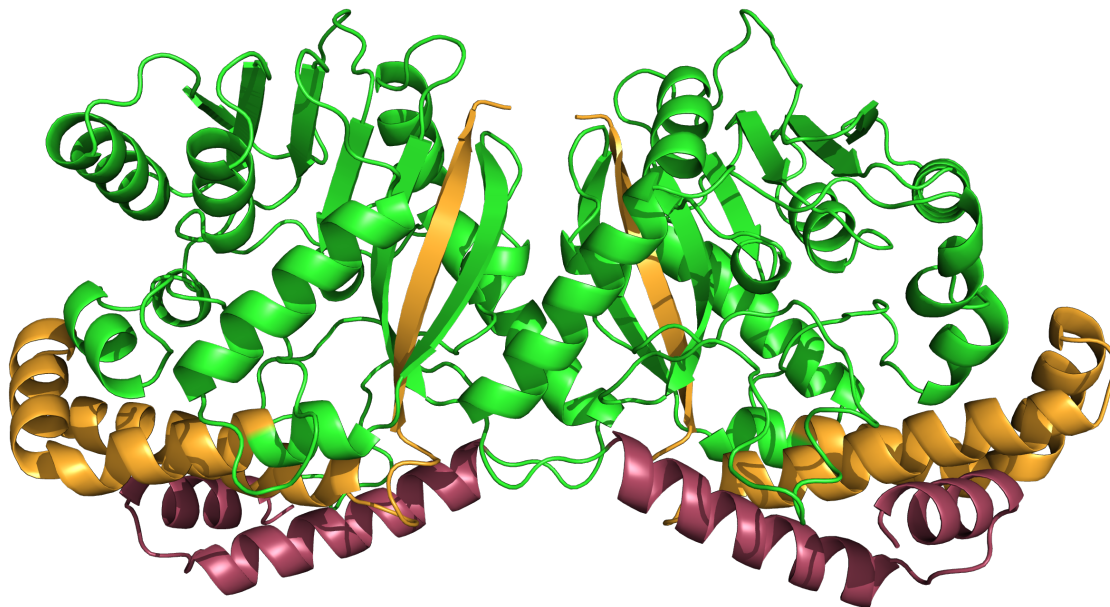
With the rise of high-confidence protein model predictions, the inclusion of putative structural features enabled new possibilities for membrane binding sequence determination. Therefore, a protein model prediction of LIME1 was computed in Deepmind's AlphaFold Colab web application (Jumper et al., 2021). Interestingly, a close spatial proximity of the putative C- and N-terminal hydrophobic regions became apparent. Therefore, a recombinant protein was designed harboring both the N-terminus (amino acids 1-45 to match the first two predicted alpha helices) and the C-terminus (residues 291-355) flanking the fluorescence protein mVenus to resemble natural distance and orientation to each other (Figure 11). This strategy led to a successful restoration of LD localization (Figure 12A), demonstrating the necessity of the combination of these residues for LD targeting. According to the predicted LIME1 model, the N-terminal region consists of two alpha-helices from residues 1-24 and 26-45.



**Figure 12: Confocal imagery of combined truncation lines.** N-terminal and C-terminal truncations of LIME1, with a mVenus fluorescence protein in between, were transiently expressed in tobacco pollen tubes. LDs were stained with Nile Red. Scale bar  $\cong 10 \mu\text{m}$ . On the left side, the protein model prediction of LIME1 (generated using AlphaFold Colab) is shown, highlighting the regions used in the corresponding truncation. Red parts correspond to the N-terminus, orange parts represent the C-terminus, and green denotes regions that have been truncated.

However, each individual alpha-helix was unable to restore localization to LDs when recombined with mVenus and the C-terminal residues 291-355 (Figure 12B,C). Instead, the respective protein shows cytosolic localization with a tendency of aggregation. This suggests that at least a partial combination of both predicted N-terminal alpha-helices are necessary for appropriate LD targeting.

According to the protein model, residues 345-353 are part of a buried beta-sheet within the protein. Thus, it is unlikely that they interact with the monolayer directly. However, construct harboring residues 1-45, mVenus, and residues 291-337, thereby truncating the beta-sheet, strongly reduced LD targeting (Figure 12D). Presumably, the beta-sheet might act as an anchor to stabilize the spatial proximity of the C-terminal site towards the N-terminal site. To specify the necessary binding site more precisely, the approved N-terminal residues 1-45 were combined either with C-terminal residues 300-355 or residues 320-355. While residues 300-355 were sufficient for LD localization (Figure 12E), the combination with residues 320-355 (Figure 12F) also led to targeting to filamentous formations, which is presumably ER. Consequently, despite the LD targeting restoration capability of residues 320-355, residues 300-319 are potentially



**Figure 13: Protein model prediction of a putative homo-dimeric confirmation of LIME1.** Prediction was performed with SWISS-MODEL (Waterhouse et al., 2018). Amino acid residues 1-45 are highlighted in red and residues 291-355 are highlighted in orange to emphasize regions that have been shown to facilitate LD localization. The putative binding domains are oriented in the same plane, supporting their assumed role as binding domains.

required to discriminate targeting to ER membrane and to specify LD targeting. The same localization was observed when the second predicted alpha-helix on the N-terminal site was halved, resulting in residues 1-35 combined with residues 320-355, demonstrating that this truncation is still sufficient to target LDs to some degree when combined with the C-terminus (Figure 12G).

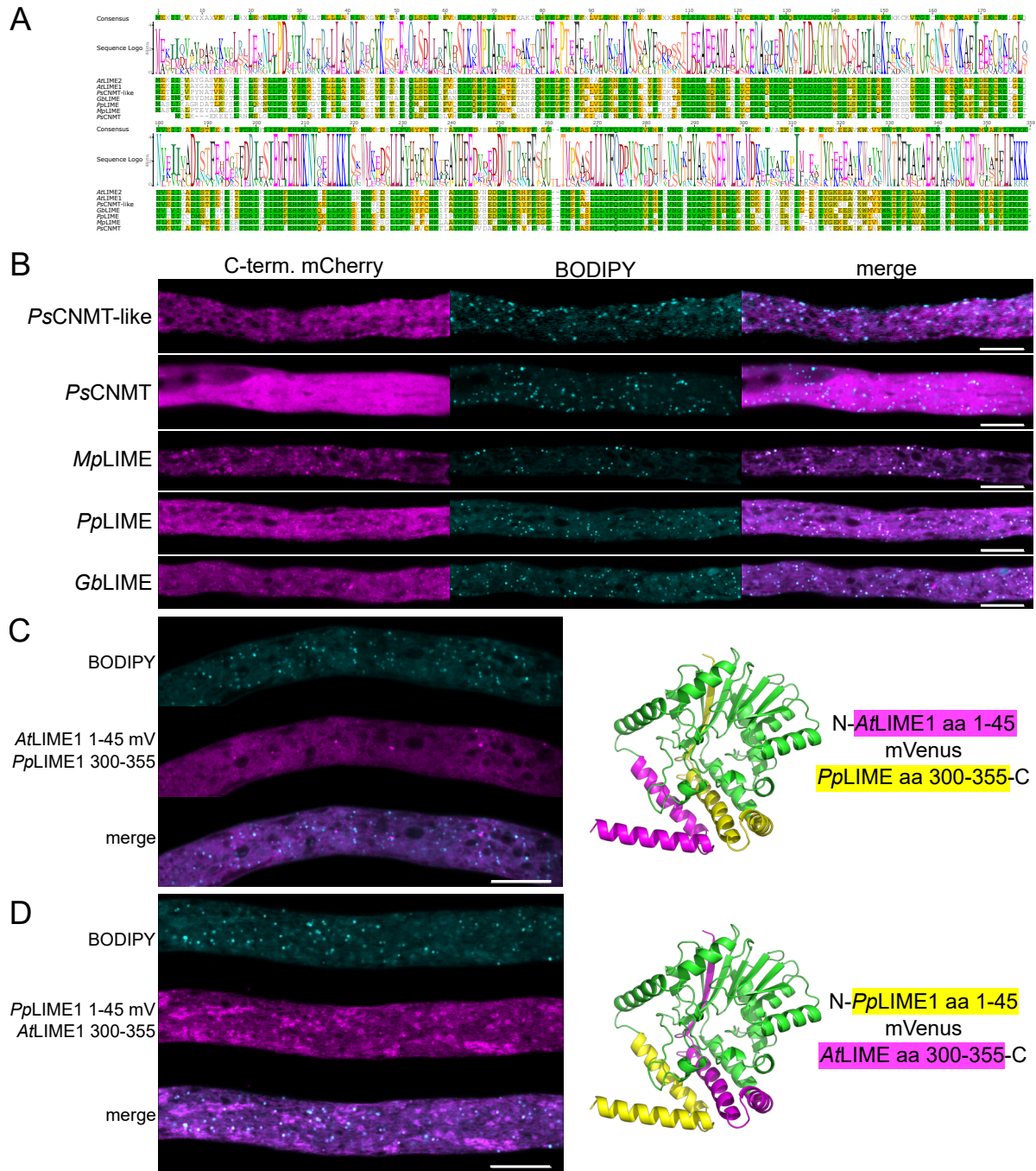
In summary, LIME1 residues 1-45 in combination with residues 300-355 are the minimum required sequences for LD-specific and sufficient targeting. Moreover, the necessary sequence section for LD and ER targeting seems to be within residues 1-35 on the N-terminal site and residues 320-337 on the C-terminal site, both within the initial prediction of the putative membrane binding regions (Figure 10C).

In addition, protein model prediction performed via SWISS-MODEL (Waterhouse et al., 2018) assumed LIME1 as a potential homo dimer (Figure 13). The potential membrane binding areas determined previously would align in the predicted dimeric configuration in a directional plane, thereby supporting their role as potential LD membrane binding sites.

### **5.2.2 LD targeting of LIME homologs is conserved**

Phylogenetic analysis revealed a possible ancient heritage of LIME-like proteins in Viridiplantae (refer to section 5.1). However, the conservation of LD targeting among homologs was unknown. Therefore, a selection of homologs from several species were selected, comprising the bryophytes *Marchantia polymorpha* and *Physcomitrium patens*, the gymnosperm *Ginkgo biloba*, and the BIAs producing angiosperm opium poppy *Papaver somniferum*. The characterized and cytosolic localized reported poppy CNMT was chosen as a control (Hagel and Facchini, 2012). All selected homologs, including the CNMT control, exhibit significant sequence similarity, implying high conservation among LIME-like proteins (Figure 14A). For the control poppy CNMT, cytosolic localization could be confirmed when transiently expressed in tobacco pollen tubes (Figure 14B). On the other hand, all other LIME-like homologs, including the uncharacterized poppy CNMT-like homolog, showed co-localization with LDs and an additional affinity to filamentous structures, which are presumably the ER. This suggests that LD localization is an ancient trait of LIME-like proteins that may have existed since embryophytes emerged or even earlier. Furthermore, it supports the assumption that poppy CNMT diverged in functional properties compared to LIME-like poppy CNMT-like protein due to the distinct localization pattern.

## 5. Unpublished results II: Characterization of lipid droplet-associated methyltransferases



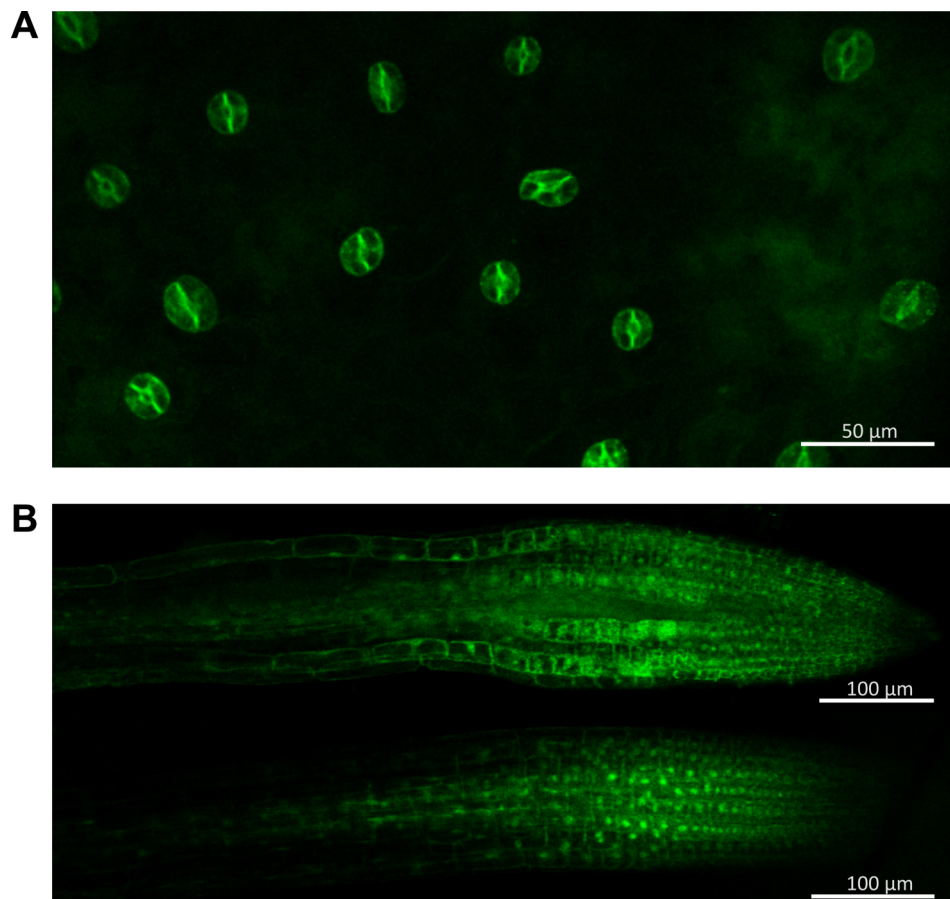
**Figure 14: Variety of LIME protein homologs.** A) Multiple homologs of *Arabidopsis thaliana* (AtLIME) proteins were identified in other plant species, including *Ginkgo biloba* GbLIME, *Marchantia polymorpha* MpLIME, *Papaver somniferum* PsCNMT-like, and *Physcomitrium patens* PpLIME. The *Papaver somniferum* PsCNMT was used as a characterized control. Multiple sequence alignment and LOGO representation of all sequences were performed. B) Truncations of LIME homologs fused with C-terminal mCherry were transiently expressed in pollen tubes, with LD stained using BODIPY 493/503, and captured via confocal microscopy. C/D) Chimeric recombination of N-terminal amino acid residues 1-45 from Arabidopsis/*P. patens* and C-terminal amino acid residues 300-355 from *P. patens*/Arabidopsis was created between a mVenus fluorescence protein. These chimeric proteins were transiently expressed in tobacco pollen tubes, LD were stained with Nile Red, images were captured by confocal microscopy. Scale bar  $\cong 10 \mu\text{m}$ . The positions of truncations within the full-length protein were highlighted using color in a protein model prediction of Arabidopsis LIME1 generated with AlphaFold Colab (Jumper et al., 2021).

Following the successful LD-targeting of Arabidopsis LIME1 through recombining its N- and C-termini flanking a fluorescence protein (refer to section 5.2.1), the question arose whether this approach would also prove effective for chimeric versions of the protein, given its high

sequence similarities. Interestingly, neither chimeric recombinant protein showed wild-type localization (Figure 14C,D). The fusion of N-terminal Arabidopsis LIME1 residues 1-45 with C-terminal *P. patens* LIME residues 300-355 yielded cytosolic localization with presumable protein aggregations. In contrast, fusing *PpLIME1* residues 1-45 with C-terminal *AtLIME* residues 300-355 led to partially restored LD co-localization but also resulted in an ER stress phenotype. Thus, mutations occurring on the binding sites during evolution needed to co-evolve to maintain targeting, which further suggests that LD (and ER) targeting is crucial for the function of LIME-like proteins.

### 5.2.3 Localization of LIME proteins expressed under the endogenous promoter

The tissue-specific expression pattern of Arabidopsis LIME proteins has been investigated solely through transcriptomic data (Klepikova et al., 2016), with the difficulty that microarray data does not distinguish between *LIME1* and *LIME2*. To confirm the subcellular localization,



**Figure 15: LIME proteins expressed under native promoter.** LIME proteins fused with C-terminal eGFP were expressed under putative native promoter in Arabidopsis. The fluorescence signal of A) *LIME1-eGFP* was detected in guard cells of 2-day-old leaves, or B) of *LIME2-eGFP* in root tips of 5-day-old plants (two independent lines are shown). Images were acquired by confocal microscopy.

and to check for tissue or cell-type-specific expression, the wild-type Arabidopsis Col-0 was stably transformed with the LIME-coding cDNA under the control of their corresponding endogenous promoter region ( $\text{Nat}_{\text{LIME}}$ ), and an additional C-terminal eGFP sequence.

At 48 h post-germination, the T2 generation  $\text{pNat}_{\text{LIME1\_LIME1\_eGFP}}$  line seedlings exhibited expression in the guard cells (Figure 15A), consistent with previous transcriptomic studies of guard cells (Obulareddy et al., 2013). Conversely, transcriptomic analysis of LIME2 suggests its predominant occurrence in the root tip (Klepikova et al., 2016), which was confirmed in 5-day-old roots of the T2 generation  $\text{pNat}_{\text{LIME2\_LIME2\_eGFP}}$  lines. The fluorescence intensity of *LIME2-eGFP* peaked at the transition zone between the meristem and elongation zone of the roots (Figure 15B). Moreover, the potential interaction of *LIME2-eGFP* with the nucleus requires validation through DAPI staining.

### 5.3 CoIP supports dimeric conformation of LIME proteins

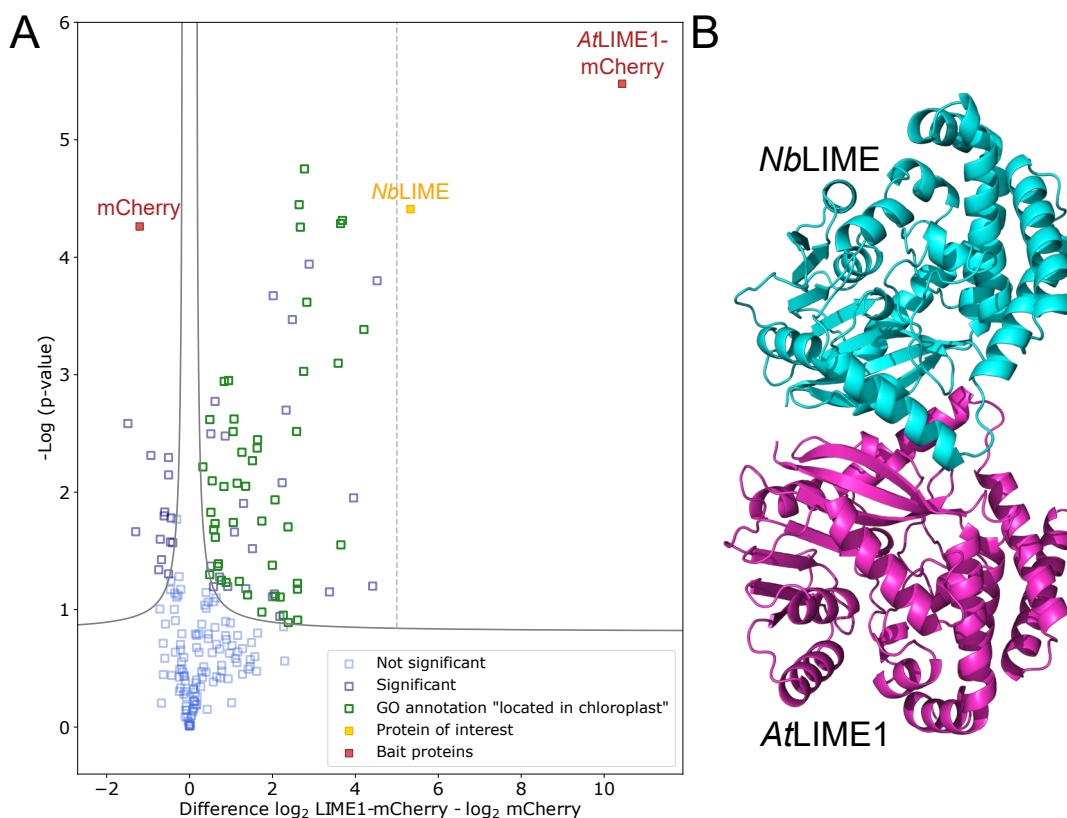
Several LD proteins have known interaction partners (Guzha et al., 2023). However, interactions of LIME1 have not been investigated so far. To determine whether LIME1 interacts with other proteins, a co-immunoprecipitation-mass spectrometry (CoIP-MS) experiment was performed. In order to conduct the CoIP, recombinant LIME1-mCherry was transiently expressed in *N. benthamiana* leaves, and the expressed protein was immobilized on beads bound to nanobodies with affinity to mCherry. A previous study has shown that LIME1-mCherry successfully targets LDs in *N. benthamiana* leaves (Kretzschmar et al., 2020), indicating that the fusion protein maintains its proper folding and targeting despite the addition of the C-terminal fluorescent protein. For exclusion of unspecific interaction with the beads and mCherry, free mCherry was transiently expressed in *N. benthamiana* leaves as a comparative control. For each condition, three biological replicates were processed to ensure the consistency of results. To enhance the chance of stabilizing protein-protein interactions, Lomant's Reagent crosslinker was used.

In the subsequent LC-MS/MS analysis, a total of 645 proteins were identified (Supplemental Dataset 2), with 777 rLFQ of all proteins in the free mCherry sample affiliated with mCherry. In the LIME1-mCherry samples, 394 rLFQ of the identified proteins belonged to LIME1 and 339 rLFQ assigned to mCherry, resulting in a combined 722 rLFQ belonging to LIME1-mCherry. This indicates successful immobilization of the bait proteins by nanobodies bound to selector beads. For the identification of potential interaction partners, proteins were filtered by two



criteria: i) identification of the protein by at least two unique peptides, and ii) identification of a protein in three out of three replicates for at least one condition. After applying the filtering criteria, a total of 205 proteins remained, which were  $\log_2$ -transformed (missing values were imputed) to determine the differences between conditions, representing the enrichment of protein co-immobilized with LIME1. Furthermore, statistical significance was calculated through a Student's *t*-test. The significance and  $\log_2$  difference values were plotted in a volcano plot (Figure 16A).

Applying an FDR threshold of 0.01, a total of 74 proteins were found to be significantly enriched in LIME1-mCherry baited samples compared to free mCherry baited samples. Of these, only a single protein was identified as a homolog of a known LD protein and showed a higher than 32-fold enrichment. Interestingly, this protein is a *N. benthamiana* (41.9-fold enrichment, *p*-value 4.4) ortholog of AtLIME1. Therefore, it is conceivable that AtLIME1 interacted with NbLIME and formed a putative chimeric heterodimer (Figure 16B), which would also imply



**Figure 16: Volcano plot of CoIP.** A) LIME1-mCherry and free mCherry were transiently expressed in *N. benthamiana* leaves for two days. A CoIP-MS approach was performed, by immobilizing bait-proteins with potential interactors with RFP-selector beads and subsequent LC-MS/MS measurements. The resulting proteins were filtered by criteria identified in at least 2 peptides and in 3 of 3 replicates in either condition.  $\log_2$  difference of each protein was plotted against  $-\log_{10} p$ -value. Bait proteins, proteins of interest, and proteins GO annotated as localized in chloroplasts were specifically highlighted. The dashed line indicates 32-fold enrichment. B) putative chimeric NbLIME AtLIME1 hetero-dimer. Protein models were computed with AlphaFold Colab (Jumper et al., 2021) and aligned against LIME1 homo-dimer model computed by SWISS-MODEL (Waterhouse et al., 2018), which served as a dimer reference.

similar localization of both proteins. Out of the remaining 73 significantly enriched proteins, 49 are located in chloroplasts according to GO term annotation, indicating strong plastidial contamination. Thus, the applied CoIP approach was considered imperfect for estimating the protein-protein interaction of LIME proteins. Despite this limitation, the CoIP approach provided some indication for the putative dimeric shape of LIME, supporting the previously postulated possibility (refer to section 5.2.1) that LIME proteins may exist as a dimer.

#### **5.4 Mutant lines were obtained for phenotype determination experiments**

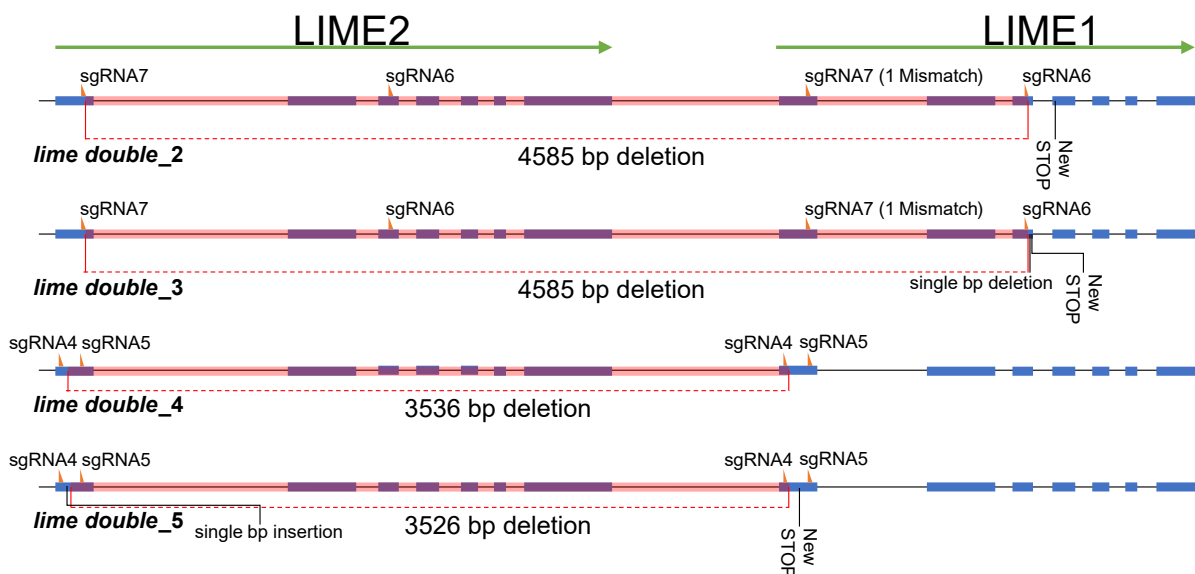
One way to collect information on the physiological and biochemical function of a putative enzyme is the study of knockout lines. In some instances, mutations of both LIME protein family members may be advantageous to eliminate potential functional redundancies and better elucidate possible phenotypic effects. One traditional approach for creating multi-mutant lines involves the utilization of line crossing. However, it is improbable to achieve allele recombination through traditional crossing of single mutants *lime1* and *lime2* due to their close physical proximity. Hence, a CRISPR/Cas9 approach was applied to generate a *lime double* mutant. The significant sequence homology between the two genes was utilized to design single-guide RNAs (sgRNAs) that simultaneously target both genes, with a maximum of one mismatched nucleotide allowed per sgRNA and gene. To increase the likelihood of a deletion event occurring either in both individual genes simultaneously or even as a single deletion of both genes, two independent sgRNAs were designed per CRISPR/Cas9 cassette.

Furthermore, two different cassettes were designed to obtain truly independent mutation patterns. The selection of mutant lines was based on the following criteria: (i) derived from independent T1 seeds, (ii) presence of deletions in both LIME genes, (iii) homozygosity for the CRISPR-induced deletions, and (iv) removal of the CRISPR/Cas9 cassette from the genome. In addition, a premature STOP codon derived from deletion-induced frameshift was considered as beneficial but not excluding criterium. In total, four *lime double* mutant lines were obtained that fulfilled the main selection criteria (Figure 17), referred to as *lime double\_2* to 5.

The *lime double\_2* mutant line carries a 4858-nucleotide deletion that spans from LIME2 exon 1 position 144 to LIME1 exon 3 position 1205, leading to a frameshift and a premature STOP codon at LIME1 positions 1339-1341. This deletion potentially results in a fusion protein of 60 amino acids (aa) in length. The *lime double\_3* mutant line also carries a 4858-nucleotide deletion that spans from LIME2 exon 1 position 144 to LIME1 exon 3 position 1205, and an

additional single nucleotide deletion in LIME1 exon 3 position 1207, leading to a frameshift and a premature STOP codon at LIME1 positions 1234-1236. A potential fusion protein of 57 aa length could be produced in this configuration. The *lime double\_4* lines derive from another CRISPR/Cas9 cassette. It has a 3536 nucleotide spanning deletion from LIME2 exon 1 position 20 to LIME1 exon 1 position 32, which does not cause a premature STOP codon. Thus, a potential fusion protein of 351 aa length can get produced, almost resembling full-length LIME1 (355 aa) with the endogenous promotor of LIME2. The *lime double\_5* derives line from the same CRISPR/Cas9 cassette as *lime double\_4*. However, it harbors the single nucleotide insertion 23+A and a 3526 nucleotide deletion spanning from LIME2 exon 1 position 31 to LIME1 exon 1 position 31, causing a frameshift, a premature stop codon in LIME1 exon 1 positions 90-92, and a potential fusion protein of 30 aa length.

In addition to CRISPR deletion lines, single T-DNA insertion lines were purchased from The European Arabidopsis Stock Center (NASC). For *lime1* lines, SALK\_024876 (*lime1-1*)(Alonso et al., 2003), GT\_3\_9311 (*lime1-2*)(Sundaresan et al., 1995), and WiscDsLox485-488C12 (*lime1-3*)(Nishal et al., 2005) were acquired. However, only for line *lime1-1*, a T-DNA insertion between LIME1 intron 3 positions 1269-1270 could be confirmed by PCR and sequencing. On the other hand, for *lime2* lines, line SM\_3\_34300 (*lime2-1*)(Tissier et al., 1999) had a confirmed T-DNA insertion between LIME2 exon 2 positions 1228-1229, and line SALK\_127412 (*lime2-2*)(Alonso et al., 2003) between LIME2 intron 2 positions 1484-1485. Thus, various single and double mutant lines could be used to investigate potential LIME protein phenotypes.

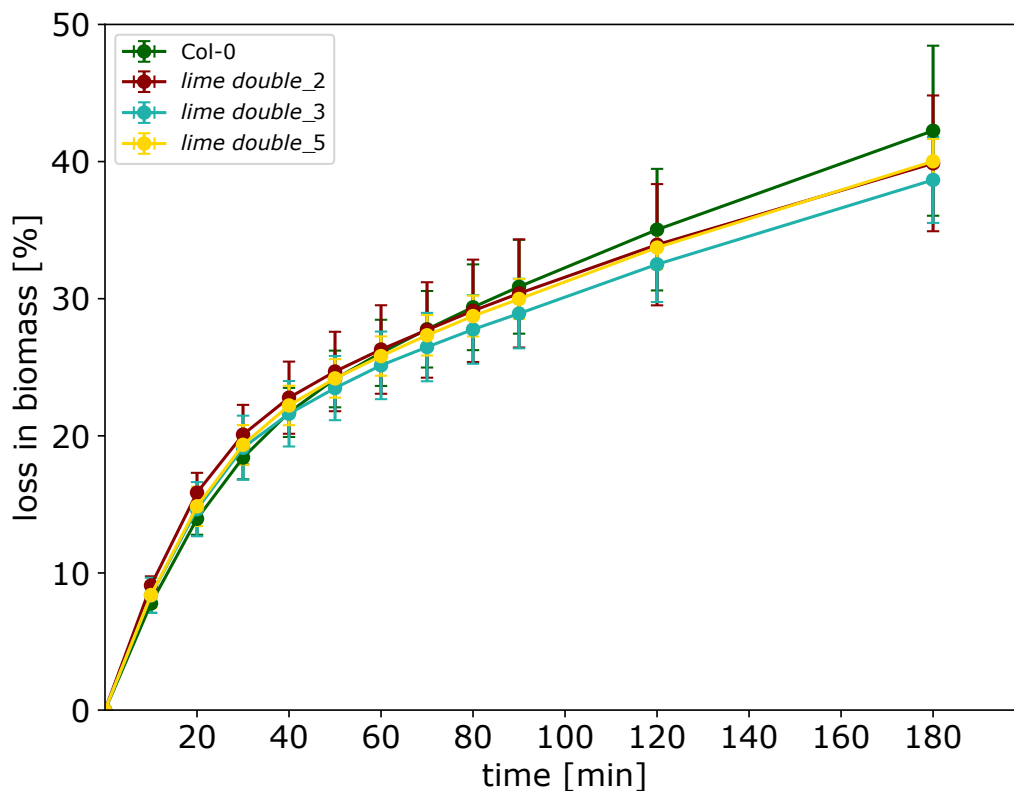


**Figure 17: Deletions caused by CRISPR/Cas9 in the *lime double* knockout mutant lines.** Confirmed lines *lime double\_2-5* are shown, with the full-length gene represented by the green arrow, exons shown in blue, and large base pair deletions marked in red. The figure indicates the sgRNA targets (including mismatch in either gene), as well as single base pair insertions/deletions and premature STOP codons.

## 5.5 LIME proteins do not affect stomata closure

A previous experiment demonstrated *LIME1* expression in guard cells, suggesting a potential function there (refer to section 5.2.3). To analyze whether LIME proteins have an impact on stomata closure, the water loss of three independent double mutant lines *lime double\_2*, *lime double\_3*, and *lime double\_5* were compared against Col-0 wild type (wt). As a result, the relative degree of stomata opening is indirectly shown. In order to carry out the comparison, whole leaf rosettes of 24-day-old plants were cut, positioned with the stomata facing upwards, and measured the relative weight loss over a period of three hours.

Analysis revealed no statistically significant effect between wt and double mutant lines over the analyzed time course (Figure 18). Consequently, LIME proteins seem to have no strong effects on stomatal opening regulation.



**Figure 18: Comparison of water loss between *lime double* mutants and Col-0 wild type.** Leaves from 24-day-old Arabidopsis plants (Col-0 wt, *lime double\_2*, *lime double\_3*, and *lime double\_5*) were cut, with the abaxial side facing upwards, and their weights were recorded at regular timeframes over 180 min. The relative water loss was documented, and the standard deviation was calculated based on 5 replicates per line.

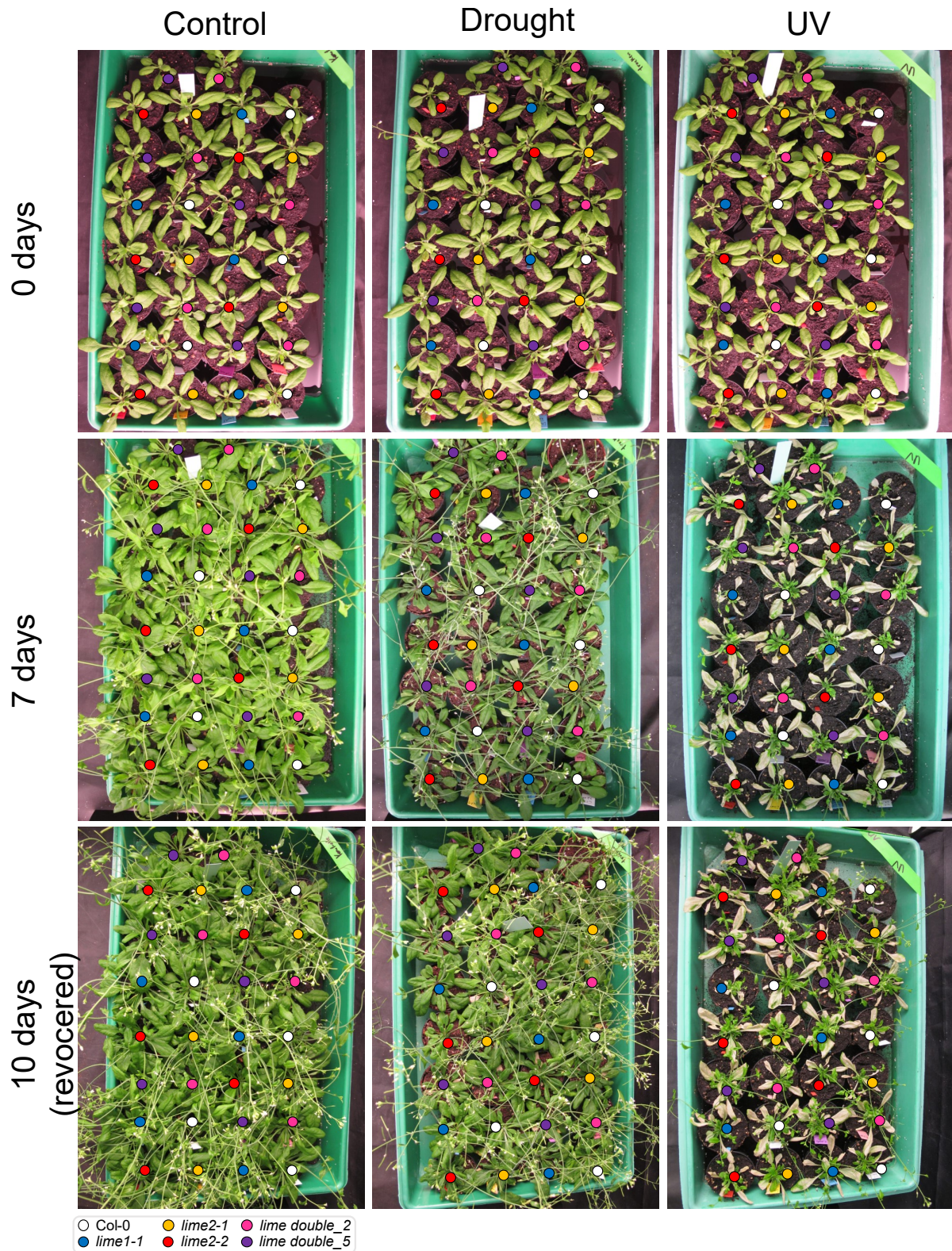
## 5.6 LIME proteins do not affect mildly dehydrated or UV-stressed plants

A previous experiment indicates that LIME proteins possibly do not affect stomata opening under harsh conditions of rapid change (refer to section 5.5). However, whether mild continuous dehydration stress causes a different result remains uncertain. *LIME1* is expressed in leaves and guard cells, and *LIME2* in roots (refer to section 5.2.3), and are thus located in the water regulation system and uptake organ, respectively. Hence, there might be a link to the water regulation systems. Furthermore, leaves are exposed to UV radiation; thus, the effect of UV stress on *lime* mutant lines compared to the wt was investigated. For the analysis, 26-day-old plants comprising Col-0 wt, *lime1* and *lime2* single mutants, as well as *lime double* mutant lines were treated with 7 days of water deprivation for mild dehydration stress, or with a 30 min treatment with UV-C light for radiation stress. Plants were documented after 7 days post-treatment and after 10 days, representing 3 days of recovery for dehydrated plants. Neither beneficial nor detrimental effects for dehydration and UV stress were observed between wt and any mutant line (Figure 19). Thus, LIME proteins seem not to affect dehydration or UV stress regulation mechanisms.

## 5.7 LIME proteins seem to play no role in root development

Based on transcriptomic data (Klepikova et al., 2016) and prior protein expression experiments (refer to section 5.2.3), it has been observed that *LIME2* is expressed in the root tip of Arabidopsis. Therefore, it is conceivable that LIME2 has an impact on root development. Furthermore, *LIME1* has also been detected low-abundantly in the roots of LD-enriched fraction of an oil-rich Arabidopsis mutant (refer to section 4), indicating a possible role in roots. In order to investigate the potential influence of LIME proteins on root length, Col-0 wt was grown vertically in parallel with *lime1* and *lime2* single mutants, as well as *lime double* mutants. The root length was determined 3, 5, and 8 days after imbibition (dai).

At 3 dai, no significant difference in root length was observed (Figure 20). Nonetheless, at 5 dai a significant decrease in root length was observed in the *lime2-2* mutant line, whereas the remaining mutant lines exhibited only a minor reduction compared to the wt, which was not statistically significant (Figure 20). This trend also persisted at 8 dai, although the decrease in root length was not statistically significant anymore (Figure 20). Additionally, the observed variability in the extent of root length reduction among the mutant lines suggests that it may result from a mutation at another locus rather than a consequence of the mutation in the *LIME2*

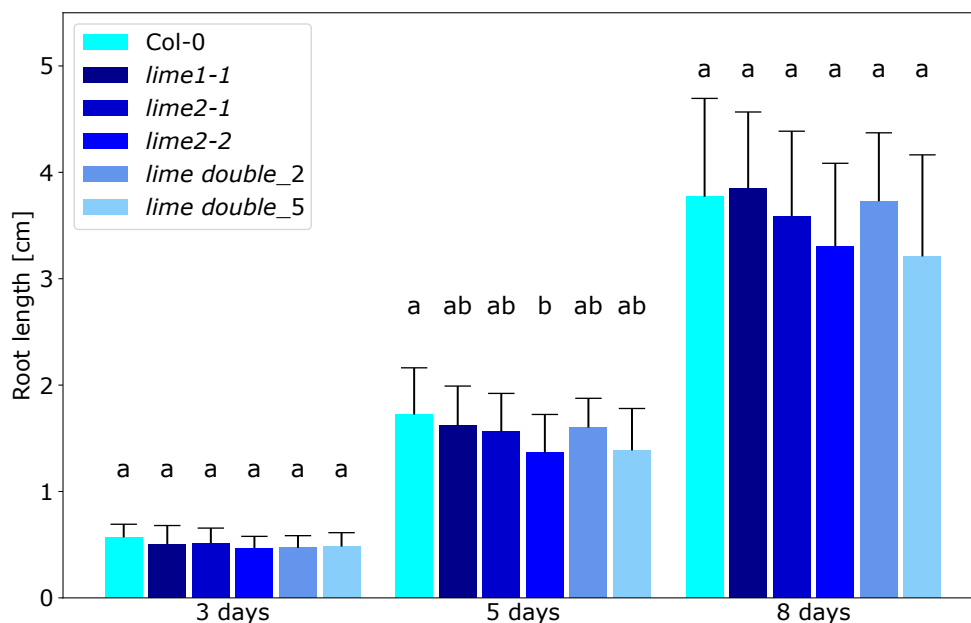


**Figure 19: Impact of *lime* mutant lines on mild dehydration and severe UV stress.** Comparison of 26-day-old *Arabidopsis* lines, including Col-0, *lime1-1*, *lime2-1*, *lime2-2*, *lime double\_2*, and *lime double\_5*, under different stress conditions. The plants were subjected to seven days of water deprivation (with recovery on the seventh day) or a 30-minute UV-C treatment, and their responses were compared to untreated control plants. The time points indicated are: 0 days (26-day-old plants), 7 days (recovery of dehydration stress) and 10 days after stress application.

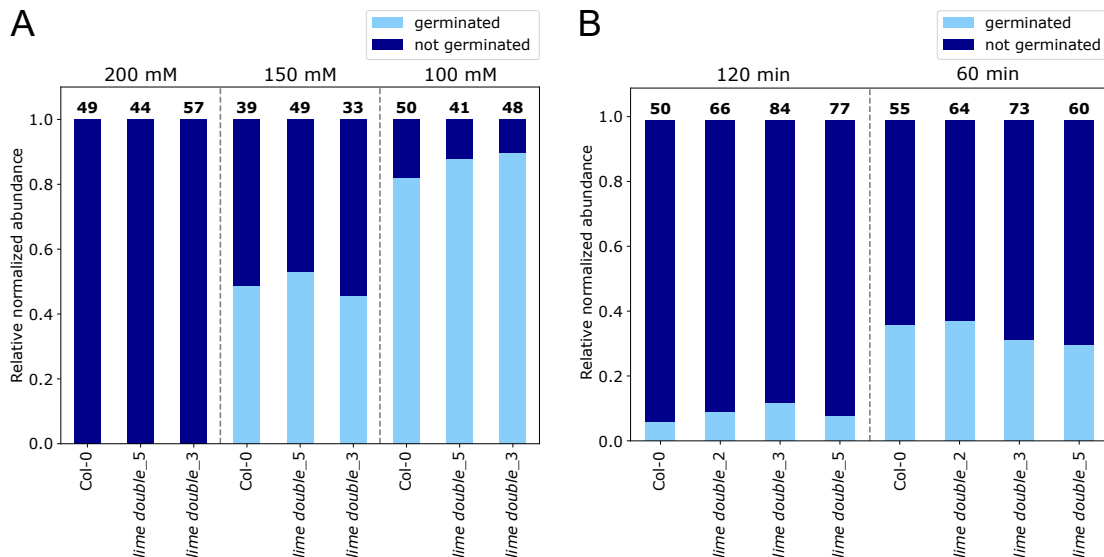
gene. Moreover, there is a possibility of undetected additional mutations within individual lines that could contribute to further effects. Thus, our findings do not give any evidence for a role of LIME proteins in root elongation under laboratory conditions. However, given the limited sample size, variability in the results, and the inherent variation among equivalent mutant lines, this finding should be interpreted cautiously and cannot be considered conclusive without further evidence.

### 5.8 LIME is not important for seed germination under salt and heat stress

Under optimal growth conditions, none of the *lime* mutant lines exhibited any developmental abnormalities or aberrational-shaped phenotypes in comparison to Col-0 wt (Supplemental Figure 2). Therefore, the responses of *lime double* mutant lines were assessed under abiotic stress conditions, apart from UV and drought. For the analysis, the seeds of *lime double* lines and Col-0 wt were exposed to heat for up to 2 h or placed on media containing increasing salt concentration. The germination efficiency was determined after 3 days (salt stress) or 10 days (heat stress). However, no significant effect was observed between wt and mutant lines at any



**Figure 20: Root length of *lime* mutants lines in comparison to Col-0 wt.** Arabidopsis line Col-0, *lime1-1*, *lime2-1*, *lime2-2*, *lime double\_2*, and *lime double\_5* were vertically grown on 1/2 ms -suc medium. Root length was assessed at 3, 5, and 8 days after imbibition. Standard deviation was calculated, and ANOVA statistical analysis followed by Tukey's test was performed. The statistical results were represented using small letter codes.



**Figure 21: Knock-out *lime double* do not have a salt or heat stress phenotype in comparison to Col-0 wt.** Arabidopsis lines Col-0, *lime double\_2* (heat stress only), *lime double\_3*, and *lime double\_5* were grown on 1/2 MS + 0.5% sucrose medium, with a salt gradient ranging from 100 to 200 mM NaCl. Germination rate was determined three days after imbibition. Additionally, seeds were treated at 50°C for 60 or 120 minutes, and germination rate on 1/2 MS - suc medium was determined after 2 weeks. The ratio of germinated to not-germinated plants were normalized to 1, and the total number of assessed plants is indicated above the respective columns.

heat stress duration or salt concentration (Figure 21). Therefore, LIME proteins probably do not act in abiotic salt and heat stress response. As LIME2 is highly abundant in 1-month-old plants (section 4), it may have an effect in the root growth of adult plants. Furthermore, LIME1 is not present in dry seeds (Kretzschmar et al., 2020), thus heat stress may be applied in different stages of development where it is actually present. However, it demonstrated that it is likely not affected by heat-induced changes to the seed.

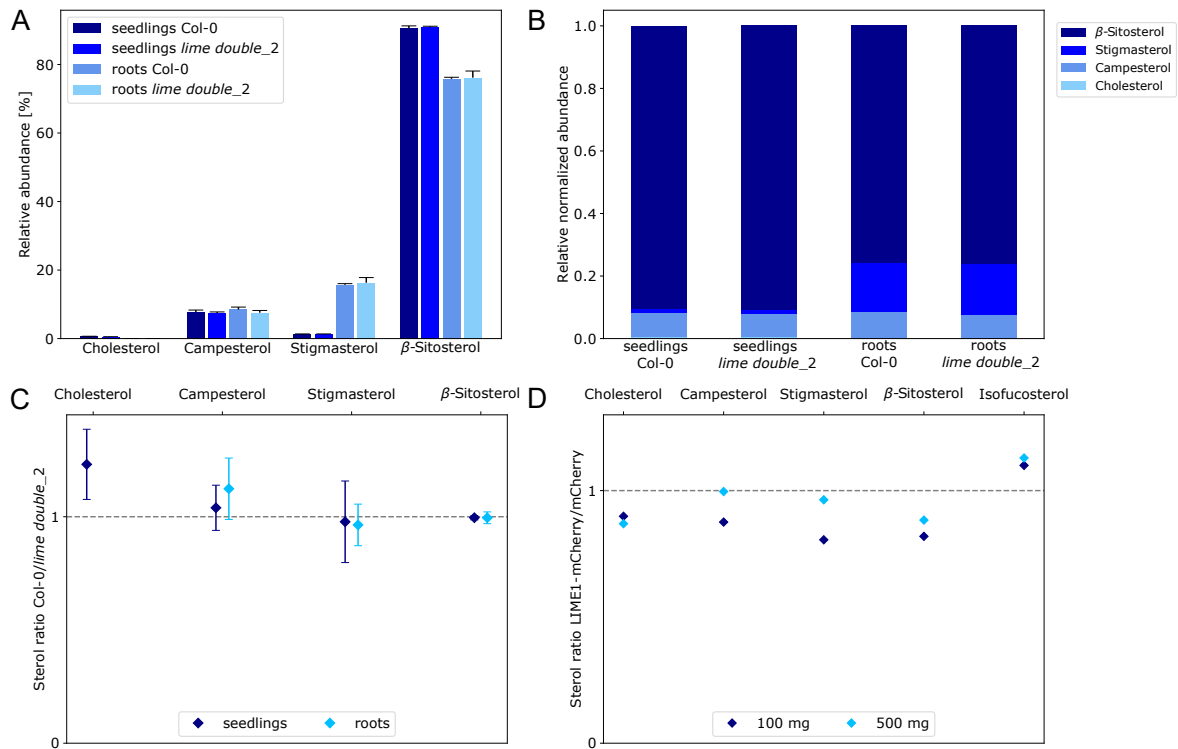
## 5.9 LIME proteins do not affect sterol composition

In the known plant LD proteome, only two methyltransferases have been reported. The STEROL METHYLTRANSFERASE 1 (SMT1) is involved in the sterol metabolism by methylating cycloartenol to 24-methylene-cycloartenol, which marks the entry of the phytosterol pathway (Carland et al., 2010). Therefore, LIME may also have a role in sterol metabolism, considering that sterols are known to be available substrates at the LD surface, as demonstrated by SMT1, which shares 25.6% sequence identity with LIME1. In order to test this possibility, free sterols were extracted from 2-day-old seedlings and axenic root cultures of *lime double\_2* Arabidopsis mutant lines and Col-0 wt from three biological replicates each, and measured with GC-MS.

For the analysis, the relative content of all identified sterol species within one sample was determined, averaged, and compared between wt and *lime double\_2* (Figure 22A-C). However,



## 5. Unpublished results II: Characterization of lipid droplet-associated methyltransferases



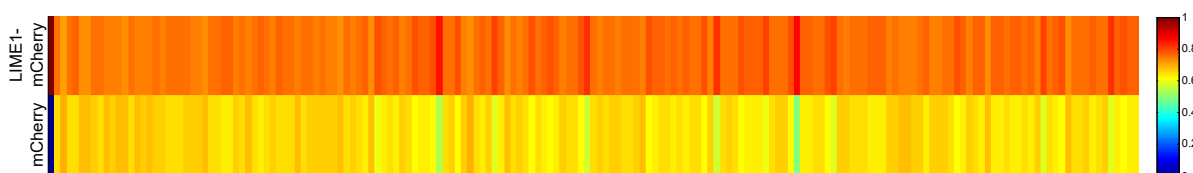
**Figure 22: Sterol distribution is not altered by LIME protein.** Free sterols were extracted from three replicates of 2-day-old seedlings or axenic root cultures of Arabidopsis Col-0 wt and *lime double\_2* and measured via GC-MS. Internal sterol-species distribution was determined and displayed A) separately for each sterol species or B) the distribution pattern within a sample type. C) The ratio between wt and mutant was determined to emphasize changes caused by the mutation. D) free sterols were extracted from 100 mg or 500 mg *N. benthamiana* leaves transiently expressing free mCherry or LIME1-mCherry. The ratio between the different sterol species was determined for comparative analysis.

no significant difference was observed (Figure 22C), suggesting that LIME proteins do not affect sterol metabolism. As an additional experiment, sterols were extracted from *N. benthamiana* 100 mg or 500 mg leaves transiently expressing either LIME1-mCherry or free mCherry. It appears that isofucosterol production is slightly increased in LIME1-expressing *N. benthamiana* leaves compared to mock-treated ones, while all other sterol species are decreased or unchanged, indicating involvement in isofucosterol production (Figure 22D). However, this result should be interpreted cautiously due to the inability of statistical validation, as only one replicate was measured per condition. Additionally, isofucosterol has not been identified in either Arabidopsis roots or seedlings, suggesting that LIME proteins may not play a role in its production in these tissues. It is a precursor of stigmasterol and  $\beta$ -sitosterol, and the levels of these compounds do not change significantly in comparison to the competing campesterol sterol pathway, further supporting that the result is likely not of physiological relevance

### 5.10 Non-targeted *ex vivo* approach does not reveal putative LIME1 substrates

SAM-dependent methyltransferases are highly versatile proteins, and predicting their preferred substrate can be challenging. For an unbiased, hypothesis-free determination of potential LIME substrates, a non-targeted *ex vivo* metabolomic approach was performed (Feussner and Feussner, 2020; Ni and Feussner, 2023). For this purpose, a quasi-native metabolite extract from *Arabidopsis lime double\_2* 48 h seedlings was processed. This stage was chosen due to its relatively high LIME protein abundance in wt plants (Kretzschmar et al., 2020), suggesting accumulation of the substrate of interest in the mutant line, in the case a redundant pathway does not utilize it. The whole metabolome extracts were spiked with excess SAM and incubated with LIME1-mCherry immobilized via CoIP on selector beads. The same process was performed as inactive control with free mCherry immobilized on selector beads. Every incubation was conducted in three technical replicates and measured with LC-MS. This method enables the active enzyme to react with all substrates it can technically convert without being affected by downstream reactions or alternative circumventing pathways.

A total of 177 metabolomic markers were identified as significantly different by comparing LIME1-mCherry against free mCherry treated samples, with an ANOVA *p*-value cut-off of 0.0098131 for 58 markers in the negative ionization mode and 0.0099733 for 119 markers in the positive ionization mode (Figure 23). However, none of the significantly different markers exhibited a methyl group mass shift between putatively active LIME1-mCherry and inactive free mCherry control. Possible causes for the lack of methylation events in the non-targeted *ex vivo* approach could include: the inactivity of LIME1-mCherry immobilized on beads, which is hard to detect without known activity; the substrate being too hydrophobic for detection by LC-MS, which is designed mainly for hydrophilic substrates in the used configuration; or the possibility that the substrate was converted via alternative metabolic pathways and was not present or accumulated in the samples. Furthermore, it is also possible that only the substrate or the product was identified. This can be attributed to the possible abundant availability of



**Figure 23: Differential metabolite species between LIME1-mCherry and mCherry treated samples.** Quasi-native metabolite extracts from two-day-old *lime double\_2* seedlings were subjected to incubation with immobilized LIME1-mCherry (active) or free mCherry (inactive) on RFP-selector beads. A total of 177 metabolomic markers that exhibited significant differences between the active and inactive conditions were determined after LC-MS measurement. The averaged results of three technical replicates per condition were analyzed.

substrate, with the large amounts that are present not being significantly consumed when the active enzyme interacts with it compared to the control. In this case, it would be excluded by the ANOVA cut-off. Conversely, the product may not be identified because it ionizes poorly and therefore falls below the detection limit. Thus, further investigations are needed to identify potential substrates for LIME1. The increased number of significantly different markers observed could potentially be attributed to contamination. Previous CoIP experiments have suggested a protein purity of 70 - 80 % on the selector beads (refer to section 5.3), and a variety of low abundant contaminants may be responsible for various changes in metabolomic markers. Due to its predominant localization in the cytosol, free mCherry is considered a sub-optimal control for the membrane-bound LIME1-mCherry, as the presence of distinct contaminants in each sample can vary (as illustrated in Figure 16). Despite being the best control option available at the time, the use of free mCherry may introduce limitations and confounding factors in the experimental analysis.

## 6 Experimental procedures of unpublished results

### 6.1 Generation of *Arabidopsis lime* mutant lines

The generation of CRISPR/Cas9 *lime double* mutant lines was performed as described in Krawczyk et al. (2022b). In short, the Cas-Designer (Park et al., 2015) and CasOFFinder (Bae et al., 2014) tools were employed to design sgRNAs spanning over 19 bp in the proximity of a *SpCas9* protospacer adjacent motif (PAM), utilizing the *Arabidopsis thaliana* (TAIR10) (Lamesch et al., 2012) genome as a reference.

The high degree of sequence similarity between LIME1 (AT4G33110) and LIME2 (AT4G33120) was taken advantage of by designing sgRNAs targeting both genes simultaneously, with a maximum of one mismatch allowed in either gene. Furthermore, two sgRNAs per CRISPR/Cas9 cassette were designed. The CRISPR/Cas9 cassette was generated as previously described (Krawczyk et al., 2022b; Xing et al., 2014; Wang et al., 2015), resulting in two sgRNAs embedded between U6 promoter and U6 terminator sequences each, accompanied by a Cas9 under egg-cell specific EC1.2 promoter and a hygromycin resistance gene within a pHEE401E vector. Two independent CRISPR/Cas9 cassettes with distinct sgRNAs were made. Used primer to generate sgRNAs amplified from *Arabidopsis Col-0* gDNA can be found in section Materials, table 9. CRISPR/Cas9 cassettes were introduced into *Arabidopsis thaliana Col-0* by floral dip transformation mediated by *Agrobacterium tumefaciens* strain GV3101, following the protocol described by Bent (2006). Positives clones were screened by hygromycin selection (on solid 1/2 MS + 1% + 25  $\mu\text{g mL}^{-1}$  hygromycin). For identification of CRISPR deletion, PCR was performed on gDNA (isolated according to Edwards et al. (1991)) with either REDTaq® (Merck, Darmstadt, Germany) or Phusion™ High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA). The CRISPR/Cas9 cassette removal was confirmed through PCR using U6 primers (Wang et al., 2015) and negative hygromycin selection. Homozygous CRISPR lines without CRISPR/Cas9 cassette were obtained in T4 (*lime double\_2*) and T5 (*lime double\_3*) generation. For (*lime double\_4*) and (*lime double\_5*), T3 generation plants were backcrossed with Col-0 wt, homozygous CRISPR lines without CRISPR/Cas9 cassette were obtained in the F3 generation. The CRISPR/Cas9 cassette 1, containing sgRNA6 and sgRNA7, was used for *lime double\_2* and *lime double\_3*, while cassette 2 with sgRNA4 and sgRNA5 was used for *lime double\_4* and *lime double\_5*.

In addition to CRISPR/Cas9 generated *lime double* mutants, SALK\_024876 (*lime1-1*), GT\_3\_9311 (*lime1-2*), WiscDsLox485-488C12 (*lime1-3*), SM\_3\_34300 (*lime2-1*), and SALK\_127412 (*lime2-2*) T-DNA lines were purchased from the Arabidopsis Stock Center.

## 6.2 Analysis of potential phenotypes

To analyze potential water loss effects of *lime* mutants, Arabidopsis Col-0, *lime double\_2*, *lime double\_3*, and *lime double\_5* seeds were imbibed for 72 h and grown for 24 days on soil. Five biological replicates were processed per line, represented by a whole leaf rosette of a single plant per replicate. Rosettes were cut with a razor blade, positioned with stomata facing upwards, and weighed after 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 120, and 180 min. The relative weight loss and standard deviations were calculated in Microsoft Excel 2019 and plotted with matplotlib package in Python 3.10.6.

For the identification of potential *lime* mutant phenotypes under moderate dehydration or intense UV stress, Arabidopsis Col-0, *lime1-1*, *lime2-1*, *lime2-2*, *lime double\_2*, and *lime double\_5* seeds were imbibed for 72 h and grown for 26 days on soil. The plants were divided into three groups: a control group, a dehydration stress group, and a UV stress group, with five plants per line in each group. For the dehydration experiments, plants were removed from excess water and not watered for seven days; starting on the seventh day, they were watered again to facilitate recovery. In the case of UV stress experiments, the plants were subjected to intense UV-C light for 30 minutes, typically used for sterile bench sterilization. Plants were documented on day 0, day 7, and day 10 following the start of stress experiments.

In order to identify potential salt stress phenotypes, seeds of Arabidopsis Col-0, *lime double\_3*, and *lime double\_5* were surface sterilized, placed in solid 1/2 MS + 0.5% suc + 100, 150, or 200 mM NaCl, and imbibed for 48 h. The germination rate was determined on the third day after imbibition, defining clear visibility of cotyledon as germinated. Data was analyzed in Microsoft Excel 2019 and plotted by matplotlib package in Python 3.10.6.

To identify potential roles of *lime* mutants in seed heat stress, Arabidopsis Col-0, *lime double\_2*, *lime double\_3*, and *lime double\_5* lines were incubated for 60 and 120 min at 50 °C. Subsequently, seeds were surfaced sterilized and placed on solid 1/2 MS -suc medium. After two weeks of recovery, the germination rate was assessed based on cotyledon visibility. The data was analyzed using Microsoft Excel 2019 and visualized with the matplotlib package in Python 3.10.6.

For the determination of root lengths of *Arabidopsis* wt and *lime* mutant lines, 20 seeds of Col-0, *lime1-1*, *lime2-1*, *lime2-2*, *lime double\_2*, and *lime double\_5* were surfaced sterilized, placed on solid 1/2 MS -suc, and imbibed for 72 h. Plants were grown vertically, and the root length was documented after 3, 5, and 8 days. Only roots that had successfully germinated and were not arrested in their development were considered. The root length was determined with Fiji 1.53t. The data was visualized using the matplotlib package in Python 3.10.6, and statistical analysis was performed using R 4.2.2. ANOVA was conducted, followed by post-hoc Tukey test, and the results were presented in compact letter output format.

All plants were grown in long-day conditions with 16 h light/8 h dark cycles (150  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  light intensity) at 22 °C.

### 6.3 Plasmid construction

A wide range of plasmids were constructed aiming to express LIME proteins and LD candidates identified in root LD-enriched fraction transiently in tobacco pollen tubes. Furthermore, plasmids were generated to stably transform *Arabidopsis* to carry LIME proteins fused with eGFP under their native promoters.

For transient tobacco pollen tube expression for localization experiments, sequences of genes of interest or LIME1 truncations were cloned into the vectors pLAT52-mVC, pLAT52-mVN, and pLAT52-mCC by classical or fast Gateway® (Thermo Fisher Scientific, Waltham, MA, USA) cloning as previously reported by Müller et al. (2017). Those vectors harbor a Gateway® cassette, a C- or N-terminal fused fluorescence protein, and a strong LAT52 pollen-specific promoter (Twell et al., 1991). All primers used for the gene of interest amplification can be found in Materials, table 6.15.4. The primer design based on published sequences with the following sequence identifier: *AtLIME1* (TAIR10 AT4G33110.1), *AtLIME2* (TAIR10 AT4G33120.1), *GbLIME* (TreeGenes Gb\_29392), *MpLIME* (Phytozome Mapoly0103s0018.1), *PpLIME* (Phytozome Pp3c17\_22760V3.1), *PsCNMT* (GenBank XP\_026398838.1), *PsCNMT-like* (GenBank XP\_026447122.1). Construct sequences were confirmed by Sanger sequencing.

For LIME1 truncations, a pMDC32-ChC/LIME1 vector (Kretzschmar et al., 2020) was used as a template, while LIME2 and root LD candidates were amplified from *Arabidopsis* Col-0 root cDNA. Furthermore, cDNA templates of *Physcomitrium patens* ecotype Gransden gametophore and *Marchantia polymorpha* female gametophyte were kindly donated by Dr.

Tegan Haslam and Dr. Alisa Keyl, respectively. For cDNA generation of poppy and ginkgo, seeds of *Papaver somniferum* (purchased from Heinrich Brüning GmbH, Hamburg, Germany) soaked for 1 h in water, and young leaves of *Ginkgo biloba* were used. The remaining soaked poppy material was immediately destroyed by autoclavation to comply with German drug regulations. The cDNAs were obtained by extracting mRNA from plant material using the Monarch Total RNA Miniprep Kit (NEB, Ipswich, MA, USA). Subsequently, cDNA synthesis was performed using Maxima Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA, USA) and oligo(dT) primer, following the manufacturer's protocols.

To create LIME truncation lines that consist of two truncated sequences flanking a mVenus fluorescence protein, vectors containing a N-terminal LIME1 truncation in pLAT52-mVC and a C-terminal truncation in pLAT52-mVN were employed. Both vectors were simultaneously digested using the restriction enzymes *Pst*I (NEB, Ipswich, MA, USA), which cleaves within the mVenus sequence, and *Sac*I (NEB, Ipswich, MA, USA), which cleaves in the vector backbone. The large resulting fragment derived from pLAT52-mVC, which includes the vector backbone, LAT52 promoter, N-terminal LIME truncation, and a partial mVenus sequence, was ligated with T4 DNA ligase (Thermo Fisher Scientific, Waltham, MA, USA) with partial mVenus, C-terminal LIME truncation, and partial vector backbone derived from the pLAT52-mVN vector.

Moreover, plasmids were constructed to investigate the native localization of LIME proteins by fusing their cDNA with the putative native promoter sequence and the coding sequence of the fluorescence protein eGFP. For this purpose, the full-length cDNA of LIME1 was amplified via PCR from pLAT52-LIME1-mV, while pLAT52-LIME2-mV was used for LIME2, respectively. This way, NotI and XhoI restriction sites were added to both ends of the amplified cDNA fragments. Equivalently, sequences 1045 bp upstream of the start codon for LIME1 and 930 bp for LIME2 were amplified from Arabidopsis Col-0 gDNA, incorporating Sall and NotI restriction sites to their ends. The primers used for amplification are listed in Materials, table 8. Following PCR amplification, all fragments were digested with the corresponding restriction enzymes for their added restriction sites. Subsequently, the digested PCR fragments were ligated into a custom pUC18-entry vector (Dr. Ellen Hornung, unpublished) linearized with Sall and XhoI. The pUC18-entry vector contains a multiple cloning site within a Gateway® cassette. Thus, the fused cDNA and putative promoter sequences were cloned by classical Gateway® LR Clonase II™ reaction (Thermo Fisher Scientific, Waltham, MA, USA) into the destination vector pGWB604 (Nakamura et al., 2010) according to the manufacturer's instructions. Using an established floral dip protocol, the resulting vector was stably transformed into Arabidopsis

Col-0 plants (Bent, 2006). Confocal microscopy was employed to examine the T2 generation plants, specifically two-day-old seedlings for Nat<sub>LIME1g</sub>-LIME1c lines and five-day-old roots for Nat<sub>LIME2g</sub>-LIME2c lines. A Zeiss LSM 780 (Carl Zeiss, Oberkochen, Germany) with an excitation wavelength of 488 nm and a detection window of 488-533 nm was used for imaging.

#### 6.4 Transient expression in tobacco pollen tubes

Tobacco pollen were harvested from *Nicotiana tabacum* L. cv. Samsun-NN plants cultivated as described in Rotsch et al. (2017). Plasmid coating of gold particles and transformation with gene gun was done as previously reported (Müller et al., 2017). After 5 h of growth, pollen tubes were microscopically examined employing Zeiss LSM 510 or a Zeiss LSM 780 confocal microscope (Carl Zeiss, Oberkochen, Germany). Prior to microscopy, pollen tubes were fixed with 1.8% (v/v) formaldehyde in tobacco pollen tube medium (5% (w/v) sucrose, 12.5% (w/v) PEG-4000, 15 mM MES-KOH pH 5.9, 1 mM CaCl<sub>2</sub>, 1 mM KCl, 0.8 mM MgSO<sub>4</sub>, 0.01% H<sub>3</sub>BO<sub>3</sub> (v/v), 30 µM CuSO<sub>4</sub>) and either stained with 0.5% (w/v) Nile Red® (Merck, Darmstadt, Germany) or 1 µg mL<sup>-1</sup> BODIPY 493/503 (Thermo Fisher Scientific, Waltham, MA, USA). The exact microscope parameters used for each construct can be found in Materials, table 10. In addition, all root LD candidates were imaged with Zeiss LSM 780 (Carl Zeiss, Oberkochen, Germany), using 488 nm as excitation wavelength and detection window of 497-542 nm for BODIPY 494/503, and 561 nm excitation wavelength and detection window of 590-640 nm for mCherry. Images were processed in Zeiss ZEN (blue edition) 3.2.

#### 6.5 *N. benthamiana* agrobacterium infiltration

In order to infiltrate *N. benthamiana* leaves for transient gene expression, overnight cultures of *Agrobacterium tumefaciens* strain GV3101 transformed with established binary vectors pMDC32-ChC/LIME1 (Kretschmar et al., 2020) and pMDC32-mCherry (Price et al., 2020), as well as pORE04-35S::P19 (Petrie et al., 2010) were grown at 28 °C and 200 rpm agitation in LB<sub>Rif,Gen,Kan</sub> medium. Bacterial suspensions were pelleted by centrifugation for 5 min at 5000 x g, resuspended in infiltration buffer (10 mM MES-KOH pH 5.6, 10 mM MgCl<sub>2</sub>, 150 µg mL<sup>-1</sup> acetosyringone), and incubated for 2 h at room temperature. Subsequently, 4-week-old *N. benthamiana* plants were infiltrated with bacterial suspension containing either pMDC32-ChC/LIME1 or pMDC32-mCherry, each diluted in infiltration buffer to OD<sub>600</sub> 0.3 per construct. To enhance gene expression and reduce gene silencing, all leaves were co-



infiltrated with pORE04-35S::P19, which encodes the tomato bushy stunt virus gene P19 (Petrie et al., 2010). The duration of transgene expression was varied based on the specific use case, with expression times ranging from two to five days.

## 6.6 CoIP for identification of potential LIME interactors

*N. benthamiana* leaves co-infiltrated either with pMDC32-ChC/LIME1 and pORE04-35S::P19 or pMDC32-mCherry and pORE04-35S::P19 were harvested 2 days after infiltration and flash frozen in liquid nitrogen. For each co-infiltration vector combination, three biological replicates were processed. The leaf material was homogenized in a nitrogen-cooled mortar, and all following steps were performed on ice. For each replicate, 2 g of homogenized plant material was transferred to 20 mL IP buffer (Na phosphate 100 mM pH 7.5, 150 mM NaCl, 1 % (v/v) Triton X-100, 400  $\mu$ M proteinase inhibitor PMSF (Carl Roth, Karlsruhe, Germany), 0.5 mM Lomant's Reagent DSP (Thermo Fisher Scientific, Waltham, MA, USA)), and incubated for 30 minutes to help stabilize protein-protein interactions by the DSP crosslinker. Subsequently, the suspension was sonicated for 10 min on ice and centrifuged for 15 min at 4000 x g and 4 °C. The supernatant was filtered through a fluting filter paper to remove floating plant remainings, resulting in the crude plant protein extract. RFP selector beads (Nanotag, Göttingen, Germany) were equilibrated by washing them three times with IP buffer and sedimentation by centrifugation for 1 min at 1000 x g. The equivalent of 40  $\mu$ L bead slurry was applied to the crude plant protein extract for each replicate and incubated for 2 h at 4 °C under constant inversion. Beads were sedimented by centrifugation for 2 min at 3000 x g, and the supernatant was discarded. Subsequently, beads were washed three times with 25 volumes of washing buffer (Na phosphate 50 mM pH 7.5, 10 mM NaCl, 200  $\mu$ M proteinase inhibitor PMSF (Carl Roth, Karlsruhe, Germany)), and once with 25 volumes tris-buffered saline (50 mM Tris-HCl pH 7.5, 150 mM NaCl). Liquids were completely removed by centrifugation for 1 min at 3000 x g. Immunoprecipitated proteins were eluted by application of 50  $\mu$ L 2x Laemmli buffer (125 mM Tris-HCl pH 6.8, 4 % (w/v) SDS, 200 mM DTT, 0.4 % (w/v) bromophenol blue, 20 % (v/v) glycerol) and heat treatment for 5 min at 95 °C.

## 6.7 CoIP to immobilize proteins for *ex vivo* assay

Three days after transformation of *N. benthamiana* leaves were infiltrated with either pMDC32-ChC/LIME1 and pORE04-35S::P19 or pMDC32-mCherry and pORE04-35S::P19, and rapidly

frozen using liquid nitrogen. The plant material was finely ground in a nitrogen-cooled mortar and 8 g of resulting material was transferred into IP buffer (100mM Tris-HCl pH 7.5, 10 mM MgCl<sub>2</sub>, 150 mM NaCl, 10 mM DTT, 1 mM proteinase inhibitor PMSF (Carl Roth, Karlsruhe, Germany)) to 50 mL total volume. All following processes were performed at 0-4 °C. Samples were treated 10 min in a cooled ultrasonic bath and centrifuged for 15 min at 4000 x g and 4 °C to sediment remaining cell debris. The supernatant was filtered through a fluting filter paper and centrifuged for 35 min at 50,000 x g and 4 °C in a fixed-angle rotor to remove remaining impurities. For each co-infiltration configuration, 80 µL RFP-selector bead slurry (Nanotag, Göttingen, Germany) was equilibrated twice in 12 volumes of IP buffer. Equilibrated beads were transferred to cleared supernatant and incubated for 5 h at 4 °C under constant inversion. Beads were sedimented through centrifugation for 1 min at 1000 x g, and washed four times with 20 volumes of reaction buffer (10 mM Tris-HCl pH 7.5, 15 mM MgCl<sub>2</sub>, 5 mM NaCl, 1 mM DTT) and bead sedimentation by centrifugation for 1 min at 1000 x g. The buffer was completely removed by centrifugation for 1 min at 3000 x g, and the beads were resuspended in an equal volume of reaction buffer prior to *ex vivo* application.

## 6.8 Axenic root cultivation

A high-yield axenic *Arabidopsis* root culture was cultivated by adapting an established protocol published by Héту et al. (2005). In short, seeds of the oil-rich mutant *tgdl-1 sdp1-4* and Col-0 for LD acquisition or *lime double\_2* for free sterol acquisition were surface sterilized, placed on sterile steel grids on top of solid 1/2 MS + 1 % suc medium, stratified for 72 h, and subsequently grown for 7 days. One-week-old seedlings were transferred to 100 mL Erlenmeyer flasks on the steel grid and supplemented with 10 mL liquid 1/2 MS + 1 % suc medium. The culture was agitated at 85 rpm for 11 days with regular exchange of the medium every three days. Afterwards, the medium was changed to 15 mL 1/2 MS + 3 % suc medium and Col-0 and *tgdl-1 sdp1-4* seedlings were grown 11 additional days before harvest, exchanging the medium every third day. In contrast, *lime double\_2* lines were grown for 8 instead of 11 days in 15 mL 1/2 MS + 3 % suc medium, before the medium was changed again to 15 mL 1/2 MS - suc for three continuous days at 85 rpm agitation. Plants were grown in long-day conditions with 16 h light/8 h dark cycles (150 µmol photons m<sup>-2</sup> s<sup>-1</sup> light intensity) at 22 °C.

### 6.9 Isolation of root LD-enriched and total protein fractions

Harvested root material was separated from other plant tissues and drained from the remaining medium by drying between paper towels and applying gentle pressure. The resulting root pads of two separate cultures were pooled into one biological replicate. Root samples of *tgdl-1 sdp1-4* and Col-0 were examined using a Zeiss LSM780 confocal microscope (Carl Zeiss, Oberkochen, Germany), after LD being stained with  $1 \mu\text{g mL}^{-1}$  BODIPY 493/503 (Thermo Fisher Scientific, Waltham, MA, USA) in 50 mM PIPES buffer to visualize lipid droplets (488 nm excitation wavelength, 488 - 533 nm detection wavelengths). Images were processed in Zeiss ZEN (blue edition) 3.2.

For the protein extraction, all processes and materials were kept on ice. Grinding buffer (50 mM Tris-HCl pH 7.5, 10 mM KCl, 0.4 M sucrose, 200  $\mu\text{M}$  proteinase inhibitor PMSF; Carl Roth, Karlsruhe, Germany) and sea sand were added to the root pads which were subsequently ground. To remove cellular debris and sea sand, the homogeneous suspension was centrifuged for 1 min at 100 x g. An aliquot was taken from the supernatant and precipitated in 96 % ethanol at  $-20^\circ\text{C}$ , representing the total protein fraction. The remaining supernatant was overlaid with washing buffer (50 mM Tris-HCl pH 7.5, 10 mM KCl, 0.2 M sucrose, 200  $\mu\text{M}$  proteinase inhibitor PMSF; Carl Roth, Karlsruhe, Germany) and centrifuged for 35 min at 100,000 x g and  $4^\circ\text{C}$  in a swing-out rotor. The floating fat pad was mechanically collected, emulsified in a small volume of washing buffer, and centrifuged in a fixed angle rotor for 35 min at 100,000 x g. The floating fat pad was collected again, dissolved in 96 % ethanol to remove fat, and stored at  $-20^\circ\text{C}$  to precipitate proteins. This is considered as LD-enriched protein fraction.

### 6.10 Protein processing and LC-MS/MS measurements

Precipitated root protein pellets were subjected to defatting by undergoing two washes with 80 % ethanol, followed by drying and an additional wash with 96 % ethanol. The obtained proteins were dissolved in 6 M urea, 5 % SDS (w/v) solution, and their concentrations were determined using the Pierce BCA protein assay kit (Thermo Fisher Scientific, Waltham, MA, USA). For root-derived proteins, 20  $\mu\text{g}$  were subjected to in gel-digest, while for the CoIP the whole protein elution was used to maximize the outcome. The in-gel digestion and desalination was executed as previously described (Shevchenko et al., 2006; Rappsilber et al., 2007).

An EASY-nLC 1200 system (Thermo Fisher Scientific, Waltham, MA, USA) in conjunction with an Exploris 480 mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) was employed for the execution of the LC-MS/MS analysis of root-derived peptides. For this purpose, peptides were separated on 20 cm frit-less silica emitters (CoAnn Technologies, Richland, WA, USA) with a 0.75  $\mu\text{m}$  inner diameter, packed in-house with ReproSil-Pur C18 AQ 1.9  $\mu\text{m}$  resin (Dr. Maisch, Ammerbuch-Entringen, Germany). The column was maintained at a constant temperature of 50 °C. Elution of peptides was carried out over 115 min using a segmented linear gradient from 0 % to 98 % solvent B (solvent A: 0 % ACN, 0.1 % formic acid; solvent B: 80 % ACN, 0.1 % formic acid) at a flow rate of 300 nL min<sup>-1</sup>. The data-dependent acquisition mode was utilized to acquire mass spectra. For full proteome samples, MS1 scans were obtained at an Orbitrap resolution of 120,000, covering a scan range of 380-1500 (m/z). The maximum injection time was set to 100 ms, and a normalized AGC target of 300 % was utilized. Precursors with charge states 2-6 were selectively chosen for fragmentation, and up to 20 dependent scans were acquired. Dynamic exclusion was enabled with an exclusion duration of 40 seconds and a mass tolerance of +/- 10 ppm. A 1.6 (m/z) isolation window with no offset was established, accompanied by the application of a normalized collision energy of 30. Acquisition of MS2 scans was performed at an Orbitrap resolution of 15,000, while maintaining a fixed First Mass (m/z) of 120. The maximum injection time was 22 ms, and the normalized AGC target was set to 50 %.

Peptides derived from ColP were measured by LC-MS/MS as described in Article I (Niemeyer et al., 2022). Obtained RAW files were processed in MaxQuant v1.6.2.17 (root proteomic data, using *Arabidopsis thaliana* TAIR10 protein release (Lamesch et al., 2012) as reference) or 2.0.3.1 (ColP, using *Nicotiana benthamiana* NbDE transcriptome derived protein dataset (Kourelis et al., 2019) as reference) (Cox and Mann, 2008) and analyzed in Perseus 1.6.2.2 (Tyanova et al., 2016) as previously described in Horn et al. (2021).

### 6.11 Phylogenetic analysis of LIME proteins

For phylogenetic analysis, the protein sequence of LIME1 (TAIR identifier AT4G33110) was used as a query for a BLASTp (2.12.0+) approach against a custom selected protein database with an *e*-value cut-off < 10<sup>-15</sup>. Sequences for the database were mined from Phytozome (Goodstein et al., 2012) for *Aquilegia coerulea* v3.1 (Filiault et al., 2018), *Arabidopsis halleri* v2.1.0 (Goodstein et al., 2012), *Arabidopsis lyrata* v2.1 (Hu et al., 2011; Rawat et al., 2015), *Brassica oleracea capitata* v1.0 (Liu et al., 2014), *Brassica rapa* FPsc v1.3 (Goodstein et al.,

2012), *Citrus sinensis* v1.1 (Wu et al., 2014), *Chlamydomonas reinhardtii* v5.6 (Merchant et al., 2007), *Cucumis sativus* v1.0 (Yang et al., 2012a), *Daucus carota* v2.0 (Iorizzo et al., 2016), *Glycine soja* v1.1 (Valliyodan et al., 2019), *Gossypium hirsutum* v1.1 (Chen et al., 2020b), *Hordeum vulgare* r1 (Beier et al., 2017), *Manihot esculenta* v6.1 (Bredeson et al., 2016), *Marchantia polymorpha* v3.1 (Bowman et al., 2017), *Oropetium thomaeum* v1.0 (VanBuren et al., 2015), *Oryza sativa* v7.0 (Ouyang et al., 2007), *Physcomitrium patens* v3.3 (Lang et al., 2018), *Populus trichocarpa* v4.1 (Ye and Zhong, 2015; Du et al., 2015), *Prunus persica* v2.1 (Verde et al., 2013), *Selaginella moellendorffii* v1.0 (Banks et al., 2011), *Solanum tuberosum* v4.03 (Xu et al., 2011), *Theobroma cacao* v2.1 (Motamayor et al., 2013); from Phycocosm (Grigoriev et al., 2021) for *Klebsormidium nitens* NIES-2285 (Hori et al., 2014), *Spirogloea muscicola* CCAC 0214 (Cheng et al., 2019); from EnsemblePlants (Cunningham et al., 2022) for *Brassica napus* (Chalhoub et al., 2014), *Camelina sativa* (Kagale et al., 2014), *Chara braunii* (Nishiyama et al., 2018), *Corchorus capsularis* (Islam et al., 2017; Parra et al., 2007), *Medicago truncatula* (Young et al., 2011; Tang et al., 2014), *Papaver somniferum* (Guo et al., 2018); from TreeGenesdb (Wegrzyn et al., 2008, 2019; Falk et al., 2018) for *Fraxinus excelsior*, *Ginkgo biloba*, *Hevea brasiliensis*, *Picea abies*; from NCBI non redundant protein database (National Center for Biotechnology Information, 2023) for *Coptis japonica* and *Litchi chinensis*; and from diverse other sources for *Anthocerotophyta agrestis* (Li et al., 2020), *Anthocerotophyta punctatus* (Li et al., 2020), *Arabidopsis thaliana* TAIR10 (Lamesch et al., 2012), *Aristolochia fimbriata* (Qin et al., 2021), *Azolla filiculoides* (Li et al., 2018), *Cyperus esculentus* (Niemeyer et al., 2022), *Eschscholzia californica* (Hori et al., 2018), *Mesotaenium endlicherianum* (Cheng et al., 2019; Dadras et al., 2022), *Nicotiana benthamiana* NbDE (Kourelis et al., 2019), *Nicotiana tabacum* (Edwards et al., 2017), and *Sequoiadendron giganteum* (Scott et al., 2020). A multiple sequence alignments of full-length sequences of selected proteins was generated with MAFFT v7.511 (Kato and Standley, 2013) using alignment mode L-INS-i. Subsequently, the maximum-likelihood phylogenetic tree was generated with W-IQ-TREE (Trifinopoulos et al., 2016), choosing the JTT+F+I+G4 model as best fitting determined by ModelFinder (Kalyaanamoorthy et al., 2017) according to bayesian information criterion. For branch support, 1000 ultrafast bootstrap (Hoang et al., 2018) replicates were calculated. The phylogenetic tree was visualized with iTOL (Letunic and Bork, 2007) as an unrooted tree. Most distinct proteins were collapsed to improve clarity. For Supplementary Figure 1, the *e* value cut-off was set to  $< 10^{-100}$  for BLASTp, alignment was performed with MAFFT L-INS-i, and the best fitting model according to bayesian information criterion in ModelFinder (Kalyaanamoorthy et al., 2017) was determined as JTT+G4, before maximum likelihood tree was computed

in W-IQ-TREE (Trifinopoulos et al., 2016). For branch support, 100 bootstrap (Felsenstein, 1985) replicates were generated. The resulting tree was pruned except species belonging to Brassicaceae.

### 6.12 Sterol extraction and GC-MS

Free sterols were extracted from 100 mg fresh weight of two-day-old seedlings or axenic root cultures (refer to axenic root culture) of Arabidopsis lines Col-0 and *lime double\_2* (three biological replicates each), and from 100 or 500 mg *N. benthamiana* leaf material transiently expressing LIME1-mCherry or free mCherry for five days, respectively.

For the extraction, the material was subjected to 500  $\mu$ L methanol, filled up with 1.5 mL MTBE and supplemented with 10  $\mu$ g heptadecanoic acid as internal standard. Subsequently, samples were agitated for 1 h at 80 rpm and 4 °C, and the phase separation was enhanced by centrifugation for 10 min at 510 x g. The nonpolar phase was recovered, and the solvent was removed using a constant stream of nitrogen. The remaining samples were redissolved in 15  $\mu$ L pyridine and additional 30  $\mu$ L MSTFA 1 h prior to GC-MS measurement. GC-MS settings for sterol detection and data analysis in MSD ChemStation was performed as previously described (Düking et al., 2022; Rotsch et al., 2017). The matplotlib package in Python 3.10.6 was used to visualize the data.

### 6.13 *Ex vivo* assay

In order to perform an *ex vivo* assay (Feussner and Feussner, 2020), a quasi-native metabolite extract was prepared from two-day-old seedlings of *lime double\_2* Arabidopsis mutant line according to a two-phase extraction method adapted from Matyash et al. (2008). In short, six times 100 mg of nitrogen cooled pulverized plant material was transferred into 0.75 mL methanol, vortexed, and subsequently 2.5 mL MTBE were added. Samples were incubated for 1 h in the dark, shaking horizontally at 80 rpm. Afterwards, 0.6 mL of water were added and mixed thoroughly. Phase separation was enhanced by centrifugation for 15 min at 2000 rpm. The upper nonpolar phase was removed and stored, and the lower polar phase was re-extracted with additional 0.7 mL methanol and 1.3 mL MTBE. The samples were mixed, centrifuged at 2000 rpm for 15 minutes, and the polar and nonpolar phases were combined with the stored previous nonpolar phase. Solvents were removed by evaporation under a

nitrogen gas stream. The remaining metabolites were dissolved in 400  $\mu\text{L}$  methanol, and all six prepared extracts were pooled to get a homogenous metabolite composition.

To perform the *ex vivo* assay, 200  $\mu\text{L}$  of methanol extract per replicate were evaporated under a constant stream of nitrogen gas and resuspended in 78  $\mu\text{L}$  reaction buffer (10 mM Tris-HCl pH 7.5, 15 mM  $\text{MgCl}_2$ , 5 mM NaCl, 1 mM DTT) supplemented with fresh 2  $\mu\text{L}$  of 50  $\mu\text{M}$  SAM. For each sample, 20  $\mu\text{L}$  of RFP-selector slurry (Nanotag, Göttingen, Germany) covered with LIME1-mCherry as active protein or mCherry as inactive protein (refer to section 6.7) was added, followed by overnight incubation at room temperature under constant inversion. A total of three technical replicates were processed for each protein that was added. The reaction was stopped by the addition of 50  $\mu\text{L}$  acetonitrile (LC-MS grade). Turbidity was removed by centrifugation twice for 20 min at 20,000  $\times$  g. LC-MS measurement and data processing in Marvis suite 2.0 was performed as previously described in Ni and Feussner (2023).

### 6.14 Additional bioinformatics

Protein model predictions were computed with AlphaFold Colab (Jumper et al., 2021), RoseTTAFold (Baek et al., 2021), and SWISS-MODEL (Waterhouse et al., 2018). Visualization of predicted models was done with PyMOL (Open Source build) v2.5.0 (Schrödinger, LLC, 2015).

Visualized sequence alignments and LOGOs were generated with Geneious 8.1 (Geneious, Auckland, New Zealand), using ClustalW algorithm.

Transmembran prediction was performed with TMHMM 2.0 (Krogh et al., 2001), hydropathicity plot was calculated in ProtScale (Gasteiger et al., 2005) using the scale of Kyte and Doolittle (1982). In both cases, raw data was retrieved and replotted with the matplotlib package in Python 3.10.6.

Venn diagrams were made with InteractiVenn (Heberle et al., 2015).

## 6.15 Materials

### 6.15.1 Software

Statement for the use of assisting AI tools. To ensure the linguistic correctness of this work, the spelling and grammar control tool Grammarly was used for the entire text. The tools DeepL and ChatGPT 3.5 were used for individual sentences as additional linguistic control. Those generative language models were used solely as tools to assist with linguistic proofreading and were not involved in the creation of entire paragraphs. The tool ChatGPT 3.5 was also used for software debugging in Python and L<sup>A</sup>T<sub>E</sub>X.

#### Stationary software:

**Table 5:** Selection of software used for analyzing or generating data. Software run either on Linux Mint 21.1 Vera, Windows 10, or Mac OS 12 Monterey.

Software	Version	Reference
BLASTp	2.12.0+	Altschul et al. (1990)
MSD ChemStation	F.01.03.2357	Merck & Co., Rahway, NJ, USA
Fiji	1.53t	Schindelin et al. (2012)
Geneious	8.1	Geneious, Auckland, NZ
MAFFT	7.511	Katoh and Standley (2013)
Marvis Suite	2.0	Kaever et al. (2009)
MaxQuant	2.0.3.1 & 1.6.2.17	Cox and Mann (2008)
Microsoft Excel	2019	Microsoft, Redmond, WA, USA
Perseus	1.6.2.2	Tyanova et al. (2016)
PyMOL (Open Source build)	2.5.0	Schrödinger, LLC (2015)
Python	3.10.6	Van Rossum and Drake Jr (1995)
R	4.2.2	R Core Team (2022)
Zeiss ZEN (blue edition)	3.2	Carl Zeiss, Oberkochen, Germany

#### Webtools:

**Table 6:** Selection of webtools used for unpublished results I and II. Hyperlinks to webpages are available in the digital version.

Webtool	Reference
AlphaFold Colab	Jumper et al. (2021)
Cas-Designer	Park et al. (2015)
Cas-OFFinder	Bae et al. (2014)
InteractiVenn	Heberle et al. (2015)
iTOL	Letunic and Bork (2007)
ProtScale	Gasteiger et al. (2005)
RoseTTAFold	Baek et al. (2021)
SWISS-MODEL	Waterhouse et al. (2018)
TMHMM 2.0	Krogh et al. (2001)
TargetP 2.0	Armenteros et al. (2019)
W-IQ-TREE	Trifinopoulos et al. (2016)



### 6.15.2 Chemicals

Unspecified chemicals used for unpublished results I and II were of analytical grade and were obtained from SigmaAldrich GmbH (Munich, Germany), Duchefa Biochemie (Haarlem, Netherlands), Thermo-Fisher Scientific (Waltham, MA, USA), Merck KGaA (Darmstadt, Germany), and Carl Roth GmbH (Karlsruhe, Germany).

### 6.15.3 Media

#### **Antibiotics final concentrations in selection media:**

Carbenicillin (Carb):  $100 \mu\text{g mL}^{-1}$

Gentamycin (Gen):  $20 \mu\text{g mL}^{-1}$

Hygromycin (Hyg):  $25 \mu\text{g mL}^{-1}$

Kanamycin (Kan):  $50 \mu\text{g mL}^{-1}$

Rifampicin (Rif):  $50 \mu\text{g mL}^{-1}$

Spectinomycin (Spec):  $100 \mu\text{g mL}^{-1}$

#### **Lysogeny Broth (LB) medium:**

Adapted from Bertani (1951)

$5 \text{ g L}^{-1}$  yeast extract

$5 \text{ g L}^{-1}$  NaCl

$20 \text{ g L}^{-1}$  peptone

$15 \text{ g L}^{-1}$  agar (for solid only)

#### **1/2 Murashige & Skoog (1/2 MS) medium :**

Adapted from Murashige and Skoog (1962)

$2.2 \text{ g L}^{-1}$  Murashige & Skoog medium (Duchefa Biochemie, Haarlem, Netherlands)

pH 5.7 adjusted with KOH

-suc, no added sucrose

+0.5 % suc, 0.5 % (w/v) added sucrose

+1 % suc, 1 % (w/v) added sucrose

+3 % suc, 3 % (w/v) added sucrose

$8 \text{ g L}^{-1}$  agar (for solid only)

## 6.15.4 Lists of primers and microscope settings

**Table 7: Primer table for LIME truncations.** Given are all forward and reverse primers (5'-3') used to amplify (partial) protein coding sequences designated for transient expression in tobacco pollen tubes. Constructs were cloned into destination vectors pLAT52-mVC (C-terminal mVenus sequence), pLAT52-mVN (N-terminal mVenus sequence), or pLAT52-mCC (C-terminal mCherry sequence), that contain a strong LAT52 pollen tube specific promoter and was firstly introduced by Müller et al. (2017). Constructs were either cloned via Gateway® (GW) or fast Gateway® method (Müller et al., 2017).

Amplicon	Primer forward	Primer reverse	Method	Destination vector
LIME1 aa 1-24	CTTTGTACAAGAAAGCTGGGTCAACCGGACGGAGACGAC	CTTTGTACAAGAAAGCTGGGTCCAGATTCATCTCCAACA-ACGTC	fast GW	pLAT52-mVC
LIME1 aa 1-35	CTTTGTACAAGAAAGCTGGGTCAACCGGACGGAGACGAC	CTTTGTACAAGAAAGCTGGGTCAACGTGTGAGCCGCC	fast GW	pLAT52-mVC
LIME1 aa 1-45	CTTTGTACAAGAAAGCTGGGTCAACCGGACGGAGACGAC	CTTTGTACAAGAAAGCTGGGTCAACCGGACGGAGACGAC	fast GW	pLAT52-mVC
LIME1 aa 48-52	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGAGAAG-ATTATTGACGTAG	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTGGGTTTGT-AACCGGACGGAG	GW	pLAT52-mVC
LIME1 aa 26-45	CTTTGTACAAAAAAGCAGGCTCATGCCGATATTGTCATT-AGGCGG	CTTTGTACAAGAAAGCTGGGTCAACCGGACGGAGACGAC	fast GW	pLAT52-mVC
LIME1 aa 291-337	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGACAAGTGA-AGAGTGGCTAAAAAG	CTTTGTACAAGAAAGCTGGGTCTCAGAAAAGCTCAGCAA-CCGC	fast GW	pLAT52-mVN
LIME1 aa 291-355	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGAC-AAGTGAAGAGTGGCTAAAAAG	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTTCTTTCTTGAAGAGGAAGTGT	GW	pLAT52-mVC
LIME1 aa 291-355	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGACAAG-TGAAGAGTGGCTAAAAAG	CTTTGTACAAGAAAGCTGGGTCTCATTCTTTCTTGAAG-AGGAAGTGT	fast GW	pLAT52-mVN
LIME1 aa 300-355	CTTTGTACAAAAAAGCAGGCTCGATAAGGAGATAGTTGC-AGTAAAAAAAT	CTTTGTACAAGAAAGCTGGGTCTCATTCTTTCTTGAAGAGG-AAGTGT	fast GW	pLAT52-mVN
LIME1 aa 320-355	CTTTGTACAAAAAAGCAGGCTCATGG-TGAAGTGGATGGTTTACTGGA	CTTTGTACAAGAAAGCTGGGTCTCATTCTTTCTTGAAGAGGA-AGTGT	fast GW	pLAT52-mVN
LIME1 aa 49-290	CTTTGTACAAAAAAGCAGGCTCATGACGGCGGAGATG-CAACTCTC	CTTTGTACAAGAAAGCTGGGTCTTTTGCATAATGCGT-CCCGTTAC	fast GW	pLAT52-mVC
LIME1 aa 49-355	CTTTGTACAAAAAAGCAGGCTCATGACGGCGGAGATGC-AACTCTC	CTTTGTACAAGAAAGCTGGGTCTTTCTTTCTTGAAGAG-GAAGTGT	fast GW	pLAT52-mCC
LIME1 aa1-290	CTTTGTACAAGAAAGCTGGGTCAACCGGACGGAGACGAC	CTTTGTACAAGAAAGCTGGGTCTTTTGCATAATGCGT-CCCGTTAC	fast GW	pLAT52-mCC
LIME1 aa1-320	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGAGAAGA-TTATTGACGTAG	GGGGACCACCTTTGTACAAGAAAGCTGGGTCCACTGCT-TCTTCTTTCC	GW	pLAT52-mCC
LIME2	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGAGAAGA-TTATTGACGTAG	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTTCTTTCTT-GAAGAGGAAGTGTGCAAC	GW	pLAT52-mVC
GbLIME	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGAGGG-CATAATGCAGGT	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTTCC-TTTTAAAGAGAAAATGTGAAACC	GW	pLAT52-mCC
MpLIME	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGCGTC-GGTGTTGGC	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTTTCTTCT-GAATAAGTAGTGCAG	GW	pLAT52-mCC
PpLIME	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGCG-TCGTTGATAAAGGC	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTTTCTT-TAAAGAGATAGTGGCACA	GW	pLAT52-mCC

Amplicon	Primer forward	Primer reverse	Method	Destination vector
PpLIME aa 1-45	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGCGTCGT-TGATAAAGGC	CTTTGTACAAGAAAGCTGGGTCAATCTCGTAAGC-CTCCCTTC	fast GW	pLAT52-mVC
PpLIME aa 300-355	CTTTGTACAAAAAAGCAGGCTCGATGCCCAAGTGAAAGTTATAAGAC	CTTTGTACAAGAAAGCTGGGTTTACTTTTTCTTAAAGAGATAGTGGCAC	fast GW	pLAT52-mVN
PsCNMT	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGCAGCTAAAGGCAAAGGAAGA	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTTTTTCTTGAAGAGAAGATGGGTGAG	GW	pLAT52-mCC
PsCNMT-like	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGAGAC-ATTAATTCAGTACCATAACAATG	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTTCTTCT-TGAATAGAAAATGTGCAACC	GW	pLAT52-mCC
AT3G52730	CTTTGTACAAAAAAGCAGGCTCATGGAGTACGCTG-CTCGGAG	CTTTGTACAAGAAAGCTGGGTCTTCTCTACTGG-CCTTTGACC	fast GW	pLAT52-mCC
AT4G13160	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGACTACCAAGAAAGTTATAGATTGAC	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTGGGAGATGTGTTGAAGATGAAGT	GW	pLAT52-mCC
AT4G38540	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGAAGA-AGAAGGCAGCCC	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTGGGACAAGG-CTTCCGC	GW	pLAT52-mCC
GPAT4	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGTCTCCGGC-GAAGAAGA	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTCCATGGAC-TTGGTCTTATTGAT	GW	pLAT52-mCC
GPAT9	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGAGCAGTACGGCAGGG	GGGGACCACCTTTGTACAAGAAAGC-TGGGTCTTCTTCCAATCTAGCCAGGA	GW	pLAT52-mCC
LPEAT1	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGGAAT-CAGAGCTCAAAGATTTGAA	GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTCTTCTTCTTGATGGAAATCACGG	GW	pLAT52-mCC
THAS1	GGGGACAAGTTTGTACAAAAAAGCAGGCTCATGTGGA-GGCTGAGAACTGG	GGGGACCACCTTTGTACAAGAAAGCTGGGTGAGGGAGGA-GACGTCGCA	GW	pLAT52-mCC

**Table 8: Primer table for generating native promoter lines.** Primers (5'-3') that were used to amplify the sequences for LIME1 and LIME2 cDNA (highlighted with "c") from vector template, and for their native promoter sequences (1045 bp upstream of LIME1, 930 bp upstream of LIME2, highlighted with "p"). Sequences were combined by restriction enzyme digestion and ligation. Ultimately, they were cloned into pGWB604 vector (Nakamura et al., 2010). The vector provided an additional C-terminal eGFP fluorescence protein.

Amplicon	Primer forward	Primer reverse	Template
<i>SalI</i> LIME1p -1045 bp <i>NotI</i>	GATCGTCGACTGATTTCTTTATTTTCTTCCACTTTTGGAA	GATCGCGGCCGCGCACTGACGCGCCTTCT	gDNA
<i>SalI</i> LIME2p -930 bp <i>NotI</i>	GATCGTCGACTAATGGACTAATGCTTGTTGATTAGAACATGA	GATCGCGGCCGCGCACTGATGCGCCTTCT	gDNA
<i>NotI</i> LIME1c <i>XhoI</i>	GCATGCGGCCGCATGGAGAAGATTATTGACGTAG	GATCCTCGAGGATTTCTTCTTGAAGAGGAAGTGTG	pLAT52-LIME1-mVC
<i>NotI</i> LIME2c <i>XhoI</i>	GCATGCGGCCGCATGGAGAAGATTATTGACGTAG	GATCCTCGAGGATTTCTTCTTGAAGAGGAAGTGTG	pLAT52-LIME2-mVC

**Table 9: Primer used to generate sgRNAs for CRISPR/Cas9 cassettes.** Two sgRNAs per cassettes were used, either sgRNA4 and sgRNA5 or sgRNA6 and sgRNA7. Cloning was performed as previously reported (Xing et al., 2014; Wang et al., 2015).

Primer name	Primer sequence 5'-3'
sgRNA6.4_DT1-F0	TGAGATTATTGACGTAGCTTAGTTTTAGAGCTAGAAATAGC
sgRNA6.5_DT2-R0	AACCGGCGGAGATGCAACTCTCCAATCTCTTAGTCGACTCTAC
sgRNA6.5_DT2-BsR	ATTATTGGTCTCGAAACCGGCGGAGATGCAACTCTCCAA
sgRNA6.6_DT1-BsF	ATATATGGTCTCGATTGGCATGAAGGGACATACGATGTT
sgRNA6.6_DT1-F0	TGGCATGAAGGGACATACGATTTTTAGAGCTAGAAATAGC
sgRNA6.7_DT2-R0	AACCCGTGGGTTTGTAACCGGACAATCTCTTAGTCGACTCTAC
sgRNA6.7_DT2-BsR	ATTATTGGTCTCGAAACCCGTGGGTTTGTAACCGGACAA

**Table 10: Microscope settings used to capture confocal images of transiently expressed tobacco pollen tubes.** Given are the coding sequences within a pLAT52 vector (Müller et al., 2017), the first excitation wavelength for the "green channel" used for BODIPY 493/503 or mVenus excitation, and the corresponding detection wavelength window. Additionally, the second excitation wavelength for the "red channel" used for Nile Red or mCherry excitation, along with its corresponding detection wavelength window, is specified. The unit of the wavelength is [nm]. The microscope used for image acquisition was either the Zeiss LSM 510 or LSM 780 (Carl Zeiss, Oberkochen, Germany)

Transiently expressed coding sequences	Excitation wavelength 1	Detection wavelengths 1	Excitation wavelength 2	Detection wavelengths 2	Microscope
LIME1 aa 1-24 mVenus LIME1 aa 291-355	488	488-524	561	568-622	LSM 780
LIME1 aa 1-290 mCherry	488	507-529	561	593-614	LSM 510
LIME1 aa 1-320 mCherry	488	507-529	561	593-614	LSM 510
LIME1 aa 1-35 mVenus LIME1 aa 320-355	488	488-524	561	568-622	LSM 780
LIME1 aa 1-45 mVenus LIME1 aa 291-337	405	450-499	514	523-599	LSM 780
LIME1 aa 1-45 mVenus LIME1 aa 291-355	488	488-524	561	568-622	LSM 780
LIME1 aa 1-45 mVenus LIME1 aa 300-355	488	488-524	561	568-622	LSM 780
LIME1 aa 1-45 mVenus LIME1 aa 320-355	488	488-524	561	568-622	LSM 780
LIME1 aa 1-48 mVenus	488	507-529	561	571-593	LSM 510
LIME1 aa 26-45 mVenus LIME1 aa 291-355	488	488-524	561	568-622	LSM 780
LIME1 aa 291-355 mVenus	488	507-529	561	571-593	LSM 510
LIME1 aa 49-290 mVenus	488	507-529	561	571-593	LSM 510
LIME1 aa 49-355 mCherry	488	507-529	561	593-614	LSM 510
LIME2 mVenus	488	488-524	561	568-622	LSM 780
AtLIME 1-45 mVenus PpLIME 300-355	488	488-524	561	568-622	LSM 780
PpLIME 1-45 mVenus AtLIME 300-355	488	488-524	561	568-622	LSM 780
GbLIME mCherry	488	488-533	561	594-648	LSM 780
MpLIME mCherry	488	488-533	561	594-648	LSM 780
PpLIME mCherry	488	488-533	561	594-648	LSM 780
PsCNMT mCherry	488	488-533	561	594-648	LSM 780
PsCNMT-like mCherry	488	507-529	561	593-614	LSM 510

**Table 11: Primers used for genotyping *lime* mutant lines.** T-DNA primer names include line designation, as well as left bound (LB) and right bound (RB) positions relative to the T-DNA insertion. T-DNA specific LB primers are given. CRISPR deletion genotyping primers indicate the target gene and positions relative to the first base pair from the start codon, along with the forward (f) or reverse (r) direction.

Name	Primer sequence 5'-3'
LB LIME1 SALK_024876/GT_3_9311/ RBWiscDsLOX	ACATTGGATCATTGGATCTCCCG
RB LIME1 SALK_024876/GT_3_9311/LB WiscDsLOX	TGATGGCATTGTTCTCCGCTG
LB LIME2 SM_3_34300/SALK_127412	ATATGTCTTATGGCAGAGATAGTGA
RB LIME2 SM_3_34300/SALK_127412	ATGCTCTTAGCGCCGCACAAC
LBb1.3 Salk	ATTTTGCCGATTTGGAAC
SM LB	TACGAATAAGAGCGTCCATTTAGAGTGA
GT LB	ACCCGACCGGATCGTATCGGT
WiscDsLox LT6	AATAGCCTTTACTTGAGTTGGCGTAAAAG
LIME1/2 +0 f	ATGGAGAAGATTATTGACGTAG
LIME2 -122bp f	GTTTTCTTGTTATTTAATTTACGTCTCTC
LIME1 -195bp f	GTAATACAGGTAAATAGGTAATTAACAATC
LIME1 +495 r	TAATATTGTGCGGCGCT
LIME1 +1250 r	AACAAATCCAAACCTCAAACATC
LIME1 +1962 r	GAAAAGCTCAGCAACCGC

## 7 Discussion

In this thesis, multiple aspects of the role of LD proteins were investigated. First, proteomic analysis was conducted for three distinct plant species and 14 separate tissue types or developmental stages comprising reproductive propagules and vegetative tissues. The resulting data unveiled the putative core proteome transition from seed-like desiccation tolerance to a vegetative state, particularly addressing the LD proteome of yellow nutsedge tubers and *P. patens* spores. Furthermore, the boundaries of the known plant LD proteome were expanded by determining the first Arabidopsis root LD proteome, revealing several novel and partially tissue-specific LD proteins. In addition, various techniques were employed to characterize the potential function, phylogeny, and targeting of the LD-associated methyltransferase (LIME). Thus, a shift in perspective was undertaken in the three projects, from a global perspective encompassing multiple tissue proteomes to identify the protein fingerprint of desiccation tolerance, through an intermediate level focusing on a local LD proteome to access potential new LD candidates, to the characterization of a single LD protein.

### 7.1 Unmasking LD proteome's role in desiccation tolerance

Desiccation tolerance describes the ability to survive prolonged periods of at least 90 % relative water loss (Alpert, 2005). This trait likely was of tremendous advantage when the first embryophytes arose, facing drought as a novel form of severe abiotic stress (Fürst-Jansen et al., 2020). However, due to evolutionary adaption, development of a vascular system to maintain water content and increased metabolomic efficiency, remnants of true desiccation tolerance remain only in seeds and pollen in most angiosperms (Gaff and Oliver, 2013; Oliver et al., 2000). Vegetative desiccation tolerance, on the other hand, is common in bryophytes that represent the extant group most closely resembling earliest land plants (Oliver, 2005).

Desiccation-tolerant seeds, also referred to as orthodox seeds, possess a synergistic interplay of important components, including late embryogenesis abundant (LEA) proteins, sugars, heat shock proteins, and antioxidants, to effectively overcome drought stress (Oliver et al., 2000; Berjak, 2006). Another hallmark of desiccation-tolerance are oleosin covered LDs (Murphy et al., 1995), typically found in most orthodox seeds (Guzha et al., 2023). Consequently, the absence or scarcity of oleosins sets desiccation-sensitive (recalcitrant) seeds apart from orthodox seeds (Leprince et al., 1997; Berjak and Pammenter, 2002). In addition, orthodox seeds tend to have a higher lipid content than recalcitrant seeds (Mello et al., 2010). Abundant

LDs can compensate the reduction in intracellular volume caused by water loss and thus protect against membrane breakdown (Matilla, 2022). In that regard, oleosins may be vital in preventing LDs from coalescing, potentially through shielding and electrostatic repulsion (Siloto et al., 2006; Huang, 1992). The adverse effects on freezing tolerance observed in oleosin-reduced *Arabidopsis* mutants emphasize their critical role in LD stability (Shimada et al., 2008). In *Arabidopsis*, oleosin occurrence has been exclusively reported at the protein level in seeds, seedlings (Kretzschmar et al., 2020), and pollen (Kim et al., 2002), while the vegetative LD proteomes derived from (senescent) leaves (Brocard et al., 2017; Fernández-Santos et al., 2020; Doner et al., 2021) or *tgd1-1 sdp1-4* roots (this thesis) demonstrate a conspicuous absence of seed type oleosins. Only trace amounts of *AtOLE8* were detected in roots. Thus, oleosin-mediated desiccation tolerance seems restricted to *Arabidopsis* reproductive tissues and abolished in vegetative tissues.

In Article I (Niemeyer et al., 2022), we investigated the LD and total proteome of oil-rich tubers of yellow nutsedge (*Cyperus esculentus*) and compared it against the total proteome of the closest related species, purple nutsedge (*Cyperus rotundus*). In Manuscript I, we accessed LD proteomic data by evaluating total protein extracts of (germinating) spores, protonema, and gametophores of *P. patens*.

Equivalent to orthodox seeds, the vegetative tubers of yellow nutsedge are desiccation tolerant (Defelice, 2002) and store approximately a third of their dry weight in lipids (Linssen et al., 1989). Conversely, the tubers of purple nutsedge are sensitive to desiccation and low in oil content, similar to recalcitrant seeds (Ji et al., 2021). We observed distinct features in yellow nutsedge tubers that resemble seeds rather than conventional tubers, such as those found in purple nutsedge, yam, or potato. These features include adaptations in the LD proteome, which may contribute to this anomaly. For instance, we detected a high abundance of several oleosin isoforms, the predominant LD proteins found in resting dry tubers. Despite identifying also an oleosin isoform homolog to *AtOLE8* in both yellow nutsedge and purple nutsedge tubers with low abundance, isoform homologs to main *Arabidopsis* seed oleosins *AtOLE1/2* (Klepikova et al., 2016; Kretzschmar et al., 2020) were exclusively found in yellow nutsedge tubers, which was supported by RNAseq results (Ji et al., 2021). Similar expressions of seed-like oleosins have been observed in vegetative desiccation-tolerant resurrection plants (Xu et al., 2018; Lyall et al., 2020; VanBuren et al., 2017; Costa et al., 2017a). Those are hypothesized to have co-opted a seed-like desiccation tolerance mechanism (VanBuren, 2017; Costa et al., 2017b), thereby suggesting the involvement of oleosins in this process. Furthermore, this also suggests

that yellow nutsedge tubers may have similarly co-opted a seed-like program, which could be the underlying cause for their seed-like characteristics. Moreover, in the desiccation-tolerant spores of *P. patens*, we observed elevated levels of oleosins (Manuscript I), consistent with previous reports (Huang et al., 2009), although they may not be the most abundant species of present LD proteins. Hence, it is plausible that the presence of these oleosins in spores contributes to their establishment of desiccation tolerance.

However, the seed-like characteristics of the LD-proteome are not restricted to the occurrence of oleosins. Another protein predominantly found at angiosperm seed LDs is steroleosin (Aziz et al., 2020). We identified a steroleosin homolog to *AtHSD1* notably enriched in LD-enriched fractions across all stages of yellow nutsedge tubers and even in the total protein fraction of fresh tubers (Article I). However, it was not detected in fresh tubers of purple nutsedge, indicating its exclusive presence in yellow nutsedge tubers and suggesting its potential involvement in the adaption of seed-like characteristics (Article I). The occurrence of *AtHSD1* homologs have also been ascertained in germinating spores of *P. patens* (Manuscript I). Steroleosins are thought to be associated with brassinosteroid metabolism, although the exact nature of their involvement is currently uncertain (Baud et al., 2009; Li et al., 2007). Brassinosteroids are plant hormones involved in growth and developmental processes (Peres et al., 2019; Clouse and Sasse, 1998). They can also occur in moss with a proposed developmental function (Yokota et al., 2017; Morikawa et al., 2009). Brassinosteroids impact spore germination in ferns, although the effect was species-dependent (Gómez-Garay et al., 2018). It remains uncertain whether this effect applies to bryophyte spores and if steroleosins are involved in this process. The occurrence and function of brassinosteroids in yellow nutsedge tubers as well as *P. patens* spores need to be elucidated. Noteworthy, steroleosins are predominantly expressed in seeds (and seed-like tissues) (Lin and Tzen, 2004) but can also occur occasionally in vegetative tissues, as observed in the LD proteome of *tgd1-1 sdp1-4* roots (this study). Nevertheless, HSDs are found only in small quantities in the LD-enriched fraction of the roots and were not detectable in the total protein fraction, indicating a negligible impact (this study). Moreover, isoforms of steroleosins, which are not typically found in seeds, were identified in that particular context (section 4). However, it has been reported that brassinosteroids can affect the size of the root meristem, and their activity in that regard may be regulated by steroleosins through brassinosteroid modification (González-García et al., 2011; Chapman et al., 2012).



Caleosin is the third integral LD protein abundant in most angiosperm seed LD proteomes identified in nutsedge and spores besides oleosin and steroleosin (Tzen, 2012). Nonetheless, caleosins can be found ubiquitously expressed in regenerative and vegetative tissues in *Arabidopsis*, showing a tissue prevalence for specific isoforms (Shimada et al., 2018; Shen et al., 2014). Caleosin isoforms homolog to *AtCLO3/4* could be found low abundantly in yellow and purple nutsedge tubers. However, a homolog of *AtCLO1* is found to be highly abundant only in yellow nutsedge tubers, surpassing even the combined quantity of all oleosins in fresh tubers, indicating high relevance (Article I). In *Arabidopsis*, *CLO3/4* isoforms are predominantly found in vegetative tissues, while *CLO1/2* isoforms are more specific to seeds (Shimada et al., 2018). However, this categorization cannot be applied based on our phylogenetic analysis, as the divergence of these isoforms likely occurred after the divergence of the organisms (Article I). Nevertheless, the predominant presence of *CeCLO1*, and to some extent *CeCLO2*, in yellow nutsedge, along with its absence in non-seed-producing purple nutsedge, suggests a seed-like adaptation and a similar function of *CeCLO1* to *AtCLO1*. On the other hand, the only identified caleosin in *P. patens* spores is a homolog of *AtCLO3*, usually found in vegetative tissues in *Arabidopsis* (Shimada et al., 2018; Park et al., 2018). However, assessing the phylogenetic characteristics solely through BLASTp analysis is insufficient and thus requires careful consideration especially since *CLO1* to *3* all belong to the same clade (Article I). It is worth noting that the high abundance of *AtCLO3* in senescent leaves suggests it is potentially sufficient in stress-related responses as well (Brocard et al., 2017; Fernández-Santos et al., 2020). The abundance of *P. patens* spore caleosin decreases upon spore germination and is not detected in vegetative gametophytes, implying their potential contribution to desiccation tolerance in spores (Manuscript I). Caleosins have a membrane-stabilizing capacity comparable to oleosins (Jiang and Tzen, 2010; Liu et al., 2009), suggesting their importance in protecting membrane integrity during water deprivation. Additionally, they demonstrate peroxygenase activity possibly utilized in their potential role in oxylipin synthesis and signaling (Hanano et al., 2023). Consequently, caleosins represent a versatile protein family involved in both abiotic and potentially biotic stress responses, and they may also play a crucial role in mediating desiccation tolerance. The presence of caleosins alongside oleosins in the LD-proteome of pine and olive pollen further highlights their potential importance in non-seed desiccation-tolerant tissues that is physiologically most analogous to spores (Zienkiewicz et al., 2010; Pasaribu et al., 2017).

LDAP is another protein that has been observed to enhance drought tolerance when overexpressed (Gidda et al., 2016; Kim et al., 2016a; Laibach et al., 2018) and reduce it when mutated (Gidda et al., 2016). Hence, LDAP contributes to the overall resilience against drought. Unlike oleosin, caleosins, and steroleosins, LDAPs are highly ubiquitous and also highly abundant in vegetative tissues such as mesocarp of avocado (Horn et al., 2013), Arabidopsis leaves (Brocard et al., 2017; Fernández-Santos et al., 2020; Doner et al., 2021), and it is the major LD protein in *tgd1-1 sdp1-4* roots (section 4). They were also detected in all investigated stages of *P. patens*, with the highest abundance 48 - 72 h post germination and low occurrence in vegetative gametophytes (Manuscript I). In general, the evaluation of LD protein abundance in vegetative gametophytes is difficult, as no LD fraction was enriched in the study. Due to the small number of LDs in vegetative gametophytes in comparison to spores Huang et al. (2009), only highly abundant LD proteins can be reliably detected. The tubers of yellow nutsedge did not demonstrate increased quantities of LDAP compared to purple nutsedge, suggesting that LDAP may not contribute to the seed-like properties of yellow nutsedge tubers (Article I). Consequently, LDAPs likely do not belong to the seed-like desiccation tolerance mediating proteins network, but might rather play a role in drought responses.

LDAP's interaction partner, LDIP, is reportedly involved in the biogenesis of LDs (Pyc et al., 2017a, 2021). Interestingly, there are striking similarities observed in the occurrence of LDIP across seeds, spores, and yellow nutsedge tubers. While relatively low abundant in all tissues, expression levels are highest in resting Arabidopsis seeds and fastly decrease after germination (Klepikova et al., 2016), while protein abundance peaked in yellow nutsedge dry tubers (Article I) and early spore stages of *P. patens* (Manuscript I). However, their function regarding desiccation tolerance needs to be elucidated.

Remarkably, in the tubers of yellow nutsedge, a homolog of Arabidopsis SLDP1 was detected, which is highly seed-specific in Arabidopsis (Klepikova et al., 2016; Kretzschmar et al., 2020). The precise role of SLDPs, apart from their involvement in an LD-plasma membrane tethering complex with LIPA (Krawczyk et al., 2022b), is still not fully understood. However, due to their potential connection with dehydration stress in leaf veins (Scholz et al., 2022; Klepikova et al., 2016), these proteins emerge as compelling candidates for the repertoire of proteins potentially enhancing desiccation resilience. Interestingly, SLDPs may have evolved in angiosperms (Krawczyk et al., 2022b), and coherently no comparable protein homolog was identified in the spores of *P. patens*, although spores harbor a LD proteome, which resembles the angiosperm

seed proteome. SLDPs in tubers provide additional evidence for potentially utilized seed-like program (including the LD proteome) that may have been co-opted by yellow nutsedge.

## 7.2 Abscisic acid and transcriptional regulators in desiccation tolerance

As conceptually proposed in the introduction, seed-like mechanisms, including LD proteins, may be regulated by transcription (master) regulators. In this regard, ABI3, and ABA signaling in general, is an essential requirement to establish (vegetative) desiccation tolerance (Stevenson et al., 2016; Matilla, 2022; Costa et al., 2016; Yang et al., 2022; Khandelwal et al., 2010). In the tubers of yellow nutsedge, we measured elevated levels of ABA in all tuber stages, but it strongly decreased in sprouted tubers. Furthermore, it was undetectable in fresh purple nutsedge tubers (Article I). In *P. patens* gametophores, exogenous ABA can mediate vegetative desiccation tolerance (Khandelwal et al., 2010; Koster et al., 2010). Thus, ABA levels may be crucial for establishing desiccation tolerance in certain tissues, and its function seems to be conserved across the land plant lineage. Furthermore, transcription factors ABI3, WRINKLED1 (WRI1), and LEAFY COTYLEDON1 (LEC1) are enriched in yellow nutsedge tubers in comparison to purple nutsedge (Article I)(Ji et al., 2021). ABI3 has been reported to upregulate the transcription of *oleosins*, *caleosins*, and *WRI1* in Arabidopsis (Yang et al., 2022; Tian et al., 2020), indicating its potential role in driving the expression of their seed-like isoforms and the elevated levels of *WRI1* in yellow nutsedge tubers. WRI1 serves as a master regulator of fatty acid synthesis, which could potentially account for the substantial neutral lipid content observed in yellow nutsedge tubers (Kong et al., 2020; Cernac and Benning, 2004). LEC1 may additionally augment this process, as it has been reported to upregulate fatty acid synthesis (Mu et al., 2008) and also enhance the expression of *oleosins* (Lotan et al., 1998). Oleosins are assumed to hold a protective role by shielding the excess oil from lipases, as a decrease in oleosin abundance correlates with a reduction in oil content (Miquel et al., 2014). Coherently, a simplified biotechnological adaption of this putative oil enhancement strategy in potato tubers, referred to as the "push-pull-protect" strategy, could boost neutral oil content about 100-fold to 3.3% dry weight (Liu et al., 2017). In this strategy, *WRI1* is overexpressed to push the production of fatty acids, which are pulled towards triacylglycerol by overexpressed *DGAT1*, and subsequently protected by overexpressed *OLEOSIN* (Liu et al., 2017). As a result, the underlying theoretical framework was validated and the strategy led to a substantial increase of neutral lipids, but yellow nutsedge tubers could store even higher quantities of lipids. Hence, it is conceivable that additional factors not accounted for in the simplified approach may

also contribute to the acquisition of higher oil contents. Nevertheless, the upregulation of ABI3, LEC1, and WR1 transcription factors may have enabled adaptations of seed-like characteristics in yellow nutsedge tubers in a relatively short time frame (Article I), which could have been an evolutionary advantage to counteract a possible rapid environmental change. To date, no studies have examined the impact of the transcription factor on the LD proteome, although transcriptional regulation has been described for oleosins, caleosins, and steroleosins (Baud et al., 2009; Yang et al., 2022; Lotan et al., 1998). Therefore, the extent of the influence of the transcription factor on the LD proteome and the specific combination of transcription factors required to co-opt a seed-like LD proteome in vegetative tissues artificially remains unclear and offers exciting approaches for future studies.

However, a question arose regarding whether yellow nutsedge has co-opted a seed-like desiccation program or if this seed-like desiccation tolerance network is derived from an even older abiotic stress program that evolved prior to plant's conquest of land (de Vries and Ischebeck, 2020; Oliver, 2005). The appearance of a seed-like desiccation-tolerance-promoting LD proteome in spores of the bryophyte *P. patens* (Manuscript I) supports the hypothesis of its preexistence before the emergence of spermatophytes (de Vries and Ischebeck, 2020). Thus, it resembles an ancient stress-related countermeasure to desiccation co-opted by propagules of most land plants to cope with harsh environments and gain survival benefits. This is coherent with the recent proteomic identification of oleosins and steroleosins homologs in LD-enriched samples of the streptophyte algae *Mesotaenium endlicherianum* (Dadras et al., 2022) sustaining the idea of a pre-embryophyte origin of the seed-like LD proteome.

LDs and their proteome constitute one component of the coping mechanism against total water loss. Several other compounds need to cooperate to maintain desiccation tolerance synergistically; most of them are also affected by the ABI3 regulon (Matilla, 2022; Lyall and Gechev, 2020; Costa et al., 2016). For instance, the Arabidopsis seed-specific antioxidant *PEROXIREDOXIN1* (Haslekås et al., 1998), known for its important role in seed dormancy (Chen et al., 2020a), has been detected in yellow nutsedge tubers (Article I) and *P. patens* spores (Manuscript I). Additionally, it has been associated with the re-induction of desiccation tolerance in resurrection plants (Costa et al., 2016, 2017b). Key components of desiccation-tolerance are LEA proteins (Artur et al., 2019), of which 82 % may be directly affected by ABI3 (Matilla, 2022). A comprehensive discussion about LEA proteins in the context of desiccation tolerance can be found in Manuscript I.

### 7.3 The shift from dormant to vegetative LD proteome

The LD proteomes of yellow nutsedge tubers and *P. patens* spores are reminiscent of angiosperm seeds in the dormant phase, potentially contributing to their desiccation tolerance and stress resilience; however, also during their respective germination, those structures share striking analogies.

For instance, the oxidosqualene cyclase CYCLOARTENOL SYNTHASE (CAS) abundance increases in *Arabidopsis* seedlings and peaks about 48 h post germination (Kretzschmar et al., 2020). Cycloartenol is the base backbone compound of sterols, such as cholesterol and phytosterols (Ohyama et al., 2009; Carland et al., 2010). The STEROL METHYLTRANSFERASE1 (SMT1) abundance increases in parallel, with a short time offset, peaking 60 h post germination (Kretzschmar et al., 2020). SMT1 is crucial for phytosterol production by catalyzing the first step of the pathway by C-24 methylation of cycloartenol (Carland et al., 2010; Diener et al., 2000). The implication of phytosterol production during germination is consistent with observations of slightly increased phytosterol content reported from various germinating seeds (Mostafa et al., 1987; Huang and Grunwald, 1988; Horbowicz and Obendorf, 1992). These observations suggest that phytosterol production serves to meet the changed sterol demand of the growing plant (Kalinowska and Wojciechowski, 1984). In germinating spores of *P. patens*, the abundance of homologs of CAS also increases after germination, peaking at 48 hours post-germination, and subsequently decreases (Manuscript I). Also, a SMT1 homolog showed a delayed response, with increased abundance observed in protonema and gametophores. The overall trend of increased abundance of phytosterol-producing enzymes with a time set off between CAS and SMT1 is thus consistent between *Arabidopsis* seeds and *P. patens* spores. In contrast, this effect was not observed in the only single investigated sprouting stage of yellow nutsedge tubers (Article I). The abundance of CAS and SMT1 in spores and seeds is low, and they might be below the detection threshold in the total protein fraction of sprouted tubers if they are present. It is also conceivable that nutsedge tubers do not rely on sterol metabolism products for sprouting. On the other hand, a homolog of SMT1 is detectable in fresh tubers, albeit at low abundance, suggesting the presence of phytosterol production in that stage.

During the germination of *Arabidopsis* seeds, lipids stored in the LDs are mobilized by SDP1 and SDP1-like lipases (Graham, 2008; D'Andrea, 2016; Kelly et al., 2011). This process is accompanied by the ubiquitination and degradation of the shielding oleosins (Deruyffelaere

et al., 2015), promoted by PUX10 (Kretzschmar et al., 2018; Deruyffelaere et al., 2018). Homologs of PUX10 have been identified in germinating seeds (Kretzschmar et al., 2020), pollen (Kretzschmar et al., 2018), spores (Manuscript I), and nutsedge tubers (Article I), suggesting its potential role in regulating ubiquitin-dependent oleosin degradation in all these tissues. Lipid mobilization in *P. patens* spores may be reminiscent of proposed mechanisms native to cucumber cotyledons (Matsui et al., 1999; Feussner et al., 1997) and olive pollen (Zienkiewicz et al., 2013) in a lipoxygenase (LOX) manner (Feussner et al., 2001), putatively in cooperation with an oil body lipase (OBL). Moreover, lipoxygenases are pivotal enzymes involved in the biosynthesis of oxylipins, which serve as important signaling molecules in plant defense mechanisms, including jasmonates (Wasternack and Feussner, 2018; Ruan et al., 2019). *PpLOX4* belongs to the highest abundant proteins in spores, showing a decreasing quantity trend during germination, indicating eminent importance in dormant spores (Manuscript I). Equivalent patterns are observed for *PpLOX3* and *PpLOX10*; however, they are not as abundant as in the previous case (Manuscript I). Although not demonstrated yet, *PpLOX4* may be an LD-localized LOX form due to several characteristics similar to angiosperm LD-associated LOX enzymes. For instance, it belongs to the type I category of 13-LOX enzymes (Anterola et al., 2009) equivalent to other LD-associated LOX (Feussner et al., 2001), which are not targeted to the chloroplasts, whereas in *Arabidopsis*, 13-LOX are commonly of type II in *Arabidopsis*, which are chloroplastic targeted (Andreou and Feussner, 2009; Shibata et al., 1994). Additionally, *PpLOX4* was present in spores prior to type II LOXs forms (Manuscript I) and had a predicted molecular weight exceeding 100 kDa (calculated with ExPASy ProtParam (Gasteiger et al., 2005)), which is typical for angiosperm LD-associated LOX enzymes (Andreou and Feussner, 2009; Feussner et al., 1996, 2001). Conversely, LOX proteins were neither detected in the LD-enriched fraction of germinating *Brassica napus* (Terp et al., 2006) nor in *Arabidopsis thaliana* seeds, seedlings (Kretzschmar et al., 2020), leaves (Brocard et al., 2017), *Nicotiana tabacum* pollen tubes (Kretzschmar et al., 2018) or yellow nutsedge tubers (Article I). Additionally, ectopically expressed *AtLOX1* did not show LD localization (Fernández-Santos et al., 2020). Consequently, LOXs are not necessarily LD proteins across all plant species.

Feussner et al. (1997) postulated an LD-associated LOX lipid catabolism model. According to this model, a 13-LOX enzyme initially produces a 13-hydroperoxide-octadecadienoic acid (13-HPODE) from linoleic acid. Subsequently, the 13-HPODE is reduced to 13-hydroxy-octadecadienoic acid (13-HODE) and released into the cytosol and glyoxysomes for  $\beta$ -oxidation

by an LD lipase, which may be *PpOBL1* in spores. While a specific reducing enzyme has not been proposed, the peroxygenase activity of caleosins could potentially fulfill this function (Shimada and Hara-Nishimura, 2015; Hanano et al., 2023). A plausible function of *PpOBL1* in spores is the liberation of LOX-oxygenated fatty acids in the LOX-dependent lipid catabolism. An indication pointing towards the potential collaboration between *PpOBL1* and *PpLOX4* is their status as the most abundant proteins observed in imbibed spores, with nearly indistinguishable expression patterns detected during both germination and gametophytic growth (Manuscript I). *AtOBL1* has reported mono-, di-, and triacylglycerol lipase activity and has been suggested to be important for providing membrane building blocks in pollen tube growth, as the growth of pollen tubes was hampered in *obl1* Arabidopsis mutants (Müller and Ischebeck, 2018). During the germination of Arabidopsis seeds, *AtOBL1* may not be very active as it cannot compensate for a knockout of the lipases SDP1 and SDP1-like, which account for approximately 90 % of the total lipase activity (Kelly et al., 2011). Furthermore, *obl1* mutants do not show any seed germination phenotypes (Müller and Ischebeck, 2018). In addition, the absence of OBL homologs in yellow nutsedge tubers suggests the presence of an alternative lipase activity in these tissues. In germinating *P. patens* spores, however, no SDP1 homolog was detected (Manuscript I), and based on the striking abundance of *PpOBL1*, it can be assumed to be the major TAG lipase in spores. There are indications that OBLs show a preference for hydroxylated fatty acids as substrates. In the case of castor beans (*Ricinus communis*), where approximately 74 % of the seed oil is composed of the hydroxylated ricinoleic acid (Yusuf et al., 2015), *RcOBL1* has been found to be abundantly expressed and may be the major LD protein (Fuchs et al., 1996; Altaf et al., 1997; Eastmond, 2004). Moreover, it has been reported that *RcOBL1* exhibits a preference for the substrate HODE over linoleic acid (Balkenhohl et al., 1998; Eastmond, 2004). Moreover, the necessity of an OBL homolog for the production of volatile compounds derived from oxylipins in tomato (*Solanum lycopersicum*) strengthens this assumption (Garbowicz et al., 2018; Ischebeck et al., 2020). Hence, it is plausible to suggest that *P. patens* spore germination heavily depends on the degradation of lipids via the LOX-dependent pathway in conjunction with the involvement of OBL. In addition, the putative presence of this lipid degradation machinery during germination suggests an origin of that pathway dating back to the emergence of embryophytes or possibly beyond. Nevertheless, the actual presence of that lipid catabolism pathway in spores remains unproven as of now. These results should be a motivation to investigate the possible nature of lipid mobilization in spores in future approaches, either through determining lipid composition in germinating spores or

through the (transient) expression of corresponding genes to replicate the mechanism in other tissues and species.

In this thesis, the analysis of those comprehensive proteomic datasets of yellow nutsedge tubers and *P. patens* focused primarily on the (putative) LD proteome and other seed-like characteristics. However, both datasets could be used as a reference data pool for various use cases for future approaches. The nutsedge dataset stands as the pioneering proteomic dataset encompassing yellow and purple nutsedge tubers, and it marks the first instance of retrieving an LD-enriched protein fraction from an underground organ. It complements published transcriptomic datasets that cover a reduced variety of tissues (Yang et al., 2016; Ji et al., 2021). Several proteomic datasets have been published for *P. patens* gametophores and protonema under various treatments (Mamaeva et al., 2022; Luo et al., 2020; Yotsui et al., 2016; Toshima et al., 2014; Cui et al., 2012; Skripnikov et al., 2009; Wang et al., 2010, 2009, 2008; Cho et al., 2006; Sarnighausen et al., 2004). In contrast, our dataset is distinct as it includes spores and several spore germination stages and provides a comparative analysis against untreated protonema and gametophores of the same ecotype (Reute). This represents the first instance where such a comprehensive approach has been undertaken for *P. patens*, surpassing the existing transcriptomic data available for spore germination stages (Fernandez-Pozo et al., 2020; Perroud et al., 2018). Moreover, the proteins intended for proteomics analysis underwent the same processing, quantification, and identification procedures, thereby assuring comparability within the datasets. Proteomic studies show a snapshot of the actual protein landscape that can substantially differ from the expected one predicted by transcriptomics (Chaturvedi et al., 2013). Furthermore, they enable the enrichment of subcellular compartments to investigate organelle compositions specifically. Therefore, they are superior for investigating LD proteomes when compared to transcriptomic data.

#### **7.4 Shedding light on the LD proteome of Arabidopsis roots**

The identification of novel LD proteins was a fundamental aim of this study. In order to maximize the probability of uncovering novel proteins, we investigated *Arabidopsis tgd1-1 sdp1-4* root tissues. This approach was undertaken due to the extensive variations observed in LD proteomes across different tissues, indicating the presence of distinct proteins within unexplored LD proteomes (Kretzschmar et al., 2020). This study represents to our knowledge the first elucidation of a root LD proteome. Following the yellow nutsedge tubers, it also



represents the second LD proteome analysis conducted on an underground organ. Arabidopsis Col-0 roots exhibit a TAG content below 1 % of dry weight, even when exposed to high sucrose concentrations that augment TAG content in roots (Kelly et al., 2013). This highlights the challenges associated with isolating LDs from this particular tissue and the necessity of the utilization of oil-rich *tgdl1-1 sdp1-4* mutants. In the LD-enriched fraction of roots, we identified 31 proteins that have been previously characterized as LD-associated proteins in Arabidopsis (section 4). Its composition reveals a distinct type of vegetative LD proteome that differs from those found in Arabidopsis leaves (Brocard et al., 2017; Fernández-Santos et al., 2020; Doner et al., 2021).

In *tgdl1-1 sdp1-4* roots, LDAP proteins are the predominant proteins found in the LD-enriched fraction, accounting for approximately 66 % of the total abundance of all known LD proteins when considering all isoforms. On the other hand, caleosins make up only about 1 % of the summed abundance. In contrast, drought-stressed Arabidopsis leaves exhibit a different distribution, with approximately 87 % of the identified LD proteome assigned to CLO3 and only 5.6 % attributed to LDAPs (Doner et al., 2021). Thus, there is an inversion in the prevalence of the major LD protein between these tissues. However, estimations of the physiological consequences of this reversal are challenging. LDAPs probably serve as the primary LD coating, similar to what has been proposed for avocado mesocarp (Horn et al., 2013; Gidda et al., 2016; Pyc et al., 2017a). On the other hand, the stabilizing function attributed to caleosins appears to be negligible in roots (Liu et al., 2009). Surprisingly,  $\alpha$ -DOX1 is the second most abundant protein, constituting about 20 % of the total known LD proteome in roots.  $\alpha$ -DOX1 is suggested to play a role in the biotic defense response by oxygenating linolenic acid, producing the unstable intermediate 2-HPOT. It is suggested that this intermediate is subsequently reduced to the antifungal compound 2-HOT through the peroxygenase activity of caleosin (Shimada et al., 2014; Shimada and Hara-Nishimura, 2015). However, it remains to be clarified whether the small quantities of caleosins are sufficient to degrade the cytotoxically active oxygenated fatty acids (Farmer and Mueller, 2013) or whether another unknown peroxidase is involved in this process. In addition, there is the possibility that  $\alpha$ -DOX1 may serve an unknown, independent function in roots. However, the culturing method itself may induce micro-wounding, possibly leading to the upregulation of  $\alpha$ -DOX1 expression. This is supported by the observation of increased  $\alpha$ -DOX1 gene expression in tomato roots following wounding (Tirajoh et al., 2004). The relative abundance of LIME proteins in roots is approximately 1.6 %, primarily attributed to LIME2. In leaves, only LIME1 is detected, which is

approximately 10 times less abundant than LIME2 in roots (Doner et al., 2021), indicating a significant, yet unknown, role of LIME2 in roots.

In addition to the distinctive distribution of previously characterized LD proteins, 12 proteins were discovered to co-localize with LDs when transiently expressed in *Nicotiana tabacum* pollen tubes or *Nicotiana benthamiana* leaves. Notably, all 12 of these proteins belong to protein families that have not been previously identified as being localized to LDs in plants. The GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 9 (GPAT9) is part of different GPAT protein family than GPAT4 and GPAT8 (Waschburger et al., 2018; Shockey et al., 2016), which belong to the GPAT1 to 8 family and were recently described as LD-localized proteins (Fernández-Santos et al., 2020). Therefore, we expanded the known Arabidopsis LD proteome from 24 protein families to 36, which represents an increase of 50 % of the known plant LD-associated protein families. This increase represents the most substantial expansion of LD-associated protein families in plants through a single study to date.

Beside of GPAT9, those further 11 novel LD localized proteins are ALDEHYDE DEHYDROGENASE 3H1, (ALDH3H1), FOREVER YOUNG 3 (FEY3), LEAF WILTING 1 (LEW1), ACYLCOA:LYSOPHOSPHATIDYLETHANOLAMINE ACYLTRANSFERASE 1 (LPEAT1), PRENYLCYSTEINE METHYLESTERASE (PCME), THALIANOL SYNTHASE 1 (THAS1), and the proteins encoded by AT1G30130, AT1G78800, AT4G13160, AT4G33180, and AT5G04070. The lack of description for the last five proteins makes them intriguing subjects for future investigation. However, due to the lack of information, it is also challenging to determine the relevance of these proteins within the scope of this study. Their average riBAQ is  $\leq 1$ , implying minor, though not necessarily unimportant, functions.

Interestingly, two distinct acyltransferases, GPAT9 and LPEAT1, were identified, suggesting lipid metabolic activities at the surface of the LDs. GPAT9 is the sole GPAT enzyme in Arabidopsis responsible for triacylglycerol production (Singer et al., 2016; Shockey et al., 2016), while GPAT4 and GPAT8, which are also associated with LDs (Fernández-Santos et al., 2020), have been described to take part in cutin biosynthesis (Li et al., 2007; Yang et al., 2012b). Thus, Arabidopsis GPAT9 parallels the function of insect GPAT4, which is also localized to LDs and participates in triacylglycerol synthesis (Wilfling et al., 2013). GPAT5 is an additional GPAT protein that exhibits significant enrichment in the LD-enriched fraction and has been assigned to the group of high-confidence LD protein candidates (Supplemental Dataset 1C). Hence, GPAT5 might also localize to LDs, but experiments confirming its LD localization are still pending. Arabidopsis *gpat5* mutants exhibited a substantial decrease in root suberin levels

(Beisson et al., 2007), which plays a critical role in maintaining the integrity of the cell wall, particularly in roots, by acting as a barrier against excess salt and water loss (de Silva et al., 2021). As a result, (putatively) LD-associated GPATs may be crucial for the stress resilience of roots. It is plausible, yet hypothetical, that LD-associated GPATs directly use acyl building blocks from the LD.

The second identified acyltransferase, LPEAT1, is suggested to be involved in the acyl-editing process of phosphatidylethanolamine (PE) within the Lands' Cycle (Lands, 1965; Jasieniecka-Gazarkiewicz et al., 2016, 2017), operating in a phospholipase-independent manner (Jasieniecka-Gazarkiewicz et al., 2016). The *Arabidopsis* knockout mutant *lpeat1* exhibited a modest decrease in growth, while the double knockout *lpeat1 lpeat2* resulted in a significant reduction in growth (Stålberg et al., 2009). On the other hand, overexpression resulted in the opposite effect, implying the crucial influence of LPEATs in PE homeostasis (Jasieniecka-Gazarkiewicz et al., 2017). The localization of LPEAT1 to the LD suggests that the acyl editing of PE may occur directly at the LD membrane. Intriguingly, homologs of LPEAT1 and GPAT9 have been detected in the LD-enriched fraction of *Chlamydomonas reinhardtii* (Nguyen et al., 2011), suggesting their association with LDs potentially being highly conserved across different organisms within Viridiplantae.

Furthermore, THAS1, a protein involved in the synthesis of the triterpenoid thalianol (Fazio et al., 2004; Field et al., 2011; Nützmänn and Osbourn, 2015; Liu et al., 2020; Bai et al., 2021), has been identified as a novel LD protein. THAS1 is likely a root-specific protein, as it has not been detected in any other tissue so far (Table 12); also, its transcript appears exclusively in roots (Klepikova et al., 2016). The enzyme catalyzes the cyclization of 2,3-oxidosqualene, resulting in the formation of the tricyclic triterpene thalianol (Fazio et al., 2004). THAS1 is part of a gene cluster including the cytochrome P450s thalianol hydroxylase (THAH) and thalian-diol desaturase (THAD), which can metabolize thalianol into other compounds (Field et al., 2011; Nützmänn and Osbourn, 2015). Remarkably, THAH and THAD were also found to be significantly enriched in the LD-enriched fractions of *tgd1-1 sdp1-4* roots. However, their localization appeared to be at the ER (unpublished work of Dr. Patricia Scholz). Thalianol and thalianol-derived metabolites have been found to affect the bacterial composition in the soil surrounding plants, indicating their potential role in shaping a beneficial root microbiome (Huang et al., 2019). Additionally, a recent study has linked thalianol content to root length. *Arabidopsis thas1* mutants exhibited longer roots while inhibiting further metabolism of thalianol led to shorter root lengths (Bai et al., 2021). It remains to be clarified if thalianol can be

sequestered into LD, as has been observed for example for carotenoids: Recent studies have provided evidence for LDs acting as storage compartments for the hydrophobic carotenoids. In a biotechnological approach, this phenomenon has been observed in the cytosolic LDs when a cytosolic retargeted carotenoid pathway was transiently expressed in *N. benthamiana* leaves, leading to the sequestration of provitamin A into LDs (Zheng et al., 2023). A similar observation has been made in the algae species *Dunaliella bardawil*, where carotenoids are also stored in cytosolic LDs (Davidi et al., 2014a). Similarly, the hydrophobic triterpenoid thalianol has the potential to be stored in LDs, possibly in an esterified form analogous to SEs, which could be advantageous for the plant considering its negative impact on root growth. Future approaches may analyze the lipid and terpenoid composition of root LDs to address this hypothesis.

LEW1 is another novel LD protein that acts in the synthesis of the poly-isoprenoid dolichol (Zhang et al., 2008). Dolichol is a lipid carrier for carbohydrates during the formation of dolichyl phosphate-linked oligosaccharides required for protein *N*-glycosylation (Burda and Aebi, 1999). Additionally, LEW1 is vital for Arabidopsis responses to drought, ER stress, and dark-induced senescence (Zhang et al., 2008). Arabidopsis *lew1* mutants showed an 85% decrease in dolichol content, resulting in glycosylation defects and compromised membrane integrity (Zhang et al., 2008). Prenylation may also enhance membrane targeting of certain proteins towards LDs (Kory et al., 2016).

PCME is an enzyme that demethylates isoprenylcysteine methyl esters of prenylated proteins (Huizinga et al., 2008). As a result, PCME may act as a positive regulator of ABA and can be induced by salt and osmotic stress, or exogenous ABA leading to a positive feedback loop (Huizinga et al., 2008; Lan et al., 2010). Due to the high sucrose concentration in the axenic root culture, the presence of PCME in the root LD proteome may be an indicator of osmotic stress.

Relatively little is currently understood about FEY3. This enzyme has been identified as an oxidoreductase, playing a vital role in the growth and development of the vegetative shoot apex (Medford et al., 1992; Callos et al., 1994). It may also be involved in developmental processes in the root meristem (Callos et al., 1994). Its role at the LD needs to be elucidated.

The ALDH3H1 is part of the NAD(P)<sup>+</sup>-dependent aldehyde dehydrogenase superfamily, which has a prominent role in detoxifying aldehydes (Kirch et al., 2004; Stiti et al., 2011, 2020). ALDHs have been reported to be LD-associated proteins in mice (Kitamura et al., 2015). This suggests that this protein family may perform similar functions across various eukaryotic organisms. *ALDH3H1s* expression levels in Arabidopsis roots are correlated with salt and

drought stress, as well as exogenous ABA (Kirch et al., 2004; Stiti et al., 2011), and potentially is a crucial part of the abiotic stress response. Furthermore, its transcription levels increased in the recovery phase following heat stress (Zhao et al., 2017). Consistently, overexpression has been reported to strengthen abiotic stress resilience (Stiti et al., 2011). Moreover, *ALDH3H1* expression has been induced particularly under osmotic stress in roots (Kirch et al., 2004). The fact that *ALDH3H1* was the fifth highest abundant LD protein in roots further supports the notion that the roots may have been under stress conditions (Supplemental Dataset 1A).

It is worth noting that previous studies have reported the subcellular localization of several of these proteins to be distinct from LDs. For instance, GPAT9 (Gidda et al., 2009) and LPEAT1 (Jasieniecka-Gazarkiewicz et al., 2021) in tobacco BY-2 cells and PCME in *N. benthamiana* protoplasts (Lan et al., 2010) are reportedly associated with the ER, and *ALDH3H1* has been reported to have a cytosolic localization (Stiti et al., 2011), respectively. The observed discrepancy may be attributed to the scarcity of LDs in the examined tissues, strong overexpression and the omission of LD staining in those studies. Furthermore, most of the examined novel LD proteins in our study also demonstrated additional ER targeting besides the LD targeting (unpublished work of Siqi Sun, Magdiel Lim Sheng Satha, and Janis Dabisch). Thus, those proteins do not have an exclusive prevalence to target LDs. In addition, these specific cell types may not provide a favorable micro-environment for appropriate LD targeting.

This study initially characterized a root LD proteome. However, it is important to interpret the results with caution. The use of sterile axenic root cultivation in a high-sucrose medium and the utilization of a *tgd1-1 sdp1-4* mutant line, which significantly affects lipid storage (Fan et al., 2014), does likely not accurately represent the LD proteome of wild-type plants grown in a natural soil environment. Furthermore, several identified novel LD proteins are potential stress markers that may be induced through osmotic stress or micro wounding caused by constant movement during cultivation. On the other hand, this experimental setup may have facilitated the identification of a greater number of LD-associated proteins than would have been possible in a wild-type setting due to the additional stress-responding proteins. Consequently, this approach successfully fulfilled the aim of identifying novel LD proteins. Finally, it is plausible that the dataset has not yet been fully utilized, and there may be additional LD proteins that could be identified using it, especially considering the presence of untested candidates exhibiting high enrichment in the LD-enriched protein fraction (Supplemental Dataset 1B,C). The vast diversity of identified LD proteins and their potential associated functions underscores

the importance of exploring LD proteomes in different tissues and under stress conditions, with the aim of expanding our understanding of lipid droplets.

### 7.5 Proteomic data reveal diversity in LD proteomes across species and tissues

This thesis includes comprehensive proteomic datasets of four distinct species, from *Arabidopsis thaliana* (section 4), *Physcomitrium patens* (Manuscript I), *Cyperus esculentus* (Article I), and to some extent *Cyperus rotundus* (Article I). From *P. patens* and *C. rotundus*, only total proteomes were examined due to sparse raw material quantity limitations and lipid scarcity in the tubers, respectively. To bring the dataset into a broader LD proteome context, a selection of LD proteomes derived from the LD-enriched fractions of *Arabidopsis thaliana* *tgdl1-1 sdp1-4* roots, *Cyperus esculentus* dry tubers, and from total protein fraction of *Physcomitrium patens* 0 h spores were compared against published datasets derived from LD-enriched fractions of *Arabidopsis thaliana* rehydrated seeds (Kretzschmar et al., 2020), *Arabidopsis thaliana* 60 h seedlings (Kretzschmar et al., 2020), and *Nicotiana tabacum* pollen tubes (Kretzschmar et al., 2018) (Figure 24A, Table 12). The datasets were chosen due to the variety of different examined tissues as well as utilization of the same sample preparation procedure, LC-MS/MS proteomic facility (except for *tgdl1-1 sdp1-4* root measurements), and data processing following the guidelines published in Horn et al. (2021). This ensured a high level of comparability between these proteomic datasets. LC-MS/MS measurements of *tgdl1-1 sdp1-4* root proteins were performed by another proteomics facility, which had a more sensitive machine with a higher measurement rate. Therefore, the detection of low quantities of newly identified LD proteins from roots may fall below the detection threshold in the other LC-MS/MS measurements. To ensure a simplified comparison, only proteins previously described as LD localized in Arabidopsis, either through publications or within the scope of this thesis, were considered.

Due to the utilization of the total protein fraction only, statistical analysis regarding the enrichment and probability of each protein in the LD proteome of *P. patens* spores were not possible. Therefore, the classification of LD proteins was based on estimations derived from sequence similarities to the closest Arabidopsis homolog and lacked any further evidence. However, LD targeting is strongly conserved across species for most major LD protein species (Lundquist et al., 2020; Kory et al., 2016). Hence, it is plausible that the identity of LD-assigned proteins in *P. patens* spores are accurate for most proteins. Due to the high abundance of LDs in spores (Huang et al., 2009), the proportion of known LD-associated proteins in the total

protein fraction of 0 h *P. patens* spores, reaching 11.9%, was even exceeding the proportional percentage found in LD-enriched fractions from Arabidopsis leaves, which was 5.1 % (Figure 24A). The tubers of yellow nutsedge, which are also rich in oil and therefore abundant in LDs, provide evidence that low-abundant LD proteins are unlikely to be detected in the total protein fraction compared to the LD-enriched protein fraction. However, if the percentage of LD proteins within the total protein fraction is high, the approximate distribution of the main LD proteins can still be estimated relatively accurately despite employing the total protein fraction (Supplemental Figure 3). Hence, the provided proteomic dataset of *P. patens* spores represent the primary LD proteins with a relatively high confidence level. Nonetheless, this approach does not allow for the identification of species-specific LD-localizing protein candidates, nor does it provide insights into whether LOX proteins may function as LD proteins in *P. patens* spores, which is a possibility as discussed in section 7.3.

The comparison reveals the extensive diversity of plant LD proteomes, which appear to be specifically adapted to meet the unique requirements of the distinct tissues (Figure 24A, Table 12). This diversity is illustrated by the broad spectrum of major LD proteins, ranging from LDAPs in Arabidopsis roots OBLs in *P. patens* spores, oleosins in yellow nutsedge tubers, Arabidopsis seeds and seedlings, and caleosins in Arabidopsis leaves and tobacco pollen tubes (Figure 24A). Only four protein families could consistently be identified in all examined LD proteomic datasets, comprising caleosins, LDAPS, LDIPs, and LIPID DROPLETS AND STOMATA 1 (LDS1) (Table 12). Therefore, these proteins may be the core LD proteins of all embryophytes. LDAPs and LDIPs have been suggested to potentially play a role in LD biogenesis (Guzha et al., 2023; Pyc et al., 2021); however, LDS1 may also serve as a crucial component in this process. LDS1, recently implicated in the morphogenetic development of stomata, reportedly interacts with the seipin complex, and knockout mutants exhibit aberrantly shaped LD phenotypes similar to the *seipin2 seipin3* double mutant, providing further support for this assumption (Ge et al., 2022). Furthermore, the pervasive presence of LDS1 in all tissues indicates its involvement in the fundamental processes of LD biogenesis. LDS1 is a member of the Rab GTPase protein family (Ge et al., 2022), previously found in various LD-enriched fractions (Kretzschmar et al., 2018; Fernández-Santos et al., 2020), and also have LD-localized homologs in mammals (Dansako et al., 2014). However, it is the first isoform demonstrating plant LD localization in a microscopic evaluation, indicating its specific association with LDs. Remarkably, the homolog of LDS1 is not significantly enriched at LDs of yellow nutsedge tubers (Article I). Similar to the interacting seipins (Ge et al., 2022; Salo et al.,

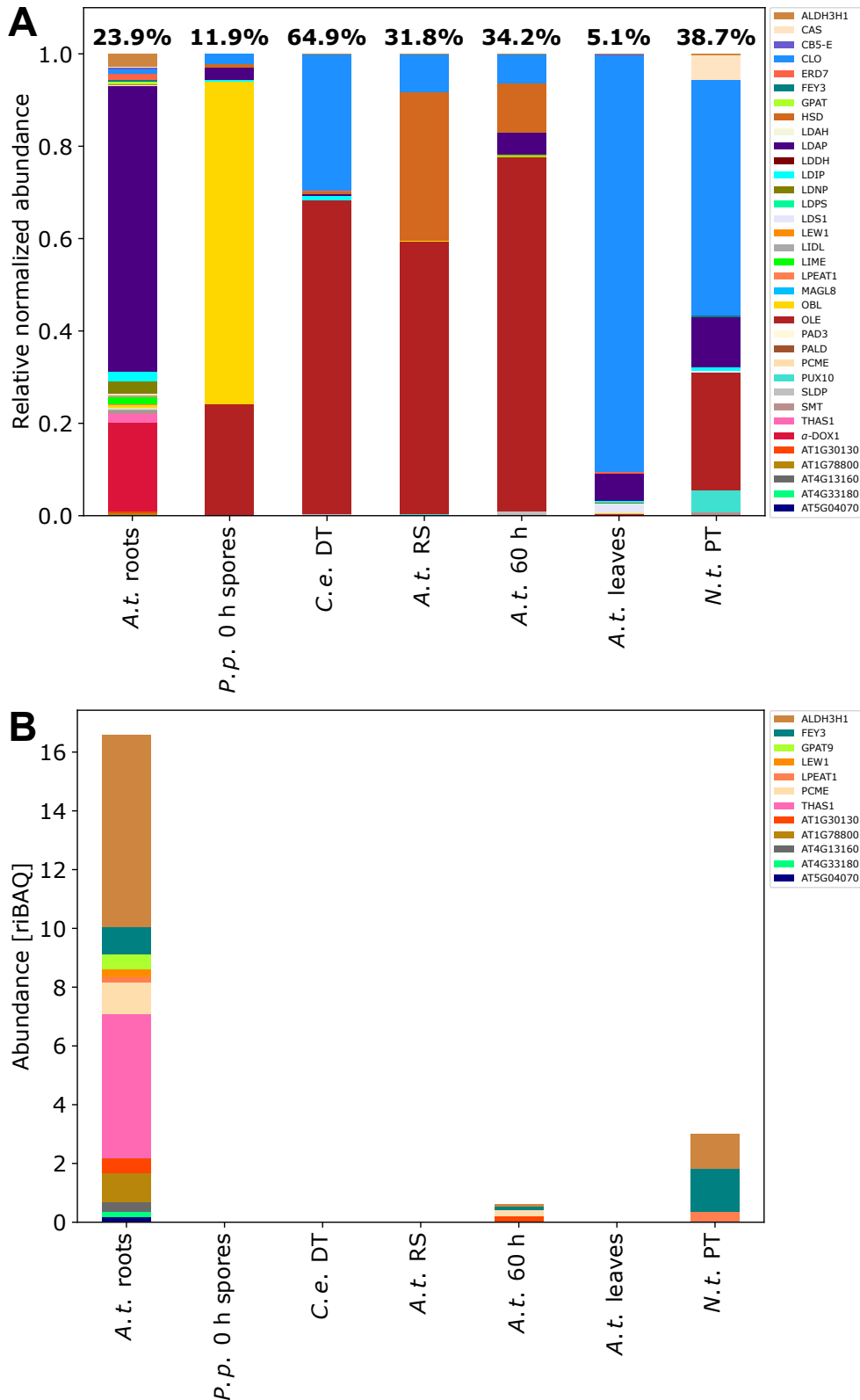
2019), they may be located at the ER-LD junction sites. Hence, their presence in all datasets could be an artifact due to the imperfect LD isolation process.

As expected, oleosins are predominantly found in desiccated tissues (Table 12). However, trace amounts of OLE8 are also present in *Arabidopsis* roots (section 4), and homologs of *AtOLE8* were also found in purple nutsedge tubers with low abundance, as well as in roots and leaves of yellow nutsedge (Article I). This raises the question of whether OLE8 may represent a vegetative tissue-expressed form of oleosin. Contrary to the proteomic finding, transcript expression of *AtOLE8* is almost exclusively in (developing) seeds and not detectable in roots (Klepikova et al., 2016). Therefore, it is possible that OLE8 is induced in response to potential osmotic stress in roots, or it may exhibit increased resistance to protein degradation, allowing it to persist in tissues in low quantities. Moreover, the confidence level for detecting OLE8 in roots is low due to its observed presence in only 3 out of 5 replicates, and it was identified based on a single peptide. As a consequence, it is unlikely that OLE8 is a vegetative tissue-expressed isoform of oleosin, at least not in *Arabidopsis*.

This comparative approach also highlights a limitation in relying solely on BLASTp for protein assignment. While OBLs have been reported in tobacco pollen tubes, the closest homologs of *Arabidopsis* OBLs do not align with their counterparts in tobacco by using BLASTp. Consequently, it appears that there are no homologs in tobacco, when in fact there are. Therefore, conducting in-depth phylogenetic analysis is essential for the thorough characterization of individual proteins when comparing across multiple species.

Interestingly, six out of the 12 newly discovered LD proteins were only found in *Arabidopsis tgd1-1 sdp1-4* roots, which emphasizes the importance of exploring LD proteomes across various tissues to gain a comprehensive understanding of their composition (Figure 24B, Table 12). It is difficult to determine whether these proteins were detected solely due to the increased sensitivity of the measurement device, if they are specific to the tissue, or if some proteins are present as a response to potential osmotic or wound stress factors. Noteworthy, the high abundance of stress associated  $\alpha$ -DOX1 and ALDH3H1, which appears to be widely present in low quantities in all angiosperm LDs, along with the presence of the stress marker PCME, implies that abiotic stress may have caused the presence of at least some novel identified root LD proteins. Similar effects were observed with PAD3, which became apparent only when leaves were subjected to biotic stress through *Pseudomonas* infiltration (Fernández-Santos et al., 2020). Therefore, it seems crucial to examine LD proteomes under different stress conditions to identify proteins that specifically localize to LDs under such circumstances.





**Figure 24: Comparison of LD protein composition of different species and tissues.** Displayed are the LD enriched fractions of *Arabidopsis thaliana tgd1-1 sdp1-4* roots (*A.t. roots*) (section 4), *Cyperus esculentus* dry tubers (*C.e. DT*) (Article I), *Arabidopsis thaliana* rehydrated seeds (*A.t. RS*) (Kretzschmar et al., 2020), *Arabidopsis thaliana* 60 h seedlings (*A.t. 60 h*) (Kretzschmar et al., 2020), *Arabidopsis thaliana* drought stressed leaves (*A.t. leaves*) (Doner et al., 2021), *Nicotiana tabacum* pollen tubes (*N.t. PT*) (Kretzschmar et al., 2018), and the total protein extract of *Physcomitrium patens* 0 h spores (*P.p. 0 h spores*). A) The abundance of all known LD proteins was normalized to 1 to show their relative distribution. The percentage of those proteins within the corresponding protein fraction is written above the columns. B) relative iBAQ in per mill of proteins newly identified in roots. The abundances of members within the same gene family were summed. Given are assigned names or encoding genes, if a name is not applicable. Only homologs of known *Arabidopsis* LD proteins are displayed.

## 7. Discussion

**Table 12: Comparison of LD protein composition of different species and tissues.** Abundance of identified proteins in relative iBAQ in per mill. Displayed are the LD-enriched fractions of *Arabidopsis thaliana* *tg1-1 sdp1-4* roots (*A.t.* roots)(section 4), *Cyperus esculentus* dry tubers (*C.e.* DT)(Article I), *Arabidopsis thaliana* rehydrated seeds (*A.t.* RS)(Kretzschmar et al., 2020), *Arabidopsis thaliana* 60 h seedlings (*A.t.* 60 h)(Kretzschmar et al., 2020), *Arabidopsis thaliana* drought stressed leaves (*A.t.* leaves)(Doner et al., 2021), *Nicotiana tabacum* pollen tubes (*N.t.* PT)(Kretzschmar et al., 2018), and the total protein extract of *Physcomitrium patens* 0 h spores (*P.p.* 0 h spores). The abundances of members within the same gene family were summed. Given are assigned names or encoding genes, if a name is not applicable. Only homologs of known *Arabidopsis* LD proteins are displayed.

Protein name	<i>A.t.</i> roots	<i>P.p.</i> 0 h spores	<i>C.e.</i> DT	<i>A.t.</i> RS	<i>A.t.</i> 60 h	<i>A.t.</i> leaves	<i>N.t.</i> PT
$\alpha$ -DOX1	45.82				0.23	0.25	
ALDH3H1	6.55		0.02	0.003	0.08	0.02	1.18
CAS	0.33			0.02	0.22	0.08	20.24
CB5-E	0.98		0.99	0.02	0.11	0.03	
CLO	2.27	2.57	191.17	26.41	21.25	46.28	197.95
ERD7	3.16		0.95			0.26	
FEY3	0.91				0.15		1.47
GPAT	1.46				0.03		
HSD	0.64	0.87	4.59	101.92	36.16		
LDAH	0.06				0.19		
LDAP	148.28	3.28	2.05	0.004	15.76	2.99	41.75
LDDH					0.31		
LDIP	4.80	0.30	5.64	0.03	0.05	0.16	2.95
LDNP	6.79			0.002	0.14	0.16	
LDPS				0.28	0.12		
LDS1	0.90	0.02	0.06	0.06	0.10	0.98	1.12
LEW1	0.26						
LIDL	0.40				0.34		
LIME	3.71				0.67	0.08	
LPEAT1	0.17		0.01				0.36
MAGL8	0.31						
OBL	1.16	83.22		0.16	0.92	0.03	
OLE	0.05	28.65	441.11	188.21	262.32		98.78
PAD3							
PALD			0.05				
PCME	1.08				0.20		
PUX10	0.21	0.08	0.40	0.02	0.15		18.13
SLDP			2.15	1.03	2.20		
SMT	1.88				0.48		3.14
THAS1	4.90						
AT1G30130	0.51				0.20		
AT1G78800	1.00						
AT4G13160	0.30						
AT4G33180	0.19						
AT5G04070	0.19						

In conclusion, these findings should serve as an encouraging stimulus for future research to investigate the LD proteomes of diverse species and tissues under various stress conditions. This will improve our understanding of LD composition and function under different stressors and facilitate the identification of new LD candidates.

### **7.6 New insights to LIME proteins, a highly conserved protein family**

Part of this thesis was the characterization of LIME proteins that were first described as LD-associated proteins in *Arabidopsis* by Kretzschmar et al. (2020) but have so far not been characterized in any detail.

The closest homologs of *Arabidopsis* LIME proteins are enzymes from the BIAs pathway (Lang et al., 2019; Bennett et al., 2018; Torres et al., 2016). However, as outlined in the introduction, LIME proteins likely have different molecular functions than BIAs-related methyltransferases. Shortly summarized, *Arabidopsis* is incapable of synthesizing BIAs (Facchini et al., 2004), the targeting of CNMTs is reported to be cytosolic instead of LD (as confirmed (Figure 14)) (Hagel and Facchini, 2012), and recombinant *AtLIME2* did not convert isoquinoline substrates (Liscombe and Facchini, 2007).

Phylogenetic analysis revealed LIME proteins being the potential origin of BIAs-related methyltransferases (Figure 8). Moreover, the phylogenetic tree supports the assumption of Liscombe et al. (2005) of a monophyletic origin of BIAs-producing species. Most of the examined BIAs-producing species exhibit additional LIME-like homologs within the putative LIME clade, and clustered homologs of these species indicate a close evolutionary relationship. This further suggests that the original function of LIME proteins remains significant in these species and is not compensated by BIAs-related methyltransferases. This finding underscores the possibility of different molecular functions attributed to LIMEs compared to BIAs-related methyltransferases. The origin of BIAs-related methyltransferases, which may be LIME proteins, is hypothesized to exhibit a highly promiscuous substrate specificity that impedes molecular characterization (Morris and Facchini, 2019). Consequently, evaluation was necessary to narrow the possibilities of substrate and product.

LIME proteins are highly conserved in Viridiplantae (Figure 8, (de Vries and Ischebeck, 2020)) and form a putative phylogenetic clade. Also, their LD localization seems to be (partially) conserved, as observed through transient expression of LIME homologs in tobacco pollen tubes (section 5.2.2), and implied by CoIP experiments transiently expressing a confirmed

LD-associating LIME1-mCherry (Kretzschmar et al., 2020) in *N. benthamiana* leaves, which successfully pulled down *NbLIME* (section 5.3), implying spatial proximity. Moreover, homologs of LIME proteins have been identified in the LD-enriched fractions of the algae *Chlamydomonas reinhardtii* (Nguyen et al., 2011) and *Mesotaenium endlicherianum* (Dadras et al., 2022), in the latter case with significant enrichment in the LD-enriched fraction compared to the total protein fraction. Therefore, LD targeting seems to be highly conserved across species. It is conceivable that LIME proteins and their inherent molecular function predate land plants. Hence, their metabolic function is more likely within primary metabolism than in specialized secondary metabolism (such as BIAs).

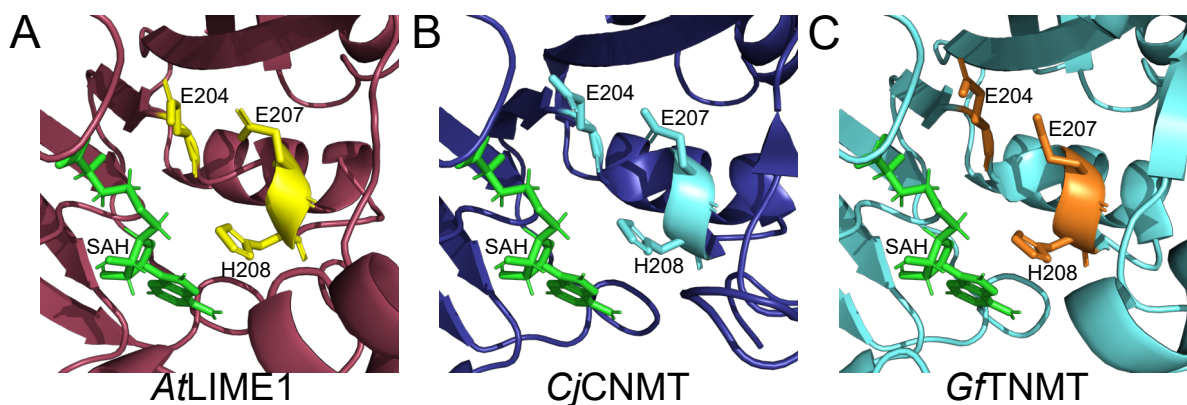
According to a protein model prediction performed with AlphaFold (Jumper et al., 2021), LIME1 features a typical class I SAM-dependent methyltransferase with SAM-binding region and  $\alpha/\beta$ -Rossmann fold, most similar to *N*-methyltransferases (NMT) (Gana et al., 2013; Schubert et al., 2003). The model prediction exhibited a remarkable confidence level, as evidenced by the quality scores of the per-residue local distance difference test (pLDDT) (Tunyasuvunakool et al., 2021). Specifically, 85.6% of the residues achieved an excellent pLDDT score of > 90, indicating high reliability. Furthermore, 95.8% of the residues attained a good pLDDT score of > 70, with the remaining residues below that threshold primarily situated in flexible loop regions. Thus, an *N*-methylation activity of LIME is plausible due to the high confidence of the prediction.

Sterol synthesis was considered a potential metabolic field for LIME. LDs are a potential sterol source (Krawczyk et al., 2022a), which are primary metabolites. Another methyltransferase, which has a known activity in sterol metabolism (SMT1, (Carland et al., 2010)), also localizes on LDs but emerged later in evolution (de Vries and Ischebeck, 2020). Therefore, the contribution of LIME proteins in sterol metabolism could not be excluded. Nevertheless, neither an *Arabidopsis lime double* mutant nor ectopic overexpression in *N. benthamiana* leaves indicated involvement in sterol metabolism (section 5.9). Thus, LIMEs may not take part in sterol metabolism directly. Noteworthy, LIME also shares sequence similarities with bacterial cyclopropane fatty acid synthases (Liscombe and Facchini, 2007; Huang et al., 2002). These enzymes have already been documented in plants, including the phylogenetic identification of homologs in *Arabidopsis* distinct to LIME proteins, and they also display notable distinctions from their bacterial counterparts (Yu et al., 2011; Shockey et al., 2018). Furthermore, this conversion does not require *N*-methylation. Hence, it seems unlikely that LIME proteins serve an identical function, although it cannot be entirely ruled out.

Sequence similarities between SAM-dependent methyltransferases are common due to the highly conserved methyltransferase domain that spans approximately 150 bp in their protein sequence (Martin and McMillan, 2002; Schubert et al., 2003). Therefore, sequence similarities do not directly imply functional analogism in SAM-dependent methyltransferases that can have a broad substrate range (Choi et al., 2002); an important reason why a stringent *e*-value cut off of  $< 10^{-15}$  was chosen for our phylogenetic analysis to identify exclusively nearly identical homologs (section 5.1).

A hypothesis-free *ex vivo* experiment has not provided further hints for the identification of substrate candidates due to the lack of detectable methylation products (section 5.10). The assay is performed in an aqueous solution, and the standard LC-MS settings are optimized for hydrophilic metabolites (Ni and Feussner, 2023). Consequently, hydrophobic compounds are sparsely detectable. Hence, the native LIME substrate may be hydrophobic, which would also be plausible considering the proximity to the LD. However, estimating whether the enzyme was active is speculative without information regarding the enzyme's activity. Nevertheless, appropriate LD targeting was reported (Kretzschmar et al., 2020) and visually approved mCherry fluorescence on immobilized protein on the beads after CoIP (data not shown) indicates a correct protein fold. Despite that, the activity could be impaired by factors such as an unknown missing co-factor, incorrect environmental requirements, or steric hindrance caused by the fluorescence tag. In addition, the *ex vivo* assay revealed the inactive protein control free mCherry as a weak point. Free mCherry is reported as a cytosolic protein (Price et al., 2020) while LIME1-mCherry targets the LD (Kretzschmar et al., 2020). In the case of CoIP, this phenomenon could have caused the relatively high abundance of significantly distinct proteins to be pulled down because the bait proteins LIME1-mCherry and free mCherry are present in different environments. The CoIP approach was selected due to the strong affinity of bacterial-expressed recombinant LIME proteins to membranes that impedes its protein purification (data not shown), despite a successful previous report that could not be replicated (Liscombe and Facchini, 2007). An appropriate control would be an enzyme inactivated by a mutation that lacks substrate conversion but still has similar physical properties.

Despite anticipated functional differences between LIME proteins and BIA-related methyltransferases, parts of the active site required for *N*-methylation seem to be conserved (Figure 25). Crystallography and biochemical studies on a *Coptis japonica* *S*-coclaurine-*N*-methyltransferase (*Cj*CNMT) (Bennett et al., 2018) and *Glaucium flavum* tetrahydroprotoberberine *N*-methyltransferase (*Gf*TNMT) (Lang et al., 2019) demonstrated crucial influence on activity



**Figure 25: Conservation of active site in methyltransferases.** Predicted protein model of *AtLIME1* (acquired from AlphaFold Protein Structure Database (Varadi et al., 2022), identifier Q8L788) in comparison to crystal structures of BIA-related methyltransferases *CjCNMT* (PDB: 6GKZ) and *GfTNMT* (PDB: 6P3O). The highly conserved residues E204, E207, and H208, demonstrated to be important for methylation activity in *CjCNMT* and *GfTNMT*, are highlighted. SAH ligand indicates a possible pocket of SAM binding. Images were visualized with PyMOL 2.5.0. Abbreviations: *AtLIME1*, *Arabidopsis thaliana* lipid droplet-associated methyltransferase; *CjCNMT*, *Coptis japonica* *S*-cocclaurine-*N*-methyltransferase; SAH, *S*-adenosyl-*L*-homocysteine; SAM, *S*-adenosyl-*L*-methionine; *GfTNMT*, *Glaucium flavum* tetrahydroprotoberberine *N*-methyltransferase.

of the residues E204, E207, and H208. Mutagenesis to alanine significantly reduced activity in both enzymes, but stronger in E204A and H208A (> 90 % reduced activity to wt) than in E207A (25 - 70 % reduced activity to wt) (Bennett et al., 2018; Lang et al., 2019); however, triple mutation E204A, E207A, H208A in *GfTNMT* abolished enzyme activity entirely (Lang et al., 2019). The substrates ammonium ion designated for methylation may get stabilized by the glutamates E204, E207, and the histidine H208 through salt or H-bridges (Bennett et al., 2018; Lang et al., 2019). The histidine H208 might act as a general base to deprotonate the substrate's ammonium group, forming a strong nucleophilic amine used for methylation from the SAM donor with an  $S_N2$  reaction (Bennett et al., 2018; Lang et al., 2019). The high confidence AlphaFold structure prediction of *AtLIME1* also assigned conserved residues E204, E207, and H208 and a confined area with excellent confidence of pDLLT > 90 at an equivalent position to the active site of *GfTNMT* and *CjCNMT* (Figure 25), suggesting a putatively analogous reaction mechanism in *AtLIME1*. The same residues are also present in *AtLIME2* (Figure 9). This further underscores the potential of LIME proteins to be *N*-methyltransferases. Thus, mutations of the residues E204, E207, and H208 to alanine may also inactivate LIMEs enzymatic activity without altering its (putative) targeting sites and, therefore, might be more suitable as a control in future approaches.

The assigned tissues localization of LIME in germinating seedlings (Kretschmar et al., 2020), the transition zone of roots and leaves (section 4 + 5.2.3)(Klepikova et al., 2016) with prevalence in guard cells (Figure 15) (Obulareddy et al., 2013) implied potential influence in development and stress response, as located close to entry sited of pathogens or

chemical toxins. However, neither developmental (Supplemental Figure 2, Figure 20) nor salt (Figure 21A), drought (Figure 18 + 19), heat (Figure 21B), or UV stress-related (Figure 19) phenotypes of single and double *lime* mutant could be observed compared to the wild type. This is coherent with transcriptomic results demonstrating no regulation upon treatment with biotic or abiotic stressors in comparison to mock treatments (Winter et al., 2007), except for selenium stress. However, SAM-dependent methyltransferases are reportedly part of selenium detoxification mechanisms (Chasteen and Bentley, 2003); thus, higher expression levels of *LIMEs* seem unsurprising under selenium stress. Consequently, LIME proteins may not be involved in developmental or stress regulatory mechanisms in laboratory conditions. However, the considerable consistency observed in the phylogenetic tree and the relatively high sequence conservation strongly suggest an essential function that does not directly impact the observed phenotype but may be essential in certain situations not covered in our assessment. Besides that, LIME proteins seem to be an exclusive part of the vegetative LD-proteome in embryophytes. So far, they have solely been described in tissues from *Arabidopsis* seedlings, leaves, and roots (Table 12, Figure 15). However, they were absent in resting seeds, pollen tubes, spores, and tubers (Table 12). This is consistent with transcriptomic data from *Arabidopsis* (Klepikova et al., 2016) and *P. patens* (Fernandez-Pozo et al., 2020). Therefore, the molecular function of LIME proteins may require an active metabolic environment of a vegetative tissue.

Summarized, the substrate of LIME needs to be elucidated. The putative compound is anticipated to be hydrophobic, nitrogen-carrying, possibly not involved in stress signaling or development, and part of the primary metabolism. However, any assumptions about specific substrates would be speculative. It also remains to be clarified whether LIME2 has the same enzymatic activity as LIME1 since their phylogenetic divergence appears to have occurred through a recent duplication event (section 5.1).

### **7.7 Putative targeting of LIMEs to the LD in the context of established mechanisms**

Our current understanding of the targeting mechanisms of proteins to the LD membrane remains limited, and no consensus LD targeting sequence has been described so far. Nevertheless, some targeting strategies have been proposed, as elaborated in the introduction (section 1.3.1). In general, LD proteins are classified in class I, comprising proteins that derive

from the ER and have to pass through the ER-LD tethering complex, and class II proteins, that bind from the cytosol to the LD directly (Kory et al., 2016).

Indications suggest that LIME proteins belong to the class I LD proteins. Firstly, *AtLIME* proteins, as well as homologs from *P. patens*, *M. polymorpha*, poppy, and ginkgo were found to exhibit additional localization to fibrous and reticulated structures, which were likely the ER (Figure 9+14). These findings suggest that LIME proteins fulfill the criterion of ER targeting, which is necessary for classifying them as class I LD proteins (Kory et al., 2016; Bersuker and Olzmann, 2017). Moreover, a conserved C-terminal tris-lysine motif, reminiscent of the eukaryotic KKXX ER retrieval motif (Teasdale and Jackson, 1996; Cosson et al., 1998; Stornaiuolo et al., 2003; Austin et al., 2007), was identified in all LIME homologs. Similar motifs can be found in *AtGPAT8* and *AtGPAT9* (Gidda et al., 2009), that also can localize to the LD ((Fernández-Santos et al., 2020), Figure 7). The C-terminal KKXX motif is canonically associated with the retrograde transport of membrane proteins from the Golgi back to the ER in a coat protein complex I (COPI) dependent manner (Teasdale and Jackson, 1996; Gao et al., 2014). In insect cells, it was observed that the translocation of insect GPAT4 protein to the LD required the involvement of COPI (Wilfling et al., 2014). Moreover, COPI can directly interact with the LD surface (Wilfling et al., 2014; Thiam et al., 2013a). These findings suggest a possible mechanism in which COPI either directly inserts class I LD proteins into the LD membrane or serves as a bridge between the ER and LD, facilitating the passage of class I LD proteins to the LD (Wilfling et al., 2014). However, the mechanisms by which ER proteins are distinguished from class I LD proteins are not yet understood. One plausible explanation could be the presence of ER membrane proteins' transmembrane-spanning domains, which cannot be incorporated into the LD monolayer.

The human protein ASSOCIATED WITH LIPID DROPLETS I (ALDI) is also an LD-associated methyltransferase that harbors a putative C-terminal ER retention signal (Turró et al., 2006; Zehmer et al., 2008). It is considered as class I LD protein (Zehmer et al., 2008; Prévost et al., 2018), supporting the assumed classification of similarly structured LIME proteins as class I LD proteins. In contrast to LIME, ALDI possesses a hydrophobic N-terminal domain that is sufficient for (weak) LD targeting (Zehmer et al., 2008), whereas LIME1 necessitates the presence of both its N-terminus and C-terminus for LD targeting (Figure 12). In contrast to typical membrane proteins, LIMEs lack a (half) membrane-spanning domain or hydrophobic stretches that would anchor it to the membrane surface (Kory et al., 2016). According to a high-confidence protein model prediction, the regions of LIME proteins that are sufficient for



targeting LDs are located within  $\alpha$ -helices. This implies a membrane targeting mechanism that resembles an amphipathic helix (Prévost et al., 2018), albeit divided into two separate components that potentially have to be fixated in spatial proximity (section 5.2.1). However, amphipathic helices are associated with class II LD proteins (Čopič et al., 2018; Dhiman et al., 2020; Prévost et al., 2018), demonstrating the limitations of the targeting classifications through prediction of putative membrane targeting motifs and the need of experimental verification.

Albeit narrowing successfully the sequence constituents likely being involved in LIMEs LD targeting (section 5.2.1), several open questions remain: i) how exactly do N- and C-terminal sequence stretches need to be positioned to facilitate targeting; ii) why does CNMT have a cytosolic localization despite high sequence and structure homology to LIME, including the C-terminal tris-lysine motif ((Liscombe et al., 2012), Figure 14B); iii) why do chimeric LIMEs, despite minimal sequence divergence, not restore same species targeting phenotypes (Figure 14C+D); iv) why do residues 300-320 increase targeting specificity towards the LD (Figure 12); v) can single mutations lead to changes in the targeting phenotypes; vi) is the predicted and experimentally supported homo-dimeric conformation essential to facilitate targeting (Figure 16); vii) are other methyltransferases programmable to target LDs? The current data would not allow any meaningful interpretation other than speculative assumptions, indicating the need of further research. This may also be relevant to explain the LD targeting of other proteins, such as the other reported plant LD-targeted methyltransferase SMT1 (Kretzschmar et al., 2018).

The COPI-mediated retrograde protein transport to the ER, or possibly directly to the LD, stands beside two other described ER membrane protein incorporation mechanisms associated with class I LD proteins. For instance, oleosins are known to be targeted to the ER membrane via the signal recognition particle pathway (Napier et al., 2001). For the incorporation of mammalian PUX8, the closest homolog to plant PUX10, a mechanism involving the peroxisome-biogenesis factor proteins PEX19 and PEX3 has been proposed (Schrul and Kopito, 2016). Those proteins may be integrally incorporated into the ER membrane by a hydrophobic hairpin (Huang and Huang, 2017) or hydrophobic stretch (Kretzschmar et al., 2018), respectively. It is worth mentioning that the integral membrane protein CLO1 relies on both its N-terminus and a hydrophobic 'proline-knot' motif to target LDs in yeast (Purkrtova et al., 2008). Thus, caleosins seem to require multiple sequence sections for LD targeting similar to LIMEs.

The observation that LIMEs require two protein constituents flanking a fluorescence protein to mimic spatial distance could aid in identifying LD targeting sequences in other proteins. For instance, when single truncations of LDAPs were examined, it was found that the entire

protein is necessary for targeting to LD (Gidda et al., 2016). However, it is probable that several smaller sequence stretches are required for targeting, and with the aforementioned approach, these stretches could be recombined and tested. When LDAPs were expressed in insect cells, they were naturally localized to LDs (Pyc et al., 2021), indicating their inherent ability for LD targeting. In contrast, LDIP required both LDAP and an amphipathic helix (Pyc et al., 2017a) to target LDs in the same experiment (Pyc et al., 2021), suggesting the necessity of a protein-protein interaction for proper LD targeting. Both, LDIP and LDAPs, are considered class II LD proteins (Pyc et al., 2021). Class II LD proteins may also target the LD through a lipid anchor (Kory et al., 2016; Dhiman et al., 2020), although this has not yet been described for plant LD proteins. Nevertheless, there are indications that prenylated proteins are present at the LDs of roots. PCME can demethylate prenylated proteins (Lan et al., 2010), indicating the presence of such proteins in close proximity. Additionally, the mammalian homolog of ALDH3H1 localizes to LDs through an attached geranylgeranyl group (Kitamura et al., 2015). Therefore, it is conceivable that also ALDH3H1 may target the LD in an equivalent way. In conclusion, amphipathic helices, protein-protein interaction, and lipid-anchor are the three elements proposed to facilitate class II LD protein targeting, which presumably occurs in plants as well. (Kory et al., 2016; Bersuker and Olzmann, 2017; Dhiman et al., 2020).

## 8 Concluding remarks

The field of molecular plant LD research prospered in the last decade, and LDs are recently recognized as *bona fide* organelles. So far, most of the gained knowledge has been acquired through studies in oil-rich angiosperm seeds, seedlings, mesocarp, and, recently, senescent or infected leaves. In this thesis, we could extend this knowledge by investigating additional tissues and species, by confirming several new LD-associated proteins, and by investigating the LIME protein family.

A broad data fundament is required to pave the path towards a deeper understanding of LD biology. To contribute to this fundament, three comprehensive proteomic datasets of previously uncharacterized tissues were generated. This includes the first and second LD proteomic datasets from an underground organ, as well as the first proteomic dataset of a germinating bryophyte spore.

The dataset enabled a more profound knowledge about the nature of a seed-like proteome in yellow nutsedge tubers, reminiscent of what was previously described for resurrection plants and orthodox seeds. It is likely, that yellow nutsedge tubers co-opted for this an ancient regulatory drought stress mechanism. In contrast to previous attempts, mainly focusing on LEA proteins, we emphasized the role of the LD proteome concerning desiccation tolerance. Especially seed-type oleosins, caleosins, steroleosins, and SLDPs were prevalent in tubers and might be the core LD proteins of the desiccation-tolerance mediation. Furthermore, proteomic data of *P. patens* spores indicate that those drought stress-related mechanisms, including the desiccation tolerance LD proteome type, predate seeds.

With the investigation of root LDs, 12 novel LD proteins were identified, leading to an increase of 50 % of the known protein families associated with LDs in plants. Among these novel LD proteins, the presence of the THALIANOL SYNTHASE 1 suggests LDs as a possible terpenoid production platform. A comparison between our proteomes and previous studies revealed high variability in the LD protein composition, which further seems to be enhanced by different stress factors.

The inaugural study of LIME proteins revealed that they exhibit high conservation across Viridiplantae and are the possible origin of BIA-related methyltransferases. They are likely class I LD proteins with a unique target sequence. Their metabolic function remains elusive.

In summary, this endeavor has the potential to act as a foundational milestone in gaining a more practical comprehension of LD's role in desiccation tolerance, their associated proteins, and a first glimpse into LIMEs.

---

## References

- Abdelraheem, E., Thair, B., Varela, R. F., Jockmann, E., Popadić, D., Hailles, H. C., Ward, J. M., Iribarren, A. M., Lewkowicz, E. S., Andexer, J. N., Hagedoorn, P.-L., and Hanefeld, U. (2022). Methyltransferases: Functions and applications. *ChemBioChem*, 23(18):e202200212.
- Abell, B. M., High, S., and Moloney, M. M. (2002). Membrane protein topology of oleosin is constrained by its long hydrophobic domain. *Journal of Biological Chemistry*, 277(10):8602–8610.
- Abell, B. M., Holbrook, L. A., Abenes, M., Murphy, D. J., Hills, M. J., and Moloney, M. M. (1997). Role of the proline knot motif in oleosin endoplasmic reticulum topology and oil body targeting. *The Plant Cell*, 9(8):1481–1493.
- Acevedo, F., Rubilar, M., Shene, C., Navarrete, P., Romero, F., Rabert, C., Jolivet, P., Valot, B., and Chardot, T. (2012). Seed oil bodies from *Gevuina avellana* and *Madia sativa*. *Journal of Agricultural and Food Chemistry*, 60(28):6994–7004.
- Aharoni, A., Jongsma, M., and Bouwmeester, H. (2005). Volatile science? Metabolic engineering of terpenoids in plants. *Trends in Plant Science*, 10(12):594–602.
- Al-Whaibi, M. H. (2011). Plant heat-shock proteins: A mini review. *Journal of King Saud University - Science*, 23(2):139–150.
- Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C. C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D. E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W. L., Berry, C. C., and Ecker, J. R. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, 301(5633):653–657.
- Alpert, P. (2005). The limits and frontiers of desiccation-tolerant life. *Integrative and Comparative Biology*, 45(5):685–695.
- Altaf, A., Ankers, T. V., Kaderbhai, N., Mercer, E. I., and Kaderbhai, M. A. (1997). Acid lipase of castor bean lipid bodies: Isolation and characterisation. *Journal of Plant Biochemistry and Biotechnology*, 6(1):13–18.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410.
- Andreou, A. and Feussner, I. (2009). Lipoxygenases – Structure and reaction mechanism. *Phytochemistry*, 70(13-14):1504–1510.
- Anterola, A., Göbel, C., Hornung, E., Sellhorn, G., Feussner, I., and Grimes, H. (2009). *Physcomitrella patens* has lipoxygenases for both eicosanoid and octadecanoid pathways. *Phytochemistry*, 70(1):40–52.
- Arlt, H., Sui, X., Folger, B., Adams, C., Chen, X., Remme, R., Hamprecht, F. A., DiMaio, F., Liao, M., Goodman, J. M., Farese, R. V., and Walther, T. C. (2022). Seipin forms a flexible cage at lipid droplet formation sites. *Nature Structural & Molecular Biology*, 29(3):194–202.
- Armenteros, J. J. A., Salvatore, M., Emanuelsson, O., Winther, O., Heijne, G. v., Elofsson, A., and Nielsen, H. (2019). Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance*, 2(5).

- Aronel, V., Lemieux, B., Hwang, I., Gibson, S., Goodman, H. M., and Somerville, C. R. (1992). Map-based cloning of a gene controlling omega-3 fatty acid desaturation in *Arabidopsis*. *Science*, 258(5086):1353–1355.
- Artur, M. A. S., Zhao, T., Ligterink, W., Schranz, E., and Hilhorst, H. W. M. (2019). Dissecting the genomic diversification of late embryogenesis abundant (LEA) protein gene families in plants. *Genome Biology and Evolution*, 11(2):459–471.
- Aubert, Y., Vile, D., Pervent, M., Aldon, D., Ranty, B., Simonneau, T., Vavasseur, A., and Galaud, J.-P. (2010). RD20, a stress-inducible caleosin, participates in stomatal control, transpiration and drought tolerance in *Arabidopsis thaliana*. *Plant and Cell Physiology*, 51(12):1975–1987.
- Austin, R. S., Provart, N. J., and Cutler, S. R. (2007). C-terminal motif prediction in eukaryotic proteomes using comparative genomics and statistical over-representation across protein families. *BMC genomics*, 8(1):1–16.
- Aymé, L., Arragain, S., Canonge, M., Baud, S., Touati, N., Bimai, O., Jagic, F., Louis-Mondésir, C., Briozzo, P., Fontecave, M., and Chardot, T. (2018). *Arabidopsis thaliana* DGAT3 is a [2Fe-2S] protein involved in TAG biosynthesis. *Scientific Reports*, 8(1):17254.
- Aziz, U., Saleem, N., Tang, T., and Zhang, M. (2020). Genomic analysis and expression investigation of steroleosin gene family in *Arabidopsis thaliana*. *Journal of Animal and Plant Sciences*, 30:133–146.
- Bae, S., Park, J., and Kim, J.-S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of cas9 RNA-guided endonucleases. *Bioinformatics*, 30(10):1473–1475.
- Baek, G. H., Cheng, H., Choe, V., Bao, X., Shao, J., Luo, S., and Rao, H. (2013). Cdc48: A swiss army knife of cell biology. *Journal of Amino Acids*, 2013:1–12.
- Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Schaeffer, R. D., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., van Dijk, A. A., Ebrecht, A. C., Opperman, D. J., Sagmeister, T., Buhlheller, C., Pavkov-Keller, T., Rathinaswamy, M. K., Dalwadi, U., Yip, C. K., Burke, J. E., Garcia, K. C., Grishin, N. V., Adams, P. D., Read, R. J., and Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, 373(6557):871–876.
- Bai, Y., Fernández-Calvo, P., Ritter, A., Huang, A. C., Morales-Herrera, S., Bicalho, K. U., Karady, M., Pauwels, L., Buyst, D., Njo, M., et al. (2021). Modulation of *Arabidopsis* root growth by specialized triterpenes. *New Phytologist*, 230(1):228–243.
- Bailey, A. P., Koster, G., Guillermier, C., Hirst, E. M., MacRae, J. I., Lechene, C. P., Postle, A. D., and Gould, A. P. (2015). Antioxidant role for lipid droplets in a stem cell niche of *Drosophila*. *Cell*, 163(2):340–353.
- Bailey-Serres, J., Parker, J. E., Ainsworth, E. A., Oldroyd, G. E. D., and Schroeder, J. I. (2019). Genetic strategies for improving crop yields. *Nature*, 575(7781):109–118.
- Baker, A., Carrier, D. J., Schaedler, T., Waterham, H. R., van Roermund, C. W., and Theodoulou, F. L. (2015). Peroxisomal ABC transporters: functions and mechanism. *Biochemical Society Transactions*, 43(5):959–965.

- Balkenhohl, T., Kuhn, H., Wasternack, C., and Feussner, I. (1998). A lipase specific for esterified oxygenated polyenoic fatty acids in lipid bodies of cucumber cotyledons. *Advances in plant lipid research. Sevilla: Secretariado de Publicaciones de la Universidad de Sevilla*, pages 320–322.
- Banaś, A., Carlsson, A. S., Huang, B., Lenman, M., Banaś, W., Lee, M., Noiriél, A., Benveniste, P., Schaller, H., Bouvier-Navé, P., and Stymne, S. (2005). Cellular sterol ester synthesis in plants is performed by an enzyme (phospholipid:sterol acyltransferase) different from the yeast and mammalian acyl-CoA:sterol acyltransferases. *Journal of Biological Chemistry*, 280(41):34626–34634.
- Banks, J. A., Nishiyama, T., Hasebe, M., Bowman, J. L., Gribskov, M., dePamphilis, C., Albert, V. A., Aono, N., Aoyama, T., Ambrose, B. A., Ashton, N. W., Axtell, M. J., Barker, E., Barker, M. S., Bennetzen, J. L., Bonawitz, N. D., Chapple, C., Cheng, C., Correa, L. G. G., Dacre, M., DeBarry, J., Dreyer, I., Elias, M., Engstrom, E. M., Estelle, M., Feng, L., Finet, C., Floyd, S. K., Frommer, W. B., Fujita, T., Gramzow, L., Gutensohn, M., Harholt, J., Hattori, M., Heyl, A., Hirai, T., Hiwatashi, Y., Ishikawa, M., Iwata, M., Karol, K. G., Koehler, B., Kolukisaoglu, U., Kubo, M., Kurata, T., Lalonde, S., Li, K., Li, Y., Litt, A., Lyons, E., Manning, G., Maruyama, T., Michael, T. P., Mikami, K., Miyazaki, S., Morinaga, S., Murata, T., Mueller-Roeber, B., Nelson, D. R., Obara, M., Oguri, Y., Olmstead, R. G., Onodera, N., Petersen, B. L., Pils, B., Prigge, M., Rensing, S. A., Riaño-Pachón, D. M., Roberts, A. W., Sato, Y., Scheller, H. V., Schulz, B., Schulz, C., Shakhov, E. V., Shibagaki, N., Shinohara, N., Shippen, D. E., Sørensen, I., Sotooka, R., Sugimoto, N., Sugita, M., Sumikawa, N., Tanurdzic, M., Theißen, G., Ulvskov, P., Wakazuki, S., Weng, J., Willats, W. W., Wipf, D., Wolf, P. G., Yang, L., Zimmer, A. D., Zhu, Q., Mitros, T., Hellsten, U., Loqué, D., Otilar, R., Salamov, A., Schmutz, J., Shapiro, H., Lindquist, E., Lucas, S., Rokhsar, D., and Grigoriev, I. V. (2011). The *Selaginella* genome identifies genetic changes associated with the evolution of vascular plants. *Science*, 332(6032):960–963.
- Bartolini, S., Leccese, A., and Andreini, L. (2014). Influence of canopy fruit location on morphological, histochemical and biochemical changes in two oil olive cultivars. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*, 148(6):1221–1230.
- Bartz, R., Li, W.-H., Venables, B., Zehmer, J. K., Roth, M. R., Welti, R., Anderson, R. G. W., Liu, P., and Chapman, K. D. (2007). Lipidomics reveals that adiposomes store ether lipids and mediate phospholipid traffic. *Journal of Lipid Research*, 48(4):837–847.
- Bates, P. D. (2016). Understanding the control of acyl flux through the lipid metabolic network of plant oil biosynthesis. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1861(9):1214–1225.
- Bates, P. D., Stymne, S., and Ohlrogge, J. (2013). Biochemical pathways in seed oil synthesis. *Current Opinion in Plant Biology*, 16(3):358–364.
- Baud, S., Dichow, N. R., Kelemen, Z., d'Andréa, S., To, A., Berger, N., Canonge, M., Kronenberger, J., Viterbo, D., Dubreucq, B., Lepiniec, L., Chardot, T., and Miquel, M. (2009). Regulation of HSD1 in seeds of *Arabidopsis thaliana*. *Plant and Cell Physiology*, 50(8):1463–1478.
- Baud, S. and Lepiniec, L. (2010). Physiological and developmental regulation of seed oil production. *Progress in Lipid Research*, 49(3):235–249.

- Beaudoin, F. and Napier, J. (2002). Targeting and membrane-insertion of a sunflower oleosin in vitro and in *Saccharomyces cerevisiae*: the central hydrophobic domain contains more than one signal sequence, and directs oleosin insertion into the endoplasmic reticulum membrane using a signal anchor sequence mechanism. *Planta*, 215(2):293–303.
- Beaudoin, F., Wilkinson, B. M., Stirling, C. J., and Napier, J. A. (2000). In vivo targeting of a sunflower oil body protein in yeast secretory (sec) mutants. *The Plant Journal*, 23(2):159–170.
- Beaudoin, G. A. W. and Facchini, P. J. (2014). Benzylisoquinoline alkaloid biosynthesis in opium poppy. *Planta*, 240(1):19–32.
- Beier, S., Himmelbach, A., Colmsee, C., Zhang, X.-Q., Barrero, R. A., Zhang, Q., Li, L., Bayer, M., Bolser, D., Taudien, S., Groth, M., Felder, M., Hastie, A., Šimková, H., Staňková, H., Vrána, J., Chan, S., Muñoz-Amatriaín, M., Ounit, R., Wanamaker, S., Schmutzer, T., Aliyeva-Schnorr, L., Grasso, S., Tanskanen, J., Sampath, D., Heavens, D., Cao, S., Chapman, B., Dai, F., Han, Y., Li, H., Li, X., Lin, C., McCooke, J. K., Tan, C., Wang, S., Yin, S., Zhou, G., Poland, J. A., Bellgard, M. I., Houben, A., Doležel, J., Ayling, S., Lonardi, S., Langridge, P., Muehlbauer, G. J., Kersey, P., Clark, M. D., Caccamo, M., Schulman, A. H., Platzer, M., Close, T. J., Hansson, M., Zhang, G., Braumann, I., Li, C., Waugh, R., Scholz, U., Stein, N., and Mascher, M. (2017). Construction of a map-based reference genome sequence for barley, *Hordeum vulgare* L. *Scientific Data*, 4(1):170044.
- Beisson, F., Li, Y., Bonaventure, G., Pollard, M., and Ohlrogge, J. B. (2007). The acyltransferase GPAT5 is required for the synthesis of suberin in seed coat and root of *Arabidopsis*. *The Plant Cell*, 19(1):351–368.
- Bennett, M. R., Thompson, M. L., Shepherd, S. A., Dunstan, M. S., Herbert, A. J., Smith, D. R. M., Cronin, V. A., Menon, B. R. K., Levy, C., and Micklefield, J. (2018). Structure and biocatalytic scope of coclaurine *N*-methyltransferase. *Angewandte Chemie*, 130(33):10760–10764.
- Bent, A. (2006). *Arabidopsis thaliana* floral dip transformation method. In *Agrobacterium Protocols*, pages 87–104. Humana Press, Totowa, NJ.
- Berjak, P. (2006). Unifying perspectives of some mechanisms basic to desiccation tolerance across life forms. *Seed Science Research*, 16(1):1–15.
- Berjak, P. and Pammenter, N. W. (2002). Orthodox and recalcitrant seeds. In *Tropical tree seed manual*, pages 137–147. USDA Forest Service, Washington, DC, USA.
- Bersuker, K. and Olzmann, J. A. (2017). Establishing the lipid droplet proteome: Mechanisms of lipid droplet protein targeting and degradation. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1862(10):1166–1177.
- Bertani, G. (1951). Studies on lysogenesis I: the mode of phage liberation by lysogenic *Escherichia coli*. *Journal of Bacteriology*, 62(3):293–300.
- Berthelot, K., Lecomte, S., Estevez, Y., and Peruch, F. (2014). *Hevea brasiliensis* REF (Hev b 1) and SRPP (Hev b 3): An overview on rubber particle proteins. *Biochimie*, 106:1–9.
- Bhatla, S., Kaushik, V., and Yadav, M. (2010). Use of oil bodies and oleosins in recombinant protein production and other biotechnological applications. *Biotechnology Advances*, 28(3):293–300.

- Bigay, J. and Antonny, B. (2012). Curvature, lipid packing, and electrostatics of membrane organelles: Defining cellular territories in determining specificity. *Developmental Cell*, 23(5):886–895.
- Blum, A. (2011). Drought resistance and its improvement. In *Plant Breeding for Water-Limited Environments*, pages 53–152. Springer New York, New York, NY.
- Blée, E., Boachon, B., Burcklen, M., Le Guédard, M., Hanano, A., Heintz, D., Ehltig, J., Herrfurth, C., Feussner, I., and Bessoule, J.-J. (2014). The reductase activity of the *Arabidopsis* caleosin RESPONSIVE TO DESSICATION20 mediates gibberellin-dependent flowering time, abscisic acid sensitivity, and tolerance to oxidative stress. *Plant Physiology*, 166(1):109–124.
- Bohman, S., Staal, J., Thomma, B. P. H. J., Wang, M., and Dixelius, C. (2004). Characterisation of an *Arabidopsis*–*Leptosphaeria maculans* pathosystem: resistance partially requires camalexin biosynthesis and is independent of salicylic acid, ethylene and jasmonic acid signalling. *The Plant Journal*, 37(1):9–20.
- Bohnert, M. (2018). Wrapping up the fats—a structure of the lipid droplet biogenesis protein seipin. *Journal of Cell Biology*, 217(12):4053–4054.
- Bouvier-Nave, P., Berna, A., Noiriél, A., Compagnon, V., Carlsson, A. S., Banas, A., Stymne, S., and Schaller, H. (2010). Involvement of the phospholipid sterol acyltransferase 1 in plant sterol homeostasis and leaf senescence. *Plant Physiology*, 152(1):107–119.
- Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S., Nishihama, R., Nakamura, Y., Berger, F., Adam, C., Aki, S. S., Althoff, F., Araki, T., Arteaga-Vazquez, M. A., Balasubramanian, S., Barry, K., Bauer, D., Boehm, C. R., Briginshaw, L., Caballero-Perez, J., Catarino, B., Chen, F., Chiyoda, S., Chovatia, M., Davies, K. M., Delmans, M., Demura, T., Dierschke, T., Dolan, L., Dorantes-Acosta, A. E., Eklund, D. M., Florent, S. N., Flores-Sandoval, E., Fujiyama, A., Fukuzawa, H., Galik, B., Grimanelli, D., Grimwood, J., Grossniklaus, U., Hamada, T., Haseloff, J., Hetherington, A. J., Higo, A., Hirakawa, Y., Hundley, H. N., Ikeda, Y., Inoue, K., Inoue, S.-i., Ishida, S., Jia, Q., Kakita, M., Kanazawa, T., Kawai, Y., Kawashima, T., Kennedy, M., Kinose, K., Kinoshita, T., Kohara, Y., Koide, E., Komatsu, K., Kopsischke, S., Kubo, M., Kyoizuka, J., Lagercrantz, U., Lin, S.-S., Lindquist, E., Lipzen, A. M., Lu, C.-W., Luna, E. D., Martienssen, R. A., Minamino, N., Mizutani, M., Mizutani, M., Mochizuki, N., Monte, I., Mosher, R., Nagasaki, H., Nakagami, H., Naramoto, S., Nishitani, K., Ohtani, M., Okamoto, T., Okumura, M., Phillips, J., Pollak, B., Reinders, A., Rövekamp, M., Sano, R., Sawa, S., Schmid, M. W., Shirakawa, M., Solano, R., Spunde, A., Suetsugu, N., Sugano, S., Sugiyama, A., Sun, R., Suzuki, Y., Takenaka, M., Takezawa, D., Tomogane, H., Tsuzuki, M., Ueda, T., Umeda, M., Ward, J. M., Watanabe, Y., Yazaki, K., Yokoyama, R., Yoshitake, Y., Yotsui, I., Zachgo, S., and Schmutz, J. (2017). Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell*, 171(2):287–304.e15.
- Bredeson, J. V., Lyons, J. B., Prochnik, S. E., Wu, G. A., Ha, C. M., Edsinger-Gonzales, E., Grimwood, J., Schmutz, J., Rabbi, I. Y., Egesi, C., Nauluvula, P., Lebot, V., Ndunguru, J., Mkamilo, G., Bart, R. S., Setter, T. L., Gleadow, R. M., Kulakow, P., Ferguson, M. E., Rounsley, S., and Rokhsar, D. S. (2016). Sequencing wild and cultivated cassava and related species reveals extensive interspecific hybridization and genetic diversity. *Nature Biotechnology*, 34(5):562–570.
- Brighigna, L., Bennici, A., Tani, C., and Tani, G. (2002). Structural and ultrastructural characterization of *Selaginella lepidophylla*, a desiccation-tolerant plant, during the



- rehydration process. *Flora - Morphology, Distribution, Functional Ecology of Plants*, 197(2):81–91.
- Brocard, L., Immel, F., Coulon, D., Esnay, N., Tuphile, K., Pascal, S., Claverol, S., Fouillen, L., Bessoule, J.-J., and Bréhélin, C. (2017). Proteomic analysis of lipid droplets from *Arabidopsis* aging leaves brings new insight into their biogenesis and functions. *Frontiers in Plant Science*, 8.
- Brown, A. P., Slabas, A. R., and Rafferty, J. B. (2009). Fatty acid biosynthesis in plants—metabolic pathways, structure and organization. *Lipids in Photosynthesis: Essential and Regulatory Functions*, pages 11–34.
- Bréhélin, C. and Nacir, H. (2013). When proteomics reveals unsuspected roles: The plastoglobule example. *Frontiers in Plant Science*, 4.
- Burda, P. and Aebi, M. (1999). The dolichol pathway of N-linked glycosylation. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 1426(2):239–257.
- Cai, Y., Goodman, J. M., Pyc, M., Mullen, R. T., Dyer, J. M., and Chapman, K. D. (2015). *Arabidopsis* seipin proteins modulate triacylglycerol accumulation and influence lipid droplet proliferation. *The Plant Cell*, 27(9):2616–2636.
- Cai, Y., McClinchie, E., Price, A., Nguyen, T. N., Gidda, S. K., Watt, S. C., Yurchenko, O., Park, S., Sturtevant, D., Mullen, R. T., Dyer, J. M., and Chapman, K. D. (2017). Mouse fat storage-inducing transmembrane protein 2 (FIT2) promotes lipid droplet accumulation in plants. *Plant Biotechnology Journal*, 15(7):824–836.
- Callos, J. D., DiRado, M., Xu, B., Behringer, F. J., Link, B. M., and Medford, J. I. (1994). The forever young gene encodes an oxidoreductase required for proper development of the *Arabidopsis* vegetative shoot apex. *The Plant Journal*, 6(6):835–847.
- Carland, F., Fujioka, S., and Nelson, T. (2010). The sterol methyltransferases SMT1, SMT2, and SMT3 influence *Arabidopsis* development through nonbrassinosteroid products. *Plant Physiology*, 153(2):741–756.
- Cartwright, B. R. and Goodman, J. M. (2012). Seipin: from human disease to molecular mechanism. *Journal of Lipid Research*, 53(6):1042–1055.
- Cernac, A. and Benning, C. (2004). WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in *Arabidopsis*. *The Plant Journal*, 40(4):575–585.
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corrêa, M., Da Silva, C., Just, J., Falentin, C., Koh, C. S., Le Clainche, I., Bernard, M., Bento, P., Noel, B., Labadie, K., Alberti, A., Charles, M., Arnaud, D., Guo, H., Daviaud, C., Alamery, S., Jabbari, K., Zhao, M., Edger, P. P., Chelaifa, H., Tack, D., Lassalle, G., Mestiri, I., Schnel, N., Le Paslier, M.-C., Fan, G., Renault, V., Bayer, P. E., Golicz, A. A., Manoli, S., Lee, T.-H., Thi, V. H. D., Chalabi, S., Hu, Q., Fan, C., Tollenaere, R., Lu, Y., Battail, C., Shen, J., Sidebottom, C. H. D., Wang, X., Canaguier, A., Chauveau, A., Bérard, A., Deniot, G., Guan, M., Liu, Z., Sun, F., Lim, Y. P., Lyons, E., Town, C. D., Bancroft, I., Wang, X., Meng, J., Ma, J., Pires, J. C., King, G. J., Brunel, D., Delourme, R., Renard, M., Aury, J.-M., Adams, K. L., Batley, J., Snowdon, R. J., Tost, J., Edwards, D., Zhou, Y., Hua, W., Sharpe, A. G., Paterson, A. H., Guan, C., and Wincker, P. (2014). Plant genetics. Early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. *Science*, 345(6199):950–953.

- Chapman, K. D., Aziz, M., Dyer, J. M., and Mullen, R. T. (2019). Mechanisms of lipid droplet biogenesis. *Biochemical Journal*, 476(13):1929–1942.
- Chapman, K. D., Dyer, J. M., and Mullen, R. T. (2012). Biogenesis and functions of lipid droplets in plants. *Journal of Lipid Research*, 53(2):215–226.
- Chasteen, T. G. and Bentley, R. (2003). Biomethylation of selenium and tellurium: microorganisms and plants. *Chemical reviews*, 103(1):1–26.
- Chaturvedi, P., Ischebeck, T., Egelhofer, V., Lichtscheidl, I., and Weckwerth, W. (2013). Cell-specific analysis of the tomato pollen proteome from pollen mother cell to mature pollen provides evidence for developmental priming. *Journal of Proteome Research*, 12(11):4892–4903.
- Che, N., Yang, Y., Li, Y., Wang, L., Huang, P., Gao, Y., and An, C. (2009). Efficient LEC2 activation of OLEOSIN expression requires two neighboring RY elements on its promoter. *Science in China Series C: Life Sciences*, 52(9):854–863.
- Chen, H., Ruan, J., Chu, P., Fu, W., Liang, Z., Li, Y., Tong, J., Xiao, L., Liu, J., Li, C., et al. (2020a). AtPER1 enhances primary seed dormancy and reduces seed germination by suppressing the ABA catabolism and GA biosynthesis in *Arabidopsis* seeds. *The Plant Journal*, 101(2):310–323.
- Chen, J. C., Tsai, C. C., and Tzen, J. T. (1999). Cloning and secondary structure analysis of caleosin, a unique calcium-binding protein in oil bodies of plant seeds. *Plant and Cell Physiology*, 40(10):1079–1086.
- Chen, Q., Steinhauer, L., Hammerlindl, J., Keller, W., and Zou, J. (2007). Biosynthesis of phytosterol esters: identification of a sterol *O*-acyltransferase in *Arabidopsis*. *Plant Physiology*, 145(3):974–984.
- Chen, Z. J., Sreedasyam, A., Ando, A., Song, Q., De Santiago, L. M., Hulse-Kemp, A. M., Ding, M., Ye, W., Kirkbride, R. C., Jenkins, J., Plott, C., Lovell, J., Lin, Y.-M., Vaughn, R., Liu, B., Simpson, S., Scheffler, B. E., Wen, L., Saski, C. A., Grover, C. E., Hu, G., Conover, J. L., Carlson, J. W., Shu, S., Boston, L. B., Williams, M., Peterson, D. G., McGee, K., Jones, D. C., Wendel, J. F., Stelly, D. M., Grimwood, J., and Schmutz, J. (2020b). Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. *Nature Genetics*, 52(5):525–533.
- Cheng, S., Xian, W., Fu, Y., Marin, B., Keller, J., Wu, T., Sun, W., Li, X., Xu, Y., Zhang, Y., Wittek, S., Reder, T., Günther, G., Gontcharov, A., Wang, S., Li, L., Liu, X., Wang, J., Yang, H., Xu, X., Delaux, P.-M., Melkonian, B., Wong, G. K.-S., and Melkonian, M. (2019). Genomes of subaerial zygnematophyceae provide insights into land plant evolution. *Cell*, 179(5):1057–1067.e14.
- Cheng, Z., Sattler, S., Maeda, H., Sakuragi, Y., Bryant, D. A., and DellaPenna, D. (2003). Highly divergent methyltransferases catalyze a conserved reaction in tocopherol and plastoquinone synthesis in cyanobacteria and photosynthetic organisms eukaryotes. *The Plant Cell*, 15(10):2343–2356.
- Chi, Y. H., Kim, S. Y., Lee, E. S., Jung, Y. J., Park, J. H., Paeng, S. K., Oh, H. T., Melencion, S. M. B., Alinapon, C. V., and Lee, S. Y. (2016). AtSRP1, SMALL RUBBER PARTICLE PROTEIN HOMOLOG, functions in pollen growth and development in *Arabidopsis*. *Biochemical and Biophysical Research Communications*, 475(2):223–229.

- Cho, S. H., Hoang, Q. T., Kim, Y. Y., Shin, H. Y., Ok, S. H., Bae, J. M., and Shin, J. S. (2006). Proteome analysis of gametophores identified a metallothionein involved in various abiotic stress responses in *Physcomitrella patens*. *Plant Cell Reports*, 25(5):475–488.
- Choi, K.-B., Morishige, T., and Sato, F. (2001). Purification and characterization of coclaurine *N*-methyltransferase from cultured *Coptis japonica* cells. *Phytochemistry*, 56(7):649–655.
- Choi, K.-B., Morishige, T., Shitan, N., Yazaki, K., and Sato, F. (2002). Molecular cloning and characterization of coclaurine *N*-methyltransferase from cultured cells of *Coptis japonica*. *Journal of Biological Chemistry*, 277(1):830–835.
- Chorlay, A., Monticelli, L., Verissimo Ferreira, J., Ben M'barek, K., Ajjaji, D., Wang, S., Johnson, E., Beck, R., Omrane, M., Beller, M., Carvalho, P., and Rachid Thiam, A. (2019). Membrane asymmetry imposes directionality on lipid droplet emergence from the ER. *Developmental Cell*, 50(1):25–42.e7.
- Chorlay, A. and Thiam, A. R. (2020). Neutral lipids regulate amphipathic helix affinity for model lipid droplets. *Journal of Cell Biology*, 219(4):e201907099.
- Choudhary, V., Golani, G., Joshi, A. S., Cottier, S., Schneiter, R., Prinz, W. A., and Kozlov, M. M. (2018). Architecture of lipid droplets in endoplasmic reticulum is determined by phospholipid intrinsic curvature. *Current Biology*, 28(6):915–926.e9.
- Choudhary, V., Ojha, N., Golden, A., and Prinz, W. A. (2015). A conserved family of proteins facilitates nascent lipid droplet budding from the ER. *Journal of Cell Biology*, 211(2):261–271.
- Clouse, S. D. and Sasse, J. M. (1998). Brassinosteroids: Essential regulators of plant growth and development. *Annual Review of Plant Physiology and Plant Molecular Biology*, 49(1):427–451.
- Coleman, R. A. (2020). The “discovery” of lipid droplets: A brief history of organelles hidden in plain sight. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1865(9):158762.
- Cosson, P., Lefkir, Y., Démollière, C., and Letourneur, F. (1998). New COP1-binding motifs involved in ER retrieval. *The EMBO journal*, 17(23):6863–6870.
- Costa, M.-C. D., Artur, M. A. S., Maia, J., Jonkheer, E., Derks, M. F. L., Nijveen, H., Williams, B., Mundree, S. G., Jiménez-Gómez, J. M., Hesselink, T., Schijlen, E. G. W. M., Ligterink, W., Oliver, M. J., Farrant, J. M., and Hilhorst, H. W. M. (2017a). A footprint of desiccation tolerance in the genome of *Xerophyta viscosa*. *Nature Plants*, 3(4):1–10.
- Costa, M.-C. D., Cooper, K., Hilhorst, H. W., and Farrant, J. M. (2017b). Orthodox seeds and resurrection plants: Two of a kind? *Plant Physiology*, 175(2):589–599.
- Costa, M.-C. D. and Farrant, J. M. (2019). Plant resistance to abiotic stresses. *Plants*, 8(12):553.
- Costa, M. C. D., Farrant, J. M., Oliver, M. J., Ligterink, W., Buitink, J., and Hilhorst, H. M. (2016). Key genes involved in desiccation tolerance and dormancy across life forms. *Plant Science*, 251:162–168.
- Coulon, D., Brocard, L., Tuphile, K., and Bréhélin, C. (2020). *Arabidopsis* LDIP protein locates at a confined area within the lipid droplet surface and favors lipid droplet formation. *Biochimie*, 169:29–40.

- Cox, J. and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26(12):1367–1372.
- Crombie, L. and Holloway, S. J. (1984). Origins of conjugated triene fatty acids. the biosynthesis of calendic acid by *Calendula officinalis*. *Journal of the Chemical Society, Chemical Communications*, (15):953–955.
- Cui, S., Hu, J., Guo, S., Wang, J., Cheng, Y., Dang, X., Wu, L., and He, Y. (2012). Proteome analysis of *Physcomitrella patens* exposed to progressive dehydration and rehydration. *Journal of Experimental Botany*, 63(2):711–726.
- Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., Donaldson, S., El Houdaigui, B., El Naboulsi, T., Fatima, R., Giron, C. G., Genez, T., Martinez, J. G., Guijarro-Clarke, C., Gymer, A., Hardy, M., Hollis, Z., Hourlier, T., Hunt, T., Juettemann, T., Kaikala, V., Kay, M., Lavidas, I., Le, T., Lemos, D., Marugán, J. C., Mohanan, S., Mushtaq, A., Naven, M., Ogeh, D. N., Parker, A., Parton, A., Perry, M., Piližota, I., Prosovetskaia, I., Sakthivel, M. P., Salam, A. I. A., Schmitt, B. M., Schuilenburg, H., Sheppard, D., Pérez-Silva, J. G., Stark, W., Steed, E., Sutinen, K., Sukumaran, R., Sumathipala, D., Suner, M.-M., Szpak, M., Thormann, A., Tricomi, F. F., Urbina-Gómez, D., Veidenberg, A., Walsh, T. A., Walts, B., Willhoft, N., Winterbottom, A., Wass, E., Chakiachvili, M., Flint, B., Frankish, A., Giorgetti, S., Haggerty, L., Hunt, S. E., Ilsley, G. R., Loveland, J. E., Martin, F. J., Moore, B., Mudge, J. M., Muffato, M., Perry, E., Ruffier, M., Tate, J., Thybert, D., Trevanion, S. J., Dyer, S., Harrison, P. W., Howe, K. L., Yates, A. D., Zerbino, D. R., and Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, 50(D1):D988–D995.
- da Silva Ramos, L. C., Tango, J. S., Savi, A., and Leal, N. R. (1984). Variability for oil and fatty acid composition in castorbean varieties. *Journal of the American Oil Chemists' Society*, 61(12):1841–1843.
- Dadras, A., Fürst-Jansen, J. M. R., Darienko, T., Krone, D., Scholz, P., Rieseberg, T. P., Irisarri, I., Steinkamp, R., Hansen, M., Buschmann, H., Valerius, O., Braus, G. H., Hoecker, U., Mutwil, M., Ischebeck, T., de Vries, S., Lorenz, M., and de Vries, J. (2022). Environmental gradients reveal stress hubs predating plant terrestrialization. preprint, bioRxiv.
- D'Andrea, S. (2016). Lipid droplet mobilization: The different ways to loosen the purse strings. *Biochimie*, 120:17–27.
- d'Andréa, S., Canonge, M., Beopoulos, A., Jolivet, P., Hartmann, M., Miquel, M., Lepiniec, L., and Chardot, T. (2007). At5g50600 encodes a member of the short-chain dehydrogenase reductase superfamily with 11 $\beta$ - and 17 $\beta$ -hydroxysteroid dehydrogenase activities associated with *Arabidopsis thaliana* seed oil bodies. *Biochimie*, 89(2):222–229.
- Dansako, H., Hiramoto, H., Ikeda, M., Wakita, T., and Kato, N. (2014). Rab18 is required for viral assembly of hepatitis C virus through trafficking of the core protein to lipid droplets. *Virology*, 462-463:166–174.
- Davidi, L., Katz, A., and Pick, U. (2012). Characterization of major lipid droplet proteins from *Dunaliella*. *Planta*, 236(1):19–33.
- Davidi, L., Levin, Y., Ben-Dor, S., and Pick, U. (2014a). Proteome analysis of cytoplasmic and plastidic  $\beta$ -carotene lipid droplets in *Dunaliella bardawil*. *Plant Physiology*, 167(1):60–79.

- Davidi, L., Shimoni, E., Khozin-Goldberg, I., Zamir, A., and Pick, U. (2014b). Origin of  $\beta$ -carotene-rich plastoglobuli in *Dunaliella bardawil*. *Plant Physiology*, 164(4):2139–2156.
- De Domenico, S., Tsesmetzis, N., Di Sansebastiano, G. P., Hughes, R. K., Casey, R., and Santino, A. (2007). Subcellular localisation of *Medicago truncatula* 9/13-hydroperoxide lyase reveals a new localisation pattern and activation mechanism for CYP74C enzymes. *BMC Plant Biology*, 7(1):58.
- De Marcos Lousa, C., van Roermund, C. W. T., Postis, V. L. G., Dietrich, D., Kerr, I. D., Wanders, R. J. A., Baldwin, S. A., Baker, A., and Theodoulou, F. L. (2013). Intrinsic acyl-CoA thioesterase activity of a peroxisomal ATP binding cassette transporter is required for transport and metabolism of fatty acids. *Proceedings of the National Academy of Sciences*, 110(4):1279–1284.
- de Silva, N. D., Murmu, J., Chabot, D., Hubbard, K., Ryser, P., Molina, I., and Rowland, O. (2021). Root suberin plays important roles in reducing water loss and sodium uptake in *Arabidopsis thaliana*. *Metabolites*, 11(11):735.
- de Vries, J. and Ischebeck, T. (2020). Ties between stress and lipid droplets pre-date seeds. *Trends in Plant Science*, 25(12):1203–1214.
- Defelice, M. S. (2002). Yellow nutsedge *Cyperus esculentus* L.—snack food of the gods. *Weed Technology*, 16(4):901–907.
- Deruyffelaere, C., Bouchez, I., Morin, H., Guillot, A., Miquel, M., Froissard, M., Chardot, T., and D'Andrea, S. (2015). Ubiquitin-mediated proteasomal degradation of oleosins is involved in oil body mobilization during post-germinative seedling growth in *Arabidopsis*. *Plant and Cell Physiology*, 56(7):1374–1387.
- Deruyffelaere, C., Purkrtova, Z., Bouchez, I., Collet, B., Cacas, J.-L., Chardot, T., Gallois, J.-L., and D'Andrea, S. (2018). PUX10 is a CDC48A adaptor protein that regulates the extraction of ubiquitinated oleosins from seed lipid droplets in *Arabidopsis*. *The Plant Cell*, 30(9):2116–2136.
- Dhiman, R., Caesar, S., Thiam, A. R., and Schrul, B. (2020). Mechanisms of protein targeting to lipid droplets: A unified cell biological and biophysical perspective. *Seminars in Cell & Developmental Biology*, 108:4–13.
- Diener, A. C., Li, H., Zhou, W.-x., Whoriskey, W. J., Nes, W. D., and Fink, G. R. (2000). STEROL METHYLTRANSFERASE 1 Controls the Level of Cholesterol in Plants. *The Plant Cell*, 12(6):853–870.
- Doner, N. M., Seay, D., Mehling, M., Sun, S., Gidda, S. K., Schmitt, K., Braus, G. H., Ischebeck, T., Chapman, K. D., Dyer, J. M., and Mullen, R. T. (2021). *Arabidopsis thaliana* EARLY RESPONSIVE TO DEHYDRATION 7 localizes to lipid droplets via its senescence domain. *Frontiers in Plant Science*, 12.
- Du, Q., Wang, L., Yang, X., Gong, C., and Zhang, D. (2015). Populus endo- $\beta$ -1,4-glucanases gene family: genomic organization, phylogenetic analysis, expression profiles and association mapping. *Planta*, 241(6):1417–1434.
- Dufourc, E. J. (2008). Sterols and membrane dynamics. *Journal of chemical biology*, 1(1-4):63–77.

- Düking, T., Spieth, L., Berghoff, S. A., Piepkorn, L., Schmidke, A. M., Mitkovski, M., Kannaiyan, N., Hosang, L., Scholz, P., Shaib, A. H., Schneider, L. V., Hesse, D., Ruhwedel, T., Sun, T., Linhoff, L., Trevisiol, A., Köhler, S., Pastor, A. M., Misgeld, T., Sereda, M., Hassouna, I., Rossner, M. J., Odoardi, F., Ischebeck, T., de Hoz, L., Hirrlinger, J., Jahn, O., and Saher, G. (2022). Ketogenic diet uncovers differential metabolic plasticity of brain cells. *Science Advances*, 8(37):eabo7639.
- Eastmond, P. J. (2004). Cloning and characterization of the acid lipase from castor beans. *Journal of Biological Chemistry*, 279(44):45540–45545.
- Eastmond, P. J. (2006). SUGAR-DEPENDENT1 encodes a patatin domain triacylglycerol lipase that initiates storage oil breakdown in germinating *Arabidopsis* seeds. *The Plant Cell*, 18(3):665–675.
- Edwards, K., Johnstone, C., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research*, 19(6):1349–1349.
- Edwards, K. D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A. D., Bombarely, A., Allen, F., Hurst, R., White, B., Kernodle, S. P., Bromley, J. R., Sanchez-Tamburrino, J. P., Lewis, R. S., and Mueller, L. A. (2017). A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics*, 18(1):448.
- Entizne, J. C., Guo, W., Calixto, C. P., Spensley, M., Tzioutziou, N., Zhang, R., and Brown, J. W. (2020). TranSuite: a software suite for accurate translation and characterization of transcripts. preprint, bioRxiv.
- Facchini, P. J., Bird, D. A., and St-Pierre, B. (2004). Can *Arabidopsis* make complex alkaloids? *Trends in Plant Science*, 9(3):116–122.
- Falk, T., Herndon, N., Grau, E., Buehler, S., Richter, P., Zaman, S., Baker, E. M., Ramnath, R., Ficklin, S., Staton, M., Feltus, F. A., Jung, S., Main, D., and Wegrzyn, J. L. (2018). Growing and cultivating the forest genomics database, TreeGenes. *Database*, 2018:bay084.
- Fan, J., Yan, C., Roston, R., Shanklin, J., and Xu, C. (2014). *Arabidopsis* lipins, PDAT1 acyltransferase, and SDP1 triacylglycerol lipase synergistically direct fatty acids toward  $\beta$ -oxidation, thereby maintaining membrane lipid homeostasis. *The Plant Cell*, 26(10):4119–4134.
- Fan, J., Yan, C., and Xu, C. (2013). Phospholipid:diacylglycerol acyltransferase-mediated triacylglycerol biosynthesis is crucial for protection against fatty acid-induced cell death in growing tissues of *Arabidopsis*. *The Plant Journal*, 76(6):930–942.
- Fan, J., Yu, L., and Xu, C. (2017). A central role for triacylglycerol in membrane lipid breakdown, fatty acid  $\beta$ -oxidation, and plant survival under extended darkness. *Plant Physiology*, 174(3):1517–1530.
- Fan, J., Yu, L., and Xu, C. (2019). Dual role for autophagy in lipid metabolism in *Arabidopsis*. *The Plant Cell*, 31(7):1598–1613.
- Farese, R. V. and Walther, T. C. (2009). Lipid droplets finally get a little R-E-S-P-E-C-T. *Cell*, 139(5):855–860.
- Farmer, E. E. and Mueller, M. J. (2013). ROS-mediated lipid peroxidation and RES-activated signaling. *Annual Review of Plant Biology*, 64:429–450.

- Farrant, J. M. and Hilhorst, H. (2022). Crops for dry environments. *Current Opinion in Biotechnology*, 74:84–91.
- Fazio, G. C., Xu, R., and Matsuda, S. P. (2004). Genome mining to identify new plant triterpenoids. *Journal of the American Chemical Society*, 126(18):5678–5679.
- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39(4):783–791.
- Fernandez-Pozo, N., Haas, F. B., Meyberg, R., Ullrich, K. K., Hiss, M., Perroud, P.-F., Hanke, S., Kratz, V., Powell, A. F., Vesty, E. F., Daum, C. G., Zane, M., Lipzen, A., Sreedasyam, A., Grimwood, J., Coates, J. C., Barry, K., Schmutz, J., Mueller, L. A., and Rensing, S. A. (2020). PEATmoss (Physcomitrella Expression Atlas Tool): a unified gene expression atlas for the model plant *Physcomitrella patens*. *The Plant Journal*, 102(1):165–177.
- Fernández-Santos, R., Izquierdo, Y., López, A., Muñoz, L., Martínez, M., Cascón, T., Hamberg, M., and Castresana, C. (2020). Protein profiles of lipid droplets during the hypersensitive defense response of *Arabidopsis* against *Pseudomonas* infection. *Plant and Cell Physiology*, 61(6):1144–1157.
- Feussner, I., Balkenhohl, T. J., Porzel, A., Kühn, H., and Wasternack, C. (1997). Structural elucidation of oxygenated storage lipids in cucumber cotyledons: Implication of lipid body lipoyxygenase in lipid mobilization during germination. *Journal of Biological Chemistry*, 272(34):21635–21641.
- Feussner, I., Hause, B., Nellen, A., Wasternack, C., and Kindl, H. (1996). Lipid-body lipoyxygenase is expressed in cotyledons during germination prior to other lipoyxygenase forms. *Planta*, 198(2):288–293.
- Feussner, I., Kühn, H., and Wasternack, C. (2001). Lipoyxygenase-dependent degradation of storage lipids. *Trends in Plant Science*, 6(6):268–273.
- Feussner, K. and Feussner, I. (2020). *Ex vivo* metabolomics: A powerful approach for functional gene annotation. *Trends in Plant Science*, 25(8):829–830.
- Feußner, I. and Kindl, H. (1992). A lipoyxygenase is the main lipid body protein in cucumber and soybean cotyledons during the stage of triglyceride mobilization. *FEBS Letters*, 298(2-3):223–225.
- Field, B., Fiston-Lavier, A.-S., Kemen, A., Geisler, K., Quesneville, H., and Osbourn, A. E. (2011). Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proceedings of the National Academy of Sciences*, 108(38):16116–16121.
- Filiault, D. L., Ballerini, E. S., Mandáková, T., Aköz, G., Derieg, N. J., Schmutz, J., Jenkins, J., Grimwood, J., Shu, S., Hayes, R. D., Hellsten, U., Barry, K., Yan, J., Mihaltcheva, S., Karafiátová, M., Nizhynska, V., Kramer, E. M., Lysak, M. A., Hodges, S. A., and Nordborg, M. (2018). The *Aquilegia* genome provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a unique history. *eLife*, 7:e36426.
- Flügge, U.-I., Hausler, R. E., Ludewig, F., and Gierth, M. (2011). The role of transporters in supplying energy to plant plastids. *Journal of Experimental Botany*, 62(7):2381–2392.
- Fuchs, C., Vine, N., and Hills, M. J. (1996). Purification and characterization of the acid lipase from the endosperm of castor oil seeds. *Journal of Plant Physiology*, 149(1):23–29.

- Fulda, M., Schnurr, J., Abbadi, A., Heinz, E., and Browse, J. (2004). Peroxisomal acyl-CoA synthetase activity is essential for seedling development in *Arabidopsis thaliana*. *The Plant Cell*, 16(2):394–405.
- Furse, S., Liddell, S., Ortori, C. A., Williams, H., Neylon, D. C., Scott, D. J., Barrett, D. A., and Gray, D. A. (2013). The lipidome and proteome of oil bodies from *Helianthus annuus* (common sunflower). *Journal of Chemical Biology*, 6(2):63–76.
- Fürst-Jansen, J. M. R., de Vries, S., and de Vries, J. (2020). Evo-physio: on stress responses and the earliest land plants. *Journal of Experimental Botany*, 71(11):3254–3269.
- Gaff, D. F. and Oliver, M. (2013). The evolution of desiccation tolerance in angiosperm plants: a rare yet common phenomenon. *Functional Plant Biology*, 40(4):315–328.
- Gana, R., Rao, S., Huang, H., Wu, C., and Vasudevan, S. (2013). Structural and functional studies of *S*-adenosyl-*L*-methionine binding proteins: a ligand-centric approach. *BMC Structural Biology*, 13(1).
- Gao, C., Cai, Y., Wang, Y., Kang, B.-H., Aniento, F., Robinson, D. G., and Jiang, L. (2014). Retention mechanisms for ER and Golgi membrane proteins. *Trends in Plant Science*, 19(8):508–515.
- Garbowicz, K., Liu, Z., Alseekh, S., Tieman, D., Taylor, M., Kuhalskaya, A., Ofner, I., Zamir, D., Klee, H. J., Fernie, A. R., and Brotman, Y. (2018). Quantitative trait loci analysis identifies a prominent gene involved in the production of fatty acid-derived flavor volatiles in tomato. *Molecular Plant*, 11(9):1147–1165.
- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. In Walker, J. M., editor, *The Proteomics Protocols Handbook*, Springer Protocols Handbooks, pages 571–607. Humana Press, Totowa, NJ.
- Ge, S., Zhang, R.-X., Wang, Y.-F., Sun, P., Chu, J., Li, J., Sun, P., Wang, J., Hetherington, A. M., and Liang, Y.-K. (2022). The *Arabidopsis* Rab protein RABC1 affects stomatal development by regulating lipid droplet dynamics. *The Plant Cell*, 34(11):4274–4292.
- Gidda, S. K., Park, S., Pyc, M., Yurchenko, O., Cai, Y., Wu, P., Andrews, D. W., Chapman, K. D., Dyer, J. M., and Mullen, R. T. (2016). Lipid droplet-associated proteins (LDAPs) are required for the dynamic regulation of neutral lipid compartmentation in plant cells. *Plant Physiology*, 170(4):2052–2071.
- Gidda, S. K., Shockey, J. M., Rothstein, S. J., Dyer, J. M., and Mullen, R. T. (2009). *Arabidopsis thaliana* GPAT8 and GPAT9 are localized to the ER and possess distinct ER retrieval signals: functional divergence of the dilysine ER retrieval motif in plant cells. *Plant Physiology and Biochemistry*, 47(10):867–879.
- Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annual Review of Phytopathology*, 43:205.
- Gómez-Garay, A., Gabriel y Galán, J. M., Cabezuelo, A., Pintos, B., Prada, C., and Martín, L. (2018). Ecological significance of brassinosteroids in three temperate ferns. In Fernández, H., editor, *Current Advances in Fern Research*, pages 453–466. Springer International Publishing, Cham.



- González-García, M.-P., Vilarrasa-Blasi, J., Zhiponova, M., Divol, F., Mora-García, S., Russinova, E., and Caño-Delgado, A. I. (2011). Brassinosteroids control meristem size by promoting cell cycle progression in *Arabidopsis* roots. *Development*, 138(5):849–859.
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, 40(D1):D1178–D1186.
- Graham, I. A. (2008). Seed storage oil mobilization. *Annual Review of Plant Biology*, 59(1):115–142.
- Greer, M. S., Cai, Y., Gidda, S. K., Esnay, N., Kretzschmar, F. K., Seay, D., McClinchie, E., Ischebeck, T., Mullen, R. T., Dyer, J. M., and Chapman, K. D. (2020). Seipin isoforms interact with the membrane-tethering protein VAP27-1 for lipid droplet formation. *The Plant Cell*, 32(9):2932–2950.
- Grigoriev, I. V., Hayes, R. D., Calhoun, S., Kamel, B., Wang, A., Ahrendt, S., Dusheyko, S., Nikitin, R., Mondo, S. J., Salamov, A., Shabalov, I., and Kuo, A. (2021). PhycoCosm, a comparative algal genomics resource. *Nucleic Acids Research*, 49(D1):D1004–D1011.
- Grillitsch, K., Connerth, M., Köfeler, H., Arrey, T. N., Rietschel, B., Wagner, B., Karas, M., and Daum, G. (2011). Lipid particles/droplets of the yeast *Saccharomyces cerevisiae* revisited: Lipidome meets proteome. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1811(12):1165–1176.
- Guo, L., Winzer, T., Yang, X., Li, Y., Ning, Z., He, Z., Teodor, R., Lu, Y., Bowser, T. A., Graham, I. A., and Ye, K. (2018). The opium poppy genome and morphinan production. *Science*, 362(6412):343–347.
- Guzha, A., Whitehead, P., Ischebeck, T., and Chapman, K. D. (2023). Lipid droplets: Packing hydrophobic molecules within the aqueous cytoplasm. *Annual Review of Plant Biology*, 74(1):195–223.
- Hagel, J. M. and Facchini, P. J. (2012). Subcellular localization of sanguinarine biosynthetic enzymes in cultured opium poppy cells. *In Vitro Cellular & Developmental Biology - Plant*, 48(2):233–240.
- Hamada, S., Kishikawa, A., and Yoshida, M. (2020). Proteomic analysis of lipid droplets in *Sesamum indicum*. *The Protein Journal*, 39(4):366–376.
- Hanano, A., Bessoule, J.-J., Heitz, T., and Blée, E. (2015). Involvement of the caleosin/peroxygenase RD20 in the control of cell death during *Arabidopsis* responses to pathogens. *Plant Signaling & Behavior*, 10(4):e991574.
- Hanano, A., Blée, E., and Murphy, D. J. (2023). Caleosin/peroxygenases: multifunctional proteins in plants. *Annals of Botany*, 131(3):387–409.
- Hanano, A., Burcklen, M., Flenet, M., Ivancich, A., Louwagie, M., Garin, J., and Blée, E. (2006). Plant seed peroxygenase is an original heme-oxygenase with an EF-hand calcium binding motif. *Journal of Biological Chemistry*, 281(44):33140–33151.
- Haslekås, C., Stacy, R. A., Nygaard, V., Culiáñez-Macià, F. A., and Aalen, R. B. (1998). The expression of a peroxiredoxin antioxidant gene, AtPer1, in *Arabidopsis thaliana* is seed-specific and related to dormancy. *Plant Molecular Biology*, 36:833–845.

- Hause, B., Weichert, H., Höhne, M., Kindl, H., and Feussner, I. (2000). Expression of cucumber lipid-body lipooxygenase in transgenic tobacco: lipid-body lipooxygenase is correctly targeted to seed lipid bodies. *Planta*, 210(5):708–714.
- Hayashi, H., De Bellis, L., Hayashi, Y., Nito, K., Kato, A., Hayashi, M., Hara-Nishimura, I., and Nishimura, M. (2002). Molecular characterization of an *Arabidopsis* acyl-coenzyme A synthetase localized on glyoxysomal membranes. *Plant Physiology*, 130(4):2019–2026.
- Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P., and Minghim, R. (2015). InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics*, 16:1–7.
- Hernández, M. L., Whitehead, L., He, Z., Gazda, V., Gilday, A., Kozhevnikova, E., Vaistij, F. E., Larson, T. R., and Graham, I. A. (2012). A cytosolic acyltransferase contributes to triacylglycerol synthesis in sucrose-rescued *Arabidopsis* seed oil catabolism mutants. *Plant Physiology*, 160(1):215–225.
- Higashi, Y., Okazaki, Y., Myouga, F., Shinozaki, K., and Saito, K. (2015). Landscape of the lipidome and transcriptome under heat stress in *Arabidopsis thaliana*. *Scientific Reports*, 5(1):10533.
- Hilhorst, H. W., Costa, M.-C. D., and Farrant, J. M. (2018). A footprint of plant desiccation tolerance: Does it exist? *Molecular Plant*, 11(8):1003–1005.
- Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, 35(2):518–522.
- Hoekstra, F., Golovina, E., and Buitink, J. (2001). Mechanisms of desiccation tolerance. *Trends in Plant Science*, 6:431–438.
- Hopkin, M. (2005). Climate change: world round-up. *Nature*.
- Horbowicz, M. and Obendorf, R. L. (1992). Changes in sterols and fatty acids of buckwheat endosperm and embryo during seed development. *Journal of Agricultural and Food Chemistry*, 40(5):745–750.
- Hori, K., Maruyama, F., Fujisawa, T., Togashi, T., Yamamoto, N., Seo, M., Sato, S., Yamada, T., Mori, H., Tajima, N., Moriyama, T., Ikeuchi, M., Watanabe, M., Wada, H., Kobayashi, K., Saito, M., Masuda, T., Sasaki-Sekimoto, Y., Mashiguchi, K., Awai, K., Shimojima, M., Masuda, S., Iwai, M., Nobusawa, T., Narise, T., Kondo, S., Saito, H., Sato, R., Murakawa, M., Ihara, Y., Oshima-Yamada, Y., Ohtaka, K., Satoh, M., Sonobe, K., Ishii, M., Ohtani, R., Kanamori-Sato, M., Honoki, R., Miyazaki, D., Mochizuki, H., Umetsu, J., Higashi, K., Shibata, D., Kamiya, Y., Sato, N., Nakamura, Y., Tabata, S., Ida, S., Kurokawa, K., and Ohta, H. (2014). *Klebsormidium flaccidum* genome reveals primary factors for plant terrestrial adaptation. *Nature Communications*, 5(1):3978.
- Hori, K., Yamada, Y., Purwanto, R., Minakuchi, Y., Toyoda, A., Hirakawa, H., and Sato, F. (2018). Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in california poppy using a draft genome sequence. *Plant and Cell Physiology*, 59(2):222–233.
- Horn, P. J., Chapman, K. D., and Ischebeck, T. (2021). Isolation of lipid droplets for protein and lipid analysis. In Bartels, D. and Dörmann, P., editors, *Plant Lipids: Methods and Protocols*, Methods in Molecular Biology, pages 295–320. Springer US, New York, NY.

- Horn, P. J., James, C. N., Gidda, S. K., Kilaru, A., Dyer, J. M., Mullen, R. T., Ohlrogge, J. B., and Chapman, K. D. (2013). Identification of a new class of lipid droplet-associated proteins in plants. *Plant Physiology*, 162(4):1926–1936.
- Hsiao, E. S. and Tzen, J. T. (2011). Ubiquitination of oleosin-H and caleosin in sesame oil bodies after seed germination. *Plant Physiology and Biochemistry*, 49(1):77–81.
- Hu, T. T., Pattyn, P., Bakker, E. G., Cao, J., Cheng, J.-F., Clark, R. M., Fahlgren, N., Fawcett, J. A., Grimwood, J., Gundlach, H., Haberer, G., Hollister, J. D., Ossowski, S., Ottillar, R. P., Salamov, A. A., Schneeberger, K., Spannagl, M., Wang, X., Yang, L., Nasrallah, M. E., Bergelson, J., Carrington, J. C., Gaut, B. S., Schmutz, J., Mayer, K. F. X., Van de Peer, Y., Grigoriev, I. V., Nordborg, M., Weigel, D., and Guo, Y.-L. (2011). The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nature Genetics*, 43(5):476–481.
- Huang, A. C., Jiang, T., Liu, Y.-X., Bai, Y.-C., Reed, J., Qu, B., Goossens, A., Nützmann, H.-W., Bai, Y., and Osbourn, A. (2019). A specialized metabolic network selectively modulates *Arabidopsis* root microbiota. *Science*, 364(6440):eaau6389.
- Huang, A. H. (2018). Plant lipid droplets and their associated proteins: Potential for rapid advances. *Plant Physiology*, 176(3):1894–1918.
- Huang, A. H. C. (1992). Oil bodies and oleosins in seeds. *Annual Review of Plant Biology*, 43(1):177–200.
- Huang, C.-c., Smith, C. V., Glickman, M. S., Jacobs, W. R., and Sacchettini, J. C. (2002). Crystal structures of mycolic acid cyclopropane synthases from *Mycobacterium tuberculosis*. *Journal of Biological Chemistry*, 277(13):11559–11569.
- Huang, C.-Y., Chung, C.-I., Lin, Y.-C., Hsing, Y.-I. C., and Huang, A. H. (2009). Oil bodies and oleosins in *Physcomitrella* possess characteristics representative of early trends in evolution. *Plant Physiology*, 150(3):1192–1203.
- Huang, C.-Y. and Huang, A. H. (2017). Unique motifs and length of hairpin in oleosin target the cytosolic side of endoplasmic reticulum and budding lipid droplet. *Plant Physiology*, 174(4):2248–2260.
- Huang, L.-S. and Grunwald, C. (1988). Sterol and phospholipid changes during alfalfa seed germination. *Phytochemistry*, 27(7):2049–2053.
- Huang, N.-L., Huang, M.-D., Chen, T.-L. L., and Huang, A. H. (2013). Oleosin of subcellular lipid droplets evolved in green algae. *Plant Physiology*, 161(4):1862–1874.
- Huizinga, D. H., Omosogbon, O., Omery, B., and Crowell, D. N. (2008). Isoprenylcysteine methylation and demethylation regulate abscisic acid signaling in *Arabidopsis*. *The Plant Cell*, 20(10):2714–2728.
- Hétu, M.-F., Tremblay, L. J., and Lefebvre, D. D. (2005). High root biomass production in anchored *Arabidopsis* plants grown in axenic sucrose supplemented liquid culture. *BioTechniques*, 39(3):345–349.
- Hözl, G. and Dörmann, P. (2019). Chloroplast lipids and their biosynthesis. *Annual Review of Plant Biology*, 70(1):51–81.

- Iorizzo, M., Ellison, S., Senalik, D., Zeng, P., Satapoomin, P., Huang, J., Bowman, M., Iovene, M., Sanseverino, W., Cavagnaro, P., Yildiz, M., Macko-Podgórní, A., Moranska, E., Grzebelus, E., Grzebelus, D., Ashrafi, H., Zheng, Z., Cheng, S., Spooner, D., Van Deynze, A., and Simon, P. (2016). A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. *Nature Genetics*, 48(6):657–666.
- Ischebeck, T. (2016). Lipids in pollen — They are different. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1861(9):1315–1328.
- Ischebeck, T., Krawczyk, H. E., Mullen, R. T., Dyer, J. M., and Chapman, K. D. (2020). Lipid droplets in plants and algae: Distribution, formation, turnover and function. *Seminars in Cell & Developmental Biology*, 108:82–93.
- Islam, M. S., Saito, J. A., Emdad, E. M., Ahmed, B., Islam, M. M., Halim, A., Hossen, Q. M. M., Hossain, M. Z., Ahmed, R., Hossain, M. S., Kabir, S. M. T., Khan, M. S. A., Khan, M. M., Hasan, R., Aktar, N., Honi, U., Islam, R., Rashid, M. M., Wan, X., Hou, S., Haque, T., Azam, M. S., Moosa, M. M., Elias, S. M., Hasan, A. M. M., Mahmood, N., Shafiuddin, M., Shahid, S., Shommu, N. S., Jahan, S., Roy, S., Chowdhury, A., Akhand, A. I., Nisho, G. M., Uddin, K. S., Rabeya, T., Hoque, S. M. E., Snigdha, A. R., Mortoza, S., Matin, S. A., Islam, M. K., Lashkar, M. Z. H., Zaman, M., Yuryev, A., Uddin, M. K., Rahman, M. S., Haque, M. S., Alam, M. M., Khan, H., and Alam, M. (2017). Comparative genomics of two jute species and insight into fibre biogenesis. *Nature Plants*, 3:16223.
- Jacquier, N., Choudhary, V., Mari, M., Toulmay, A., Reggiori, F., and Schneiter, R. (2011). Lipid droplets are functionally connected to the endoplasmic reticulum in *Saccharomyces cerevisiae*. *Journal of Cell Science*, 124(14):2424–2437.
- Jasieniecka-Gazarkiewicz, K., Demski, K., Gidda, S. K., Klińska, S., Niedojadło, J., Lager, I., Carlsson, A. S., Minina, E. A., Mullen, R. T., Bozhkov, P. V., Stymne, S., and Banaś, A. (2021). Subcellular localization of acyl-CoA: Lysophosphatidylethanolamine acyltransferases (LPEATs) and the effects of knocking-out and overexpression of their genes on autophagy markers level and life span of *A. thaliana*. *International Journal of Molecular Sciences*, 22(6):3006.
- Jasieniecka-Gazarkiewicz, K., Demski, K., Lager, I., Stymne, S., and Banaś, A. (2016). Possible role of different yeast and plant lysophospholipid: acyl-CoA acyltransferases (LPLATs) in acyl remodelling of phospholipids. *Lipids*, 51:15–23.
- Jasieniecka-Gazarkiewicz, K., Lager, I., Carlsson, A. S., Gutbrod, K., Peisker, H., Dörmann, P., Stymne, S., and Banaś, A. (2017). Acyl-CoA: Lysophosphatidylethanolamine acyltransferase activity regulates growth of *Arabidopsis*. *Plant Physiology*, 174(2):986–998.
- Javee, A., Sulochana, S. B., Pallisery, S. J., and Arumugam, M. (2016). Major lipid body protein: A conserved structural component of lipid body accumulated during abiotic stress in *S. quadricauda casa-cc202*. *Frontiers in Energy Research*, 4.
- Ji, H., Liu, D., and Yang, Z. (2021). High oil accumulation in tuber of yellow nutsedge compared to purple nutsedge is associated with more abundant expression of genes involved in fatty acid synthesis and triacylglycerol storage. *Biotechnology for Biofuels*, 14(1):54.
- Jiang, P.-L. and Tzen, J. T. (2010). Caleosin serves as the major structural protein as efficient as oleosin on the surface of seed oil bodies. *Plant Signaling & Behavior*, 5(4):447–449.
- Jiang, P.-L., Wang, C.-S., Hsu, C.-M., Jauh, G.-Y., and Tzen, J. T. C. (2007). Stable oil bodies sheltered by a unique oleosin in lily pollen. *Plant and Cell Physiology*, 48(6):812–821.

- Jolivet, P., Acevedo, F., Boulard, C., d'Andréa, S., Faure, J.-D., Kohli, A., Nesi, N., Valot, B., and Chardot, T. (2013). Crop seed oil bodies: From challenges in protein identification to an emerging picture of the oil body proteome. *PROTEOMICS*, 13(12-13):1836–1849.
- Jolivet, P., Aymé, L., Giuliani, A., Wien, F., Chardot, T., and Gohon, Y. (2017). Structural proteomics: Topology and relative accessibility of plant lipid droplet associated proteins. *Journal of Proteomics*, 169:87–98.
- Jolivet, P., Boulard, C., Bellamy, A., Larré, C., Barre, M., Rogniaux, H., d'Andréa, S., Chardot, T., and Nesi, N. (2009). Protein composition of oil bodies from mature *Brassica napus* seeds. *PROTEOMICS*, 9(12):3268–3284.
- Jolivet, P., Boulard, C., Bellamy, A., Valot, B., d'Andréa, S., Zivy, M., Nesi, N., and Chardot, T. (2011). Oil body proteins sequentially accumulate throughout seed development in *Brassica napus*. *Journal of Plant Physiology*, 168(17):2015–2020.
- Jolivet, P., Roux, E., d'Andrea, S., Davanture, M., Negroni, L., Zivy, M., and Chardot, T. (2004). Protein composition of oil bodies in *Arabidopsis thaliana* ecotype WS. *Plant Physiology and Biochemistry*, 42(6):501–509.
- Joshi, A. S., Nebenfuehr, B., Choudhary, V., Satpute-Krishnan, P., Levine, T. P., Golden, A., and Prinz, W. A. (2018). Lipid droplet and peroxisome biogenesis occur at the same ER subdomains. *Nature Communications*, 9(1):2940.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873):583–589.
- Kaever, A., Lingner, T., Feussner, K., Göbel, C., Feussner, I., and Meinicke, P. (2009). MarVis: a tool for clustering and visualization of metabolic biomarkers. *BMC Bioinformatics*, 10(1).
- Kagale, S., Koh, C., Nixon, J., Bollina, V., Clarke, W. E., Tuteja, R., Spillane, C., Robinson, S. J., Links, M. G., Clarke, C., Higgins, E. E., Huebert, T., Sharpe, A. G., and Parkin, I. A. P. (2014). The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. *Nature communications*, 5:3706.
- Kalinowska, M. and Wojciechowski, Z. A. (1984). Sterol conjugate interconversions during germination of white mustard (*Sinapis alba*). *Phytochemistry*, 23(11):2485–2488.
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6):587–589.
- Kang, B.-H., Anderson, C. T., Arimura, S.-i., Bayer, E., Bezanilla, M., Botella, M. A., Brandizzi, F., Burch-Smith, T. M., Chapman, K. D., Dünser, K., Gu, Y., Jaillais, Y., Kirchhoff, H., Otegui, M. S., Rosado, A., Tang, Y., Kleine-Vehn, J., Wang, P., and Zolman, B. K. (2022). A glossary of plant cell structures: Current insights and future questions. *The Plant Cell*, 34(1):10–52.
- Katavic, V., Agrawal, G. K., Hajduch, M., Harris, S. L., and Thelen, J. J. (2006). Protein and lipid composition analysis of oil bodies from two *Brassica napus* cultivars. *PROTEOMICS*, 6(16):4586–4598.
- Katoh, K. and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4):772–780.

- Katz, J. E., Dlakić, M., and Clarke, S. (2003). Automated identification of putative methyltransferases from genomic open reading frames. *Molecular & Cellular Proteomics*, 2(8):525–540.
- Kelly, A. A., Quettier, A.-L., Shaw, E., and Eastmond, P. J. (2011). Seed storage oil mobilization is important but not essential for germination or seedling establishment in *Arabidopsis*. *Plant Physiology*, 157(2):866–875.
- Kelly, A. A., van Erp, H., Quettier, A.-L., Shaw, E., Menard, G., Kurup, S., and Eastmond, P. J. (2013). The SUGAR-DEPENDENT1 lipase limits triacylglycerol accumulation in vegetative tissues of *Arabidopsis*. *Plant Physiology*, 162(3):1282–1289.
- Khandelwal, A., Cho, S. H., Marella, H., Sakata, Y., Perroud, P.-F., Pan, A., and Quatrano, R. S. (2010). Role of ABA and ABI3 in desiccation tolerance. *Science*, 327(5965):546–546.
- Kim, E. Y., Park, K. Y., Seo, Y. S., and Kim, W. T. (2016a). *Arabidopsis* small rubber particle protein homolog SRPs play dual roles as positive factors for tissue growth and development and in drought stress responses. *Plant Physiology*, 170(4):2494–2510.
- Kim, H. U., Hsieh, K., Ratnayake, C., and Huang, A. H. C. (2002). A novel group of oleosins is found within the pollen of *Arabidopsis*. *Journal of Biological Chemistry*, 277(25):22677–22684.
- Kim, R. J., Kim, H. J., Shim, D., and Suh, M. C. (2016b). Molecular and biochemical characterizations of the monoacylglycerol lipase gene family of *Arabidopsis thaliana*. *The Plant Journal*, 85(6):758–771.
- Kirch, H.-H., Bartels, D., Wei, Y., Schnable, P. S., and Wood, A. J. (2004). The ALDH gene superfamily of *Arabidopsis*. *Trends in Plant Science*, 9(8):371–377.
- Kitamura, T., Takagi, S., Naganuma, T., and Kihara, A. (2015). Mouse aldehyde dehydrogenase ALDH3B2 is localized to lipid droplets via two C-terminal tryptophan residues and lipid modification. *Biochemical Journal*, 465(1):79–87.
- Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D., and Penin, A. A. (2016). A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *The Plant Journal*, 88(6):1058–1070.
- Kong, Q., Yang, Y., Guo, L., Yuan, L., and Ma, W. (2020). Molecular basis of plant oil biosynthesis: Insights gained from studying the WRINKLED1 transcription factor. *Frontiers in Plant Science*, 11:24.
- Kooyers, N. J. (2015). The evolution of drought escape and avoidance in natural herbaceous populations. *Plant Science*, 234:155–162.
- Kory, N., Farese, R. V., and Walther, T. C. (2016). Targeting fat: Mechanisms of protein localization to lipid droplets. *Trends in Cell Biology*, 26(7):535–546.
- Koster, K. L., Balsamo, R. A., Espinoza, C., and Oliver, M. J. (2010). Desiccation sensitivity and tolerance in the moss *Physcomitrella patens*: assessing limits and damage. *Plant Growth Regulation*, 62(3):293–302.
- Kourelis, J., Kaschani, F., Grosse-Holz, F. M., Homma, F., Kaiser, M., and van der Hoorn, R. A. L. (2019). A homology-guided, genome-based proteome for improved proteomics in the allopolyploid *Nicotiana benthamiana*. *BMC Genomics*, 20(1):722.

- Kranner, I. and Birtić, S. (2005). The modulating role of antioxidants in desiccation tolerance. *Integrative and Comparative Biology*, 45(5):734–740.
- Krawczyk, H. E., Rotsch, A. H., Herrfurth, C., Scholz, P., Shomroni, O., Salinas-Riester, G., Feussner, I., and Ischebeck, T. (2022a). Heat stress leads to rapid lipid remodeling and transcriptional adaptations in *Nicotiana tabacum* pollen tubes. *Plant Physiology*, 189(2):490–515.
- Krawczyk, H. E., Sun, S., Doner, N. M., Yan, Q., Lim, M. S. S., Scholz, P., Niemeyer, P. W., Schmitt, K., Valerius, O., Pleskot, R., Hillmer, S., Braus, G. H., Wiermer, M., Mullen, R. T., and Ischebeck, T. (2022b). SEED LIPID DROPLET PROTEIN1, SEED LIPID DROPLET PROTEIN2, and LIPID DROPLET PLASMA MEMBRANE ADAPTOR mediate lipid droplet–plasma membrane tethering. *The Plant Cell*, 34(6):2424–2448.
- Kretzschmar, F. K., Doner, N. M., Krawczyk, H. E., Scholz, P., Schmitt, K., Valerius, O., Braus, G. H., Mullen, R. T., and Ischebeck, T. (2020). Identification of low-abundance lipid droplet proteins in seeds and seedlings. *Plant Physiology*, 182(3):1326–1345.
- Kretzschmar, F. K., Mengel, L. A., Müller, A. O., Schmitt, K., Blersch, K. F., Valerius, O., Braus, G. H., and Ischebeck, T. (2018). PUX10: A lipid droplet-localized scaffold protein that interacts with CELL DIVISION CYCLE48 and is involved in the degradation of lipid droplet proteins. *The Plant Cell*, 30(9):2137–2160.
- Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology*, 305(3):567–580.
- Kunz, H.-H., Scharnewski, M., Feussner, K., Feussner, I., Flügge, U.-I., Fulda, M., and Gierth, M. (2009). The ABC transporter PXA1 and peroxisomal  $\beta$ -oxidation are vital for metabolism in mature leaves of *Arabidopsis* during extended darkness. *The Plant Cell*, 21(9):2733–2749.
- Kurusu, T., Koyano, T., Hanamata, S., Kubo, T., Noguchi, Y., Yagi, C., Nagata, N., Yamamoto, T., Ohnishi, T., Okazaki, Y., Kitahata, N., Ando, D., Ishikawa, M., Wada, S., Miyao, A., Hirochika, H., Shimada, H., Makino, A., Saito, K., Ishida, H., Kinoshita, T., Kurata, N., and Kuchitsu, K. (2014). OsATG7 is required for autophagy-dependent lipid metabolism in rice postmeiotic anther development. *Autophagy*, 10(5):878–888.
- Kyte, J. and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157(1):105–132.
- Laibach, N., Hillebrand, A., Twyman, R. M., Prüfer, D., and Schulze Gronover, C. (2015). Identification of a *Taraxacum brevicorniculatum* rubber elongation factor protein that is localized on rubber particles and promotes rubber biosynthesis. *The Plant Journal*, 82(4):609–620.
- Laibach, N., Schmidl, S., Müller, B., Bergmann, M., Prüfer, D., and Schulze Gronover, C. (2018). Small rubber particle proteins from *Taraxacum brevicorniculatum* promote stress tolerance and influence the size and distribution of lipid droplets and artificial poly( *cis* -1,4-isoprene) bodies. *The Plant Journal*, 93(6):1045–1061.
- Lamberti, C., Nebbia, S., Balestrini, R., Marengo, E., Manfredi, M., Pavese, V., Cirrincione, S., Giuffrida, M. G., Cavallarin, L., Acquadro, A., and Abbà, S. (2020). Identification of a caleosin associated with hazelnut (*Corylus avellana* L.) oil bodies. *Plant Biology*, 22(3):404–409.

- 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., and Huala, E. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, 40(D1):D1202–D1210.
- Lan, P., Li, W., Wang, H., and Ma, W. (2010). Characterization, sub-cellular localization and expression profiling of the isoprenylcysteine methyltransferase gene family in *Arabidopsis thaliana*. *BMC plant biology*, 10(1):1–17.
- Lands, W. E. M. (1965). Lipid metabolism. *Annual Review of Biochemistry*, 34(1):313–346.
- Lang, D., Ullrich, K. K., Murat, F., Fuchs, J., Jenkins, J., Haas, F. B., Piednoel, M., Gundlach, H., Van Bel, M., Meyberg, R., Vives, C., Morata, J., Symeonidi, A., Hiss, M., Muchero, W., Kamisugi, Y., Saleh, O., Blanc, G., Decker, E. L., van Gessel, N., Grimwood, J., Hayes, R. D., Graham, S. W., Gunter, L. E., McDaniel, S. F., Hoernstein, S. N., Larsson, A., Li, F.-W., Perroud, P.-F., Phillips, J., Ranjan, P., Rokshar, D. S., Rothfels, C. J., Schneider, L., Shu, S., Stevenson, D. W., Thümmel, F., Tillich, M., Villarreal Aguilar, J. C., Widiez, T., Wong, G. K.-S., Wymore, A., Zhang, Y., Zimmer, A. D., Quatrano, R. S., Mayer, K. F., Goodstein, D., Casacuberta, J. M., Vandepoele, K., Reski, R., Cuming, A. C., Tuskan, G. A., Maumus, F., Salse, J., Schmutz, J., and Rensing, S. A. (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal*, 93(3):515–533.
- Lang, D. E., Morris, J. S., Rowley, M., Torres, M. A., Maksimovich, V. A., Facchini, P. J., and Ng, K. K. S. (2019). Structure–function studies of tetrahydroprotoberberine *N*-methyltransferase reveal the molecular basis of stereoselective substrate recognition. *The Journal of Biological Chemistry*, 294(40):14482–14498.
- Lara, J. A., Burciaga-Monge, A., Chávez, A., Revés, M., Lavilla, R., Arró, M., Boronat, A., Altabella, T., and Ferrer, A. (2018). Identification and characterization of sterol acyltransferases responsible for steryl ester biosynthesis in tomato. *Frontiers in Plant Science*, 9.
- Layerenza, J. P., González, P., García de Bravo, M. M., Polo, M. P., Sisti, M. S., and Ves-Losada, A. (2013). Nuclear lipid droplets: A novel nuclear domain. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1831(2):327–340.
- Le Moigne, D., Guéguen, N., and Salvaing, J. (2022). Lipid droplets in plants: More than a simple fat storage. In *Advances in Botanical Research*, volume 101, pages 191–223. Elsevier.
- Leprince, O., van Aelst, A. C., Pritchard, H. W., and Murphy, D. J. (1997). Oleosins prevent oil-body coalescence during seed imbibition as suggested by a low-temperature scanning electron microscope study of desiccation-tolerant and -sensitive oilseeds. *Planta*, 204(1):109–119.
- Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature*, 529(7584):84–87.
- Letunic, I. and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, 23(1):127–128.
- Leyland, B., Boussiba, S., and Khozin-Goldberg, I. (2020). A review of diatom lipid droplets. *Biology*, 9(2):38.



- Li, F., Han, X., Guan, H., Xu, M. C., Dong, Y. X., and Gao, X.-Q. (2022). PALD encoding a lipid droplet-associated protein is critical for the accumulation of lipid droplets and pollen longevity in *Arabidopsis*. *New Phytologist*, 235(1):204–219.
- Li, F.-W., Brouwer, P., Carretero-Paulet, L., Cheng, S., de Vries, J., Delaux, P.-M., Eily, A., Koppers, N., Kuo, L.-Y., Li, Z., Simenc, M., Small, I., Wafula, E., Angarita, S., Barker, M. S., Bräutigam, A., dePamphilis, C., Gould, S., Hosmani, P. S., Huang, Y.-M., Huettel, B., Kato, Y., Liu, X., Maere, S., McDowell, R., Mueller, L. A., Nierop, K. G. J., Rensing, S. A., Robison, T., Rothfels, C. J., Sigel, E. M., Song, Y., Timilsena, P. R., Van de Peer, Y., Wang, H., Wilhelmsson, P. K. I., Wolf, P. G., Xu, X., Der, J. P., Schlupepmann, H., Wong, G. K.-S., and Pryer, K. M. (2018). Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature Plants*, 4(7):460–472.
- Li, F.-W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N., Barker, M. S., Bennett, T., Blázquez, M. A., Cheng, S., Cuming, A. C., de Vries, J., de Vries, S., Delaux, P.-M., Diop, I. S., Harrison, C. J., Hauser, D., Hernández-García, J., Kirbis, A., Meeks, J. C., Monte, I., Mutte, S. K., Neubauer, A., Quandt, D., Robison, T., Shimamura, M., Rensing, S. A., Villarreal, J. C., Weijers, D., Wicke, S., Wong, G. K.-S., Sakakibara, K., and Szövényi, P. (2020). *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. *Nature Plants*, 6(3):259–272.
- Li, N., Gügel, I. L., Giavalisco, P., Zeisler, V., Schreiber, L., Soll, J., and Philippar, K. (2015). FAX1, a novel membrane protein mediating plastid fatty acid export. *PLOS Biology*, 13(2):e1002053.
- Li, R., Yu, K., and Hildebrand, D. F. (2010). *DGAT1*, *DGAT2* and *PDAT* expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants. *Lipids*, 45(2):145–157.
- Li, Y., Beisson, F., Koo, A. J. K., Molina, I., Pollard, M., and Ohlrogge, J. (2007). Identification of acyltransferases required for cutin biosynthesis and production of cutin with suberin-like monomers. *Proceedings of the National Academy of Sciences*, 104(46):18339–18344.
- Lin, L.-J., Tai, S. S., Peng, C.-C., and Tzen, J. T. (2002). Steroleosin, a sterol-binding dehydrogenase in seed oil bodies. *Plant Physiology*, 128(4):1200–1211.
- Lin, L.-J. and Tzen, J. T. (2004). Two distinct steroleosins are present in seed oil bodies. *Plant Physiology and Biochemistry*, 42(7-8):601–608.
- Linssen, J. P. H., Cozijnsen, J. L., and Pilnik, W. (1989). Chufa (*Cyperus esculentus*): A new source of dietary fibre. *Journal of the Science of Food and Agriculture*, 49(3):291–296.
- Liscombe, D. K. and Facchini, P. J. (2007). Molecular cloning and characterization of tetrahydroprotoberberine cis-*N*-methyltransferase, an enzyme involved in alkaloid biosynthesis in opium poppy. *Journal of Biological Chemistry*, 282(20):14741–14751.
- Liscombe, D. K., Louie, G. V., and Noel, J. P. (2012). Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Natural Product Reports*, 29(10):1238.
- Liscombe, D. K., MacLeod, B. P., Loukanina, N., Nandi, O. I., and Facchini, P. J. (2005). Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms. *Phytochemistry*, 66(11):1374–1393.
- Listenberger, L., Townsend, E., Rickertsen, C., Hains, A., Brown, E., Inwards, E. G., Stoeckman, A. K., Matis, M. P., Sampathkumar, R. S., Osna, N. A., and Kharbanda, K. K. (2018). Decreasing phosphatidylcholine on the surface of the lipid droplet correlates with altered protein binding and steatosis. *Cells*, 7(12):230.

- Liu, C.-J., Deavours, B. E., Richard, S. B., Ferrer, J.-L., Blount, J. W., Huhman, D., Dixon, R. A., and Noel, J. P. (2006). Structural basis for dual functionality of isoflavonoid *O*-methyltransferases in the evolution of plant defense responses. *The Plant Cell*, 18(12):3656–3669.
- Liu, H., Wang, C., Chen, F., and Shen, S. (2015). Proteomic analysis of oil bodies in mature *Jatropha curcas* seeds with different lipid content. *Journal of Proteomics*, 113:403–414.
- Liu, Q., Guo, Q., Akbar, S., Zhi, Y., El Tahchy, A., Mitchell, M., Li, Z., Shrestha, P., Vanhercke, T., Ral, J.-P., et al. (2017). Genetic enhancement of oil content in potato tuber (*Solanum tuberosum* L.) through an integrated metabolic engineering strategy. *Plant Biotechnology Journal*, 15(1):56–67.
- Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I. A. P., Zhao, M., Ma, J., Yu, J., Huang, S., Wang, X., Wang, J., Lu, K., Fang, Z., Bancroft, I., Yang, T.-J., Hu, Q., Wang, X., Yue, Z., Li, H., Yang, L., Wu, J., Zhou, Q., Wang, W., King, G. J., Pires, J. C., Lu, C., Wu, Z., Sampath, P., Wang, Z., Guo, H., Pan, S., Yang, L., Min, J., Zhang, D., Jin, D., Li, W., Belcram, H., Tu, J., Guan, M., Qi, C., Du, D., Li, J., Jiang, L., Batley, J., Sharpe, A. G., Park, B.-S., Ruperao, P., Cheng, F., Waminal, N. E., Huang, Y., Dong, C., Wang, L., Li, J., Hu, Z., Zhuang, M., Huang, Y., Huang, J., Shi, J., Mei, D., Liu, J., Lee, T.-H., Wang, J., Jin, H., Li, Z., Li, X., Zhang, J., Xiao, L., Zhou, Y., Liu, Z., Liu, X., Qin, R., Tang, X., Liu, W., Wang, Y., Zhang, Y., Lee, J., Kim, H. H., Denoeud, F., Xu, X., Liang, X., Hua, W., Wang, X., Wang, J., Chalhoub, B., and Paterson, A. H. (2014). The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications*, 5(1):3930.
- Liu, T.-h., Chyan, C.-l., Li, F.-y., and Tzen, J. T. C. (2009). Stability of artificial oil bodies constituted with recombinant caleosins. *Journal of Agricultural and Food Chemistry*, 57(6):2308–2313.
- Liu, Z., Cheema, J., Vigouroux, M., Hill, L., Reed, J., Paajanen, P., Yant, L., and Osbourn, A. (2020). Formation and diversification of a paradigm biosynthetic gene cluster in plants. *Nature Communications*, 11(1):5354.
- Loer, D. S. and Herman, E. M. (1993). Cotranslational integration of soybean (*Glycine max*) oil body membrane protein oleosin into microsomal membranes. *Plant Physiology*, 101(3):993–998.
- Lotan, T., Ohto, M.-a., Yee, K. M., West, M. A., Lo, R., Kwong, R. W., Yamagishi, K., Fischer, R. L., Goldberg, R. B., and Harada, J. J. (1998). Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell*, 93(7):1195–1205.
- Lundquist, P. K., Poliakov, A., Bhuiyan, N. H., Zybailov, B., Sun, Q., and van Wijk, K. J. (2012). The functional network of the *Arabidopsis* plastoglobule proteome based on quantitative proteomics and genome-wide coexpression analysis. *Plant Physiology*, 158(3):1172–1192.
- Lundquist, P. K., Shivaiah, K.-K., and Espinoza-Corral, R. (2020). Lipid droplets throughout the evolutionary tree. *Progress in Lipid Research*, 78:101029.
- Luo, W., Komatsu, S., Abe, T., Matsuura, H., and Takahashi, K. (2020). Comparative proteomic analysis of wild-type *Physcomitrella patens* and an OPDA-deficient *Physcomitrella patens* mutant with disrupted PpAOS1 and PpAOS2 genes after wounding. *International Journal of Molecular Sciences*, 21(4):1417.
- Lyall, R. and Gechev, T. (2020). Multi-omics insights into the evolution of angiosperm resurrection plants. In *Annual Plant Reviews online*, pages 77–110. John Wiley & Sons, Ltd.

- Lyall, R., Schlebusch, S. A., Proctor, J., Prag, M., Hussey, S. G., Ingle, R. A., and Illing, N. (2020). Vegetative desiccation tolerance in the resurrection plant *Xerophyta humilis* has not evolved through reactivation of the seed canonical LAFL regulatory network. *The Plant Journal*, 101(6):1349–1367.
- Lyon, C., Saupe, E. E., Smith, C. J., Hill, D. J., Beckerman, A. P., Stringer, L. C., Marchant, R., McKay, J., Burke, A., O'Higgins, P., Dunhill, A. M., Allen, B. J., Riel-Salvatore, J., and Aze, T. (2022). Climate change research and action must look beyond 2100. *Global Change Biology*, 28(2):349–361.
- Mamaeva, A., Knyazev, A., Glushkevich, A., and Fesenko, I. (2022). Quantitative proteomic dataset of the moss *Physcomitrium patens* PSEP3 KO and OE mutant lines. *Data in Brief*, 40:107715.
- Martin, J. L. and McMillan, F. M. (2002). SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Current opinion in structural biology*, 12(6):783–793.
- Matilla, A. J. (2022). The orthodox dry seeds are alive: A clear example of desiccation tolerance. *Plants*, 11(1).
- Matsui, K., Hijiya, K., Tabuchi, Y., and Kajiwara, T. (1999). Cucumber cotyledon lipoxygenase during postgerminative growth: Expression and action on lipid bodies. *Plant Physiology*, 119(4):1279–1288.
- Matyash, V., Liebisch, G., Kurzchalia, T. V., Shevchenko, A., and Schwudke, D. (2008). Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics. *Journal of Lipid Research*, 49(5):1137–1146.
- McLachlan, D. H., Lan, J., Geilfus, C.-M., Dodd, A. N., Larson, T., Baker, A., Hōrak, H., Kollist, H., He, Z., Graham, I., Mickelbart, M. V., and Hetherington, A. M. (2016). The breakdown of stored triacylglycerols is required during light-induced stomatal opening. *Current Biology*, 26(5):707–712.
- Medford, J. I., Behringer, F. J., Callos, J. D., and Feldmann, K. A. (1992). Normal and abnormal development in the *Arabidopsis* vegetative shoot apex. *The Plant Cell*, 4(6):631–643.
- Meinke, D. W., Franzmann, L. H., Nickle, T. C., and Yeung, E. C. (1994). Leafy cotyledon mutants of *Arabidopsis*. *The Plant Cell*, 6(8):1049–1064.
- Mello, J. I. d. O., Barbedo, C. J., Salatino, A., and Figueiredo-Ribeiro, R. d. C. L. (2010). Reserve carbohydrates and lipids from the seeds of four tropical tree species with different sensitivity to desiccation. *Brazilian Archives of Biology and Technology*, 53(4):889–899.
- Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Maréchal-Drouard, L., Marshall, W. F., Qu, L.-H., Nelson, D. R., Sanderfoot, A. A., Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.-L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. A., Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M., Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J.-P., Riaño-Pachón, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J.,

- Bisova, K., Chen, C.-J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H., Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S., Bowler, C., Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W. C. A., Otiillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*, 318(5848):245–250.
- Miller, D. J., Ouellette, N., Evdokimova, E., Savchenko, A., Edwards, A., and Anderson, W. F. (2003). Crystal complexes of a predicted S-adenosylmethionine-dependent methyltransferase reveal a typical AdoMet binding domain and a substrate recognition domain. *Protein Science*, 12(7):1432–1442.
- Miquel, M., Trigui, G., d'Andréa, S., Kelemen, Z., Baud, S., Berger, A., Deruyffelaere, C., Trubuil, A., Lepiniec, L., and Dubreucq, B. (2014). Specialization of oleosins in oil body dynamics during seed development in *Arabidopsis* seeds. *Plant Physiology*, 164(4):1866–1878.
- Mock, T. and Kroon, B. M. (2002). Photosynthetic energy conversion under extreme conditions—I: important role of lipids as structural modulators and energy sink under N-limited growth in Antarctic sea ice diatoms. *Phytochemistry*, 61(1):41–51.
- Moellering, E. R. and Benning, C. (2010). RNA interference silencing of a major lipid droplet protein affects lipid droplet size in *Chlamydomonas reinhardtii*. *Eukaryotic Cell*, 9(1):97–106.
- Morikawa, T., Saga, H., Hashizume, H., and Ohta, D. (2009). CYP710A genes encoding sterol C22-desaturase in *Physcomitrella patens* as molecular evidence for the evolutionary conservation of a sterol biosynthetic pathway in plants. *Planta*, 229(6):1311–1322.
- Morris, J. S. and Facchini, P. J. (2019). Molecular origins of functional diversity in benzyloisoquinoline alkaloid methyltransferases. *Frontiers in Plant Science*, 10.
- Mostafa, M., Rahma, E., and Rady, A. (1987). Chemical and nutritional changes in soybean during germination. *Food Chemistry*, 23(4):257–275.
- Motamayor, J. C., Mockaitis, K., Schmutz, J., Haiminen, N., III, D. L., Cornejo, O., Findley, S. D., Zheng, P., Utro, F., Royaert, S., Saski, C., Jenkins, J., Podicheti, R., Zhao, M., Scheffler, B. E., Stack, J. C., Feltus, F. A., Mustiga, G. M., Amores, F., Phillips, W., Marelli, J. P., May, G. D., Shapiro, H., Ma, J., Bustamante, C. D., Schnell, R. J., Main, D., Gilbert, D., Parida, L., and Kuhn, D. N. (2013). The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. *Genome Biology*, 14(6):r53.
- Mu, J., Tan, H., Zheng, Q., Fu, F., Liang, Y., Zhang, J., Yang, X., Wang, T., Chong, K., Wang, X.-J., and Zuo, J. (2008). LEAFY COTYLEDON1 is a key regulator of fatty acid biosynthesis in *Arabidopsis*. *Plant Physiology*, 148(2):1042–1054.
- Mueller, S. P., Krause, D. M., Mueller, M. J., and Fekete, A. (2015). Accumulation of extra-chloroplastic triacylglycerols in *Arabidopsis* seedlings during heat acclimation. *Journal of Experimental Botany*, 66(15):4517–4526.
- Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia plantarum*, 15(3):473–497.
- Murata, N. and Los, D. A. (1997). Membrane fluidity and temperature perception. *Plant Physiology*, 115(3):875–879.

- Murphy, D. J. (2012). The dynamic roles of intracellular lipid droplets: from archaea to mammals. *Protoplasma*, 249(3):541–585.
- Murphy, D. J., Ross, J. H. E., and Pritchard, H. W. (1995). Are oleosins only associated with oil bodies from desiccation tolerant plant tissues? In *Plant Lipid Metabolism*, pages 558–560. Springer Netherlands.
- Mönke, G., Seifert, M., Keilwagen, J., Mohr, M., Grosse, I., Hähnel, U., Junker, A., Weisshaar, B., Conrad, U., Bäumlein, H., and Altschmied, L. (2012). Toward the identification and regulation of the *Arabidopsis thaliana* ABI3 regulon. *Nucleic Acids Research*, 40(17):8240–8254.
- Müller, A. O., Blersch, K. F., Gippert, A. L., and Ischebeck, T. (2017). Tobacco pollen tubes – a fast and easy tool for studying lipid droplet association of plant proteins. *The Plant Journal*, 89(5):1055–1064.
- Müller, A. O. and Ischebeck, T. (2018). Characterization of the enzymatic activity and physiological function of the lipid droplet-associated triacylglycerol lipase AtOBL1. *New Phytologist*, 217(3):1062–1076.
- Nafisi, M., Goregaoker, S., Botanga, C. J., Glawischnig, E., Olsen, C. E., Halkier, B. A., and Glazebrook, J. (2007). *Arabidopsis* cytochrome P450 monooxygenase 71A13 catalyzes the conversion of indole-3-acetaldoxime in camalexin synthesis. *The Plant Cell*, 19(6):2039–2052.
- Nakamura, S., Mano, S., Tanaka, Y., Ohnishi, M., Nakamori, C., Araki, M., Niwa, T., Nishimura, M., Kaminaka, H., Nakagawa, T., Sato, Y., and Ishiguro, S. (2010). Gateway binary vectors with the bialaphos resistance gene, *bar*, as a selection marker for plant transformation. *Bioscience, Biotechnology, and Biochemistry*, 74(6):1315–1319.
- Nambara, E., Nambara, E., McCourt, P., and Naito, S. (1995). A regulatory role for the ABI3 gene in the establishment of embryo maturation in *Arabidopsis thaliana*. *Development*, 121(3):629–636.
- Napier, J. A., Beaudoin, F., Tatham, A. S., Alexander, L. G., and Shewry, P. R. (2001). The seed oleosins: Structure, properties and biological role. In *Advances in Botanical Research*, volume 35, pages 111–138. Academic Press.
- National Center for Biotechnology Information (2023). NCBI NR protein database. <https://www.ncbi.nlm.nih.gov/protein/>. [Accessed Feb 6, 2023].
- Nettebrock, N. T. and Bohnert, M. (2020). Born this way – Biogenesis of lipid droplets from specialized ER subdomains. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1865(1):158448.
- Nguyen, H. M., Baudet, M., Cuiné, S., Adriano, J.-M., Barthe, D., Billon, E., Bruley, C., Beisson, F., Peltier, G., Ferro, M., and Li-Beisson, Y. (2011). Proteomic profiling of oil bodies isolated from the unicellular green microalga *Chlamydomonas reinhardtii*: With focus on proteins involved in lipid metabolism. *PROTEOMICS*, 11(21):4266–4273.
- Ni, B. and Feussner, K. (2023). *Ex vivo* metabolomics—a hypothesis-free approach to identify native substrate(s) and product(s) of orphan enzymes. In *Biochemical Pathways and Environmental Responses in Plants: Part B*, pages 303–323. Elsevier.

- Niemeyer, P. W., Irisarri, I., Scholz, P., Schmitt, K., Valerius, O., Braus, G. H., Herrfurth, C., Feussner, I., Sharma, S., Carlsson, A. S., de Vries, J., Hofvander, P., and Ischebeck, T. (2022). A seed-like proteome in oil-rich tubers. *The Plant Journal*, 112(2):518–534.
- Nishal, B., Tantikanjana, T., and Sundaresan, V. (2005). An inducible targeted tagging system for localized saturation mutagenesis in *Arabidopsis*. *Plant Physiology*, 137(1):3–12.
- Nishiyama, T., Sakayama, H., Vries, J. d., Buschmann, H., Saint-Marcoux, D., Ullrich, K. K., Haas, F. B., Vanderstraeten, L., Becker, D., Lang, D., Vosolsobě, S., Rombauts, S., Wilhelmsson, P. K. I., Janitza, P., Kern, R., Heyl, A., Rümpler, F., Villalobos, L. I. A. C., Clay, J. M., Skokan, R., Toyoda, A., Suzuki, Y., Kagoshima, H., Schijlen, E., Tajeshwar, N., Catarino, B., Hetherington, A. J., Saltykova, A., Bonnot, C., Breuninger, H., Symeonidi, A., Radhakrishnan, G. V., Nieuwerburgh, F. V., Deforce, D., Chang, C., Karol, K. G., Hedrich, R., Ulvskov, P., Glöckner, G., Delwiche, C. F., Petrášek, J., Peer, Y. V. d., Friml, J., Beilby, M., Dolan, L., Kohara, Y., Sugano, S., Fujiyama, A., Delaux, P.-M., Quint, M., Theißen, G., Hagemann, M., Harholt, J., Dunand, C., Zachgo, S., Langdale, J., Maumus, F., Straeten, D. V. D., Gould, S. B., and Rensing, S. A. (2018). The *Chara* genome: Secondary complexity and implications for plant terrestrialization. *Cell*, 174(2):448–464.e24.
- Nojima, D., Yoshino, T., Maeda, Y., Tanaka, M., Nemoto, M., and Tanaka, T. (2013). Proteomics analysis of oil body-associated proteins in the oleaginous diatom. *Journal of Proteome Research*, 12(11):5293–5301.
- Nützmann, H.-W. and Osbourn, A. (2015). Regulation of metabolic gene clusters in *Arabidopsis thaliana*. *New Phytologist*, 205(2):503–510.
- Næsted, H., Frandsen, G. I., Jauh, G.-Y., Hernandez-Pinzon, I., Nielsen, H. B., Murphy, D. J., Rogers, J. C., and Mundy, J. (2000). Caleosins: Ca<sup>2+</sup>-binding proteins associated with lipid bodies. *Plant Molecular Biology*, 44(4):463–476.
- Obulareddy, N., Panchal, S., and Melotto, M. (2013). Guard cell purification and rna isolation suitable for high-throughput transcriptional analysis of cell-type responses to biotic stresses. *Molecular Plant-Microbe Interactions*, 26(8):844–849.
- Ohyama, K., Suzuki, M., Kikuchi, J., Saito, K., and Muranaka, T. (2009). Dual biosynthetic pathways to phytosterol via cycloartenol and lanosterol in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, 106(3):725–730.
- Okuley, J., Lightner, J., Feldmann, K., Yadav, N., Lark, E., and Browse, J. (1994). Arabidopsis FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *The Plant Cell*, 6(1):147–158.
- Oliver, M. J. (2005). Desiccation tolerance in bryophytes: A reflection of the primitive strategy for plant survival in dehydrating habitats? *Integrative and Comparative Biology*, 45(5):788–799.
- Oliver, M. J., Farrant, J. M., Hilhorst, H. W., Mundree, S., Williams, B., and Bewley, J. D. (2020). Desiccation tolerance: Avoiding cellular damage during drying and rehydration. *Annual Review of Plant Biology*, 71(1):435–460.
- Oliver, M. J., Tuba, Z., and Mishler, B. D. (2000). The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology*, 151(1):85–100.
- Olzmann, J. A. and Carvalho, P. (2019). Dynamics and functions of lipid droplets. *Nature Reviews Molecular Cell Biology*, 20(3):137–155.

- 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., and Buell, C. R. (2007). The TIGR Rice Genome Annotation Resource: improvements and new features. *Nucleic Acids Research*, 35(suppl\_1):D883–D887.
- Park, J., Bae, S., and Kim, J.-S. (2015). Cas-Designer: a web-based tool for choice of CRISPR-Cas9 target sites. *Bioinformatics*, 31(24):4014–4016.
- Park, K. Y., Kim, W. T., and Kim, E. Y. (2018). The proper localization of RESPONSIVE TO DESICCATION 20 in lipid droplets depends on their biogenesis induced by STRESS-RELATED PROTEINS in vegetative tissues. *Biochemical and Biophysical Research Communications*, 495(2):1885–1889.
- Park, S., Rancour, D. M., and Bednarek, S. Y. (2007). Protein domain-domain interactions and requirements for the negative regulation of *Arabidopsis* CDC48/p97 by the plant ubiquitin regulatory x (UBX) domain-containing protein, PUX1. *Journal of Biological Chemistry*, 282(8):5217–5224.
- Parra, G., Bradnam, K., and Korf, I. (2007). CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics*, 23(9):1061–1067.
- Parthibane, V., Rajakumari, S., Venkateshwari, V., Iyappan, R., and Rajasekharan, R. (2012). Oleosin: A bifunctional enzyme with both monoacylglycerol acyltransferase and phospholipase activities. *Journal of Biological Chemistry*, 287(3):1946–1954.
- Pasaribu, B., Chen, C.-S., Liao, Y. K., Jiang, P.-L., and Tzen, J. T. (2017). Identification of caleosin and oleosin in oil bodies of pine pollen. *Plant Physiology and Biochemistry*, 111:20–29.
- Peres, A. L. G. L., Soares, J. S., Tavares, R. G., Righetto, G., Zullo, M. A. T., Mandava, N. B., and Menossi, M. (2019). Brassinosteroids: The sixth class of phytohormones - a molecular view from discovery to hormonal interactions in plant development and stress adaptation. *International Journal of Molecular Sciences*, 20(2):331.
- Perroud, P.-F., Haas, F. B., Hiss, M., Ullrich, K. K., Alboresi, A., Amirebrahimi, M., Barry, K., Bassi, R., Bonhomme, S., Chen, H., Coates, J. C., Fujita, T., Guyon-Debast, A., Lang, D., Lin, J., Lipzen, A., Nogué, F., Oliver, M. J., Ponce de León, I., Quatrano, R. S., Rameau, C., Reiss, B., Reski, R., Ricca, M., Saidi, Y., Sun, N., Szövényi, P., Sreedasyam, A., Grimwood, J., Stacey, G., Schmutz, J., and Rensing, S. A. (2018). The *Physcomitrella patens* gene atlas project: large-scale RNA-seq based expression data. *The Plant Journal*, 95(1):168–182.
- Petan, T., Jarc, E., and Jusović, M. (2018). Lipid droplets in cancer: Guardians of fat in a stressful world. *Molecules*, 23(8):1941.
- Petrie, J. R., Shrestha, P., Liu, Q., Mansour, M. P., Wood, C. C., Zhou, X.-R., Nichols, P. D., Green, A. G., and Singh, S. P. (2010). Rapid expression of transgenes driven by seed-specific constructs in leaf tissue: DHA production. *Plant Methods*, 6(1).
- Popluechai, S., Froissard, M., Jolivet, P., Breviaro, D., Gatehouse, A. M., O'Donnell, A. G., Chardot, T., and Kohli, A. (2011). *Jatropha curcas* oil body proteome and oleosins: L-form JcOle3 as a potential phylogenetic marker. *Plant Physiology and Biochemistry*, 49(3):352–356.
- Prasanna, X., Salo, V. T., Li, S., Ven, K., Vihinen, H., Jokitalo, E., Vattulainen, I., and Ikonen, E. (2021). Seipin traps triacylglycerols to facilitate their nanoscale clustering in the endoplasmic reticulum membrane. *PLOS Biology*, 19(1):e3000998.

- Price, A. M., Doner, N. M., Gidda, S. K., Jambunathan, S., James, C. N., Schami, A., Yurchenko, O., Mullen, R. T., Dyer, J. M., Puri, V., and Chapman, K. D. (2020). MOUSE FAT-SPECIFIC PROTEIN 27 (FSP27) expressed in plant cells localizes to lipid droplets and promotes lipid droplet accumulation and fusion. *Biochimie*, 169:41–53.
- Prost, I., Dhondt, S., Rothe, G., Vicente, J., Rodriguez, M. J., Kift, N., Carbonne, F., Griffiths, G., Esquerré-Tugayé, M.-T., Rosahl, S., Castresana, C., Hamberg, M., and Fournier, J. (2005). Evaluation of the antimicrobial activities of plant oxylipins supports their involvement in defense against pathogens. *Plant Physiology*, 139(4):1902–1913.
- Prévost, C., Sharp, M. E., Kory, N., Lin, Q., Voth, G. A., Farese, R. V., and Walther, T. C. (2018). Mechanism and determinants of amphipathic helix-containing protein targeting to lipid droplets. *Developmental Cell*, 44(1):73–86.e4.
- Purkrtova, Z., Le Bon, C., Kralova, B., Ropers, M.-H., Anton, M., and Chardot, T. (2008). Effect of calcium on the functional and structural properties of caleosin in *Arabidopsis thaliana*. *Journal of Agricultural and Food Chemistry*, 56(23):11217–11224.
- Pyc, M., Cai, Y., Gidda, S. K., Yurchenko, O., Park, S., Kretzschmar, F. K., Ischebeck, T., Valerius, O., Braus, G. H., Chapman, K. D., Dyer, J. M., and Mullen, R. T. (2017a). *Arabidopsis* lipid droplet-associated protein (LDAP) – interacting protein (LDIP) influences lipid droplet size and neutral lipid homeostasis in both leaves and seeds. *The Plant Journal*, 92(6):1182–1201.
- Pyc, M., Cai, Y., Greer, M. S., Yurchenko, O., Chapman, K. D., Dyer, J. M., and Mullen, R. T. (2017b). Turning Over a New Leaf in Lipid Droplet Biology. *Trends in Plant Science*, 22(7):596–609.
- Pyc, M., Gidda, S. K., Seay, D., Esnay, N., Kretzschmar, F. K., Cai, Y., Doner, N. M., Greer, M. S., Hull, J. J., Coulon, D., Bréhélin, C., Yurchenko, O., de Vries, J., Valerius, O., Braus, G. H., Ischebeck, T., Chapman, K. D., Dyer, J. M., and Mullen, R. T. (2021). LDIP cooperates with SEIPIN and LDAP to facilitate lipid droplet biogenesis in *Arabidopsis*. *The Plant Cell*, 33(9):3076–3103.
- Qin, L., Hu, Y., Wang, J., Wang, X., Zhao, R., Shan, H., Li, K., Xu, P., Wu, H., Yan, X., Liu, L., Yi, X., Wanke, S., Bowers, J. E., Leebens-Mack, J. H., dePamphilis, C. W., Soltis, P. S., Soltis, D. E., Kong, H., and Jiao, Y. (2021). Insights into angiosperm evolution, floral development and chemical biosynthesis from the *Aristolochia fimbriata* genome. *Nature Plants*, 7(9):1239–1253.
- Qu, R., Wang, S. M., Lin, Y. H., Vance, V. B., and Huang, A. H. (1986). Characteristics and biosynthesis of membrane proteins of lipid bodies in the scutella of maize (*Zea mays* L.). *Biochemical Journal*, 235(1):57–65.
- R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, 2(8):1896–1906.
- Rawat, V., Abdelsamad, A., Pietzenek, B., Seymour, D. K., Koenig, D., Weigel, D., Pecinka, A., and Schneeberger, K. (2015). Improving the annotation of *Arabidopsis lyrata* using rna-seq data. *PLOS ONE*, 10(9):e0137391.



- Renne, M. F., Klug, Y. A., and Carvalho, P. (2020). Lipid droplet biogenesis: A mystery “unmixing”? *Seminars in Cell & Developmental Biology*, 108:14–23.
- Rodríguez-Rosales, M. P., Kerkeb, L., Ferrol, N., and Douaire, J. P. (1998). Lipoxygenase activity and lipid composition of cotyledons and oil bodies of two sunflower hybrids. *Plant Physiology and Biochemistry*, 36(4):285–291.
- Romani, F., Banić, E., Florent, S. N., Kanazawa, T., Goodger, J. Q., Mentink, R. A., Dierschke, T., Zachgo, S., Ueda, T., Bowman, J. L., Tsiantis, M., and Moreno, J. E. (2020). Oil body formation in *Marchantia polymorpha* is controlled by MpC1HDZ and serves as a defense against arthropod herbivores. *Current Biology*, 30(14):2815–2828.e8.
- Ross, J. H. and Murphy, D. J. (1996). Characterization of anther-expressed genes encoding a major class of extracellular oleosin-like proteins in the pollen coat of Brassicaceae. *The Plant Journal*, 9(5):625–637.
- Rotsch, A. H., Kopka, J., Feussner, I., and Ischebeck, T. (2017). Central metabolite and sterol profiling divides tobacco male gametophyte development and pollen tube growth into eight metabolic phases. *The Plant Journal*, 92(1):129–146.
- Ruan, J., Zhou, Y., Zhou, M., Yan, J., Khurshid, M., Weng, W., Cheng, J., and Zhang, K. (2019). Jasmonic acid signaling pathway in plants. *International Journal of Molecular Sciences*, 20(10):2479.
- Rudolph, M., Schlereth, A., Körner, M., Feussner, K., Berndt, E., Melzer, M., Hornung, E., and Feussner, I. (2011). The lipoxygenase-dependent oxygenation of lipid body membranes is promoted by a patatin-type phospholipase in cucumber cotyledons. *Journal of Experimental Botany*, 62(2):749–760.
- Salas, J. J. and Ohlrogge, J. B. (2002). Characterization of substrate specificity of plant FatA and FatB acyl-ACP thioesterases. *Archives of Biochemistry and Biophysics*, 403(1):25–34.
- Salo, V. T., Li, S., Vihinen, H., Hölttä-Vuori, M., Szkalitsy, A., Horvath, P., Belevich, I., Peränen, J., Thiele, C., Somerharju, P., Zhao, H., Santinho, A., Thiam, A. R., Jokitalo, E., and Ikonen, E. (2019). Seipin facilitates triglyceride flow to lipid droplet and counteracts droplet ripening via endoplasmic reticulum contact. *Developmental Cell*, 50(4):478–493.e9.
- Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., and Lepiniec, L. (2008). Deciphering gene regulatory networks that control seed development and maturation in *Arabidopsis*. *The Plant Journal*, 54(4):608–620.
- Sarnighausen, E., Wurtz, V., Heintz, D., Van Dorsselaer, A., and Reski, R. (2004). Mapping of the *Physcomitrella patens* proteome. *Phytochemistry*, 65(11):1589–1607.
- Schepers, J. and Behl, C. (2021). Lipid droplets and autophagy—links and regulations from yeast to humans. *Journal of Cellular Biochemistry*, 122(6):602–611.
- 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., and Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9(7):676–682.
- Scholz, P. (2022). *Lipids in plant development and stress responses*. PhD thesis, Georg-August-Universität Göttingen, Göttingen, Germany.

- Scholz, P., Chapman, K. D., Mullen, R. T., and Ischebeck, T. (2022). Finding new friends and revisiting old ones – how plant lipid droplets connect with other subcellular structures. *New Phytologist*, 236(3):833–838.
- Schrödinger, LLC (2015). The PyMOL molecular graphics system, version 1.8.
- Schrul, B. and Kopito, R. R. (2016). Peroxin-dependent targeting of a lipid-droplet-destined membrane protein to ER subdomains. *Nature Cell Biology*, 18(7):740–751.
- Schubert, H. L., Blumenthal, R. M., and Cheng, X. (2003). Many paths to methyltransfer: a chronicle of convergence. *Trends in Biochemical Sciences*, 28(6):329–335.
- Scott, A. D., Zimin, A. V., Puiu, D., Workman, R., Britton, M., Zaman, S., Caballero, M., Read, A. C., Bogdanove, A. J., Burns, E., Wegrzyn, J., Timp, W., Salzberg, S. L., and Neale, D. B. (2020). A reference genome sequence for giant *Sequoia*. *G3 Genes/Genomes/Genetics*, 10(11):3907–3919.
- Sharma, M. and Laxmi, A. (2016). Jasmonates: Emerging players in controlling temperature stress tolerance. *Frontiers in Plant Science*, 6.
- Shen, Y., Xie, J., Liu, R.-d., Ni, X.-f., Wang, X.-h., Li, Z.-x., and Zhang, M. (2014). Genomic analysis and expression investigation of caleosin gene family in *Arabidopsis*. *Biochemical and Biophysical Research Communications*, 448(4):365–371.
- Shevchenko, A., Tomas, H., Havli, J., Olsen, J. V., and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, 1(6):2856–2860.
- Shibata, D., Slusarenko, A., Casey, R., Hildebrand, D., and Bell, E. (1994). Lipoxygenases. *Plant Molecular Biology Reporter*, 12(2):S41–S42.
- Shimada, T. L. and Hara-Nishimura, I. (2015). Leaf oil bodies are subcellular factories producing antifungal oxylipins. *Current Opinion in Plant Biology*, 25:145–150.
- Shimada, T. L., Hayashi, M., and Hara-Nishimura, I. (2018). Membrane dynamics and multiple functions of oil bodies in seeds and leaves. *Plant Physiology*, 176(1):199–207.
- Shimada, T. L., Shimada, T., Takahashi, H., Fukao, Y., and Hara-Nishimura, I. (2008). A novel role for oleosins in freezing tolerance of oilseeds in *Arabidopsis thaliana*. *The Plant Journal*, 55(5):798–809.
- Shimada, T. L., Takano, Y., Shimada, T., Fujiwara, M., Fukao, Y., Mori, M., Okazaki, Y., Saito, K., Sasaki, R., Aoki, K., and Hara-Nishimura, I. (2014). Leaf oil body functions as a subcellular factory for the production of a phytoalexin in *Arabidopsis*. *Plant Physiology*, 164(1):105–118.
- Shockey, J., Kuhn, D., Chen, T., Cao, H., Freeman, B., and Mason, C. (2018). Cyclopropane fatty acid biosynthesis in plants: phylogenetic and biochemical analysis of *Litchi* Kennedy pathway and acyl editing cycle genes. *Plant Cell Reports*, 37(11):1571–1583.
- Shockey, J., Regmi, A., Cotton, K., Adhikari, N., Browse, J., and Bates, P. D. (2016). Identification of *Arabidopsis* GPAT9 (At5g60620) as an essential gene involved in triacylglycerol biosynthesis. *Plant Physiology*, 170(1):163–179.
- Siedlecki, P. and Zielenkiewicz, P. (2006). Mammalian DNA methyltransferases. *Acta Biochimica Polonica*, 53(2):245–256.

- Siegler, H., Valerius, O., Ischebeck, T., Popko, J., Tourasse, N. J., Vallon, O., Khozin-Goldberg, I., Braus, G. H., and Feussner, I. (2017). Analysis of the lipid body proteome of the oleaginous alga *Lobosphaera incisa*. *BMC Plant Biology*, 17(1):98.
- Siloto, R. M., Findlay, K., Lopez-Villalobos, A., Yeung, E. C., Nykiforuk, C. L., and Moloney, M. M. (2006). The accumulation of oleosins determines the size of seed oil bodies in *Arabidopsis*. *The Plant Cell*, 18(8):1961–1974.
- Singer, S. D., Chen, G., Mietkiewska, E., Tomasi, P., Jayawardhane, K., Dyer, J. M., and Weselake, R. J. (2016). *Arabidopsis* GPAT9 contributes to synthesis of intracellular glycerolipids but not surface lipids. *Journal of Experimental Botany*, 67(15):4627–4638.
- Singh, A. P. and Savaldi-Goldstein, S. (2015). Growth control: brassinosteroid activity gets context. *Journal of Experimental Botany*, 66(4):1123–1132.
- Skripnikov, A. Y., Polyakov, N. B., Tolcheva, E. V., Velikodvorskaya, V. V., Dolgov, S. V., Demina, I. A., Rogova, M. A., and Govorun, V. M. (2009). Proteome analysis of the moss *Physcomitrella patens* (Hedw.) B.S.G. *Biochemistry*, 74(5):480–490.
- Slocombe, S. P., Cornah, J., Pinfield-Wells, H., Soady, K., Zhang, Q., Gilday, A., Dyer, J. M., and Graham, I. A. (2009). Oil accumulation in leaves directed by modification of fatty acid breakdown and lipid synthesis pathways. *Plant Biotechnology Journal*, 7(7):694–703.
- Song, J., Mizrak, A., Lee, C.-W., Cicconet, M., Lai, Z. W., Tang, W.-C., Lu, C.-H., Mohr, S. E., Farese, R. V., and Walther, T. C. (2022). Identification of two pathways mediating protein targeting from ER to lipid droplets. *Nature Cell Biology*, 24(9):1364–1377.
- Soto-Burgos, J., Zhuang, X., Jiang, L., and Bassham, D. C. (2018). Dynamics of autophagosome formation. *Plant Physiology*, 176(1):219–229.
- Softysik, K., Ohsaki, Y., and Fujimoto, T. (2019). Duo in a mystical realm—nuclear lipid droplets and the inner nuclear membrane. *Contact*, 2:251525641989696.
- Softysik, K., Ohsaki, Y., Tatematsu, T., Cheng, J., Maeda, A., Morita, S.-y., and Fujimoto, T. (2021). Nuclear lipid droplets form in the inner nuclear membrane in a seipin-independent manner. *Journal of Cell Biology*, 220(1):e202005026.
- Stålberg, K., Ståhl, U., Stymne, S., and Ohlrogge, J. (2009). Characterization of two *Arabidopsis thaliana* acyltransferases with preference for lysophosphatidylethanolamine. *BMC plant biology*, 9:1–8.
- Stevenson, S. R., Kamisugi, Y., Trinh, C. H., Schmutz, J., Jenkins, J. W., Grimwood, J., Muchero, W., Tuskan, G. A., Rensing, S. A., Lang, D., Reski, R., Melkonian, M., Rothfels, C. J., Li, F.-W., Larsson, A., Wong, G. K.-S., Edwards, T. A., and Cuming, A. C. (2016). Genetic analysis of *Physcomitrella patens* identifies ABSCISIC ACID NON-RESPONSIVE (ANR), a regulator of ABA responses unique to basal land plants and required for desiccation tolerance. *The Plant Cell*, 28(6):1310–1327.
- Stiti, N., Missihoun, T. D., Kotchoni, S. O., Kirch, H.-H., and Bartels, D. (2011). Aldehyde dehydrogenases in *Arabidopsis thaliana*: biochemical requirements, metabolic pathways, and functional analysis. *Frontiers in Plant Science*, 2:65.
- Stiti, N., Podgórska, K. A., and Bartels, D. (2020). S-nitrosation impairs activity of stress-inducible aldehyde dehydrogenases from *Arabidopsis thaliana*. *Plant Science*, 292:110389.

- Stornaiuolo, M., Lotti, L. V., Borgese, N., Torrisi, M.-R., Mottola, G., Martire, G., and Bonatti, S. (2003). KDEL and KKXX retrieval signals appended to the same reporter protein determine different trafficking between endoplasmic reticulum, intermediate compartment, and Golgi complex. *Molecular Biology of the Cell*, 14(3):889–902.
- Sturtevant, D., Lu, S., Zhou, Z.-W., Shen, Y., Wang, S., Song, J.-M., Zhong, J., Burks, D. J., Yang, Z.-Q., Yang, Q.-Y., Cannon, A. E., Herrfurth, C., Feussner, I., Borisjuk, L., Munz, E., Verbeck, G. F., Wang, X., Azad, R. K., Singleton, B., Dyer, J. M., Chen, L.-L., Chapman, K. D., and Guo, L. (2020). The genome of jojoba (*Simmondsia chinensis*): A taxonomically isolated species that directs wax ester accumulation in its seeds. *Science Advances*, 6(11):eaay3240.
- Sui, X., Arlt, H., Brock, K. P., Lai, Z. W., DiMaio, F., Marks, D. S., Liao, M., Farese, Jr., R. V., and Walther, T. C. (2018). Cryo–electron microscopy structure of the lipid droplet–formation protein seipin. *Journal of Cell Biology*, 217(12):4080–4091.
- Sundaresan, V., Springer, P., Volpe, T., Haward, S., Jones, J. D., Dean, C., Ma, H., and Martienssen, R. (1995). Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements. *Genes & Development*, 9(14):1797–1810.
- Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gentzbittel, L., Childs, K. L., Yandell, M., Gundlach, H., Mayer, K. F. X., Schwartz, D. C., and Town, C. D. (2014). An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC Genomics*, 15:312.
- Taurino, M., Costantini, S., De Domenico, S., Stefanelli, F., Ruano, G., Delgadillo, M. O., Sánchez-Serrano, J. J., Sanmartín, M., Santino, A., and Rojo, E. (2018). Seipin proteins mediate lipid droplet biogenesis to promote pollen transmission and reduce seed dormancy. *Plant Physiology*, 176(2):1531–1546.
- Teasdale, R. D. and Jackson, M. R. (1996). Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the Golgi apparatus. *Annual Review of Cell and Developmental Biology*, 12(1):27–54.
- Terp, N., Göbel, C., Brandt, A., and Feussner, I. (2006). Lipoxygenases during *Brassica napus* seed germination. *Phytochemistry*, 67(18):2030–2040.
- Terrasson, E., Buitink, J., Righetti, K., Ly Vu, B., Pelletier, S., Lalanne, D., Zinsmeister, J., and Leprince, O. (2013). An emerging picture of the seed desiccome: confirmed regulators and newcomers identified using transcriptome comparison. *Frontiers in Plant Science*, 4.
- Thakur, A. and Bhatla, S. C. (2015). Proteomic analysis of oil body membrane proteins accompanying the onset of desiccation phase during sunflower seed development. *Plant Signaling & Behavior*, 10(12):e1030100.
- Thazar-Poulot, N., Miquel, M., Fobis-Loisy, I., and Gaude, T. (2015). Peroxisome extensions deliver the *Arabidopsis* SDP1 lipase to oil bodies. *Proceedings of the National Academy of Sciences*, 112(13):4158–4163.
- Thiam, A. R., Antonny, B., Wang, J., Delacotte, J., Wilfling, F., Walther, T. C., Beck, R., Rothman, J. E., and Pincet, F. (2013a). COPI buds 60-nm lipid droplets from reconstituted water–phospholipid–triacylglyceride interfaces, suggesting a tension clamp function. *Proceedings of the National Academy of Sciences*, 110(33):13244–13249.

- Thiam, A. R. and Dugail, I. (2019). Lipid droplet–membrane contact sites – from protein binding to function. *Journal of Cell Science*, 132(12):jcs230169.
- Thiam, A. R., Farese Jr, R. V., and Walther, T. C. (2013b). The biophysics and cell biology of lipid droplets. *Nature Reviews Molecular Cell Biology*, 14(12):775–786.
- Thiam, A. R. and Forêt, L. (2016). The physics of lipid droplet nucleation, growth and budding. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, 1861(8):715–722.
- Thiam, A. R. and Ikonen, E. (2021). Lipid droplet nucleation. *Trends in Cell Biology*, 31(2):108–118.
- Tian, R., Wang, F., Zheng, Q., Niza, V. M. A. G. E., Downie, A. B., and Perry, S. E. (2020). Direct and indirect targets of the *Arabidopsis* seed transcription factor ABSCISIC ACID INSENSITIVE3. *The Plant Journal*, 103(5):1679–1694.
- Tirajoh, A., Aung, T. S. T., McKay, A. B., and Plant, A. L. (2004). Stress-responsive  $\alpha$ -dioxxygenase expression in tomato roots. *Journal of Experimental Botany*, 56(412):713–723.
- Tissier, A. F., Marillonnet, S., Klimyuk, V., Patel, K., Torres, M. A., Murphy, G., and Jones, J. D. G. (1999). Multiple independent defective *Suppressor-mutator* transposon insertions in arabidopsis: A tool for functional genomics. *The Plant Cell*, 11(10):1841–1852.
- Tnani, H., López, I., Jouenne, T., and Vicient, C. M. (2011). Protein composition analysis of oil bodies from maize embryos during germination. *Journal of Plant Physiology*, 168(5):510–513.
- Torres, M. A., Hoffarth, E., Eugenio, L., Savtchouk, J., Chen, X., Morris, J. S., Facchini, P. J., and Ng, K. K.-S. (2016). Structural and functional studies of pavinic *N*-methyltransferase from *Thalictrum flavum* reveal novel insights into substrate recognition and catalytic mechanism. *The Journal of Biological Chemistry*, 291(45):23403–23415.
- Toshima, E., Nanjo, Y., Komatsu, S., Abe, T., Matsuura, H., and Takahashi, K. (2014). Proteomic analysis of *Physcomitrella patens* treated with 12-oxo-phytodienoic acid, an important oxylipin in plants. *Bioscience, Biotechnology, and Biochemistry*, 78(6):946–953.
- Tranbarger, T. J., Dussert, S., Joët, T., Argout, X., Summo, M., Champion, A., Cros, D., Omere, A., Nouy, B., and Morcillo, F. (2011). Regulatory mechanisms underlying oil palm fruit mesocarp maturation, ripening, and functional specialization in lipid and carotenoid metabolism. *Plant Physiology*, 156(2):564–584.
- Trifinopoulos, J., Nguyen, L.-T., von Haeseler, A., and Minh, B. Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, 44(W1):W232–W235.
- Troncoso-Ponce, M. A., Cao, X., Yang, Z., and Ohlrogge, J. B. (2013). Lipid turnover during senescence. *Plant Science*, 205–206:13–19.
- Tuba, Z., Protor, C., and Csintalan, Z. (1998). Ecophysiological responses of homoiochlorophyllous and poikilochlorophyllous desiccation tolerant plants: a comparison and an ecological perspective. *Plant Growth Regulation*, 24(3):211–217.
- Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Židek, A., Bridgland, A., Cowie, A., Meyer, C., Laydon, A., Velankar, S., Kleywegt, G. J., Bateman, A., Evans, R., Pritzel,

- A., Figurnov, M., Ronneberger, O., Bates, R., Kohl, S. A. A., Potapenko, A., Ballard, A. J., Romera-Paredes, B., Nikolov, S., Jain, R., Clancy, E., Reiman, D., Petersen, S., Senior, A. W., Kavukcuoglu, K., Birney, E., Kohli, P., Jumper, J., and Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. *Nature*, 596(7873):590–596.
- Turesson, H., Marttila, S., Gustavsson, K.-E., Hofvander, P., Olsson, M. E., Bülow, L., Stymne, S., and Carlsson, A. S. (2010). Characterization of oil and starch accumulation in tubers of *Cyperus esculentus* var. *sativus* (Cyperaceae): A novel model system to study oil reserves in nonseed tissues. *American Journal of Botany*, 97(11):1884–1893.
- Turró, S., Ingelmo-Torres, M., Estanyol, J. M., Tebar, F., Fernández, M. A., Albor, C. V., Gaus, K., Grewal, T., Enrich, C., and Pol, A. (2006). Identification and characterization of ASSOCIATED WITH LIPID DROPLET PROTEIN 1: A novel membrane-associated protein that resides on hepatic lipid droplets. *Traffic*, 7(9):1254–1269.
- Twell, D., Yamaguchi, J., Wing, R. A., Ushiba, J., and McCormick, S. (1991). Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. *Genes & Development*, 5(3):496–507.
- Tyanova, S., Temu, T., and Cox, J. (2016). The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols*, 11(12):2301–2319.
- Tzen, J., Cao, Y., Laurent, P., Ratnayake, C., and Huang, A. (1993). Lipids, proteins, and structure of seed oil bodies from diverse species. *Plant Physiology*, 101(1):267–276.
- Tzen, J. T. C. (2012). Integral proteins within plant oil bodies. *ISRN Botany*, 2012:1–16.
- Uzbekov, R. and Roingeard, P. (2013). Nuclear lipid droplets identified by electron microscopy of serial sections. *BMC Research Notes*, 6(1):386.
- Valliyodan, B., Cannon, S. B., Bayer, P. E., Shu, S., Brown, A. V., Ren, L., Jenkins, J., Chung, C. Y.-L., Chan, T.-F., Daum, C. G., Plott, C., Hastie, A., Baruch, K., Barry, K. W., Huang, W., Patil, G., Varshney, R. K., Hu, H., Batley, J., Yuan, Y., Song, Q., Stupar, R. M., Goodstein, D. M., Stacey, G., Lam, H.-M., Jackson, S. A., Schmutz, J., Grimwood, J., Edwards, D., and Nguyen, H. T. (2019). Construction and comparison of three reference-quality genome assemblies for soybean. *The Plant Journal*, 100(5):1066–1082.
- Van Rossum, G. and Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.
- van Zutphen, T., Todde, V., de Boer, R., Kreim, M., Hofbauer, H. F., Wolinski, H., Veenhuis, M., van der Klei, I. J., and Kohlwein, S. D. (2014). Lipid droplet autophagy in the yeast *Saccharomyces cerevisiae*. *Molecular Biology of the Cell*, 25(2):290–301.
- VanBuren, R. (2017). Desiccation tolerance: Seedy origins of resurrection. *Nature Plants*, 3(4):17046.
- VanBuren, R., Bryant, D., Edger, P. P., Tang, H., Burgess, D., Challabathula, D., Spittle, K., Hall, R., Gu, J., Lyons, E., Freeling, M., Bartels, D., Ten Hallers, B., Hastie, A., Michael, T. P., and Mockler, T. C. (2015). Single-molecule sequencing of the desiccation-tolerant grass *Oropetium thomaeum*. *Nature*, 527(7579):508–511.
- VanBuren, R., Man Wai, C., Pardo, J., Giarola, V., Ambrosini, S., Song, X., and Bartels, D. (2018a). Desiccation tolerance evolved through gene duplication and network rewiring in *Lindernia*. *The Plant Cell*, 30(12):2943–2958.

- VanBuren, R., Wai, C. M., Ou, S., Pardo, J., Bryant, D., Jiang, N., Mockler, T. C., Edger, P., and Michael, T. P. (2018b). Extreme haplotype variation in the desiccation-tolerant clubmoss *Selaginella lepidophylla*. *Nature Communications*, 9(1):13.
- VanBuren, R., Wai, C. M., Zhang, Q., Song, X., Edger, P. P., Bryant, D., Michael, T. P., Mockler, T. C., and Bartels, D. (2017). Seed desiccation mechanisms co-opted for vegetative desiccation in the resurrection grass *Oropetium thomaeum*. *Plant, Cell & Environment*, 40(10):2292–2306.
- Vanni, S., Vamparys, L., Gautier, R., Drin, G., Etchebest, C., Fuchs, P. F., and Antonny, B. (2013). Amphipathic lipid packing sensor motifs: Probing bilayer defects with hydrophobic residues. *Biophysical Journal*, 104(3):575–584.
- Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022). AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*, 50(D1):D439–D444.
- Veerabagu, M., Rinne, P. L. H., Skaugen, M., Paul, L. K., and van der Schoot, C. (2021). Lipid body dynamics in shoot meristems: Production, enlargement, putative organellar interactions, and plasmodesmal targeting. *Frontiers in Plant Science*, 12.
- Verde, I., Abbott, A. G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M. T., Grimwood, J., Cattonaro, F., Zuccolo, A., Rossini, L., Jenkins, J., Vendramin, E., Meisel, L. A., Decroocq, V., Sosinski, B., Prochnik, S., Mitros, T., Policriti, A., Cipriani, G., Dondini, L., Ficklin, S., Goodstein, D. M., Xuan, P., Fabbro, C. D., Aramini, V., Copetti, D., Gonzalez, S., Horner, D. S., Falchi, R., Lucas, S., Mica, E., Maldonado, J., Lazzari, B., Bielenberg, D., Pirona, R., Miculan, M., Barakat, A., Testolin, R., Stella, A., Tartarini, S., Tonutti, P., Arús, P., Orellana, A., Wells, C., Main, D., Vizzotto, G., Silva, H., Salamini, F., Schmutz, J., Morgante, M., and Rokhsar, D. S. (2013). The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nature Genetics*, 45(5):487–494.
- Vermachova, M., Purkrtova, Z., Santrucek, J., Jolivet, P., Chardot, T., and Kodicek, M. (2011). New protein isoforms identified within *Arabidopsis thaliana* seed oil bodies combining chymotrypsin/trypsin digestion and peptide fragmentation analysis. *PROTEOMICS*, 11(16):3430–3434.
- Vermachova, M., Purkrtova, Z., Santrucek, J., Jolivet, P., Chardot, T., and Kodicek, M. (2014). Combining chymotrypsin/trypsin digestion to identify hydrophobic proteins from oil bodies. In Jorin-Novo, J. V., Komatsu, S., Weckwerth, W., and Wienkoop, S., editors, *Plant Proteomics*, volume 1072, pages 185–198. Humana Press, Totowa, NJ.
- Vieler, A., Brubaker, S. B., Vick, B., and Benning, C. (2012). A lipid droplet protein of *Nannochloropsis* with functions partially analogous to plant oleosins. *Plant Physiology*, 158(4):1562–1569.
- Vigeolas, H., Waldeck, P., Zank, T., and Geigenberger, P. (2007). Increasing seed oil content in oil-seed rape (*Brassica napus* L.) by over-expression of a yeast glycerol-3-phosphate dehydrogenase under the control of a seed-specific promoter. *Plant Biotechnology Journal*, 5(3):431–441.
- Vimolmangkang, S., Deng, X., Owiti, A., Meelaph, T., Ogutu, C., and Han, Y. (2016). Evolutionary origin of the NCSI gene subfamily encoding norcochlorine synthase is

- associated with the biosynthesis of benzyloisoquinoline alkaloids in plants. *Scientific Reports*, 6(1):26323.
- Wahlroos, T., Soukka, J., Denesyuk, A., Wahlroos, R., Korpela, T., and Kilby, N. J. (2003). Oleosin expression and trafficking during oil body biogenesis in tobacco leaf cells. *genesis*, 35(2):125–132.
- Walker, K. A. and Harwood, J. L. (1985). Localization of chloroplastic fatty acid synthesis *de novo* in the stroma. *Biochemical Journal*, 226(2):551–556.
- Walther, T. C., Chung, J., and Farese, R. V. (2017). Lipid droplet biogenesis. *Annual Review of Cell and Developmental Biology*, 33:491–510.
- Wang, X., Kuang, T., and He, Y. (2010). Conservation between higher plants and the moss *Physcomitrella patens* in response to the phytohormone abscisic acid: a proteomics analysis. *BMC Plant Biology*, 10(1):192.
- Wang, X., Wei, H., Mao, X., and Liu, J. (2019). Proteomics analysis of lipid droplets from the oleaginous alga *Chromochloris zofingiensis* reveals novel proteins for lipid metabolism. *Genomics, Proteomics & Bioinformatics*, 17(3):260–272.
- Wang, X., Yang, P., Gao, Q., Liu, X., Kuang, T., Shen, S., and He, Y. (2008). Proteomic analysis of the response to high-salinity stress in *Physcomitrella patens*. *Planta*, 228(1):167–177.
- Wang, X., Yang, P., Zhang, X., Xu, Y., Kuang, T., Shen, S., and He, Y. (2009). Proteomic analysis of the cold stress response in the moss, *Physcomitrella patens*. *PROTEOMICS*, 9(19):4529–4538.
- Wang, Z.-P., Xing, H.-L., Dong, L., Zhang, H.-Y., Han, C.-Y., Wang, X.-C., and Chen, Q.-J. (2015). Egg cell-specific promoter-controlled CRISPR/Cas9 efficiently generates homozygous mutants for multiple target genes in *Arabidopsis* in a single generation. *Genome Biology*, 16(1).
- Waschburger, E., Kulcheski, F. R., Veto, N. M., Margis, R., Margis-Pinheiro, M., and Turchetto-Zolet, A. C. (2018). Genome-wide analysis of the glycerol-3-phosphate acyltransferase (GPAT) gene family reveals the evolution and diversification of plant GPATs. *Genetics and Molecular Biology*, 41(1 suppl 1):355–370.
- Wasternack, C. and Feussner, I. (2018). The oxylipin pathways: Biochemistry and function. *Annual Review of Plant Biology*, 69(1):363–386.
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., et al. (2018). Swiss-model: homology modelling of protein structures and complexes. *Nucleic acids research*, 46(W1):W296–W303.
- Wegrzyn, J. L., Lee, J. M., Tearse, B. R., and Neale, D. B. (2008). TreeGenes: A forest tree genome database. *International Journal of Plant Genomics*, 2008:1–7.
- Wegrzyn, J. L., Staton, M. A., Street, N. R., Main, D., Grau, E., Herndon, N., Buehler, S., Falk, T., Zaman, S., Ramnath, R., Richter, P., Sun, L., Condon, B., Almsaeed, A., Chen, M., Mannapperuma, C., Jung, S., and Ficklin, S. (2019). Cyberinfrastructure to improve forest health and productivity: The role of tree databases in connecting genomes, phenomes, and the environment. *Frontiers in Plant Science*, 10.
- Wilfling, F., Haas, J. T., Walther, T. C., and Jr, R. V. F. (2014). Lipid droplet biogenesis. *Current Opinion in Cell Biology*, 29:39–45.



- Wilfling, F., Wang, H., Haas, J. T., Krahmer, N., Gould, T. J., Uchida, A., Cheng, J.-X., Graham, M., Christiano, R., Fröhlich, F., Liu, X., Buhman, K. K., Coleman, R. A., Bewersdorf, J., Farese, R. V., and Walther, T. C. (2013). Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocalizing from the ER to lipid droplets. *Developmental Cell*, 24(4):384–399.
- Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An “electronic fluorescent pictograph” browser for exploring and analyzing large-scale biological data sets. *PLoS ONE*, 2(8):e718.
- Wu, G. A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., Perrier, X., Ruiz, M., Scalabrin, S., Terol, J., Takita, M. A., Labadie, K., Poulain, J., Couloux, A., Jabbari, K., Cattonaro, F., Del Fabbro, C., Pinosio, S., Zuccolo, A., Chapman, J., Grimwood, J., Tadeo, F. R., Estornell, L. H., Muñoz-Sanz, J. V., Ibanez, V., Herrero-Ortega, A., Aleza, P., Pérez-Pérez, J., Ramón, D., Brunel, D., Luro, F., Chen, C., Farmerie, W. G., Desany, B., Kodira, C., Mohiuddin, M., Harkins, T., Fredrikson, K., Burns, P., Lomsadze, A., Borodovsky, M., Reforgiato, G., Freitas-Astúa, J., Quetier, F., Navarro, L., Roose, M., Wincker, P., Schmutz, J., Morgante, M., Machado, M. A., Talon, M., Jaillon, O., Ollitrault, P., Gmitter, F., and Rokhsar, D. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature Biotechnology*, 32(7):656–662.
- Wu, L. S., Hong, G. H., Hou, R. F., and Tzen, J. T. (1999). Classification of the single oleosin isoform and characterization of seed oil bodies in gymnosperms. *Plant and Cell Physiology*, 40(3):326–334.
- Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., Wang, X.-C., and Chen, Q.-J. (2014). A CRISPR/cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biology*, 14(1).
- Xu, C., Fan, J., Froehlich, J. E., Awai, K., and Benning, C. (2005). Mutation of the TGD1 chloroplast envelope protein affects phosphatidate metabolism in *Arabidopsis*. *The Plant Cell*, 17(11):3094–3110.
- Xu, X., Pan, S., Cheng, S., Zhang, B., Mu, D., Ni, P., Zhang, G., Yang, S., Li, R., Wang, J., Orjeda, G., Guzman, F., Torres, M., Lozano, R., Ponce, O., Martinez, D., De la Cruz, G., Chakrabarti, S. K., Patil, V. U., Skryabin, K. G., Kuznetsov, B. B., Ravin, N. V., Kolganova, T. V., Beletsky, A. V., Mardanov, A. V., Di Genova, A., Bolser, D. M., Martin, D. M. A., Li, G., Yang, Y., Kuang, H., Hu, Q., Xiong, X., Bishop, G. J., Sagredo, B., Mejía, N., Zagorski, W., Gromadka, R., Gawor, J., Szczesny, P., Huang, S., Zhang, Z., Liang, C., He, J., Li, Y., He, Y., Xu, J., Zhang, Y., Xie, B., Du, Y., Qu, D., Bonierbale, M., Ghislain, M., del Rosario Herrera, M., Giuliano, G., Pietrella, M., Perrotta, G., Facella, P., O’Brien, K., Feingold, S. E., Barreiro, L. E., Massa, G. A., Diambra, L., Whitty, B. R., Vaillancourt, B., Lin, H., Massa, A. N., Geoffroy, M., Lundback, S., DellaPenna, D., Robin Buell, C., Sharma, S. K., Marshall, D. F., Waugh, R., Bryan, G. J., Destefanis, M., Nagy, I., Milbourne, D., Thomson, S. J., Fiers, M., Jacobs, J. M. E., Nielsen, K. L., Sønderkær, M., Iovene, M., Torres, G. A., Jiang, J., Veilleux, R. E., Bachem, C. W. B., de Boer, J., Borm, T., Kloosterman, B., van Eck, H., Datema, E., te Lintel Hekkert, B., Goverse, A., van Ham, R. C. H. J., Visser, R. G. F., The Potato Genome Sequencing Consortium, The Potato Genome Consortium (Participants are listed alphabetically by institution.), BGI-Shenzhen, Cayetano Heredia University, Central Potato Research Institute, Centre Bioengineering RAS, CGR-CMM, U. d. C., College of Life Sciences, U. o. D., High Technology Research Center, S. A. o. A. S., Huazhong Agriculture University, Hunan Agricultural University, Imperial College London, Instituto de Investigaciones Agropecuarias, Institute of Biochemistry & Biophysics, Institute of

- Vegetables & Flowers, C. A. o. A. S., International Potato Center, Italian National Agency for New Technologies, E. . S. D., J Craig Venter Institute, Laboratorio de Agrobiotecnología, I. N. d. T. A., Laboratorio de Biología de Sistemas, U. N. d. L., Michigan State University, Scottish Crop Research Institute, Teagasc Crops Research Centre, The New Zealand Institute for Plant & Food Research Ltd, University of Aalborg, University of Wisconsin, Virginia Polytechnic Institute & State University, and Wageningen University & Research Centre (2011). Genome sequence and analysis of the tuber crop potato. *Nature*, 475(7355):189–195.
- Xu, Z., Xin, T., Bartels, D., Li, Y., Gu, W., Yao, H., Liu, S., Yu, H., Pu, X., Zhou, J., Xu, J., Xi, C., Lei, H., Song, J., and Chen, S. (2018). Genome analysis of the ancient tracheophyte *Selaginella tamariscina* reveals evolutionary characteristics relevant to the acquisition of desiccation tolerance. *Molecular Plant*, 11(7):983–994.
- Yadav, M. K. and Bhatla, S. C. (2011). Localization of lipoxygenase activity on the oil bodies and in protoplasts using a novel fluorescence imaging method. *Plant Physiology and Biochemistry*, 49(2):230–234.
- Yan, R., Qian, H., Lukmantara, I., Gao, M., Du, X., Yan, N., and Yang, H. (2018). Human SEIPIN binds anionic phospholipids. *Developmental Cell*, 47(2):248–256.e4.
- Yang, L., Koo, D.-H., Li, Y., Zhang, X., Luan, F., Havey, M. J., Jiang, J., and Weng, Y. (2012a). Chromosome rearrangements during domestication of cucumber as revealed by high-density genetic mapping and draft genome assembly. *The Plant Journal*, 71(6):895–906.
- Yang, W., Simpson, J. P., Li-Beisson, Y., Beisson, F., Pollard, M., and Ohlrogge, J. B. (2012b). A land-plant-specific glycerol-3-phosphate acyltransferase family in *Arabidopsis*: substrate specificity, *sn-2* preference, and evolution. *Plant Physiology*, 160(2):638–652.
- Yang, Y. and Benning, C. (2018). Functions of triacylglycerols during plant development and stress. *Current Opinion in Biotechnology*, 49:191–198.
- Yang, Z., Ji, H., and Liu, D. (2016). Oil biosynthesis in underground oil-rich storage vegetative tissue: Comparison of *Cyperus esculentus* tuber with oil seeds and fruits. *Plant and Cell Physiology*, 57(12):2519–2540.
- Yang, Z., Liu, X., Wang, K., Li, Z., Jia, Q., Zhao, C., and Zhang, M. (2022). ABA-INSENSITIVE 3 with or without FUSCA3 highly up-regulates lipid droplet proteins and activates oil accumulation. *Journal of Experimental Botany*, 73(7):2077–2092.
- Ye, Z.-H. and Zhong, R. (2015). Molecular control of wood formation in trees. *Journal of Experimental Botany*, 66(14):4119–4131.
- Yokota, T., Ohnishi, T., Shibata, K., Asahina, M., Nomura, T., Fujita, T., Ishizaki, K., and Kohchi, T. (2017). Occurrence of brassinosteroids in non-flowering land plants, liverwort, moss, lycophyte and fern. *Phytochemistry*, 136:46–55.
- Yoshimoto, K. and Ohsumi, Y. (2018). Unveiling the molecular mechanisms of plant autophagy—from autophagosomes to vacuoles in plants. *Plant and Cell Physiology*, 59(7):1337–1344.
- Yotsui, I., Serada, S., Naka, T., Saruhashi, M., Taji, T., Hayashi, T., Quatrano, R. S., and Sakata, Y. (2016). Large-scale proteome analysis of abscisic acid and ABSCISIC ACID INSENSITIVE3-dependent proteins related to desiccation tolerance in *Physcomitrella patens*. *Biochemical and Biophysical Research Communications*, 471(4):589–595.

- Young, N. D., Debelle, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F. X., Gouzy, J., Schoof, H., Van de Peer, Y., Proost, S., Cook, D. R., Meyers, B. C., Spannagl, M., Cheung, F., De Mita, S., Krishnakumar, V., Gundlach, H., Zhou, S., Mudge, J., Bharti, A. K., Murray, J. D., Naoumkina, M. A., Rosen, B., Silverstein, K. A. T., Tang, H., Rombauts, S., Zhao, P. X., Zhou, P., Barbe, V., Bardou, P., Bechner, M., Bellec, A., Berger, A., Bergès, H., Bidwell, S., Bisseling, T., Choise, N., Couloux, A., Denny, R., Deshpande, S., Dai, X., Doyle, J. J., Dudez, A.-M., Farmer, A. D., Fouteau, S., Franken, C., Gibelin, C., Gish, J., Goldstein, S., González, A. J., Green, P. J., Hallab, A., Hartog, M., Hua, A., Humphray, S. J., Jeong, D.-H., Jing, Y., Jöcker, A., Kenton, S. M., Kim, D.-J., Klee, K., Lai, H., Lang, C., Lin, S., Macmil, S. L., Magdelenat, G., Matthews, L., McCarrison, J., Monaghan, E. L., Mun, J.-H., Najjar, F. Z., Nicholson, C., Noirot, C., O'Bleness, M., Paule, C. R., Poulain, J., Prion, F., Qin, B., Qu, C., Retzel, E. F., Riddle, C., Sallet, E., Samain, S., Samson, N., Sanders, I., Saurat, O., Scarpelli, C., Schiex, T., Segurens, B., Severin, A. J., Sherrier, D. J., Shi, R., Sims, S., Singer, S. R., Sinharoy, S., Sterck, L., Viollet, A., Wang, B.-B., Wang, K., Wang, M., Wang, X., Warfsmann, J., Weissenbach, J., White, D. D., White, J. D., Wiley, G. B., Wincker, P., Xing, Y., Yang, L., Yao, Z., Ying, F., Zhai, J., Zhou, L., Zuber, A., Dénarié, J., Dixon, R. A., May, G. D., Schwartz, D. C., Rogers, J., Quétier, F., Town, C. D., and Roe, B. A. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature*, 480(7378):520–524.
- Yu, X.-H., Rawat, R., and Shanklin, J. (2011). Characterization and analysis of the cotton cyclopropane fatty acid synthase family and their contribution to cyclopropane fatty acid synthesis. *BMC Plant Biology*, 11(1):97.
- Yusuf, A., Mamza, P., Ahmed, A., and Agunwa, U. (2015). Extraction and characterization of castor seed oil from wild *Ricinus communis* linn. *International Journal of Science, Environment and Technology*, 4(5):1392–1404.
- Zaaboul, F., Raza, H., Chen, C., and Liu, Y. (2018). Characterization of peanut oil bodies integral proteins, lipids, and their associated phytochemicals. *Journal of Food Science*, 83(1):93–100.
- Zehmer, J. K., Bartz, R., Liu, P., and Anderson, R. G. W. (2008). Identification of a novel N-terminal hydrophobic sequence that targets proteins to lipid droplets. *Journal of Cell Science*, 121(11):1852–1860.
- Zeng, X., Jiang, J., Wang, F., Liu, W., Zhang, S., Du, J., and Yang, C. (2022). Rice OsClo5, a caleosin protein, negatively regulates cold tolerance through the jasmonate signalling pathway. *Plant Biology*, 24(1):52–61.
- Zhang, H., Ohyama, K., Boudet, J., Chen, Z., Yang, J., Zhang, M., Muranaka, T., Maurel, C., Zhu, J.-K., and Gong, Z. (2008). Dolichol biosynthesis and its effects on the unfolded protein response and abiotic stress resistance in *Arabidopsis*. *The Plant Cell*, 20(7):1879–1898.
- Zhang, M., Fan, J., Taylor, D. C., and Ohlrogge, J. B. (2009). DGAT1 and PDAT1 acyltransferases have overlapping functions in *Arabidopsis* triacylglycerol biosynthesis and are essential for normal pollen and seed development. *The Plant Cell*, 21(12):3885–3901.
- Zhang, Q. and Bartels, D. (2018). Molecular responses to dehydration and desiccation in desiccation-tolerant angiosperm plants. *Journal of Experimental Botany*, 69(13):3211–3222.
- Zhang, S., Wang, Y., Cui, L., Deng, Y., Xu, S., Yu, J., Cichello, S., Serrero, G., Ying, Y., and Liu, P. (2016). Morphologically and functionally distinct lipid droplet subpopulations. *Scientific Reports*, 6(1):29539.

- Zhang, X., Ma, L., Tian, Y., Zhang, G., and Luo, Y. (2013). Molecular cloning and heterologous expression of putative (S)-norcoclaurine synthases from *Arabidopsis thaliana*. *Chinese Journal of Applied Environmental Biology*, 19(1):61–68.
- Zhao, C., Kim, Y., Zeng, Y., Li, M., Wang, X., Hu, C., Gorman, C., Dai, S. Y., Ding, S.-Y., and Yuan, J. S. (2018). Co-compartmentation of terpene biosynthesis and storage via synthetic droplet. *ACS Synthetic Biology*, 7(3):774–781.
- Zhao, J., Missihoun, T. D., and Bartels, D. (2017). The role of *Arabidopsis* aldehyde dehydrogenase genes in response to high temperature and stress combinations. *Journal of Experimental Botany*, 68(15):4295–4308.
- Zheng, X., Zhang, Y., Balakrishna, A., Liew, K. X., Kuijjer, H. N., Xiao, T. T., Blilou, I., and Al-Babili, S. (2023). Installing the *Neurospora* carotenoid pathway in plants enables cytosolic formation of provitamin A and its sequestration in lipid droplets. *Molecular Plant*.
- Zhi, Y., Taylor, M. C., Campbell, P. M., Warden, A. C., Shrestha, P., El Tahchy, A., Rolland, V., Vanhercke, T., Petrie, J. R., White, R. G., Chen, W., Singh, S. P., and Liu, Q. (2017). Comparative lipidomics and proteomics of lipid droplets in the mesocarp and seed tissues of chinese tallow (*Triadica sebifera*). *Frontiers in Plant Science*, 8.
- Zhou, N., Tootle, T. L., and Glazebrook, J. (1999). Arabidopsis PAD3, a gene required for camalexin biosynthesis, encodes a putative cytochrome P450 monooxygenase. *The Plant Cell*, 11(12):2419–2428.
- Zienkiewicz, A., Zienkiewicz, K., Rejón, J. D., Rodríguez-García, M. I., and Castro, A. J. (2013). New insights into the early steps of oil body mobilization during pollen germination. *Journal of Experimental Botany*, 64(1):293–302.
- Zienkiewicz, K., Castro, A. J., Alché, J. d. D., Zienkiewicz, A., Suárez, C., and Rodríguez-García, M. I. (2010). Identification and localization of a caleosin in olive (*Olea europaea* L.) pollen during in vitro germination. *Journal of Experimental Botany*, 61(5):1537–1546.
- Zienkiewicz, K. and Zienkiewicz, A. (2020). Degradation of lipid droplets in plants and algae—right time, many paths, one goal. *Frontiers in Plant Science*, 11.
- Zolman, B. K., Silva, I. D., and Bartel, B. (2001). The *Arabidopsis pxa1* mutant is defective in an atp-binding cassette transporter-like protein required for peroxisomal fatty acid  $\beta$ -oxidation. *Plant Physiology*, 127(3):1266–1278.
- Zoni, V., Khaddaj, R., Lukmantara, I., Shinoda, W., Yang, H., Schneiter, R., and Vanni, S. (2021). Seipin accumulates and traps diacylglycerols and triglycerides in its ring-like structure. *Proceedings of the National Academy of Sciences*, 118(10):e2017205118.
- Zubieta, C., Ross, J. R., Koscheski, P., Yang, Y., Pichersky, E., and Noel, J. P. (2003). Structural basis for substrate recognition in the salicylic acid carboxyl methyltransferase family. *The Plant Cell*, 15(8):1704–1716.
- Zuidmeer-Jongejan, L., Fernández-Rivas, M., Winter, M. G., Akkerdaas, J. H., Summers, C., Lebens, A., Knulst, A. C., Schilte, P., Briza, P., Gadermaier, G., and van Ree, R. (2014). Oil body-associated hazelnut allergens including oleosins are underrepresented in diagnostic extracts but associated with severe symptoms. *Clinical and Translational Allergy*, 4(1):4.
- Čopič, A., Antoine-Bally, S., Giménez-Andrés, M., La Torre Garay, C., Antonny, B., Manni, M. M., Pagnotta, S., Guihot, J., and Jackson, C. L. (2018). A giant amphipathic helix from a perilipin that is adapted for coating lipid droplets. *Nature Communications*, 9(1):1332.

## Supplementary materials

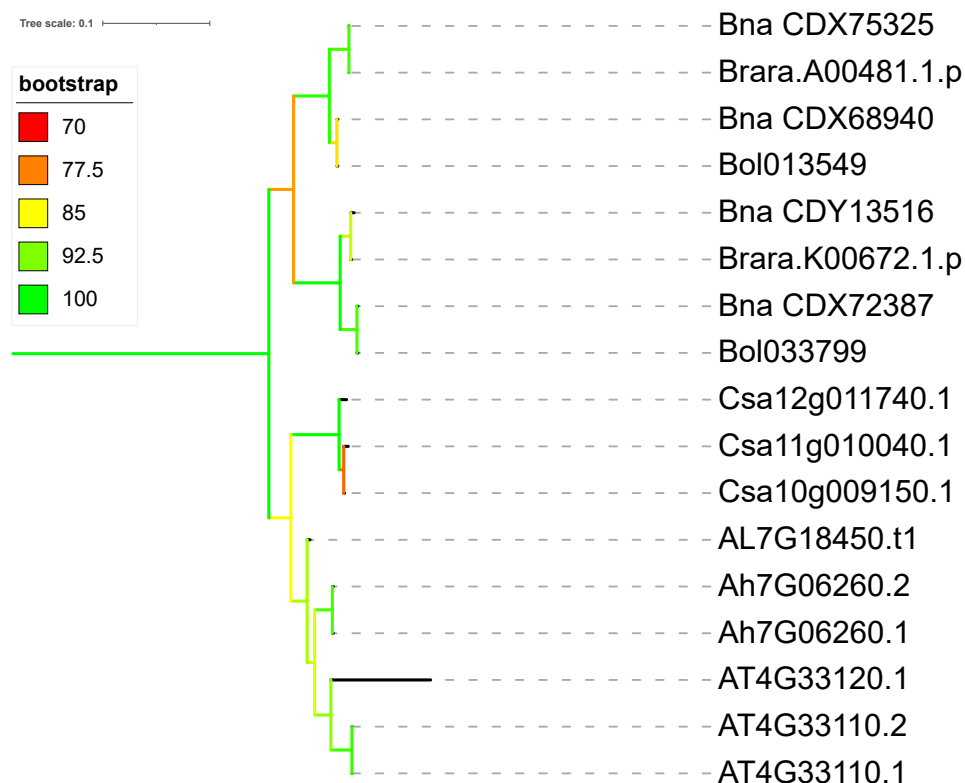
### Supplementary files

Additional files are stored on physical data drive. Files belonging to Manuscript I are included in a compressed .ZIP file.

**Supplemental Dataset 1A,B,C:** A) Proteomic datasets of root TE and LD-enriched protein fractions, B) RAW data from volcano plot, C) root LD-candidate list

**Supplemental Dataset 2:** Proteomic dataset of ColP using free mCherry and LIME1-mCherry as a bait transiently expressed in *N. benthamiana*.

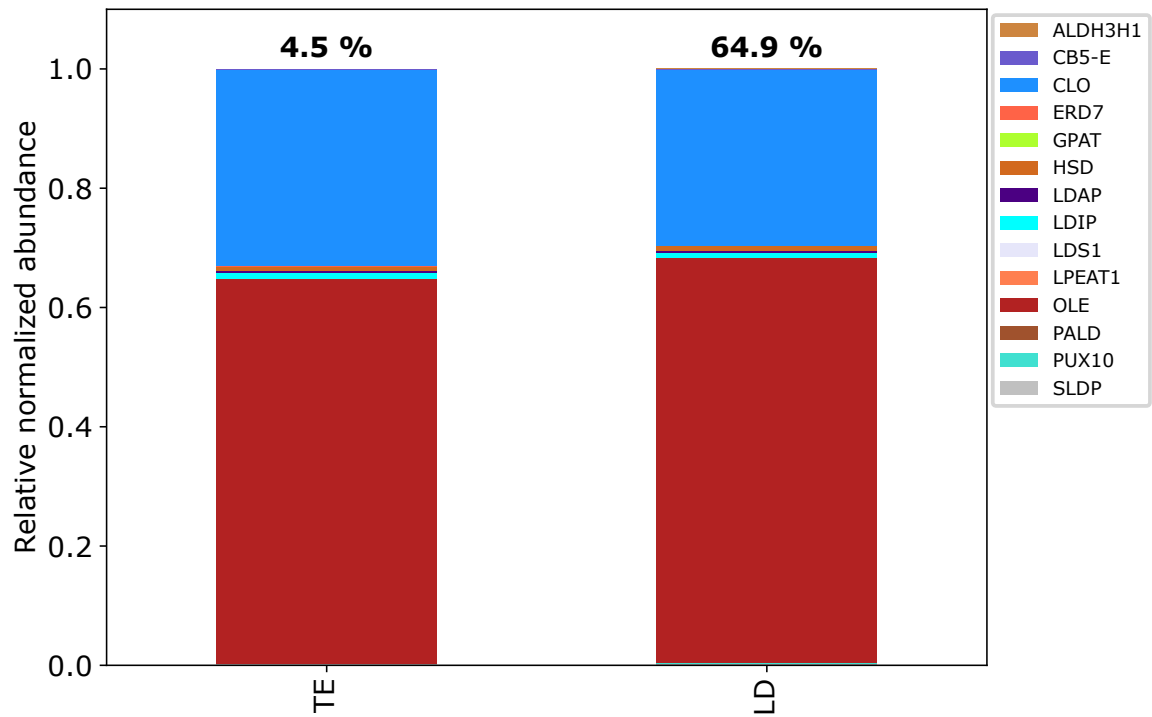
### Supplementary figures



**Supplemental Figure 1:** Maximum likelihood phylogenetic tree of AtLIME1 homologs found in selected transcriptome/genome derived proteomes from Viridiplantae via BLASTp search ( $e$  value cut off  $10^{-100}$ ). BLASTp resulting proteins were aligned with MAFFT L-INS-i and the phylogenetic tree was calculated with W-IQ-TREE. Branch validation was performed by computing 100 bootstrap replicates. Only bootstraps in the range 70-100 are displayed. Tree was pruned except for sequences belonging to Brassicaceae. Bna, *Brassica napus*; Brara, *Brassica rapa*; Bol, *Brassica oleracea*; Csa, *Camelina sativa*; AL, *Arabidopsis lyrata*; Ah, *Arabidopsis halleri*; AT, *Arabidopsis thaliana*.



**Supplemental Figure 2:** Seven-week-old Arabidopsis lines Col-0, *lime1-1*, *lime2-1*, *lime2-2*, *lime double\_2*, *lime double\_5*. The mutations were not associated with any developmental defects.



**Supplemental Figure 3:** Normalized distribution of known LD proteins in total (TE) and LD-enriched (LD) protein fractions of dry tubers of yellow nutsedge (*Cyperus esculentus*). The numbers displayed above each column indicate the percentage of the total sum of assigned LD proteins within the corresponding protein fraction.

## Acknowledgements

Das Anfertigen einer Doktorarbeit ist ein langer Prozess, in meinem Fall 3,5 Jahre. Eine Erfahrung, an der ich gewachsen bin. Und natürlich würde es ohne die Unterstützung von vielen Leuten dieser Arbeit nicht geben. Diesen möchte an dieser Stelle gerne herzlichst danken.

Zuallererst will ich meinem Doktorvater, Till Ischebeck, Danke sagen. Till hat mich mit offenen Armen in seiner Abteilung willkommen geheißen und mir immer viele Freiheiten gelassen, wie ich meine Projekte gestalten möchte. Dabei hatte er im Hintergrund immer eine stützende Funktion innegehabt. Und obwohl es durch COVID und Umzug der Abteilung nach Münster turbulente Zeiten waren, konnten wir am Ende doch ansehbare Resultate erreichen. Außerdem konnte ich so an seinen legendären Grillabenden mit dem sehr leckeren Essen teilhaben, an die ich mich gerne erinnern werde. Es war so, als wäre man ein vollwertiges Mitglied der Ischebeck Familie. Und dafür nochmal, herzlichsten Dank, Till!

Außerdem möchte ich gerne Prof. Dr. Ivo Feußner danken, nicht nur, da er ohne zu zögern ja gesagt hat, zu meinem Examination Board zu zählen, sondern auch, da er immer eine schützende Hand über seine (assoziierten) Mitarbeiter hielt. Das hat sich besonders durch den hervorragenden Infektionsschutz verdeutlicht wird, was in den Pandemie Jahren für eine sichere Arbeitsatmosphäre sorgte. Ich habe mich in den Räumlichkeiten der Abteilung Biochemie der Pflanze immer sehr wohlgeföhlt und bin dankbar, dass ich dort sein und die Gerätschaften benutzen durfte.

Ein großen Dank an meine weiteren Mitglieder des Thesis Advisory Committees, Prof. Dr. Andrea Polle, Prof. Harry Brumer, PhD, und Prof. Dr. Jan de Vries. Zum einen für die regelmäßige konstruktive Kritik des Fortschritts, zum anderen für die Teilnahme am Examination Boards. Darüber hinaus möchte ich Prof. Dr. Gerhard Braus und Prof. Dr. Kai Heimel für die Teilnahme meines Examinations Boards bedanken.

A special thanks to Harry Brumer. I had the opportunity for a small stay in his lab at the UBC, Vancouver, in the beautiful British Columbia. It was a warm, friendly atmosphere and a great experience. Usually, the first day of work does not start with a hike on a remote Canadian island, but I enjoyed that memorable day. Also I want to thank Jessica Fong for her time, her experience, and her patience with me in the lab. Thanks for having me, Harry and Jess, you were fantastic hosts.



Natürlich lässt es sich sehr gut arbeiten, wenn das Umfeld stimmt. Danke an die nette Atmosphäre an meine (ehemaligen) Kollegen aus der Biochemie der Pflanze und der Arbeitsgruppe Ischebeck, in Göttingen und Münster, als auch aus der Gruppe de Vries. Ihr habt jeden Tag etwas besser gemacht.

Es waren große Fußstapfen, die Anna, Franzi, Athanas, Elisa und Patricia hinterlassen haben in der AG Ischebeck und ich hoffe, ich kann ein weiteres Teilstück überreichen für die neue Generation von Doktoranden, Siqi, Magdiel und Janis. Ihr werdet es auch packen! Und auch den beiden PostDocs Lea und Kathi möchte ich vom Herzen danken, ihr habt den Laden am laufen gehalten und für eine gute Atmosphäre gesorgt, auch wenn ich das nur aus der Ferne sehen konnte.

Besonders hervorheben möchte ich Patricia und Benedikt, der nie zum Spaß da war 😊. Ohne euch hätte es nicht halb so viel Spaß gemacht. Und ohne eure produktiven (oder eher konspirativen?) Kuchen Pausen hätte es das Wurzel Projekt nie so erfolgreich gegeben. Außerdem wusste ich eure Bereitschaft, jede Frage zu diskutieren, und bei Problemen zu helfen, sehr zu schätzen. Danke euch! Ich möchte den beiden Lab-Rotation Studenten Ivana Gavrilova und Fabienne Dreier für ihr Mitarbeit an meinen Projekten bedanken. Fabienne war eine Glücksbringerin, bei der bereits der erste gescreente LD Protein Kandidat positiv war. Das war der Grundstein des Erfolgs für das Wurzel LD Projekt. Außerdem möchte ich noch gerne Susanne Mesters danken, die immer gut im Hintergrund die Pflanzen gepflegt hat und immer auch mal für unkonventionelle Anbauversuche (z.B. Erdmandeln) bereit war. Ein besonderen Dank an Dr. Ellen Hornung, die immer ein offenes Ohr hatte und sehr viele technische Fragen zu beantworten wusste und außerdem immer mit sehr leckerem Essen und Kuchen aufwarten konnte. Außerdem danke ich Dr. Kirstin Feußner, Dr. Tegan Haslam und Sabine Freitag für die immer sehr nette Bereitstellung ihrer wissenschaftlichen und ausführenden Expertisen. Und ohne den hervorragenden technischen Support von Malte und Tarek hätte das Labor vermutlich still gestanden. Ihr seid unentbehrlich.

Es war mir eine besondere Freude für sämtlicher meiner Projekte mit Dr. Kerstin Schmitt und Dr. Oliver Valerius zusammen arbeiten zu können für die LC-MS/MS Messungen. Nicht nur wurden Proben in rekordverdächtiger Zeit gemessen, auf die viele Leute neidisch waren, sondern ihr habt mit eurer Expertise für eine stetige Verbesserung meiner Arbeitsabläufe gesorgt.

Außerdem möchte ich sämtlichen Kollaborationspartnern und Co-Autoren danken. Also Jan, Jacco, Iker, Per, Kerstin, Ole, Prof. Braus, Janis, Dennis, Lea, Patricia, Conni, Prof. Feußner, Shrikant, Anders, und natürlich Till. Denn Wissenschaft ist Teamwork.

Des Weiteren will ich den Mitgliedern des IRTGs danken, die eine starke Gemeinschaft sind. Allen voran danke an Steven, der ein hervorragender Student Representative war und mit einem ungeheurem Organisationstalent aufwarten konnte. Außerdem Judith und Janina, die vieles der Bürokratie im Hintergrund erledigt haben. Ihr seid Gold wert. Des weiteren wollte ich mich an dieser Stelle an Lina bedanken für die nette Zeit in Vancouver als Mitbewohnerin. Ich verspreche, ich werde nie wieder 5 Kg Äpfel auf Vorrat kaufen 😊.

Ich möchte auch Anton Farr danken, ohne den mein  $\LaTeX$ Dokument vermutlich immer noch einen schwerwiegenden Kompilierungs Fehler hätte und meine Literatur seltsam gegliedert wäre. Mein Computer-Experte des Vertrauens.

Einen großen Dank an meine Familie, meinen Großeltern Anne, Gerd †, Waltraud †, Walter †, meinen Eltern Eckhard und Susanne, und meinem großen Bruder Nico und allen weiteren Mitgliedern der Familie für ihre für mich lebenslange Unterstützung in jeglicher Hinsicht.

Meiner Freundin Anni gebührt großen Respekt und meinen Dank, denn sie hat doch einige verzweifelte Momente am Ende der Schreibphase miterleben müssen. Ich kann dir gar nicht dankbar genug sein, dass du mich trotzdem die ganze Zeit liebevoll unterstützt hast. Ohne deine konstruktive Kritik hätte ich vermutlich niemals alle Plots in Python gemacht oder das Dokument in  $\LaTeX$  erstellt. Und das hat dem Layout tatsächlich gut getan.

Ich möchte Anni, Patricia, Tegan, Nico und Cece für das sorgfältige Korrekturlesen in ihrer Freizeit bedanken. Es war sicherlich nicht das Spannendste was man hätte machen können, und dennoch war es für mich sehr hilfreich. Ich hoffe, ich kann euch diesen Gefallen einmal erwidern.

Das letzte Dankeschön widme ich allen Freunden, Kollegen und Bekannten, die nicht namentlich erwähnt, aber mit gemeint waren. Ihr seid eine tragende Säule, ohne euch wäre ich nicht so weit gekommen. DANKESCHÖN!