

# PROTEIN TURNOVER ON PLANT LIPID DROPLETS

Dissertation

For the award of the degree

*“Doctor rerum naturalium”*

of the Georg-August-Universität Göttingen

within the doctoral program

“Molecular Biology”

of the Georg-August University School of Science (GAUSS)

Submitted by

Franziska Kerstin Kretzschmar

From Leipzig, Germany

Göttingen, April 2019

Department of Plant Biochemistry

Albrecht-von-Haller-Institute for Plant Sciences

Georg-August-Universität Göttingen

## Thesis Advisory Committee

### **PD Dr. Till Ischebeck**

Department of Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences, Georg-August-Universität Göttingen

### **Dr. Alexander Stein**

Membrane Protein Biochemistry, Max-Planck-Institute for Biophysical Chemistry

### **Prof. Dr. Christiane Gatz**

Department of Plant Molecular Biology and Physiology, Albrecht-von-Haller-Institute for Plant Sciences, Georg-August-Universität Göttingen

## Members of the Examination Board

Referee: **PD Dr. Till Ischebeck**, Department of Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences, Georg-August-Universität Göttingen

2<sup>nd</sup> Referee: **Dr. Alexander Stein**, Membrane Protein Biochemistry, Max-Planck-Institute for Biophysical Chemistry

Other members of the Examination Board:

**Prof. Dr. Christiane Gatz**, Department of Plant Molecular Biology and Physiology, Albrecht-von-Haller-Institute for Plant Sciences, Georg-August-Universität Göttingen

**Prof. Dr. Gerhard Braus**, Department for Molecular Microbiology and Genetics, Institute for Microbiology and Genetics, Georg-August-Universität Göttingen

**Prof. Dr. Jörg Stülke**, Department for General Microbiology, Institute for Microbiology and Genetics, Georg-August-Universität Göttingen

**PD Dr. Marcel Wiermer**, RG Molecular Biology of Plant-Microbe Interaction, Albrecht-von-Haller-Institute for Plant Sciences, Georg-August-Universität Göttingen

**Date of oral examination: 05.06.2019**

## Affidavit

Herewith, I declare that that I prepared this doctoral thesis “Protein Turnover on Plant Lipid Droplets” on my own and with no sources and aids other than quoted.

Franziska K. Kretzschmar

Göttingen, April 2019

# Table of Contents

List of Abbreviations.....	I
Abstract.....	III
1. Introduction.....	1
1.1. <i>Arabidopsis thaliana</i> – an Oil-Seed Model Organism .....	1
1.1.1. The Seed.....	1
1.1.2. Seed Storage Reserves .....	1
1.1.3. Oil Seed Plants.....	2
1.2. The Lipid Droplet – a Neutral Lipid Storage Organelle.....	4
1.3. The Biogenesis of Lipid Droplets .....	8
1.4. The Degradation of Lipid Droplets.....	12
1.4.1. The Regulation and Function of Neutral Lipid Breakdown.....	12
1.4.2. Degradation of other Lipid Droplet Components .....	13
1.4.2.1. The Degradation of Coat Proteins .....	13
1.4.2.2. The Degradation of Phospholipids.....	14
1.5. Functions of Lipid Droplets in Non-Lipid-Storing Tissues.....	15
1.5.1. Lipid Droplet Function in Anther Development .....	15
1.5.2. Functions of Lipid Droplets in Pollen Tube Growth.....	15
1.5.3. Lipid Droplets in Leaves.....	16
1.5.4. Production Site for Antimicrobial Compounds .....	17
1.6. Aims of this Study.....	18
2. Article I: PUX10 is a lipid droplet-localized scaffold protein that interacts with CELL DIVISION CYCLE48 and is involved in the degradation of lipid droplet proteins .....	19
3. Manuscript II: Calculation of enrichment factors identifies low abundant lipid droplet-associated proteins from <i>Arabidopsis thaliana</i> seeds and seedlings.....	75
4. Article III: Arabidopsis lipid droplet-associated protein (LDAP) – interacting protein (LDIP) influences lipid droplet size and neutral lipid homeostasis in both leaves and seeds.....	128
5. Discussion.....	149
5.1. The Proteome of Lipid Droplets in Plants and Beyond.....	149



5.1.1.	The Composition of the Plant Lipid Droplet Proteome .....	149
5.1.2.	Conservation of Plant Lipid Droplet-associated Proteins throughout Evolution .....	153
5.1.3.	Targeting of Lipid Droplet-associated Proteins .....	156
5.2.	Degradation of LDs and LD proteins .....	159
5.2.1.	The Degradation of Lipid Droplet Proteins by the Ubiquitin Proteasome System .....	159
5.2.2.	The Degradation of Lipid Droplets by Autophagy .....	162
6.	Concluding Remarks .....	164
7.	References .....	165
	Acknowledgements .....	180

## List of Abbreviations

<b>ATGL</b>	Adipose triglyceride lipase
<b>CDC48</b>	Cell Division Cycle 48
<b>CIDE</b>	CELL DEATH-INDUCING DFF45-LIKE EFFECTOR
<b>CLO</b>	Caleosin
<b>COP</b>	Coat protein complex
<b>CRISPR</b>	Clustered regularly interspaced short palindromic repeats
<b>DAG</b>	Diacylglycerol
<b>DGAT</b>	Diacylglycerol-Acyltransferase
<b>DOX</b>	Dioxygenase
<b>ER</b>	Endoplasmic reticulum
<b>ERAD</b>	ER-associated degradation
<b>FA</b>	Fatty acid
<b>FIT</b>	Fat storage-inducing transmembrane protein
<b>GFP</b>	Green fluorescent protein
<b>HOT</b>	hydroxide
<b>HPOT</b>	hydroperoxyd
<b>HSD</b>	Hydroxysteroid dehydroxygenase (Steroleosin)
<b>HSL</b>	Hormone-sensitive lipase
<b>LC</b>	Liquid Chromatography
<b>LD</b>	Lipid droplet
<b>LDAP</b>	LD-associated protein
<b>LDIP</b>	LDAP-interacting protein
<b>LDPS</b>	Lipid droplet-associated protein of seeds
<b>LEC2</b>	LEAFY COTELYDON2
<b>LIDL</b>	Lipid droplet-associated lipase
<b>LIME</b>	Lipid droplet-associated methyltransferase
<b>LOX</b>	Lipoxygenase
<b>MAG</b>	Monoacylglycerol
<b>MLDP</b>	Major lipid droplet protein
<b>MS</b>	Mass spectrometry
<b>OBL</b>	Oil body lipase
<b>OLE</b>	Oleosin
<b>PA</b>	Phosphatidic acid
<b>PC</b>	Phosphatidylcholine
<b>PE</b>	Phosphatidylethanolamine

<b>PL</b>	Phospholipid
<b>PLA</b>	Phospholipase A
<b>PLIN</b>	Perilipin
<b>PSV</b>	Protein storage vacuole
<b>PTLD</b>	Pollen tube lipid droplet protein
<b>PUX</b>	Plant UBX-domain containing protein
<b>SDP1</b>	SUGAR-DEPENDENT1
<b>SDP1-L</b>	SUGAR-DEPENDENT1-LIKE
<b>SE</b>	Sterol estera
<b>SLDP</b>	Seed lipid droplet-associated protein
<b>TAG</b>	Triacylglycerol
<b>Ub</b>	Ubiquitin
<b>UBX</b>	Ubiquitin regulatory X domain
<b>UPS</b>	Ubiquitin Proteasome System
<b>WR11</b>	WRINKLED1
<b>WT</b>	Wildtype

## Abstract

Lipid droplets (LDs) are lipid storage organelles found across all kingdoms of life. In recent years, the perception of the organelle has shifted from inactive lipid storage sites to dynamic organelles central to the lipid metabolism. In plants, LDs are best studied in the seed of oil seed plants like the model plant *Arabidopsis thaliana* or the related crop *Brassica napus*.

LDs have a unique membrane topology, because they consist of a phospholipid monolayer that shields the neutral lipid core from the aqueous environment of the cytosol. This special topology requires a unique set of proteins to associate with the organelle.

The most abundant LD proteins in both plants and animals are coat proteins that are not conserved between the two kingdoms. In plants, oleosins, steroleosins and caleosins are embedded in the phospholipid monolayer and are thought to be anchored into the neutral lipid core through hydrophobic domains. Enzymatic activities have been observed for steroleosins and caleosins, and oleosins have been shown to shield LDs from each other to keep them from coalescing. However, these three proteins alone are not able to fully describe the dynamic role of LDs in different tissues and different developmental stages. Therefore, efforts have been taken to expand the LD proteome to help investigate the many open questions that remain about LD biology: their biogenesis, functions, interactions in the cellular environment, and the breakdown of their component.

In this thesis, I present a bottom-up proteomics approach of LD-enriched fractions of tobacco pollen tubes, and *Arabidopsis* siliques, seeds and seedlings. By quantitative comparison to total cellular extracts followed by cell biological studies, I could contribute to the discovery of eight new plant LD proteins or protein families. Within these are three protein families annotated as unknown, three proteins or protein families with putative enzymatic activity in the lipid metabolism, one protein family conserved in plants except in Brassicaceae, and a scaffold protein whose homologs in other systems are involved in protein degradation pathways.

For this scaffold protein, PUX10, we could confirm its involvement in protein degradation; specifically at the LD. Mutants of this protein are delayed in the degradation of LD coat proteins, mainly oleosins, during seedling establishment.

# 1. Introduction

## 1.1. *Arabidopsis thaliana* – an Oil-Seed Model Organism

### 1.1.1. The Seed

The seed is one of the most crucial adaptations to land life during plant evolution. Together with the development of pollen, the seed was responsible for the rapid spread of seed plants, or Spermatophytes, which began at the end of the Carbon age when the climate started to get drier (Campbell, 1997). Today, seeds are of enormous economic importance: 70 % of the food we eat comes directly from seeds; most of the remaining 30 % are derived from animals that are also predominantly fed with seeds. In addition, seed-derived products are also important for industrial applications like lubricants and oils for both cosmetics, machinery and pharmaceutical industry (Bewley et al., 2013). The seeds of angiosperms, while varying greatly in size and appearance between the numerous species, all have the same principle components: the embryo, non-embryonic storage tissue, and a seed coat (Figure 1). The seed coat (or testa) is the final protective barrier between the embryo and the environment. Seed coats, too, are highly variable between species, but often contain higher concentration of protective molecules like waxes or pectin. Two major forms of non-embryonic storage tissue, usually mutually exclusive in matured seeds, can be distinguished: the perisperm and the endosperm. The diploid perisperm will not be discussed here in any more detail. The endosperm, which is triploid due to the fusion of one pollen tube nucleus with the two polar nuclei of the central cell, can be the major storage sites for reserve compounds. Depending on the endosperm to embryo size ratio, two seed variants can be distinguished: endospermic and non-endospermic seeds. In endospermic seeds, the endosperm serves as major storage site, and the embryo remains small, as it is the case for cereals like wheat and rice. In non-endospermic seeds, for example of many legumes, the embryo itself contains the storage compounds. In all of the cases, however, the embryo inside the seeds is always composed of one or two cotyledons, the shoot apex, and an embryonic axis, which contains the embryonic root (also called radicle) and the hypocotyl.



**Figure 1: Components of the mature *Arabidopsis thaliana* seed.** The mature embryo (yellow), which constitutes the majority of the volume in the oil seed *Arabidopsis thaliana*, is surrounded by a single-cell but robust endosperm (orange). The seed coat (green) protects the embryo from the environment but is also the first barrier the embryo has to overcome to germinate.

Figure from (Sharma et al., 2012)

### 1.1.2. Seed Storage Reserves

The aspect that makes seeds so economically interesting is that they store high amounts of nutritious, energy dense biomolecules: carbohydrates, proteins, and neutral lipids mostly in the

form of triacylglycerols (TAGs). Both the embryo and the endosperm may contain storage reserves, in varying composition and concentration. While all three major reserves can be found in the seeds of most economically important species, there is usually a bias towards one or not more than two of them. Cereals, for example, all store majorly carbohydrates in their endosperm, usually around three quarters of their dry weight (Table 1). Additionally, they contain around 10 % of dry weight as protein reserves, and only minor amounts of oil. Legumes and other crops may contain carbohydrates (beans and peas), proteins (soybean), or oil (peanut, rape seed) as their major storage compound.

**Table 1:** Composition of major storage compounds (in percent) of selected crops and the model plant *Arabidopsis thaliana*. Table adapted from (Bewley et al., 2013).

Plant	Protein	Oil	Carbohydrates	Major storage site
Barley	12	3	76	Endosperm
Rice	7	3	75	Endosperm
Garden pea	25	6	52	Cotyledons
Soybean	37	17	26	Cotyledons
Rape ( <i>Brassica napus</i> )	21	48	19	Cotyledons
<i>Arabidopsis thaliana</i> (Baud et al., 2008)	30 – 40	30 – 40	2	Cotyledons

While the localization of the storage reserve within the seed may vary between endosperm and embryo, their subcellular localizations are more conserved. Storage proteins, which can be divided into 4 classes based on their solubility (Osborne, 1924), are usually deposited in a specialized organelle called the protein storage vacuole (PSV). The most common carbohydrate stored in seeds is starch. It is synthesized in the amyloplast, a non-photosynthetic plastid, and stored in granules inside this plastid. Another specialized organelle serves as neutral lipid storage: the lipid droplet (LD, also known as lipid body, oil body, spherosome or oleosome). LDs can be found in both endosperm (castor bean) and the embryo (rapeseed) as major sites of oil storage.

### 1.1.3. Oil Seed Plants

Oil-seed rape (*Brassica napus*) is the second most important oil seed crop in the world, behind soybean (Sharma et al., 2012). Its relevance, and therefore world production, increased in recent years, and it is the major oil-seed grown in the European Union. Despite its economic importance, genomic studies on *Brassica napus* are notoriously challenging. This is due to fact that the species has a allopolyploid genome, created through its origin as a hybrid of *Brassica rapa* and *Brassica oleacea*, followed by a genome duplication event (Chalhoub et al., 2014). Consequently, Chalhoub et al., 2014 identified more than 100,000 putative genes in the *B. napus* genome. However, the model organism of plant biology, *Arabidopsis thaliana*, is from the same family, Brassicaceae, like *Brassica napus*. The *Arabidopsis* genome was the first plant genome sequenced (Arabidopsis Genome Initiative, 2000). Researchers have long been able to manipulate the diploid, small, and gene-rich *Arabidopsis* genome via Agrobacterial-mediated transformation (Clough and Bent,

1998), and recently, advances in the CRISPR-Cas9 gene editing technology have also been expanded to this model organism (Wang et al. 2015). Like *Brassica napus*, *Arabidopsis* stores high concentrations of neutral lipids in its embryo (Table 1). While the amount of protein reserves are equally high, the lipids and the ability of the embryo to mobilize them during germination are determining germination success (Eastmond, 2006).

As mentioned earlier, neutral lipids are generally stored in a specialized organelle called LDs. The following chapters will summarize relevant findings on the biogenesis, functions, and turnover of LDs in plants, with many of the studies being conducted on *Arabidopsis thaliana*. Furthermore, these results are compared to findings in yeast and animals.

## 1.2. The Lipid Droplet – a Neutral Lipid Storage Organelle

The ability to store hydrophobic substances in specialized, membrane-separated compartments can be found in all domains of life. Many prokaryotes and some archaea species form some type of LDs where they store liquid or semi-solid hydrophobic biomolecules, most commonly polyhydroxybutyrates (Murphy, 2012).

In eukaryotes, it becomes more and more clear that most, if not all, cells are able to synthesize LDs and also contain them at some time or under certain conditions, if one “look(s) hard enough” (Murphy, 2001). In mammals, LDs are most prominent, and best studied, in white and brown adipose tissue, steroidogenic cells, mammary gland epithelial cells, different leukocytes including macrophages, hepatocytes and enterocytes, and the pigmented epithelial cells in the retina (Murphy, 2012). In plants, in addition to oil seeds containing large amounts of LDs, the organelle can be found in high numbers in pollen and pollen tubes, and the tapetum. To a lesser extent, LDs are also found in vegetative tissues like roots and leaves. In normal leaves, the highest number of LDs are found in stomatal cells, but they accumulate generally in leaf cells during heat and drought stress, and leaf senescence (Gidda et al., 2016; Kim et al., 2016a; Shimada et al., 2015). Additionally, LDs can be found in the oil-rich mesocarp of some fruits, like avocado (*Persea americana*) and olive (*Olea europaea*).

In plants, LD-like particles can also be found in plastids, where they are referred to as plastoglobules or plastoglobuli. Plastoglobules can vary in shape, size, and composition depending on the type of plastid in which they are formed (van Wijk and Kessler, 2017).

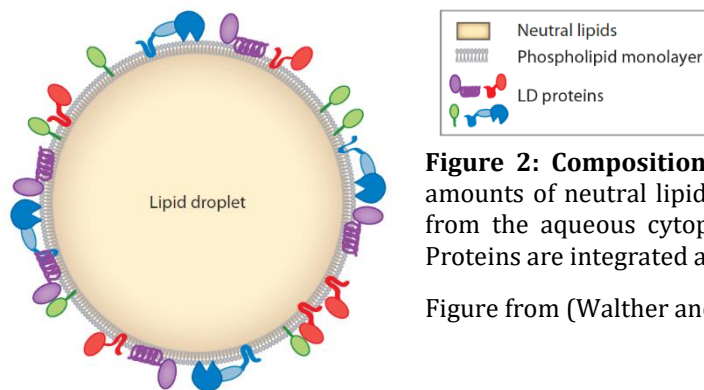
In animals, plants and fungi, the most common neutral compounds stored in LDs are either TAGs or sterol esters (SEs), or a combination of the two. In mammals, for example, LDs of adipocytes store mostly TAG, but those in steroidogenic cells contain predominantly cholesterol-esters needed for hormone production (Onal et al., 2017). In Arabidopsis seeds, SEs make up less than one percent of all neutral lipids (Bouvier-Nave et al., 2010), whereas in tobacco pollen, SEs contribute 24 % to neutral lipid weight (Rotsch et al., 2017). Within LDs with mixed neutral lipid composition, the different neutral lipids probably do not occur as a homogenous mass. Rather, it seems like different SEs form distinct shells around a TAG core, like it was observed in yeast (Czabany et al., 2008).

In the plant kingdom, other interesting hydrophobic compounds can be found as the main storage molecule in LDs. Jojoba (*Simmondsia chinensis*) seeds are filled with LDs that contain wax esters as their hydrophobic matrix. This unique composition makes jojoba interesting for both cosmetic, pharmaceutical, and chemical industry, where wax esters find a range of applications (Al-Obaidi et al., 2017). Another commercially important hydrophobic compound stored in LD-like particles



is natural rubber of the rubber tree (*Hevea brasiliensis*), which consists mainly of polyisoprenes (Nawamawat et al., 2011). In the liverwort *Marchantia polymorpha*, the LDs contain mainly isoprenoids, many of which have been shown to have interesting pharmacological functions (He et al., 2013).

In all cases, the hydrophobic core is separated from the aqueous cytoplasm by a phospholipid (PL) monolayer (Figure 2).



**Figure 2: Composition of Lipid Droplets.** LDs store large amounts of neutral lipids. The neutral lipid core is separated from the aqueous cytoplasm via a phospholipid monolayer. Proteins are integrated and associated with this monolayer.

Figure from (Walther and Farese Jr., 2012)

The most abundant PL of this monolayer is in most cases phosphatidylcholine (PC). The exact composition of the phospholipidome, just like that of the neutral lipids, can vary. In both yeast and humans, for example, phosphatidylserine and phosphatidic acid (PA) are mostly absent from LD membranes, while lyso-PC and lyso-phosphoethanolamine (PE) are found to be enriched in LDs isolated from cultured cells (Bartz et al., 2007; Grillitsch et al., 2011; Onal et al., 2017). In LDs of *Brassica napus* seeds, PA is one of three most abundant PL classes (Katavic et al., 2006).

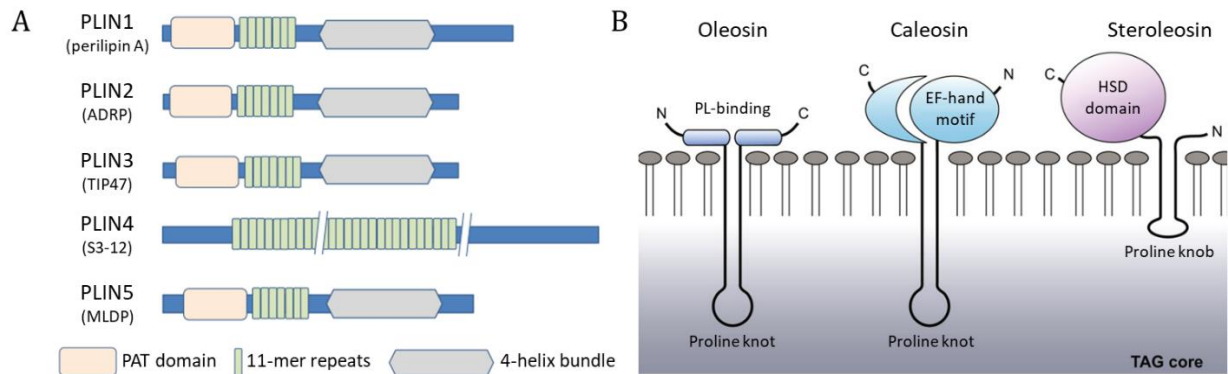
Like other membranes in the cell, the PL monolayer of the LDs is decorated with a selection of proteins. LDs in all systems are collectively covered with coat proteins, a principle otherwise known from cellular vesicles (Murphy, 2012). Between plants and animals, the most abundant LD “coat” proteins are not conserved.

In mammals, LD membranes harbor a family of five proteins called Perilipins (PLINs, Kimmel et al., 2010). PLINs are conserved from slime molds to humans, but not found in the plant lineage (Murphy, 2012). Although PLINs lack a classical transmembrane domain, there is mounting evidence that a conserved 11-mer repeats domain of the proteins, which is proposed to form amphipathic helices, contributes to LD targeting (Figure 3, Čopič et al., 2018; Rowe et al., 2016). While PLINs are expressed in all LD-enriched mammalian tissues (white adipocytes – PLIN1 and PLIN4, liver – PLIN2, brown adipocytes – PLIN5), and PLIN3 is even expressed ubiquitously, all these LDs contain TAG as major storage lipid. Interestingly, on cholesterol ester-storing LDs from mouse tumor cells two enzymes were found as major proteins (Yamaguchi et al., 2015).

Another family of proteins found on mammalian LDs is also among the major LD coat proteins in plants: hydroxysteroid dehydrogenases (HSD, or steroleosins in plants (Lin et al., 2002)). Together with the oleosin and caleosin protein families, steroleosins constitute the majority of the seed LD proteome (Chapman et al., 2012). Steroleosins contain the HSD domain C-terminally to a hydrophobic domain responsible for LD localization (Lin et al., 2002). Due to the conservation between plant and mammalian HSDs, related enzymatic activities are assumed between the enzymes. For example, AtHSD1 can convert estradiol to its biologically less active, dehydrogenated form estrone *in vitro* (Lin et al., 2002). Also in plants, steroid-based hormones (brassinosteroids) play important roles in developmental processes (Clouse and Sasse, 2002). Alterations in AtHSD1 expression resulted in phenotypes that mimic those of plants with altered brassinosteroid metabolism (Li et al., 2007).

While oleosins are only found in plants, caleosin homologs have been detected in both plants and fungi (Partridge and Murphy, 2009). Oleosins possibly derived from evolutionary older caleosin genes after genome duplication events. As opposed to PLINs, specific hydrophobic domains responsible for LD anchoring have been identified in all three major plant LD coat protein families. The hydrophobic domains are quite long and contain a proline knot, in case of oleosins and caleosins, or a proline knob, in steroleosin. For oleosin, the proline knot is necessary for LD targeting, variants where the prolines were replaced by leucines do not retarget from the ER membrane, where the proteins are co-translationally inserted, to the growing LD (Abell et al., 1997). The two cytosolic domains of oleosins are proposed to fold into amphipathic helices, integrating into the PL monolayer via the hydrophobic side while exposing charged residues towards the cytosol (Tzen et al., 1992). The exact three-dimensional structure of any of these protein families is yet to be elucidated, but it is proposed that both N-terminal and C-terminal regions of the proteins face the cytosol while the hydrophobic middle part reaches into the lipid droplet, where a hairpin structure is formed by the proline residues (Figure 3). Circular dichroism structural analysis of an Arabidopsis oleosin heterologously expressed in *S. cerevisiae* revealed  $\beta$ -sheets as the dominant fold of the hydrophobic sequence of the protein (Vindigni et al., 2013). Oleosins are thought to be mere structural proteins. They are specific to land plants, and within those, to organs affected by desiccation, like seeds, pollen, the tapetum and the spores of mosses (Huang et al., 2009). They are less abundant in tropical plants where seeds are less effected by desiccation and do not require a desiccation phase for successful germination. Additionally to increasing LD stability during dessication, a protective effect of oleosins has also been found during cold stress. Single mutants of three abundant seed oleosins germinated at a much lower rate after freezing treatment (Shimada et al., 2008). The observed effect was even more severe in different double mutant combinations. The germination phenotype was combined with the

observation that freeze-treated mutant seeds contained much larger LDs than WT or mutants that had not been exposed to cold stress.



**Figure 3: Domain structures of LD coat proteins in mammals and plants.** The major LD proteins in mammals are the five-member Perilipin (PLIN) protein family (A). They all share a region of 11-mer repeats that is important for LD targeting. Four of the five PLINs additionally contain the PAT domain, and a 4-helix bundle also implicated in LD localization. In plants, there are three families of major coat proteins: Oleosins, caleosins, and steroleosins (B). All of them are anchored into the hydrophobic LD core through a hydrophobic domain containing a specific proline motif. Caleosins and steroleosins additionally harbor protein domains with enzymatic functions.

Figures adapted from Itabe et al., 2017 for Perilipins, and Chapman et al., 2012 for plant coat proteins.

So far, oleosins have not been found on leaf LDs. Caleosins (CLOs), however, are conserved here, but in different isoforms than in seeds. While CLO1 and CLO2 are the coating LDs in seeds, CLO3 and CLO4 are expressed in leaves, and their expression there is even increasing during pathogen attack and senescence (Aubert et al., 2010; Kim et al., 2011).

In the model alga *Chlamydomonas reinhardtii*, LDs are coated by a protein unrelated to the major plant LD coat proteins or perilipins named major LD protein (MLDP, Moellering and Benning, 2010). Knockdown of this protein had similar effects like oleosin knockouts in plant seeds: A decrease in protein on the surface of the LDs led to a significant increase in LD size (Moellering and Benning, 2010). The MLDP also has orthologs in other green algae species (Murphy, 2012).

In the yeast *Saccharomyces cerevisiae*, coat proteins of both plants and animals are absent, though this observation is not true for all fungi which might contain both PLIN proteins and caleosins (Rahman et al., 2018; Wang and St. Leger, 2007). The most abundant yeast LD proteins are enzymes associated with lipid metabolism, though it is questionable whether they can be considered coat proteins like PLINs and oleosins (Athenstaedt et al., 1999).

### 1.3. The Biogenesis of Lipid Droplets

The eukaryotic LDs discussed in this thesis all have a common place of biosynthesis: the endoplasmic reticulum (ER).

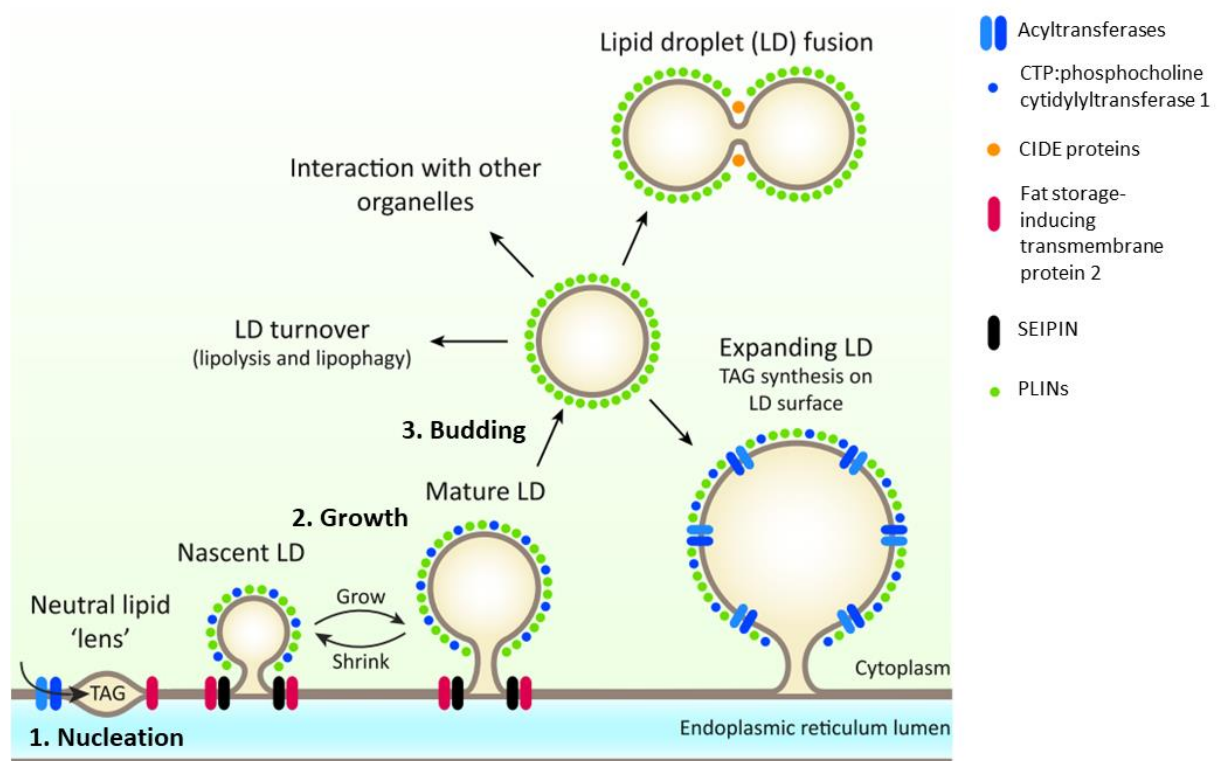
In plant seeds, the genetic control of LD formation has been subject of intensive research. During seed maturation, the action of the transcription factor LEAFY COTYLEDON2 (LEC2), which is considered a key regulator of seed development, leads to the induction of another transcription factor, WRINKLED1 (WRI1, Baud et al., 2007; Cernac & Benning, 2004; Santos Mendoza et al., 2005). Overexpression of both transcription factors can lead to the reprogramming of a vegetative tissue like leaves to “seed-like” state, hallmarked by the accumulation of TAGs in oleosin-coated LDs (Che et al., 2009; Shen et al., 2010; Stone et al., 2008). WRI1 mainly activates genes involved in the biosynthesis of fatty acids (FA), the precursors for TAG production. Given that *WRI1* knockout seeds contain up to 80 % less oil than WT, the availability of FA is one of the determining factors for successful seed oil production (Focks and Benning, 2002). FAs, which are synthesized in the chloroplast, are converted into acyl-CoAs and in subsequent reactions used for the acylation of glycerol-3-phosphate, lyso-PA, and PA to diacylglycerol (DAG). DAG can in a last acylation reaction, by the action of an enzyme called Diacylglycerol-Acyltransferase (DGAT), be used as a direct TAG precursor, or it can be converted to PC, which itself can be converted into TAG by additional enzymes (Chapman and Ohlrogge, 2012; Kennedy, 1961). These conserved TAG biosynthesis pathways take place in the ER membrane.

The exact biophysics and molecular biology of the processes happening next are yet to be elucidated. What is clear is that for a LD to form, newly synthesized TAG needs to accumulate between the two leaflets of the ER membrane, forming a separate phase from the bilayer, a process termed nucleation (Figure 4). Then, the nascent LD needs to grow, but not just by accumulating more TAG but also by bending one part of the membrane leaflet (the cytosolic) more than the other, to achieve directionality. Finally, the mature LD becomes almost completely spherical, enclosed by the cytosolic ER membrane leaflet (Thiam and Forêt, 2016). Under this condition, termed budding, the LD is also able to dissociate completely from the mother bilayer membrane.

It was shown recently that with specific combinations of PLs and TAGs, LDs bud of a bilayer membrane spontaneously and without the help of proteins (M'barek et al., 2017).

*In vivo*, proteins are likely supporting and regulating this process. Knockout mutants of a variety of proteins show changes in LD size, implicating roles during LD biogenesis, budding, or growth. In plants, coat proteins seem to influence LD biogenesis. Oleosins, which are first co-translationally inserted into the ER membrane, display a higher affinity for growing LDs in the membrane (Abell et al., 2004). Double knockout of OLEOSIN1 and OLEOSIN2 creates seeds with

very large LDs (Shimada et al., 2008). Because LD biogenesis in plants is yet to be observed *in vivo*, it is unclear whether oleosins regulate LD size directly during growth and budding, or indirectly by inhibiting coalescence. Experimental evidence, while not excluding the first option, supports the second hypothesis. Oleosins are positively charged at pH 7, creating charge repulsion from different LD surfaces to each other, inhibiting coalescence (Tzen et al., 1992). Trypsin treatment of isolated LDs from maize embryos lead to coalescence, while treatment with phospholipases A2 and C did not. After trypsin treatment, only a small peptide remained, probably corresponding to the hydrophobic domain inaccessible to the protease (Tzen and Huang, 1992).



**Figure 4: LD biogenesis in mammalian cells.** The action of ER-resident acyltransferases leads to lens-like accumulation of the neutral lipid TAG between the two membrane leaflets. Through regulated addition of neutral lipids, PLs, and proteins, the nascent LD grows and matures. The access to the growing LD is presumably regulated by SEIPIN and Fat storage-inducing transmembrane proteins localized at ER/LD junction sites. CTP:phosphocholine cytidyltransferase synthesizes CDP-choline needed for membrane lipid synthesis. The mature LD can bud off into the cytoplasm or remain in constant connection with the ER. In the cytoplasm, LDs engage in multiple interaction with other organelles like mitochondria, peroxisomes, and the autophagosome. Through the action of CIDE proteins, LD can fuse with each other.

Figure adapted from (Pyc et al., 2017a).

A class of conserved ER-localized proteins has a major influence on LD size during biogenesis, with implications on human health: SEIPINs (Szymanski et al., 2007). Mutations of SEIPIN lead to Berardinelli-Seip congenital lipodystrophy in humans, and to severe LD phenotypes in yeast and Arabidopsis (Cai et al., 2015; Fei et al., 2008, 2011; Magré et al., 2001; Taurino et al., 2018). SEIPINs have been found to localize to ER-LD junction sites, and are thought to regulate the access of lipids and/or proteins to the growing LD. In Arabidopsis, where the genome contains three

variants of the single gene found in humans and yeast, single knockdowns lead to decreased seed size and seed oil content (Cai et al., 2015). Combinations of multiple knockouts decrease fertilization efficiency because the LDs are so large that they are no longer able to enter the growing pollen tube (Taurino et al., 2018). In yeast, SEIPIN was identified in a mutant screen through a 50-fold size increase of LDs (Fei et al., 2008, 2011). In the same screen, mutants of enzymes of PL synthesis were identified, indicating the importance of the synthesis of TAG precursors and the action of PL themselves on LD growth or coalescence behavior after budding (Fei et al., 2011). Interestingly, in SEIPIN knockout mutants of both yeast and Arabidopsis, the formation of nuclear LDs was observed, indicating a possible involvement of SEIPIN in the directionality of LD growth (Cartwright et al., 2015; Taurino et al., 2018).

Another class of ER-resident proteins influences LD size in animal cells are fat storage-inducing transmembrane proteins (FIT) (Kadereit et al., 2008). Knockouts of the adipose tissue-specific isoform FIT2 resulted in the inability of mice to develop normal white adipose tissue (Miranda et al., 2014). In plants, no homologs of these proteins are known. However, heterologous expression of FIT2 in Arabidopsis led to increased neutral lipid contents of both seeds and leaves (Cai et al., 2017).

The composition of the PL monolayer also contributes to LD budding and size (Krahmer et al., 2011; M'barek et al., 2017). In animals, the rate limiting enzyme in the synthesis of PC, CTP:phosphocholine cytidyltransferase, was found to be reversibly retargeted from the ER to growing LDs (Krahmer et al., 2011), and therefore actively regulating PL composition of the LD monolayer.

Consequently, multiple factors contribute to LD size within a cell, and some of those factors can vary in different cell types or organisms, while others, like the impact of SEIPINs, are conserved. Different cells and organisms also deal differently with mature LDs. In yeast, there is compelling evidence that at least a subpopulation of LDs stays in continuous membrane connection with the ER, indicated by energy independent exchange of proteins between the ER and LDs (Jacquier et al., 2011). The opposite was observed in a study on HeLa cells, where the recruitment of the adipocyte triglyceride lipase (ATGL) from the ER to the LDs was COP-dependent, which the authors explained by a lack of continuous membrane between the two organelles (Soni et al., 2009). However, how exactly a COP protein-coated vesicle would fuse or otherwise interact with a LD is still unclear. It has been reported that the COPI machinery can indeed assemble at the LD PL monolayer and alter the PL surface tension by removing 60 nm “nano-LDs” from the mother LD (Thiam et al., 2013a).

Even after the detachment from the ER membrane, LDs can continue to grow, meaning that they can acquire more neutral lipids, PLs and attached proteins. The mammalian TAG synthesis

enzyme DGAT2 was observed on LD surfaces (Kuerschner et al., 2008). However, its topology indicates two distinct putative transmembrane regions. This suggest that it still resides at least partially in the ER membrane, proposing it is one of the proteinaceous mediators of ER/LD contact sites (Schuldiner and Bohnert, 2017). Independent of ER membrane proteins is the growth of LDs by fusion. In mammals, a family of proteins called “CELL DEATH-INDUCING DFF45-LIKE EFFECTOR” (CIDE, also FAT-SPECIFIC PROTEIN 27) have been identified as important players mediating LD fusion (Gao et al., 2017). CIDE proteins are thought to create a pore- or channel-like structure for the transfer of neutral lipids between two lipid droplets. No homologs or analogs of CIDE proteins have been detected in plants. LDs can also fuse spontaneously, if their surface tension is low enough (Thiam et al., 2013b). *In vivo*, this is usually avoided through coating of LDs with proteins and controlling PL composition of the monolayer, in return requiring proteins to mediate fusion when it is wanted.

In recent years, it has become more and more evident that LDs are in center of the lipid metabolism of the cell (Barbosa et al., 2015). They therefore must closely interact with other organelles involved in these metabolic pathways, like the ER, where lipid synthesis takes place, and mitochondria or peroxisomes, the site of  $\beta$ -oxidation in animals and plants, respectively. Such connections of LDs with the ER, peroxisomes, or mitochondria, have been shown in many cells (Binns et al., 2006; Tarnopolsky et al., 2007). It is a still matter of ongoing research how those interactions are mediated, as well as their extent in space and time.

## 1.4. The Degradation of Lipid Droplets

As described in chapter 1, LDs are one of the three major seed storage sites. The degradation of storages during germination and seedling establishment supplies the embryo with the energy necessary to break out of the seed coat, reach the light, and start photosynthesis. In seeds where LDs are the major storage reserve, like in *Arabidopsis thaliana* or *Brassica napus*, the degradation of LD components are determining for species survival. In the following chapter, I will present findings that have contributed to our understanding of seed germination based on TAG degradation, again comparing this process to findings in other species. I will also touch upon our limited knowledge concerning the degradation of the other LD components, namely PLs and proteins.

### 1.4.1. The Regulation and Function of Neutral Lipid Breakdown

The breakdown of TAGs in lipid-storing tissues or cells of both plants and animals is a highly regulated process. Misregulation would lead either to an energy shortage, in the case of inhibited lipolysis, or to an accumulation of toxic free FAs, if lipolysis rates are higher than the metabolism of the lipolysis products.

Germination *sensu stricto* is defined as the time period between initial water uptake by the dry seed until the emergence of the radicle or any other part of the embryo (Nonogaki et al., 2010). The success of this initial phase in a plant's life is mostly determined during seed maturation (Rajjou et al., 2012), as it depends on an initial set of metabolic proteins and conserved mRNAs deposited in the seed, and the ability of the seed to determine the ideal germination conditions. However, the seedling establishment, which follows germination, is of equal importance for the success of the plant species. This phase is hallmarked by the mobilization of storage reserves (Bewley et al., 2013). In *Arabidopsis thaliana*, the energy required for successful and fast seedling establishment come from the degradation of the TAG stored in the LDs. The initial hydrolysis of TAG into DAG and one FA is performed by two lipases, SUGAR-DEPENDENT1 (SDP1) and SUGAR-DEPENDENT1-LIKE (SDP1L), which together account for 95 % of TAG lipolysis in seedlings (Kelly et al., 2011). Double mutants of the lipases complete germination *sensu stricto* slower, but at almost the same rate as WT seeds. During post-germinative growth, however, only 20 % of seedlings can be established, a phenotype that can be completely rescued by the supply of an external carbon source (Kelly et al., 2011). Originally, it was assumed that SDP1 is localized at the LD surface (Eastmond, 2006). However, newer data suggests that the lipase resides in the peroxisomal membrane, which encircles LDs via membrane extensions (peroxules) during post-germinative growth, therefore allowing a close proximity of FA release and further processing via  $\beta$ -oxidation (Thazar-Poulot et al., 2015). Both SDP1 and SDP1L can hydrolyze TAG and DAG, with a strong preference for TAG (Eastmond, 2006; Kelly et al., 2011), indicating the necessity for one



or more separate DAG lipases involved in LD breakdown. No such enzymes has been identified on LDs, so far. Recently, a major seed and pollen MAG lipase could be identified (Kim et al., 2016b), and there is evidence that it localizes to the LD in Arabidopsis seedlings. The FAs hydrolyzed from TAG, DAG, and MAG are imported into the seed's specialized peroxisomes, which are also called glyoxysomes, where the activated acyl-CoAs are degraded by  $\beta$ -oxidation to acetyl-CoA. Acetyl-CoA is then utilized in the glyoxylate cycle to generate the organic acid precursors of gluconeogenesis, ATP production, amino acid synthesis, and membrane lipid synthesis (Theodoulou and Eastmond, 2012).

In mammalian adipocytes, the lipases responsible for the complete degradation of TAG as well as their regulation are better understood than in plant seeds. The activation of TAG degradation in adipocytes is hormone-dependent: Upon  $\beta$ -adrenergic stimulation, PLIN1, which normally shields LD from lipase activity, becomes phosphorylated (Brasaemle et al., 2009). This leads to the release of a PLIN1-bound effector protein that binds and activates the ATGL (Granneman et al., 2009). In addition, ATGL itself is phosphorylated upon adipocyte stimulation (Pagnon et al., 2012) and recruited to the LD surface (Yang et al., 2010). In addition to ATGL, also the adipocyte DAG lipase, called hormone-sensitive lipase (HSL), is hormone-activated (Lass et al., 2011). Upon adipocyte stimulation, HSL is phosphorylated and interacts with phosphorylated PLIN1, both events boosting its activity. Only the activity of the enzyme that completes the hydrolysis of TAG to glycerol and free FA by hydrolyzing MAG seems to be hormone-independent (Lass et al., 2011).

#### 1.4.2. Degradation of other Lipid Droplet Components

For a LD to be degraded, it is not only necessary to degrade the neutral lipids but also the PL monolayer and the coat proteins embedded in it. In fact, there is evidence that, at least partially, the degradation of PLs and proteins is a prerequisite for the efficient degradation of neutral lipids.

##### 1.4.2.1. The Degradation of Coat Proteins

In mammals, the degradation of the LD coat proteins PLINs is mediated by both the major cellular degradation pathway: the ubiquitin-proteasome system and autophagy. As described above, PLIN1 does not only passively shield the LDs and therefore blocks lipase access to the lipids, it additionally actively participates in the regulation of ATGL activity. Therefore, the degradation of PLINs might be required for efficient lipolysis. PLIN1 has been identified as a substrate of both the proteasome and the autophagy pathway (Kovsan et al., 2007; Xu et al., 2006). It was observed that the degradation of PLIN1 happened before neutral lipid breakdown. Similarly, also PLIN2 has been identified as a substrate of autophagic degradation, preceding lipolysis (Kaushik and Cuervo, 2015). In a different study, pharmacological inhibition of the proteasome could protect PLIN2 levels, leading to an accumulation of polyubiquitinated PLIN2 in the cells (Masuda et al., 2006).

In plants, the first indications on the degradation of LD coat proteins were found in sesame (*Sesamum indicum* L.) seedlings (Hsiao and Tzen, 2011). Ubiquitination sites could be identified on both oleosins and caleosins. Additionally, it was shown that the abundance of these proteins decreased over time after germination. However, the authors also hypothesized possible other signals transmitted by the seemingly random ubiquitination patterns on the proteins, for example to establish contacts between the peroxisomes and LDs (Hsiao and Tzen, 2011; Thazar-Poulot et al., 2015). In *Arabidopsis*, oleosins marked by K48 di-ubiquitination were indeed degraded by the proteasome, and the efficiency of this process could be blocked by the addition of pharmacological proteasome inhibitors (Deruyffelaere et al., 2015).

#### 1.4.2.2. The Degradation of Phospholipids

Similarly to proteins, PLs somewhat limit the access of lipases to the neutral lipid core, especially when the lipases lack hydrophobic domains with which they can get access to the neutral lipids. However, very little is known about how and when the amount of PL on a degrading LD is being reduced.

One possibility for the reduction of PL content is the budding of nano-LD in a COPI-dependent manner (Thiam et al., 2013a). This exposed hydrophobic patches and altered the LD surface tension. Both these processes could be used by proteins to bind to the LDs and execute their enzymatic function on the neutral lipid core.

In plants, the activity of Phospholipase A (PLA) has been detected on the surface of LDs, creating holes in the PL monolayer of LD from cucumber (*Cucumis sativus*) seeds (Noll et al., 2000). These hydrophobic patches could again be used to grant access to TAG lipases. PLA was one of the major enzymatic activities detected on LDs from olive pollen tubes (Zienkiewicz et al., 2013). The activity of the enzyme increased when the pollen tubes were grown without external carbon source, indicating that the demand for rapid and efficient TAG degradation can be a regulator of enzymatic activity.

## 1.5. Functions of Lipid Droplets in Non-Lipid-Storing Tissues

So far, I have presented the basic function, composition and life cycle of LDs in the major lipid-storing tissues, like oil seeds in plants and adipocytes in humans. However, as research on LDs intensified and covered different organisms, organs, and cells, it became clear that the functions of LDs can be as diverse as their occurrence. In the following, I would therefore like to present findings that support the notion that LDs are multi-functional organelles contributing to fitness and survival of an organism.

### 1.5.1. Lipid Droplet Function in Anther Development

The stamen, consisting of a filament and the anther, is the male part of the flower. It harbors the developing pollen grains, the strongly reduced male gametophyte of seed plants. Inside the anther, there is a highly metabolically active layer of cells called the tapetum. During pollen development in *Brassica* species, tapetum cells accumulate two kinds of specialized organelles: the elaioplast and the tapetosome (Platt et al., 1998). The elaioplast is a plastid that instead of thylakoids contains plastoglobuli rich in SEs. The tapetosome is an ER-derived, TAG-containing and oleosin-coated LD (Hsieh and Huang, 2004). However, the oleosin genes expressed in the tapetum are very different from the ones found in seed and pollen (Hsieh and Huang, 2004). While non-tapetum oleosins are small (15-30 kDa), the tapetal oleosins can be much larger, one of them being more than 100 kDa. The function of the specialized organelles comes into place at the end of pollen maturation: Then, the tapetum cells die and selectively release the oleosins of the tapetosomes and the SEs from the elaioplast onto the pollen grains. There, they form a hydrophobic barrier, the pollen coat. The pollen coat functions in different aspects ensuring successful fertilization: It protects the pollen grain from dehydration, helps in successful transmission from the anther to the stigma, and allows communication with the stigma to transition into the next phase of fertilization – the pollen tube growth (Murphy, 2006).

### 1.5.2. Functions of Lipid Droplets in Pollen Tube Growth

Pollen grains and pollen tubes are another plant organ that harbor comparably large amounts of TAG stored in LDs (Rotsch et al., 2017). When the pollen tube germinates, the LDs contained in the grain enter into the tube, where they are transported along the cytoplasmic stream (Müller et al., 2017). However, the fate and purpose of LDs in the pollen tubes are not completely clear. Whether the FAs released during TAG degradation are used for pollen tube growth or not, depends on the species. In olive, no additional carbon source is needed in the medium for tube growth *in vitro* (Zienkiewicz et al., 2013), while pollen tubes from *Nicotiana tabacum* (tobacco) or *Arabidopsis* are not viable for long during *in vitro* experiments without external carbon source (Boavida and McCormick, 2007; Rotsch et al., 2017). However, it was shown recently that a lipase does contribute significantly to pollen tube growth in *Arabidopsis* and tobacco (Müller and

Ischebeck, 2018). In tobacco, there is also evidence that no glyoxylate cycle takes place (Bucher et al., 1995; Mellema et al., 2002; Müller et al., 2017). At the same time, TAG does not only seem to be degraded but also synthesized in tobacco pollen tubes (Mellema et al., 2002). These observations could indicate that in tobacco pollen tubes, LDs and the TAG stored within them serves as a sink and source of membrane building blocks which are needed in large amounts for the rapid apical growth of the pollen tube (Müller and Ischebeck, 2018).

### 1.5.3. Lipid Droplets in Leaves

In recent years, LDs in leaves have shifted more into focus of the plant LD research community. While leaves contain only comparably small amounts of neutral lipids stored in few LDs, those LDs seem to have important functions in the metabolic coupling inside the leaf. Leaves are the major site of photosynthetic activity in the plant. The carbon that is reduced during the day is stored as starch in the leaf chloroplasts. It can be degraded to maltose or glucose to supply non-photosynthetic tissues or to be used directly for growth during the night (Chapman et al., 2013). Although TAG is the most energy-dense form of reduced carbon, starch is the preferred carbon storage compound in leaves. Via genetic manipulation, this balance can be shift more towards the oil side, for example by overexpression of seed transcription factors like WRI1 or LEC2 (Che et al., 2009; Shen et al., 2010; Stone et al., 2008).

Within leaves, a rather high and consistent amount of LDs are found in guard cells (McLachlan et al., 2016). These LDs are degraded after a blue-light stimulus and the energy generated from the TAG breakdown is used for the opening of the stomata (McLachlan et al., 2016).

The TAG concentration and LD abundance in leaves can also be increased by environmental factors. Senescence, drought and temperature stress have all been shown to induce LD formation (Gidda et al., 2016; Kim et al., 2016a; Shimada et al., 2015). One family of proteins has been observed to correlate with increased LD abundance during these conditions: LD-associated proteins (LDAP; Gidda et al., 2016; Kim et al., 2016a). Originally discovered on the oleosin-devoid LD of the avocado mesocarp (Horn et al., 2013), they are homologs of coat proteins of LD-like organelles of the rubber tree (Oh et al., 1999). Further analysis of the gene family showed that they are conserved in plants, and that Arabidopsis contains three isoforms (Gidda et al., 2016; Kim et al., 2016a). Different from typically seed LD proteins, the expression of LDAPs is rather ubiquitous (Gidda et al., 2016). Changes in LDAP expression, whether it is through knockout or overexpression, changes the LD abundance in direct correlation. For example, overexpressors of LDAPs contain more leaf LDs, larger seed LD and are more resistant to drought stress (Kim et al., 2016a). However, the functions of LDAPs remain still unknown. Pyc et al., (2017a) proposed a role for LDAP in the lipid metabolism through membrane remodeling or lipid signaling.

Possibly, the molecular functions of LDAPs could be solved with the help of a recently discovered interactor of this protein family: the LDAP-interacting protein (LDIP) (Pyc et al., 2017b). Until now, not much is known about the function of LDIP, either. It is a LD-localized protein that has an influence on LD size in both seeds and leaves (Pyc et al., 2017b).

The recent discovery of previously unknown LD proteins has opened new possibilities for the elucidation of LD functions in non-seed tissues. The ubiquitous presence of the organelle throughout the plants indicates an important role in the lipid homeostasis of all cells.

#### 1.5.4. Production Site for Antimicrobial Compounds

Plants lack features that are essential in animals for fighting off infection. Therefore, plant microbial defense has developed to be multilayered and complex, consisting in one part of the production of specific or non-specific secondary metabolites.

Pathogen attack has been shown to induce expression of the LD-localized proteins CLO3 and  $\alpha$ -dioxygenase ( $\alpha$ -DOX, Shimada et al., 2015). Caleosins are peroxygenases, and their activity is calcium- and heme-dependent (Hanano et al., 2006; Næsted et al., 2000). On the LDs in Arabidopsis, caleosins work in close coordination with  $\alpha$ -DOX1 (Shimada et al., 2014).  $\alpha$ -DOX1 catalyzes the oxidation of  $\alpha$ -linolenic acid to 2-hydroperoxy-octadecatrienoic acid, which is highly unstable. Then, CLO3 metabolizes this compound to the stable 2-hydroxy-octadecatrienoic acid (2-HOT). 2-HOT is an antimicrobial compound that has been shown to be active against the fungus *Collectotrichum higginsianum* (Shimada et al., 2014). Furthermore, this compound has been found in increased amounts during leaf senescence.

In addition to dioxygenases, LDs have long been known to harbor lipoxygenases that also require the presence of polyunsaturated FAs for their activity (Feussner and Kindl, 1992). LOX produces hydroperoxides and a conjugated double bond system next to it, for example 13-hydroperoxyoctadecatrienic acid (13-HPOT), produced by 13-LOX. In Arabidopsis, two 13-LOX enzymes are expressed in the cytosol of leaf cells and could therefore potentially produce this compound (Bannenberg et al., 2009). CLO3 has been implicated in the production of 13-hydroxyoctadecatrienic acid (13-HOT) and 15,16-epoxy-13-HOT (Blée et al., 2014). The production of both compounds requires 13-HPOT as a substrate. 13-HOT itself is active against oomycete and fungal infections, for example by *Botrytis cinerea* (Prost et al., 2005).

## 1.6. Aims of this Study

The limitations in the study of plant LD function, biogenesis, degradation, and interaction with other cellular compartments arise, at least partially, from the lack of proteins known to associate with this organelle.

Therefore, one aim of this work was the identification of new LD-localized proteins and protein families. For this, we employ a bottom-up proteomics approach. What sets our approach apart from similar attempts already present in the literature is a two-step process to ensure that the candidates we identify are reliable. First, in addition to a LD-enriched fraction, we perform bottom-up proteomics on the total protein extract sampled before LD enrichment. This way, we can calculate enrichment factors, allowing the separation of true candidates from contaminant. Secondly, LD candidate proteins are subjected to a cell biology approach to analyze their subcellular localization. Only proteins that are confirmed to localize to LD with this approach are then studied in further detail.

For the successful processing of label-free bottom-up proteomics data in our group, my first aim was to establish the use of MaxQuant as data processing software. In a first application of the software, I reprocessed an older LD dataset from *Nicotiana tabacum* pollen tubes to allow label-free quantification of the data with a newly published, improved proteome for this species.

Furthermore, I was aiming to identify new LD proteins from different tissues of *Arabidopsis thaliana* where LD are particularly prominent. The total and LD proteome of eight different developmental phases during Arabidopsis seed development, seed germination, and seedling establishment were to be sampled. We hoped to not only identify new LD proteins, including proteins specific to a certain phase, but also to create an extensive bottom-up proteomics study of those delicate phases of plant development. Furthermore, we aimed to be able to study the dynamics of LD proteins with the data collected during this study.

The detailed characterization of LD proteins was another aim of this thesis. I have employed a combination of classical genetics, cell biology and modern proteomics approaches to assess the function of the LD-localized protein PUX10. The impact of the knockout of PUX10 was studied by a comparative proteomics approach. These data were correlated with the results of cell biological studies of the knockout mutants.

In summary, the aims of this thesis were to extent the LD proteome and with that, lay the ground work for further studies on the function of LD proteins, and LD themselves.

## 2. Article 1: PUX10 is a lipid droplet-localized scaffold protein that interacts with CELL DIVISION CYCLE48 and is involved in the degradation of lipid droplet proteins

This article was published online in the journal *The Plant Cell* in August 2018. The supplemental figures are attached to the main article. The supplemental tables containing processed mass spectrometry data can be found online together with the full article:

<https://doi.org/10.1105/tpc.18.00276>

### **Author contribution:**

Franziska K. Kretzschmar reprocessed the proteomics data of *Nicotiana tabacum* pollen tubes that lead to the identification of the correct isoforms of caleosin and oleosins present in this tissue, and to the new identification of PTLD. She cloned, expressed and performed the microscopy presented in Figure 1A-G, 2B and 3A-C. She performed some of the cloning and all the expression and microscopy for the images presented in Figure 4. She performed microscopy on *pux10-1*, *pux10-3*, C#1, C#2, Ws-4 and *qrt PUX10* seedlings, analyzed and statistically analyzed the data presented in Figure 6. She performed the proteomics experiment, mass spectrometry data processing and data analysis presented in Figures 7 and 9 and all the supplemental tables. She cloned, expressed and performed microscopy presented in Figure 8 C-F. She designed Figures 4, 6-10 as well as the supplemental figures, wrote the introduction, material and methods, half of the results, and the discussion.



# PUX10 Is a Lipid Droplet-Localized Scaffold Protein That Interacts with CELL DIVISION CYCLE48 and Is Involved in the Degradation of Lipid Droplet Proteins

Franziska K. Kretschmar,<sup>a</sup> Laura A. Mengel,<sup>a,1</sup> Anna O. Müller,<sup>a</sup> Kerstin Schmitt,<sup>b</sup> Katharina F. Biersch,<sup>a,2</sup> Oliver Valerius,<sup>b</sup> Gerhard H. Braus,<sup>b</sup> and Till Ischebeck<sup>a,3</sup>

<sup>a</sup>Department of Plant Biochemistry, Georg-August-University, Albrecht-von-Haller-Institute for Plant Sciences, 37077 Göttingen, Germany

<sup>b</sup>Department of Molecular Microbiology and Genetics, Georg-August-University, Institute for Microbiology and Genetics, 37077 Göttingen, Germany

ORCID IDs: 0000-0001-8875-6336 (F.K.K.); 0000-0002-0510-0663 (L.A.M.); 0000-0001-7376-7841 (A.O.M.); 0000-0001-9627-031X (K.S.); 0000-0001-7020-5587 (K.F.B.); 0000-0003-4430-819X (O.V.); 0000-0002-3117-5626 (G.H.B.); 0000-0003-0737-3822 (T.I.)

The number of known proteins associated with plant lipid droplets (LDs) is small compared with other organelles. Many aspects of LD biosynthesis and degradation are unknown, and identifying and characterizing candidate LD proteins could help elucidate these processes. Here, we analyzed the proteome of LD-enriched fractions isolated from tobacco (*Nicotiana tabacum*) pollen tubes. Proteins that were highly enriched in comparison with the total or cytosolic fraction were further tested for LD localization via transient expression in pollen tubes. One of these proteins, PLANT UBX DOMAIN-CONTAINING PROTEIN10 (PUX10), is a member of the plant UBX domain-containing (PUX) protein family. This protein localizes to LDs via a unique hydrophobic polypeptide sequence and can recruit the AAA-type ATPase CELL DIVISION CYCLE48 (CDC48) protein via its UBX domain. PUX10 is conserved in *Arabidopsis thaliana* and expressed in embryos, pollen tubes, and seedlings. In *pux10* knockout mutants in *Arabidopsis*, LD size is significantly increased. Proteomic analysis of *pux10* mutants revealed a delayed degradation of known LD proteins, some of which possessed ubiquitination sites. We propose that PUX10 is involved in a protein degradation pathway at LDs, mediating an interaction between polyubiquitinated proteins targeted for degradation and downstream effectors such as CDC48.

## INTRODUCTION

Lipid droplets (LDs; also referred to as oleosomes and oil or lipid bodies) are specialized organelles mainly occurring in lipid-rich seeds, tapetal cells, and pollen grains (Chapman et al., 2012; Ischebeck, 2016; Huang, 2018). Additionally, all other plant cell types are considered either to contain a relatively small number of LDs or to form them only under certain environmental conditions, such as in response to abiotic stress (Gidida et al., 2016; Brocard et al., 2017; VanBuren et al., 2017). Structurally, LDs consist of a core of hydrophobic compounds, in most cases triacylglycerols (TAGs) and sterol esters, surrounded by a phospholipid monolayer (Pyc et al., 2017a). The monolayer derives from the outer layer of the endoplasmic reticulum (ER), where the LDs are initially formed, before they are released into the cytoplasm (they are distinct from plastid-derived plastoglobuli that also store neutral lipids; van Wijk and Kessler, 2017).

Several proteins are known to be involved in the proper formation of LDs (Huang, 2018). In pollen and seeds, diacylglycerol O-acyltransferase 1 and phospholipid:diacylglycerol acyltransferase 1, localized at the ER, are the main enzymes catalyzing the final step of TAG biosynthesis (Zhang et al., 2009). Situated at the ER-LD junction sites, SEIPINs are important for TAG accumulation and LD proliferation (Cai et al., 2015; Taurino et al., 2018). Other proteins are directly associated with LDs. For instance, caleosin and steroleosin (also referred to as hydroxysteroid dehydrogenases [HSDs]) are anchored to LDs by a hydrophobic hairpin sequence that contains conserved proline residues that form a so-called proline “knot” and “knob,” respectively (Shimada and Hara-Nishimura, 2010). Caleosins are considered to have a structural role in maintaining LD integrity, but some isoforms have also been reported to have peroxxygenase activity (Hanano et al., 2006; Blée et al., 2014) and are thought to act after the oxidation of fatty acids by LD-localized lipoxygenases (Rudolph et al., 2011; Zienkiewicz et al., 2013). Based on genetic evidence, steroleosins are associated with brassinosteroid metabolism (Li et al., 2007; Baud et al., 2009) and have been shown to be able to act on steroid substrates as dehydrogenases, interconverting hydroxyl-groups of the steroid backbone to keto-groups (d’Andréa et al., 2007). Rather unknown on the molecular level is the function of LDAPs (LD-associated proteins; Gidida et al., 2013). These proteins, first discovered on LDs of avocado (*Persea americana*) mesocarp (Horn et al., 2013), are also present in *Arabidopsis thaliana*, where they are needed for the proper

<sup>1</sup>Current address: ZIEL-Institute for Food and Health, Technical University of Munich, Gregor-Mendel-Straße 2, 85354 Freising, Germany.

<sup>2</sup>Current address: Department for Molecular Structural Biology, Georg-August-University, Institute for Microbiology and Genetics, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany.

<sup>3</sup>Address correspondence to tischeb@gwdg.de.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Till Ischebeck (tischeb@gwdg.de).

www.plantcell.org/cgi/doi/10.1105/tpc.18.00276



**Background:** Lipid droplets (LDs) are structures inside cells that store oil or fat, for example, in oil seeds. They were once considered plain storage organelles. Ongoing research in yeast and animal systems, however, has revealed that LDs are dynamic organelles with complex functions. In plants, new roles for LDs are emerging, too. Nevertheless, knowledge about the proteins embedded in the surface of plant LDs is limited. As our first model to study, we chose tobacco pollen tubes, which contain a high number of LDs. Pollen tubes are formed by pollen grains during plant reproduction, but can also form in growth medium.

**Question:** We first aimed to identify previously unknown LD-associated proteins. We then chose one of the newly identified LD proteins to study its role in LD biology.

**Findings:** We compared the protein composition—the proteome—of LDs isolated from tobacco pollen tubes to the total proteome of pollen tubes. Based on these data, we identified candidates that were highly enriched in the LD fractions. From this list, we chose several proteins and verified their LD localization by confocal laser scanning microscopy. Thereby, we identified three previously unknown LD proteins. We chose one of these proteins, named PUX10, for an in-depth characterization and also extended our work on this protein to the model system *Arabidopsis thaliana*. We discovered that PUX10 is involved in the degradation of other proteins sitting on the surface of LDs during seed germination, when LDs are broken down to fuel the germination process. PUX10 appears to do so by binding proteins that need to be degraded and also binding another protein, CDC48, known to be important for protein degradation. PUX10 functions as a bridge, or scaffold, between these proteins and brings them together.

**Next steps:** Our findings extend the plant LD proteome by three proteins and reveal a player involved in the regulation of the plant LD proteome. However, many open questions concerning various processes of LD biology remain. Therefore, the identification and characterization of further proteins involved in these processes is needed.

compartmentation of neutral lipids during postgerminative seedling growth and also in mature leaves, especially under drought stress conditions (Gidda et al., 2016; Kim et al., 2016). A similar role has been proposed for the LDAP interaction partner LDIP (LDAP-interacting protein; Pyc et al., 2017b).

The best studied and most prominent LD proteins, at least in seeds, tapetum, and pollen, are oleosins (Tzen et al., 1990; Lee et al., 1994; Huang, 1996; Wang et al., 1997). Similar to caleosins and steroleosins, oleosins are tightly anchored to the LD by a hydrophobic proline knot structure. Oleosins are considered LD coat proteins, which function to prevent the coalescence of LDs, especially during seed desiccation (Tzen, 2012). Oleosins might also be involved in the formation of LDs and their budding from the ER (Chapman et al., 2012). It is furthermore speculated that they might act as binding sites for other proteins (Quettier and Eastmond, 2009). In germinating seeds, when TAG, but also LD proteins are degraded, oleosins have been found to be ubiquitinated (Hsiao and Tzen, 2011; Deruyffelaere et al., 2015), which might be crucial for their breakdown.

Apart from the aforementioned proteins, only a few other LD proteins have been clearly identified in higher plants (Chapman et al., 2012; Pyc et al., 2017a; Huang, 2018), making the total number of known LD proteins relatively small in comparison to other organelles. Moreover, while hundreds of proteins have been found by proteomic screens in LD-enriched fractions from various higher plants (Jolivet et al., 2004; Liu et al., 2015), these fractions are prone to contaminations and only few of these proteins have been confirmed to be LD-associated by additional lines of evidence. Nevertheless, knowledge of LDs and their proteome has steadily increased in recent years and, in doing so, has raised a number of new and important questions about the organelle's protein homeostasis: How is LD protein turnover regulated, and especially how are proteins removed from the LD when the organelle is being degraded, for instance, during lipolysis or in the case of protein misfolding?

Several protein degradation pathways exist in plants including autophagy (Wang et al., 2018) and degradation by the 26S proteasome (Sharma et al., 2016). In the latter case, the majority of proteins are ubiquitinated prior to their breakdown. The proteasome is localized in the cytoplasm and hence has no access to membrane proteins or proteins inside organelles (with the exception of the nucleus; Peters et al., 1994). Therefore, proteins derived, for example, from the ER lumen or membrane have to be transported to the cytoplasm before they can be degraded by the proteasome. One mechanism for the elimination of misfolded proteins from the ER lumen and membrane is the ER-associated protein degradation (ERAD) pathway. This pathway requires a unique set of proteins to recognize, ubiquitinate, retrotranslocate, and ultimately degrade misassembled proteins, and its components are mostly conserved among eukaryotes. In terms of the cytoplasmic events that take place in the ERAD pathway, a variety of proteins are known (Meusser et al., 2005). Among these proteins are the UBX (ubiquitin regulatory X domain) domain-containing proteins that can act as a scaffold to mediate contact between the ubiquitinated substrates of the ERAD pathway and downstream effector proteins like the AAA-type ATPase CELL DIVISION CYCLE48 (CDC48) protein, which is also known (in mammals) as p97/VCP (Schubert and Buchberger, 2008).

The Arabidopsis genome encodes 16 PLANT UBX DOMAIN-CONTAINING (PUX) proteins with a diverse domain structure (Liu and Li, 2014). We identified one member of this family, PUX10, in the proteome of LD-enriched fractions isolated from tobacco (*Nicotiana tabacum*) pollen tubes and confirmed that the protein is localized to LDs in both pollen tubes and seeds via a hydrophobic region. Furthermore, we show that disruption of *PUX10* expression influences LD size and number and that PUX10 is able to recruit the p97/VCP homolog At-CDC48A to LDs, implicating a role for PUX10 and At-CDC48A in LD protein turnover.

## RESULTS

### Identification of Candidate Proteins Associated with LDs from Tobacco Pollen Tubes

One strategy for improving our understanding of LDs with regard to their synthesis, degradation, and/or function is to identify proteins specifically associated with this organelle. Therefore, the first goal in this study was to investigate the proteome of LDs to extend the number of known LD-bound proteins. Tobacco pollen tubes were chosen, as LDs of this tissue have not been previously investigated by proteomics, in contrast to other studies that focused on LDs in seeds (Jolivet et al., 2004; Liu et al., 2015), seedlings (Pyc et al., 2017b), mesocarp of avocado (*Persea americana*) (Horn et al., 2013), and chinese tallow (*Triadica sebifera*; Zhi et al., 2017). In addition, tobacco pollen tubes can be easily transformed to verify the association of candidate proteins with the LDs (Müller et al., 2017), eliminating the need to switch systems from identification to verification. LD-enriched fractions were obtained together with cytosolic and total fractions, and the protein composition of all samples was analyzed by tryptic digestion and liquid chromatography-tandem mass spectrometry (LC-MS/MS). Protein composition was subsequently investigated using the label-free intensity-based absolute quantification (iBAQ) algorithm (Schwanhäusser et al., 2011) from MaxQuant (Supplemental Data Set 1), which identified a total of 1314 protein groups. The relative iBAQ values given by the software represent an estimate of the abundance of a protein in a mixture and were calculated as % of all proteins in one sample (Shin et al., 2013). The enrichment was determined from the relative LD iBAQ values divided by the higher value of the total and cytosolic fractions (Supplemental Data Set 2).

The two most abundant proteins found in the tobacco pollen tube LD fraction were a caleosin (Nt-CLO1a) with an iBAQ % of 158 and an enrichment factor of 1234 and an oleosin (Nt-OLE6b) with an iBAQ % of 64 and an enrichment factor of 2137. The LD fraction, however, also contained contaminants from other organelles. For example, an ATP synthase subunit (Uniprot ID A0A1S4CK04) was the third most abundant protein in the LD fraction with an iBAQ % of 56, but it was only enriched by a factor of 2.6 and thereby ruled out to be a LD protein, since all known LD proteins that were identified displayed much higher enrichment factors (Supplemental Data Set 3). To determine if another organelle was also strongly enriched in the LD fraction, various marker proteins that could be assigned to a specific organelle and that showed a high abundance in the total fraction were chosen (Supplemental Data Set 4). Overall, the selected marker proteins for mitochondria, plastids, peroxisomes, or the cytosol were only slightly enriched or depleted in the LD fraction (Supplemental Data Set 4). On the other hand, the three chosen marker proteins from the ER were enriched in the LD fraction by a factor of 4 to 21, indicating that the ER, to some extent, copurified with LDs, albeit the relative abundance of the ER marker proteins in the LD fraction was relatively minor compared with enrichment of homologs of known LD proteins in the LD fraction (Supplemental Data Set 3).

We also compared the isoforms of known LD proteins found in the pollen tubes on a qualitative level to the isoforms found in

tobacco seeds (Supplemental Data Sets 5 and 6). This revealed that two of the oleosin and caleosin isoforms were only found in pollen tubes (Supplemental Figures 1 and 2 and Supplemental Data Sets 7 and 8). Steroleosins, however, were only detected in seeds but not in pollen tubes. For an overview of all LD-localized protein families and the protein members detected, see Supplemental Figures 1 to 10 and Supplemental Data Sets 7 to 16; all proteins were named according to their closest homolog from Arabidopsis.

### Verification of LD Association by Transient Expression in Tobacco Pollen Tubes

In the pollen tube data set, 18 proteins had an enrichment factor higher than 100 and an iBAQ % higher than 1 (Table 1). From these proteins, at least one member of the corresponding nine gene families was cloned from tobacco pollen tube cDNA, except for Nt-*OBL1*, which was previously characterized (Müller and Ischebeck, 2018). All constructs were transiently expressed in tobacco pollen tubes as fusion proteins with mVenus, and Nile Red was used to stain LDs. In total, seven proteins were confirmed to be localized, at least partially, to LDs (Figure 1), including Nt-LDAP1a, Nt-LDIPa, which, notably, appeared to localize only to some LDs, the oleosin Nt-OLE6b, the caleosin Nt-CLO1b, a cycloartenol synthase Nt-CAS1b (and its Arabidopsis homolog, At-CAS1; Supplemental Figure 11), a sterol methyltransferase 1 (Nt-SMT1c), a protein formerly annotated as unknown (named here as Nt-PTLD2b for POLLEN TUBE LIPID DROPLET PROTEIN 2b), and a member of the family of PUX proteins (Nt-PUX10c), which also partially localized to the ER (Figures 2A and 2B). A homolog of Arabidopsis retinol dehydrogenase, Nt-FEY1, did not localize to LDs, but instead localized exclusively to the ER (Figure 3).

We also tested other selected proteins that had either high iBAQ % scores or a strong enrichment and that were potentially interesting candidates based on their known cellular function(s) (Table 1, lower part). None of these proteins, however, showed any obvious localization, partial or otherwise, to LDs (Figure 3), reinforcing the premise that only the combination of a high iBAQ % and a strong enrichment reflected a protein's association with LDs. Nonetheless, one protein with unknown function termed here ARFAPTIN DOMAIN-CONTAINING PROTEIN (ADCP) localized mostly to the ER, but appeared enriched in domains interacting with LDs (Figures 3A and 3B). Other selected proteins localized to either the cytosol (the small G proteins Nt-RABB1b and Nt-RAB2c; Figures 3G and 3H), plasma membrane (carotenoid cleavage dioxygenase Nt-CCD; Figure 3I), or to unknown punctate structures (a PLAT-domain containing protein, Nt-PDCP).

Taken together, the addition of a second line of evidence for LD localization via transient expression in tobacco pollen tubes confirmed many of the positive candidates identified in our proteomics screen. One of these, Nt-PUX10c, and its homolog from Arabidopsis (Supplemental Figure 4 and Supplemental Data Set 10), were subsequently chosen for a more in-depth investigation (see below), as these proteins have previously not been described as LD proteins in plants, yet other members of the Arabidopsis family of PUX proteins are known or implied to play a central role in protein degradation (Liu and Li, 2014).

**Table 1.** Proteins Enriched in the LDs of Tobacco Pollen Tubes

Protein Type	Name	Gene ID	iBAQ %	Enrichment
Caleosin <sup>1</sup>	NtCLO1a	LOC107783728	158 ± 10	1234
Oleosin <sup>1</sup>	NtOLE6b	LOC107824536	64 ± 30	2137
Caleosin	NtCLO1b	LOC107817909	40 ± 5	993
Oleosin	NtOLE6a	LOC107780677	35 ± 16	6746
LDAP <sup>1</sup>	NtLDAP1a	LOC107765167	28 ± 3	741
Cycloartenol synthase <sup>1</sup>	NtCASb	LOC107762593	16 ± 4	1517
Plant UBX domain-containing protein <sup>1</sup>	NtPUX10c	LOC107769345	15 ± 3	341
LDAP	NtLDAP1b	LOC107804268	7.0 ± 0.5	n.d. in controls
LDAP	NtLDAP3a	LOC107827582	5.0 ± 0.4	557
Cycloartenol synthase	NtCASA	LOC107826198	4.1 ± 1.2	2605
Sterol methyltransferase <sup>1</sup>	NtSMT1c	LOC107813194	3.1 ± 0.3	219
Unknown protein	NtPTLD1	LOC107815104	2.8 ± 0.02	n.d. in controls
Oil body lipase	NtOBL1	LOC107788962	2.6 ± 0.9	307
Plant UBX domain-containing protein	NtPUX10a	LOC107799924	2.2 ± 0.3	n.d. in controls
LDIP <sup>1</sup>	NtLDIPa	LOC107763626	1.8 ± 0.8	n.d. in controls
Unknown protein	NtPTLD3	LOC107796437	1.7 ± 0.8	249
Retinol dehydrogenase <sup>2</sup>	NtFEY1	LOC107800303	1.2 ± 4	1087
LDIP	NtLDIPb	LOC107831283	1.1 ± 1.6	n.d. in controls
Arfaptin domain-containing protein <sup>2</sup>	NtADCP	LOC107827250	3.0 ± 0.6	9
PLAT domain-containing protein <sup>3</sup>	NtPDCP	LOC107793109	2.2 ± 0.6	10
Small G protein <sup>3</sup>	NtRAB1b	LOC107771805	1.12 ± 0.16	36
Unknown protein <sup>1</sup>	NtPTLD4	LOC107778489	0.7 ± 0.2	n.d. in controls
Carotenoid cleavage dioxygenase <sup>3</sup>	NtCCD	LOC107830214	0.45 ± 0.08	245
Small G protein <sup>3</sup>	NtRAB2c	LOC107794510	0.13 ± 0.11	15

LDs were isolated from tobacco pollen tubes. Their protein composition ( $n = 3$  biological replicates) was compared to total extracts and a cytosolic fraction (both  $n = 4$ ) by LC-MS/MS after a tryptic in-gel digest. iBAQ values were determined using MaxQuant software. The relative iBAQ values were calculated as % of all proteins in one sample. The enrichments were each determined from the LD iBAQ % value divided by the higher value of the total and cytosolic fractions. Only proteins with an enrichment factor of  $>100$  and making up more than 1 % of the total iBAQ values were considered for the upper part of the list. Further proteins analyzed in respect to their localization are depicted in the lower part. For a list of all identified proteins, see Supplemental Data Set 1. Several corresponding cDNAs were cloned and transiently expressed as mVenus fusions in tobacco pollen tubes: protein localized to LDs (1), protein localized to the ER (2), and protein that did not localize to the LDs or the ER (3). The “±” indicates sd. n.d., not detected.

### Nt-PUX10 Is Targeted to LDs by a Short, Hydrophobic Polypeptide Sequence

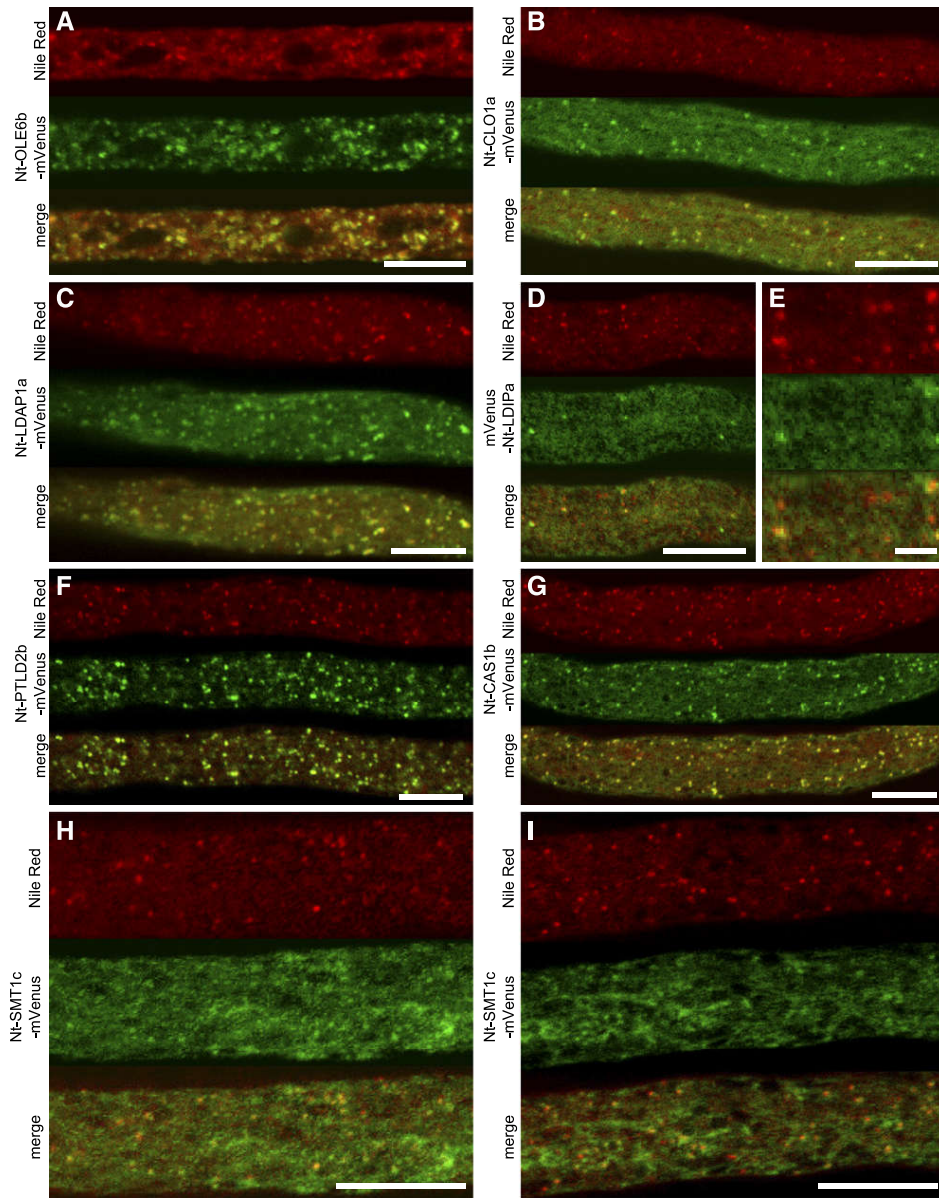
PUX proteins are a diverse family of proteins that share the eponymous UBX domain, but can differ in their overall domain structure (Supplemental Figure 12; Liu and Li, 2014). In Arabidopsis, 16 PUX proteins have been identified (Supplemental Figure 12) and, apart from the UBX domain, the isoforms At-PUX5, At-PUX7 to At-PUX11, and At-PUX13 to At-PUX16 also possess a putative ubiquitin-associated (UBA) domain. Additionally, At-PUX10 harbors a unique hydrophobic polypeptide sequence that is predicted to be a transmembrane domain (TMHMM Server v. 2.0; Krogh et al., 2001; Supplemental Figure 13) and is not present in any of the other members of the Arabidopsis PUX protein family (Supplemental Figure 13). Tobacco possesses six isoforms most closely related to At-PUX10 (Supplemental Figure 4 and Supplemental Data Set 10), which we named Nt-PUX10a-f, and among these, only Nt-PUX10a-d contain a hydrophobic region similar to that in At-PUX10 (Supplemental Figure 4 and Supplemental Data Set 10).

To test whether the hydrophobic region in the various PUX10 proteins that contain this sequence is necessary for LD targeting, a series of truncated constructs missing portions of the N terminus and/or the C terminus of Nt-PUX10c were generated and then transiently expressed as fusion proteins to mVenus in tobacco pollen tubes (Figure 2D; see Supplemental Figure 14

for magnified images). Overall, all the Nt-PUX10c-mVenus fusion constructs lacking the hydrophobic region (i.e., amino acids 89–119) were mislocalized to the cytoplasm, while all other fusion constructs containing this region targeted, at least partially, to LDs, including one that comprised only the 50 amino acids (amino acids 81–130; Figure 2D) spanning the hydrophobic region. An even shorter variant (amino acids 89–119) was not able to target the fluorophore to LDs. These results indicate that the novel hydrophobic region and the adjacent amino acids in Nt-PUX10 are not only necessary, but also sufficient for targeting to LDs.

### Nt-PUX10 Can Recruit At-CDC48A to LDs

One of the described functions of UBX domain-containing proteins is to interact and recruit CDC48 (Liu and Li, 2014), an AAA-type ATPase involved in the proteome homeostasis of various cellular processes and an important player for the degradation of ubiquitinated proteins by the proteasome (Baek et al., 2013). To test if PUX10 can influence the subcellular localization of CDC48, we employed again the transient expression system in tobacco pollen tubes. We chose At-CDC48A as a potential interaction partner, as this is the best-described isoform from plants (Park et al., 2008; Copeland et al., 2016). At-CDC48A expressed alone as an mVenus fusion accumulated in the cytoplasm, with no obvious localization to any other intracellular compartment(s)



**Figure 1.** Proteins Localizing to LDs.

The proteins were transiently expressed in tobacco pollen tubes, as C-terminal ([D] and [E]) or N-terminal ([A] to [C] and [F] to [I]) fusions to the fluorescent protein mVenus. The tubes were cultivated for 5 to 8 h, fixed with formaldehyde, and stained with Nile Red. Then, they were monitored by confocal microscopy. Colocalization was observed in all of 7 (A), 9 (B), 11 (C), 14 ([D] and [E]), 13 (F), and 7 (G) pollen tubes. (E) is a magnified section of (D). Partial association of mVenus structures and Nile Red-stained LDs was observed in 17 of 23 ([H] and [I]) pollen tubes. Bars = 10  $\mu$ m in (A) to (D) and (F) to (I) and 2  $\mu$ m in (E).

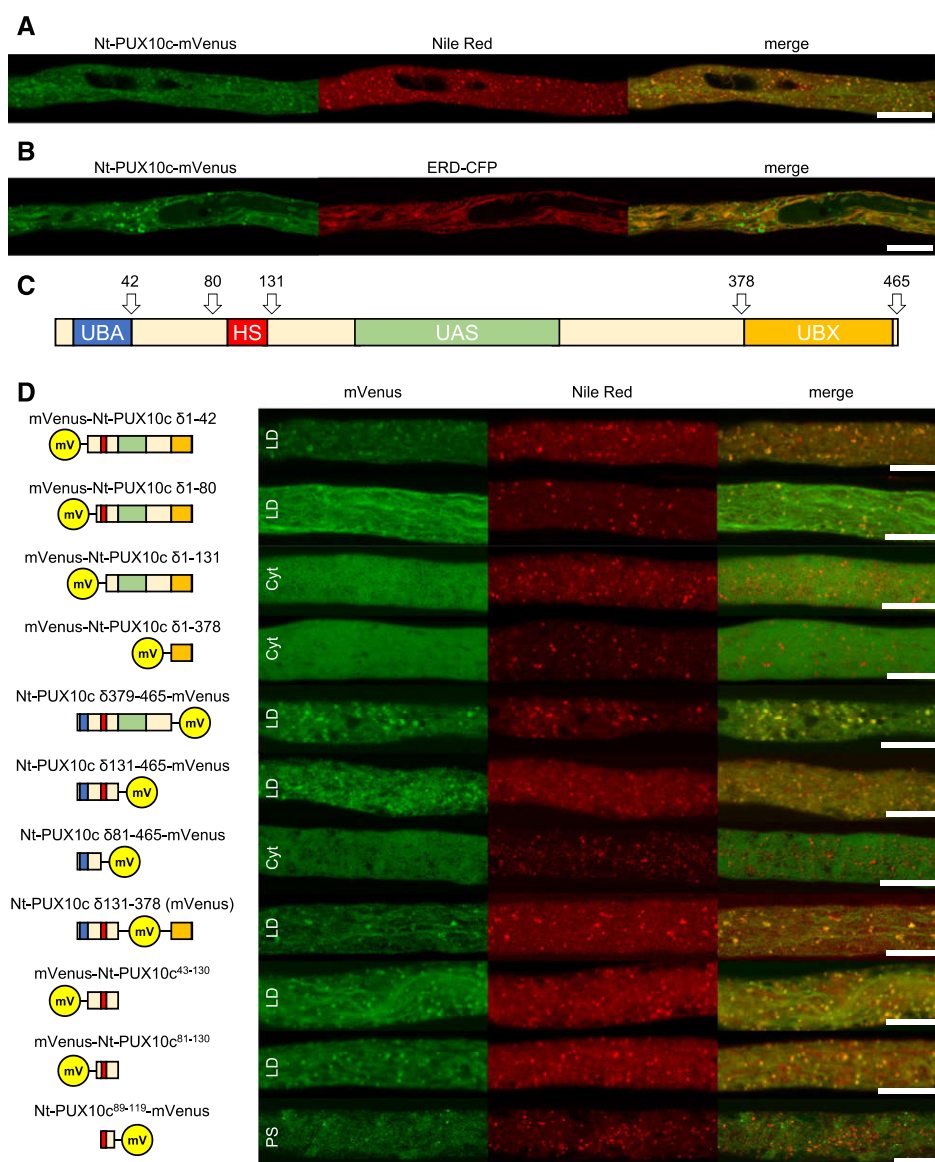
(Figure 4A). However, when At-CDC48A was coexpressed with Nt-PUX10c-mCherry, both proteins clearly colocalized to LDs stained with the LD marker stain monodansylpentane (Figure 4B). Notably, this relocalization of At-CDC48A to LDs was not observed when At-CDC48A was coexpressed with Nt-PUX10c lacking its UBX domain (Figure 4C), indicating that the recruitment of At-CDC48A to LDs by Nt-PUX10c is mediated by its UBX domain. Similarly, At-CDC48A was relocalized to LDs when

coexpressed with the Arabidopsis homolog of Nt-PUX10c, i.e., At-PUX10 (Figures 4D and 4E).

#### At-PUX10 Localizes to LDs in Pollen Tubes, Developing Embryos, and Seedlings

To further investigate the tissue-specific expression of At-PUX10 and its subcellular localization in a native system, the entire





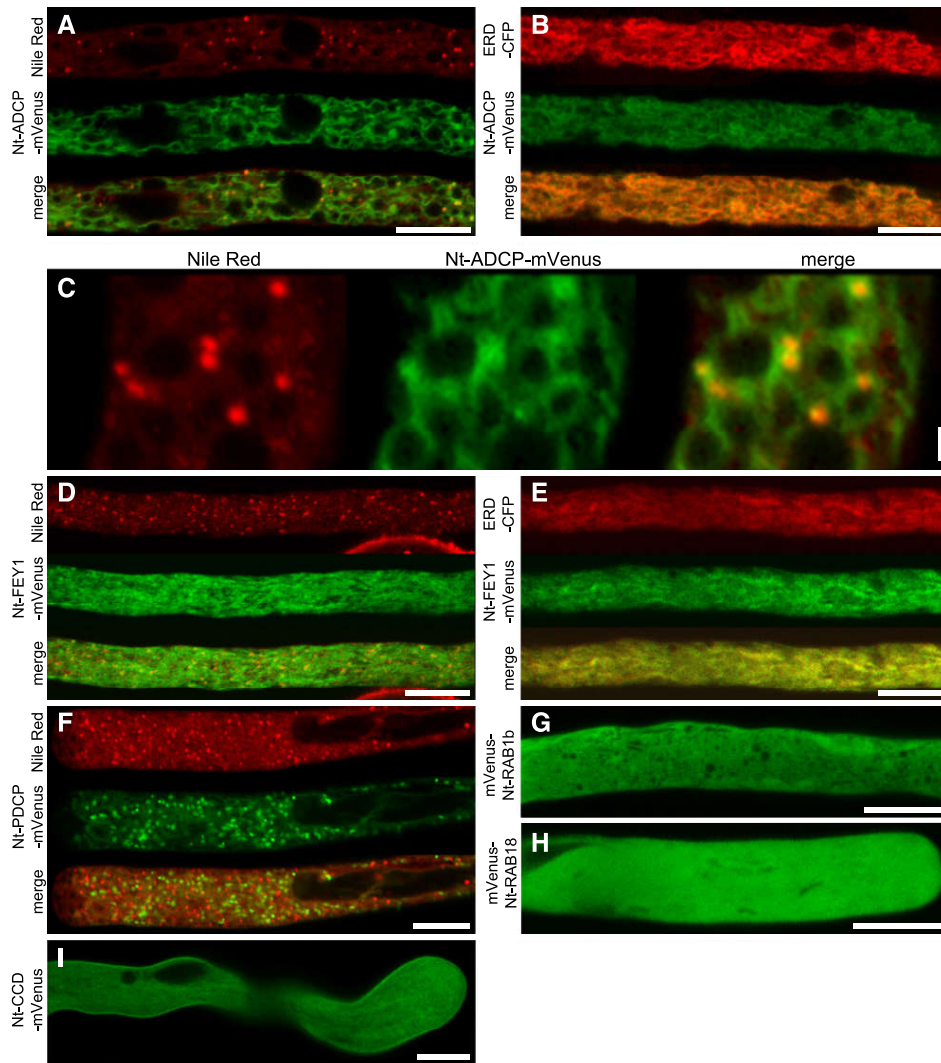
**Figure 2.** The Hydrophobic Stretch Is Both Essential and Sufficient for LD Association.

**(A)** and **(B)** The full-length Nt-PUX10c protein with a C-terminal mVenus fusion localizes to LDs and to the ER highlighted by the ER marker ERD-CFP. **(C)** The overall domain structure of Nt-PUX10. The protein contains a hydrophobic stretch (HS; hydrophobic region predicted with TMHMM; see Supplemental Figure 13) a UBA domain predicted to interact with ubiquitinated proteins, a UAS domain of unknown function, and a UBX domain commonly considered to interact with CDC48-type proteins. Domain prediction made with CD-Search.

**(D)** All variants including the region comprising amino acids 81 to 131 colocalized at least in part with LDs, while all constructs excluding this region were cytosolic (indicated by Cyt). A construct missing amino acids 1 to 88 and 120 to 465 localized to punctate structures (indicated by PS) not identical with LDs. The proteins were transiently expressed in tobacco pollen tubes, as C- or N-terminal fusions to the fluorescent protein mVenus. The tubes were cultivated for 5 to 8 h, fixed with formaldehyde, and stained with Nile Red **(A)** and **(D)** or not fixed and stained **(B)**. Similar localization was observed in  $n = 6$  to 11 independent pollen tubes per construct. Bars = 10  $\mu$ m. See Supplemental Figure 14 for magnified images.

genomic sequence of At-PUX10, including a 642-bp promoter region in frame with an eGFP-encoding sequence (i.e., *At-PUX10<sub>pro</sub>:At-PUX10-eGFP*), was cloned and 10 stable transgenic lines were generated, from which three were chosen for further analysis (based on the relative strength of the fluorescence signal attributable to *At-PUX10<sub>pro</sub>:At-PUX10-eGFP* in transgenic seedlings).

In all lines, At-PUX10-eGFP fluorescence was observed in pollen tubes grown in vitro (Figure 5) and in developing embryos starting at the heart stage through to mature embryos (Figures 5B and 5C). In the latter case, At-PUX10-eGFP fluorescence levels were somewhat weaker, but during seed germination they increased again (Figures 5H to 5J). At all stages, At-PUX10-eGFP was found



**Figure 3.** Proteins Not Localizing to LDs.

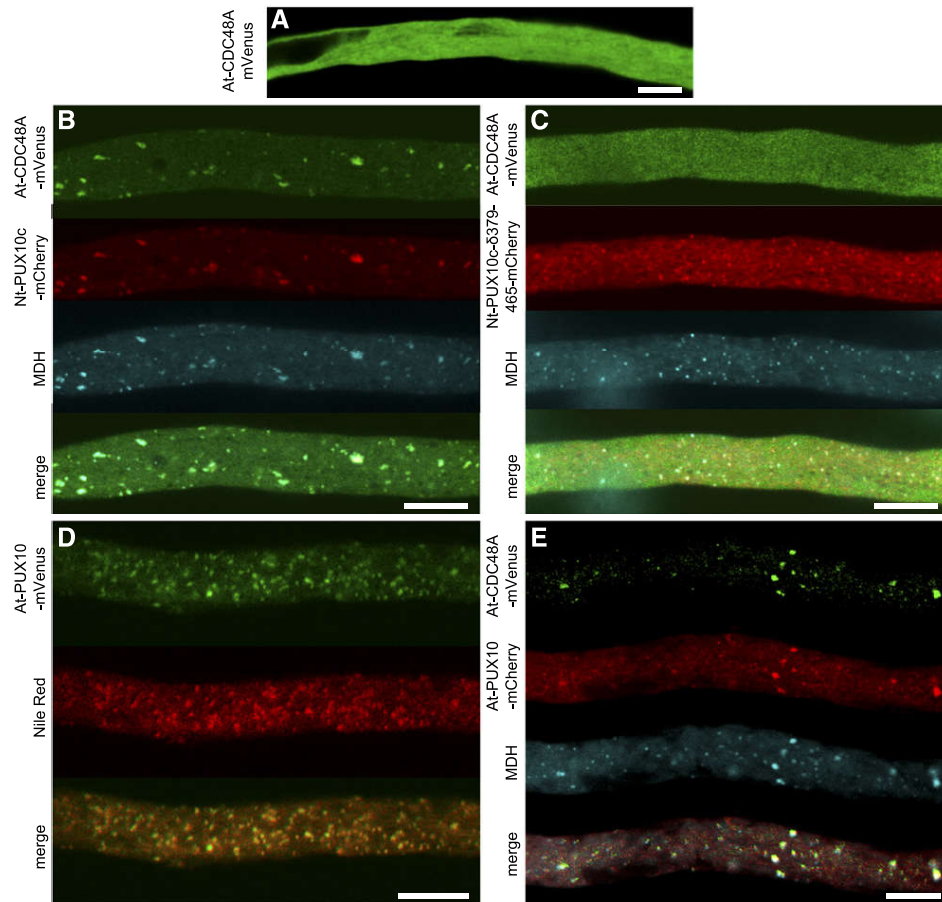
The proteins were transiently expressed in tobacco pollen tubes, as N-terminal ([A] to [F] and [I]) or C-terminal ([G] and [H]) fusions to the fluorescent protein mVenus. The tubes were cultivated for 5 to 8 h, fixed with formaldehyde, and stained with Nile Red ([A], [C], [D], and [F]), only fixed ([B] and [E]), or not treated ([H] and [I]). Then, they were monitored by confocal microscopy. (C) is a magnified section of (A), showing partial association of mVenus and Nile Red fluorescence. Images are representative of 10 ([A], [C] to [E], and [H]), 14 ([B] and [I]), 7 ([C] and [F]), and 13 (G) pollen tubes. Bars = 10  $\mu$ m in (A), (B), and (D) to (I) and 1  $\mu$ m (C).

predominantly on LDs (Figures 5D to 5G). Interestingly, in seedlings, At-PUX10-eGFP fluorescence was not equally distributed between all LDs, but, instead, appears to be concentrated on only a subpopulation of LDs (Figure 5K). This uneven distribution persisted in isolated LDs (Supplemental Figure 15). On the other hand, in comparison to Nt-PUX10-mVenus transiently expressed in tobacco pollen tubes (Figure 2A), no localization of At-PUX10-eGFP to ER-like structures was observed.

#### LD Size Is Altered in Arabidopsis *pux10* T-DNA Insertion Mutants

For functional characterization of At-PUX10, T-DNA insertion lines of Arabidopsis were obtained from the NASC seed stock

center and the insertions of the T-DNA in *PUX10* were determined to be in the promoter region (SALK\_139056 in the Col-0 background; referred to here as *pux10-2*) and the first exon (SAIL\_1187 B06 in the *qrt* [quartet] background; *qrt pux10-1*). Another mutant line was obtained from INRA Versailles-Grignon Center with an insertion in the second exon of *PUX10* (FST EAT-TV209T3 in the Ws-4 background; *pux10-3*). All insertions were confirmed by genotyping PCR and sequencing (Figure 6A; Supplemental Data Set 17). *qrt PUX10* and *qrt pux10-1* plants are offspring of the same heterozygous plant (*qrt pux10-1/PUX10*). Additionally, RT-qPCR was performed to assess transcript levels in each of the *pux10* mutants and their respective parental backgrounds. Toward that end, RNA was isolated from dry seeds of each line and, as shown in Supplemental Figure 16, for *qrt*



**Figure 4.** PUX10 Can Recruit At-CDC48A to LDs in Pollen Tubes with the Help of Its UBX Domain.

The proteins were transiently expressed in tobacco pollen tubes as N-terminal fusions to the fluorescent protein mVenus or mCherry as indicated. The tubes were cultivated for 5 to 8 h, fixed with formaldehyde, and stained with the blue fluorescent LD dye monodansylpentane (MDH; **[B]**, **[C]**, and **[E]**) or Nile Red (**[D]**). Then, they were monitored by confocal microscopy. At-CDC48A expressed alone is localized in the cytosol (**A**) but is recruited to LDs by Nt-PUX10c (**B**). Nt-PUX10c missing the C-terminal UBX domain (Δ379-465) still localizes to LDs but can no longer recruit At-CDC48A (**C**). At-PUX10 alone localizes to LDs (**D**) and can also recruit At-CDC48A. Images are representative of 11 (**A**), 9 (**B**), 10 (**C**), 12 (**D**), and 10 (**E**) pollen tubes. Bars = 10  $\mu$ m.

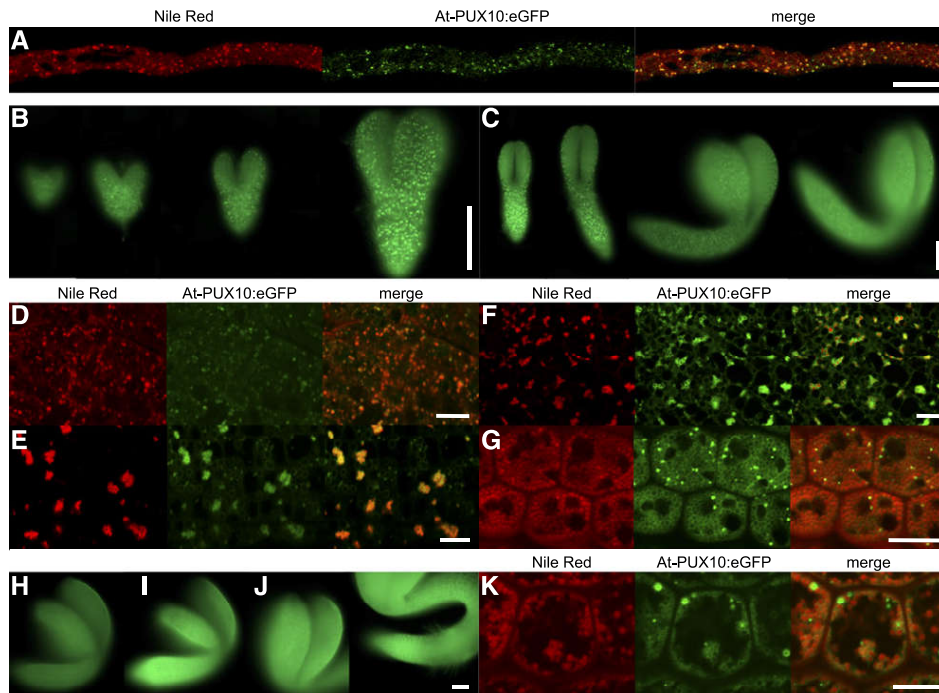
*pux10-1* and *pux10-2*, *PUX10* transcripts were reduced to 10% and 25%, respectively. For *pux10-3*, no *PUX10* transcript was detected (Supplemental Figure 16). Despite the fact that *PUX10* expression in the *qrt pux10-1* mutant was not abolished completely, no PUX10 peptides were detected in all but one sample of LDs isolated from the *qrt pux10-1* mutant (Supplemental Figure 17).

In all three *pux10* mutant lines, LD morphology during germination was altered when compared with their respective nonmutated backgrounds. For instance, in Col-0 (background to *pux10-2*) and *qrt PUX10* (background to *qrt pux10-1*) seedlings, LD size increased while their number decreased during the progress of germination (Figure 6B). However, in both of the corresponding *pux10* mutant lines, LDs appeared to stay smaller and were higher in number at later stages of germination (36 and 48 h). Indeed, quantification of LD sizes in five different seedlings per line (at 36 h) revealed that LDs in *qrt pux10-1* and *pux10-2*

seedlings were significantly smaller than in the respective backgrounds (Figure 6C, left panel), and in particular the number of LDs larger than 1  $\mu$ m were reduced (Figure 6D, left panel).

At a similar time point (40 h after imbibition), the LDs in the hypocotyl cells of two complemented lines (complemented with *At-PUX10<sub>pro</sub>::At-PUX10-eGFP*, termed complemented 1 [C#1] and complemented 2 [C#2]) were significantly larger than the LDs in the *qrt pux10-1* mutant, but did not differ significantly from the *qrt PUX10* background (Figure 6C, middle panel). Furthermore and similar to the results described above, the proportion of LDs with a diameter of more than 1  $\mu$ m was similar in the *qrt PUX10* background and the two complementation lines, but significantly reduced in the *qrt pux10-1* mutant (Figure 6D, middle panel). These results were confirmed using the third independent *pux10* T-DNA insertion mutant line, *pux10-3*. That is, LDs in *pux10-3* seedlings at 40 h after imbibition were significantly smaller than those in the Ws-4 background (Figure





**Figure 5.** PUX10-eGFP Is Present at LDs during Embryo Development, Seed Germination, and in Pollen Tubes When Expression Is Driven under the Native Promoter.

The construct *At-PUX10<sub>pro</sub>::gAt-PUX10:eGFP* was used to stably transform Arabidopsis plants. Three of originally 10 independent lines were chosen for further investigation. eGFP fluorescence colocalized with Nile Red in Arabidopsis pollen tubes. The tubes were cultivated for 2 h and then fixed with formaldehyde and stained with Nile Red. Fluorescence was also detected by epifluorescence microscopy throughout embryo development ([B] and [C]) and in germinating seedlings ([H], imbibed seed; [I], 24 h; and [J], 48 h after imbibition; both images are taken from the same seedling). All images were obtained with identical settings, under which wild-type plants did not show any significant signal. Colocalization with Nile Red-stained LDs was observed by confocal microscopy in the heart (D), torpedo (E), walking-stick (F), and mature stage (G) as well as in the seedling 24 h after imbibition (K). In the mature stage and in seedlings, fluorescence was increased around individual LDs. All images are representative of at least five embryos, seedlings per stage, or five pollen tubes for each of three independent lines. Bars = 100  $\mu$ m in (B), (C), and (H) to (J) and 10  $\mu$ m in (A), (D) to (G), and (K).

6C, right panel). Furthermore, almost no LDs with a diameter above 1  $\mu$ m were observed in the *pux10-3* mutant, whereas ~20% of the LDs of that size were seen in *Ws-4* seedlings (Figure 6D, right panel). Overall, these results confirm that the observed phenotype in LD size and number is due to a disruption of *PUX10* expression.

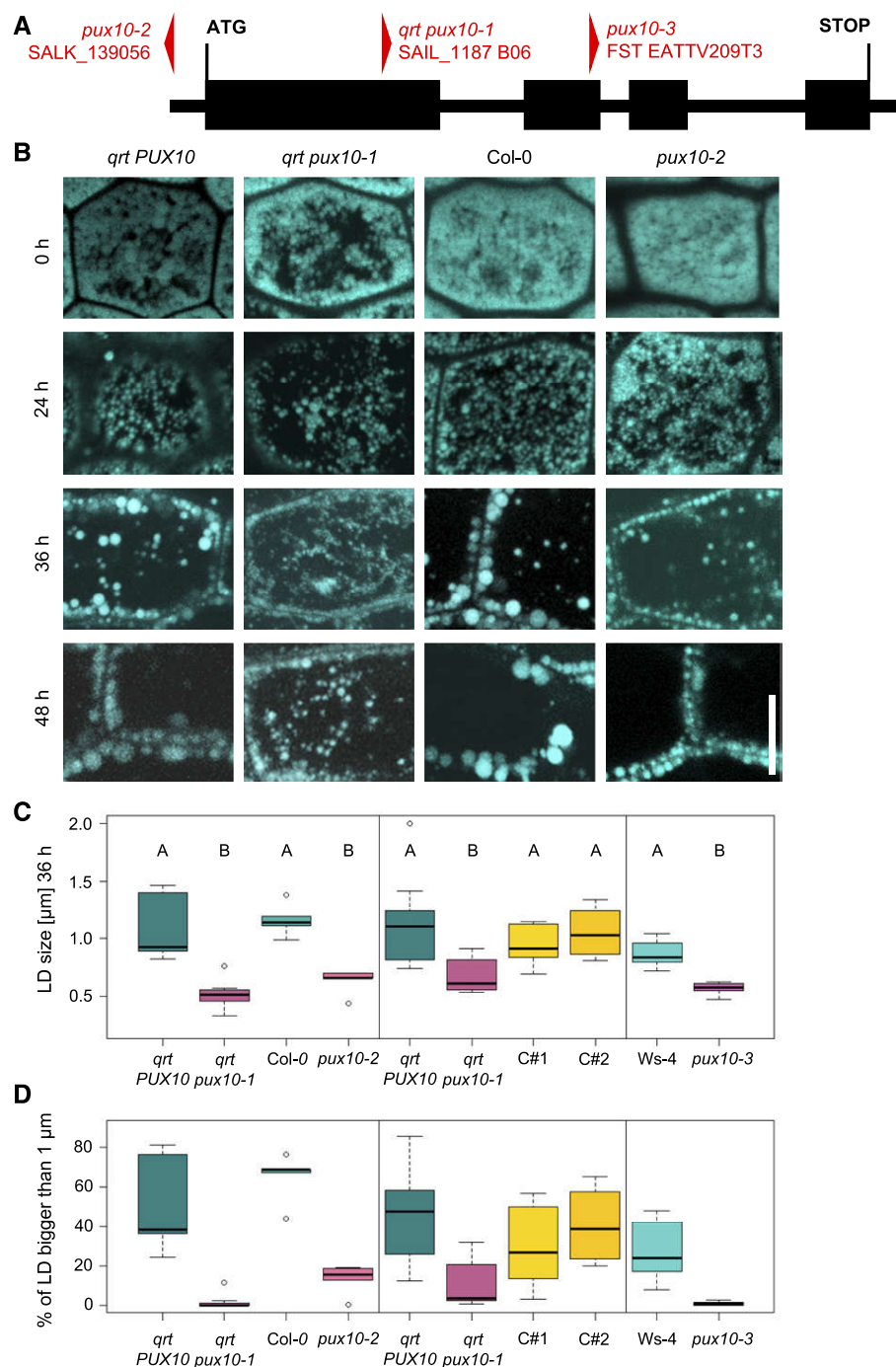
#### LD Proteins Are Degraded More Slowly in the Arabidopsis *pux10* Mutants

It has been shown previously that the removal of a LD protein affects LD morphology in plants (Pyc et al., 2017a; Huang, 2018). For instance, Siloto et al. (2006) reported that the disruption of expression of the major structural LD protein in Arabidopsis seeds, OLEOSIN1, results in a dramatic increase in LD size, indicating a link between the abundance of structural LD proteins and LD size. As such, PUX10, by recruiting CDC48 to LDs and thereby possibly regulating LD degradation, might influence the abundance of LD proteins. To test this hypothesis, we took a proteomic approach, whereby proteins isolated from homogenized

seedlings of different lines (*qrt PUX10*, *qrt pux10-1*, C#1, and C#2, as well as *Ws-4* and *pux10-3*) and different developmental stages (rehydrated seeds, and seedlings 1 to 3 d after imbibition [DAI]) were subjected to tryptic digestion and LC-MS/MS analysis (Figure 7; Supplemental Figure 16). The resulting data were processed with the MaxQuant label-free quantification (LFQ) algorithm for each time point (Supplemental Data Sets 18 to 21) and calculated as the % value of total LFQ intensities (Supplemental Data Sets 22 to 27).

In total, 1690 proteins were detected and quantified with the complexity of the proteome strongly increasing during germination. The sum of these relative LFQ intensities for 13 major LD proteins (i.e., the oleosins OLE1-8, HSD1, and HSD5, LDAP2, and the caleosins CLO1 and CLO2) was calculated for every time point and genetic background (Supplemental Data Set 28). While the relative abundance of the LD proteins in the total fractions strongly decreased in all lines within 3 d, at 1 and 2 DAI, significantly more of these proteins were present in the *qrt pux10-1* mutant when compared with the *qrt PUX10* background (Figure 7A). This effect was rescued in C#2 at 1 DAI. At 2 DAI, especially





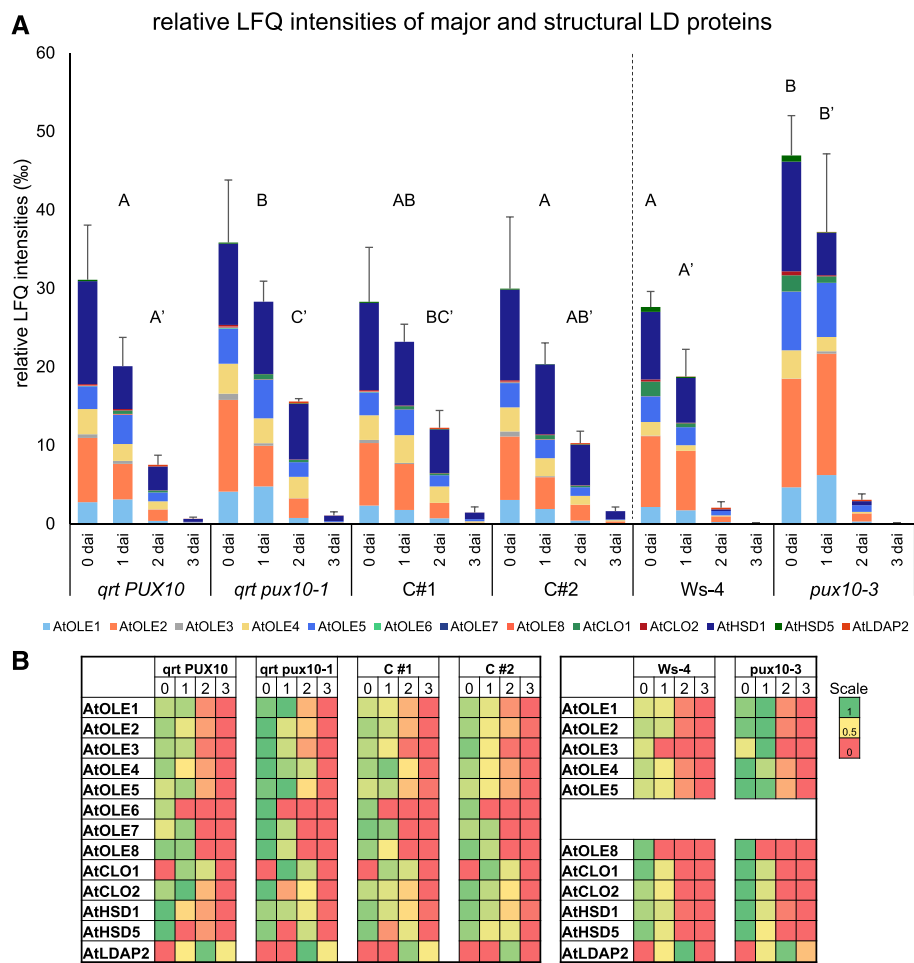
**Figure 6.** LDs in Seedling Hypocotyls Are Smaller in the *pux10* Mutants.

**(A)** Localization of the T-DNA insertions in *PUX10*. The gene consists of four exons (bold line) and three introns (thin lines). The SALK\_139056 (*pux10-2*) insertion is localized 100 bp in 5' direction from the start codon in the promoter region, the SAIL\_1187 B06 (*qrt pux10-1*) insertion is 637 bp into the first exon, and the FST EATTV209T3 (*pux10-3*) insertion 387 bp into the second exon.

**(B)** to **(D)** Seeds were imbibed for 16 h at 4°C and subsequently grown under continuous light (**[B]** and **[C]**, left panel) or long-day conditions (**[C]**, middle and right panels).

**(B)** Confocal images of the hypocotyl represent single slides (0 and 24 h) or stacks (36 and 48 h). Images are excerpts from larger images and representative of 10 seedlings (from two independent experiments). Bar = 10  $\mu\text{m}$ .

**(C)** and **(D)** Confocal scans from seedlings (36 h) were used for quantification and the average diameter of the LDs determined from one image of each plant ( $n = 4$ –11 independent seedlings per line, 20–188 LDs per seedling). Absolute average sizes of LDs (**[C]**) and the percentage of LDs bigger than



**Figure 7.** LD Proteins Are Enriched in the *pux10* Mutant Seedlings.

**(A)** In the *qrt pux10-1* line, LD proteins are significantly more abundant at 1 and 2 DAI when compared with the *qrt PUX10* line. This effect is partially rescued in the complemented lines. Proteins were isolated from total seedling homogenate of rehydrated seeds (0 DAI) and seedlings at 1, 2, and 3 DAI. The same experiment was independently conducted on seeds and seedlings of Ws-4 and *pux10-3*. Here, the observed difference is greatest at 0 and 1 DAI. LC-MS/MS data were processed with MaxQuant software's LFQ algorithm. Values are given as % of total LFQ intensities. Error bars correspond to SD, when comparing the summed up intensities of the 13 given proteins between the three biological replicates. The same summed up intensities for each biological replicate were used for one-factor ANOVA followed by Tukey's post-hoc test; lowercase letters indicate statistical significance,  $P < 0.05$ .

**(B)** The same values as in **(A)** were normalized for each protein to the highest value in all compared lines and time points. Green shows high abundance and dark red not detectable. For separate graphs, see Supplemental Figure 18, and for all individual values, see Supplemental Data Sets 18 and 29.

the abundances of several individual proteins (namely, OLE4, OLE5, and HSD1; Figure 7B; Supplemental Figure 18A) were higher in the *qrt pux10-1* mutant than in the wild type. Similar analysis of the *pux10-3* mutant line (Supplemental Data Sets 29 to 39) also showed a higher abundance of LD proteins at 0 and 1 DAI (Figure 7A). Furthermore, the degradation of LD proteins was relatively faster in comparison to that in *pux10* mutant lines in the Col-0 background, presumably due to the difference in the ecotype. In the *pux10-3* line, again several individual proteins showed a higher abundance, especially at 1 DAI, where the oleosins

**Figure 6.** (continued).

1  $\mu\text{m}$  are presented in **(C)** and **(D)**, respectively. LD size was quantified with ImageJ and statistically analyzed by one-factor ANOVA followed by Tukey's post-hoc test. The test indicated a significant difference ( $P < 0.02$ ) for all *pux10* mutants in comparison to all of the respective control lines. Plots were created in R.

Results presented in **(B)**, and in the left panels in **(C)** and **(D)**, originate from the same experiment. Different panels represent independent experiments using seed material of independently grown plants.

OLE1, OLE4, and OLE5 were significantly increased (Figure 7B; Supplemental Figure 18B). Overall, however, LD protein degradation was not completely abolished in any of the *pux10* mutants, but merely slowed down. In contrast to protein degradation, TAG degradation and postgerminative growth were not changed as measured by the decrease of the marker fatty acid C20:1 (Supplemental Figure 19).

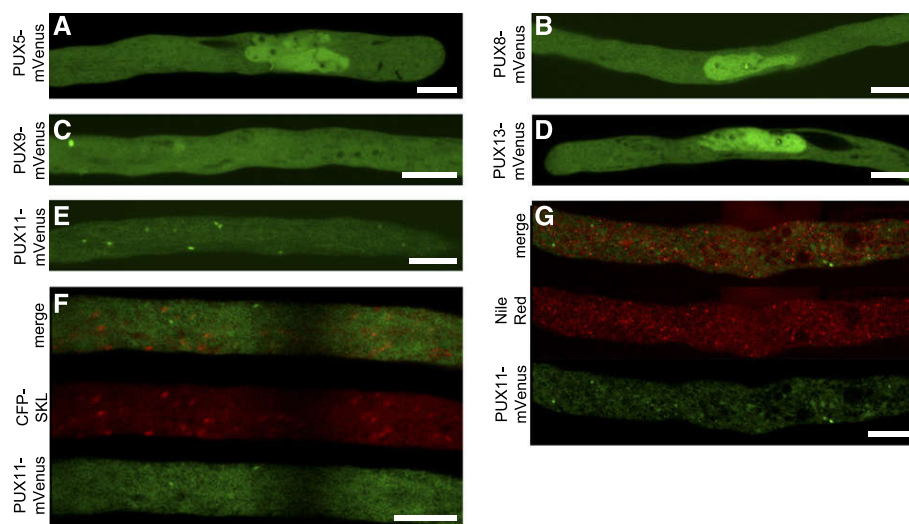
To investigate whether one of the other proteins of the *PUX* gene family containing the UBA and UBX domains might have redundant functions with *PUX10* (Supplemental Figure 12), we cloned and analyzed the subcellular localization of the Arabidopsis proteins *PUX5*, *PUX8*, *PUX9*, *PUX11*, and *PUX13* (fused to mVenus) by transient expression in tobacco pollen tubes. As shown in Figure 8, *PUX5*, *PUX8*, and *PUX13* localized to the nucleus, similar to what has been described previously for *PUX7* (Gallois et al., 2013), whereas *PUX9* was cytosolic. *PUX11* localized also to the cytoplasm, as well as to punctate structures, which were ruled out to be LDs or peroxisomes (Figure 8).

#### Ubiquitin Is Enriched in LD-Enriched Fractions of the *qrt pux10-1* Mutant

To gain more detailed insight into the protein composition of LDs of the *qrt pux10-1* mutant line and to increase the chances of detecting posttranslationally modified peptides, LD-enriched fractions were analyzed in another proteomics-based approach. Here, the LD fraction from homogenized seedlings of *qrt PUX10*,

*qrt pux10-1*, C#1, and C#2 at 38 h after imbibition (Supplemental Data Sets 40 to 43) was isolated, but prior to fractionation by centrifugation, aliquots of total protein were taken for comparison. Again, the different samples were tryptically digested and analyzed by LC-MS/MS. The data were processed as described above, and, overall, there was an enrichment of LD proteins in the LD fraction, as expected (Supplemental Data Set 42). In addition, the relative LFQ intensities of ubiquitin in the LD fraction of *qrt pux10-1* seedlings was significantly increased compared with all control plants (Figure 9A). Ubiquitin was also 2-fold enriched in the LD fraction compared with the total fraction ( $P < 0.02$ ). Moreover, an immunoblot analysis against ubiquitin indicated stronger ubiquitination of proteins in the LD fraction of the *qrt pux10-1* line in comparison to the wild type and complemented lines (Supplemental Figure 20).

To detect potential ubiquitination sites on a qualitative level, the proteomic data were analyzed for peptides containing lysine residues with two attached glycines, which remain at the C terminus of the attached ubiquitin after tryptic digestion. More specifically, data obtained for *qrt PUX10* and *qrt pux10-1* seedlings at 1 and 2 DAI were analyzed (Supplemental Data Sets 44 to 46), and, based on this, several ubiquitination sites were identified. The ubiquitination position within the ubiquitin protein, Lys-48 (K48) was identified (Figure 9B). Ubiquitination positions were also detected for HSD1 (K269) and OLE4 (K157, K159, and K168) (Figure 9B; see Supplemental Figures 21 to 24 for fragmentation spectra). These proteins are among the proteins significantly enriched in *qrt pux10-1* seedlings at 2 DAI



**Figure 8.** Subcellular Localization of Other Members of the Arabidopsis *PUX* Gene Family.

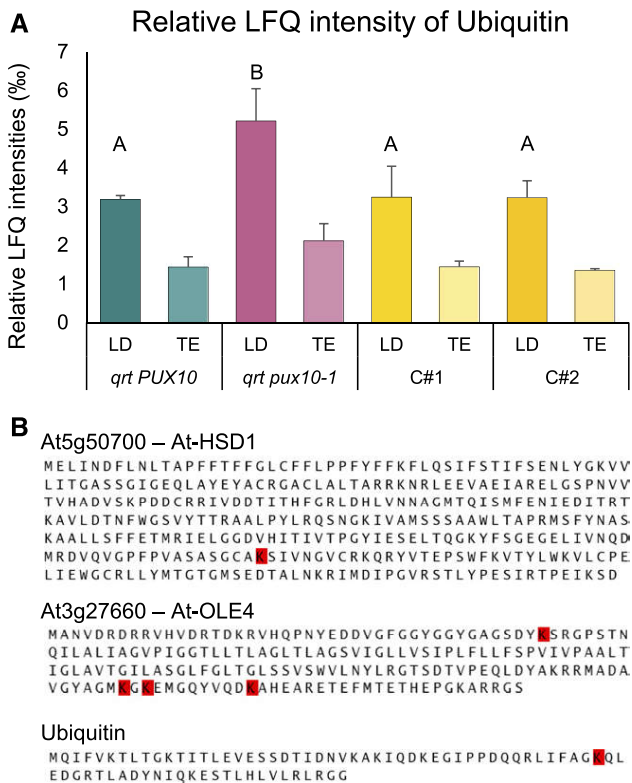
To analyze whether *PUX10* has functionally redundant gene family members that also localize to LD and/or the ER, the following genes were cloned, transiently expressed as N-terminal fusions to mVenus in tobacco pollen tubes, fixed, and observed by confocal microscopy. Images are representative of 6 (**A**), 11 (**E**), and 8 (**F**), 11 (**B**), 8 (**C**), 9 (**D**), and 5 (**G**) pollen tubes. Bars = 10  $\mu$ m.

(**A**), (**B**), and (**D**) *PUX5*, *PUX8*, and *PUX13* localize preferentially to the nucleus and the cytosol.

(**C**) *PUX9* is localized in the cytosol.

(**E**) *PUX11* localizes to punctate structures in the cytosol.

(**F**) and (**G**) The structures were excluded to be LDs or peroxisomes by staining with Nile Red (**G**) or by coexpression with the peroxisomal marker CFP-SKL (**F**), respectively.



**Figure 9.** Ubiquitin Accumulates in LD-Enriched Fractions of the *qrt pux10-1* Mutant Seedlings.

(A) Proteins were isolated from LD fractions and the respective total protein fractions (TE) of seedlings at 2 DAI. Peptides derived from these fractions were analyzed via LC-MS/MS. The LFQ algorithm was applied, and values are presented as averages of three biological replicates in % of total LFQ. Error bars indicate sd. Statistical analysis was performed by one-factor ANOVA followed by Tukey's post-hoc test. Lowercase letters indicate significant differences ( $P < 0.05$ ).

(B) Ubiquitination sites were detected in the LD proteins HSD1, OLE4, and ubiquitin via LC-MS/MS analysis of their tryptic peptides (see Supplemental Figures 22 to 24 for spectra). Ubiquitination sites are highlighted in red.

(Supplemental Figure 18 and Supplemental Data Sets 28 and 39). All proteomic data of LDs isolated from *qrt PUX10*, *qrt pux10-1*, at 1 and 2 DAI as well as from *qrt PUX10*, *qrt pux10-1*, C#1, and C#2 at 38 h after imbibition are displayed in the supplements (Supplemental Data Sets 40, 41, 43, and 44). All ubiquitinated peptides are displayed in Supplemental Data Sets 43 and 46.

## DISCUSSION

### The Composition of LDs Differs from Pollen Tubes to Seeds

Tobacco pollen grains store smaller amounts of neutral lipids (2% of the dry weight; Rotsch et al., 2017) in comparison to many dicotyledon oilseeds like Arabidopsis (30% of the dry weight; Cai et al., 2017), and the LDs appear smaller (Rotsch et al.,

2017). In addition, the LDs in pollen tubes and oilseeds differ in their neutral lipid composition: Sterol esters make up a much larger proportion of neutral lipids (24% on a weight basis; Rotsch et al., 2017) in pollen tubes in comparison to seeds from Arabidopsis (~0.3% of the level of TAG; Bouvier-Navé et al., 2010). While seeds degrade their neutral lipids following germination, tobacco pollen tubes synthesize TAG (Mellema et al., 2002) and sterol esters (Müller and Ischebeck, 2018), and pollen tubes cultivated for 14 h still contain large amounts of these neutral lipids (Müller and Ischebeck, 2018). By contrast, pollen tubes from olive (*Olea europaea*) degrade their LDs during pollen tube growth (Zienkiewicz et al., 2013). It remains unclear, though, if LDs are synthesized de novo in tobacco pollen tubes. Evidence for the importance of LDs for pollen tube growth comes in part from recent observations of Arabidopsis pollen disrupted in all three *SEIPIN* genes (Taurino et al., 2018). These pollen grains contain conspicuously enlarged LDs that cannot enter the pollen tube, leading to reduced fertility. LD degradation appears to be important to some degree, as mutants disrupted in a LD-associated TAG lipase show reduced pollen tube growth (Müller and Ischebeck, 2018). Hence, while these data support the importance of LDs during pollen tube growth, it remains to be determined what specific function(s) these organelles serve. Differences in protein and/or lipid composition of LDs in different tissue types might both reflect different function(s) of the organelle. For instance, similar to other studies on pollen LDs (Kim et al., 2002; Jiang et al., 2007), we found two oleosins as two of the most abundant proteins (Table 1). Their presence could be explained by the desiccation phase that pollen grains undergo before rehydration, similar to seed desiccation, where oleosins are also present in relatively high amounts (Lee et al., 1994; Huang, 1996; Wang et al., 1997). On the other hand, we also detected steroleosins in tobacco seeds (Supplemental Data Set 6) and Arabidopsis seedlings (Supplemental Data Sets 39, 42, and 46), but not in pollen tubes (Supplemental Data Set 2), indicating that their potential function in brassinosteroid synthesis is missing in pollen tubes (Li et al., 2007). By contrast, at least three caleosin isoforms were detected in pollen tubes (Supplemental Data Set 4). Two of these caleosins, Nt-CLO1a and Nt-CLO1b, were the most and third most abundant LD proteins based on the relative iBAQ score (Table 1), indicating that caleosins are more predominant in the LDs of pollen tubes than in seeds and seedlings, where oleosins predominate (Pyc et al., 2017b; Supplemental Data Sets 42 and 46). Nt-CLO1a and Nt-CLO1b are most closely related to the isoforms CLO1 and CLO2 from Arabidopsis (Supplemental Figure 4), which display peroxxygenase activity (Hanano et al., 2006; Poxleitner et al., 2006), indicating that such an activity could also be of importance in pollen tubes.

Similar to caleosins, the recently described LDAPs (Horn et al., 2013) and LDIP (Pyc et al., 2017b) showed a much higher relative abundance in tobacco pollen tubes (Supplemental Data Set 4) compared with Arabidopsis seedlings (Pyc et al., 2017b) and were not detected in tobacco seeds (Supplemental Data Set 6). Interestingly, Nt-LDIPa did not localize to all LDs, when transiently expressed in tobacco pollen tubes (Figures 1D and 1E) as its Arabidopsis homolog does in *Nicotiana benthamiana* leaves (Pyc et al., 2017b), suggesting the existence of different subpopulations of LDs in pollen tubes. Future studies using

pollen tubes as model systems should help elucidate the function of LDIP and its sorting to distinct LDs.

Our proteomic and cell biological data indicated that sterol synthesis is a potential novel function for LDs, since we detected both a putative cycloartenol synthase (Nt-CAS1) and a cycloartenol-C-24-methyltransferase (Nt-SMT1) enriched in LD fractions (Table 1). While both Nt-CAS1 and At-CAS1 clearly localized to LDs in tobacco pollen tubes (Figure 1G; Supplemental Figure 11), the localization of Nt-SMT1 was less clear (Figure 1H and 1I), indicating that it is not exclusively localized at LDs. In contrast to At-CAS1, At-SMT1 is also less strongly enriched in Arabidopsis seedling LDs based on proteomic data (Pyc et al., 2017b; Supplemental Data Sets 42 and 46). As such, while we found on LDs two of the four enzymes required for the truncated sterol synthesis pathway from squalene to cycloeucalenol described for tobacco pollen tubes (Villette et al., 2015; Rotsch et al., 2017), it remains to be determined if sterols are synthesized on LDs in tobacco pollen tubes or on LDs in other tissues. Interestingly, some enzymes of the newly discovered alternative cholesterol synthesis pathway have been speculated to be localized at LDs (Sonawane et al., 2016), further reinforcing the growing functional repertoire of this organelle.

Another enzyme we identified enriched in the LD fraction (Table 1) is a tobacco homolog of the castor bean (*Ricinus communis*) oil body lipase (Rc-OBL; Eastmond, 2004). Recently, we confirmed that both Nt-OBL and its Arabidopsis homolog At-OBL1 are LD-associated and that At-OBL1 is a TAG, diacylglycerol, and 1-monoacylglycerol lipase and that the activity of At-OBL1 is important for pollen tube growth (Müller and Ischebeck, 2018). We also found a previously unknown LD protein, PTLDb (Table 1, Figure 1F), and via BLAST analysis, we identified four additional tobacco homologs (Supplemental Figure 10), all of which were enriched in LD fractions isolated from tobacco pollen tubes (Supplemental Data Set 4). PTLD proteins are evolutionarily conserved with homologs present in many species, including plants important for oil production such as olive, oil palm (*Elaeis guineensis*), soybean (*Glycine max*), and maize (*Zea mays*), but not in Brassicaceae. It remains to be resolved, however, if this protein is only specific for pollen tubes or if it is also found in seeds. Finally, one of the most notable findings from our proteomics analysis of the LD fraction of tobacco pollen tubes in terms of potentially novel LD proteins were Nt-PUX10 proteins, although At-PUX10 was found to be enriched in the LD fraction of Arabidopsis seedlings (Pyc et al., 2017b; Supplemental Data Set 42). In addition, in humans, the UBX domain-containing protein UBXD8, which is normally ER localized, relocates to LDs upon treatment of cultured cells with the fatty acid oleate (Olzmann et al., 2013).

Overall, the verification of LD localization via an appropriate screening method as presented here with transient expression in tobacco pollen tubes efficiently eliminates false positive candidates determined by proteomic approaches and identified several novel LD proteins.

### Determinants of LD Localization

One culprit that limits the discovery and characterization of LD proteins is the lack of a known consensus signal sequence for

targeting the organelle. Targeting sequences are a conserved concept for the correct distribution of proteins to subcellular compartments and in past years, in silico analyses of sequenced genomes allowed for the reliable prediction of the subcellular localization of putative proteins (Meinken and Min, 2012). However, such a conserved sequence has not been described for known LD proteins, complicating both the discovery of additional LD proteins and the confirmation of their subcellular localization.

Nonetheless, one common feature among at least some plant LD proteins is a hydrophobic domain that presumably anchors the protein to the LD. For instance, oleosins are embedded in the TAG core of the LD through their protruding hydrophobic sequence that is predicted to form a hairpin-like structure with a proline knot motif (Jolivet et al., 2017). Similarly, caleosins and steroleosins have elongated hydrophobic sequences with a proline knot and knob, respectively, that anchors them to the LD (Tzen, 2012).

Both Nt-PUX10c and Rc-OBL contain N-terminal hydrophobic sequences, which are necessary for LD localization (Figure 2; Eastmond, 2004). However, the hydrophobic sequence of Nt-PUX10c alone (amino acids 94–116, Supplemental Figure 26) is not sufficient for the targeting to LDs (Nt-PUX10c<sup>89–119</sup>; Figure 2; Supplemental Figure 14). On the other hand, the inclusion of additional amino acid residues immediately adjacent to the hydrophobic sequence (as in Nt-PUX10c<sup>81–130</sup>; Figure 2D; Supplemental Figure 14) conveys proper localization to LDs. Using in silico studies, we also showed that this latter sequence, i.e., Nt-PUX10c<sup>81–130</sup>, is predicted to form an amphipathic helix that is evidently necessary for correct targeting (Supplemental Figure 26). Similarly, the PTLDs described here also contain a prominent hydrophobic region that might insert into the membrane (Supplemental Figure 27). Furthermore, Pyc et al. (2017b) recently reported that the proper targeting of At-LDIP to LDs relies on a discrete region of the protein that is predicted to form an amphipathic sequence and includes a hydrophobic sequence and adjacent residues.

How LD proteins with hydrophobic sequences are properly folded in the aqueous environment of the cytosol and finally find their correct target membrane remains to be investigated.

Other LD proteins, for example LDAPs (Gidda et al., 2016) and At-CAS1 (Supplemental Figure 25), have no obvious predicted hydrophobic regions and yet localize efficiently to the LD surface. Hence, their localization is thought to be mediated in other ways, perhaps via protein-protein or protein-lipid interactions. Evidence for protein-protein interactions mediating LD targeting comes from our experiments with PUX10 and CDC48A. The AAA-type ATPase CDC48 is a conserved ubiquitous enzyme that achieves spatial and temporal specificity through interaction with scaffold proteins. Scaffold proteins that interact with CDC48 are characterized by the presence of specific domains that mediate this interaction (Baek et al., 2013). One of those domains is the evolutionarily conserved UBX domain (Schuberth and Buchberger, 2008). Previously, it was shown for At-PUX1 and At-PUX7 via yeast two-hybrid experiments that their UBX domain can interact with CDC48 (Park et al., 2007; Gallois et al., 2013). Our transient pollen tube expression system showed that At-CDC48A relocated from the cytoplasm when expressed



on its own to LDs when coexpressed with either Nt-PUX10c or At-PUX10 (Figure 4). In a similar manner, other cytosolic factors needed to exert different LD functions at certain times may be recruited, extending the number of proteins potentially interacting with this organelle.

#### Differential Localization of LD Proteins within the Same Cell

While At-PUX10 could be confirmed to localize to LDs in embryos, seeds and pollen tubes (Figure 5), it does not appear to do so in a homogenous manner in all tissues. That is, following germination, PUX10 was observed to localize to specific LDs (Figure 5B). Similarly, Nt-LDIPa did not localize to all LDs when expressed transiently in tobacco pollen tubes (Figures 1D and 1E). Subpopulations of LDs were previously observed in yeast (Eisenberg-Bord et al., 2018), human cell culture (Wolins et al., 2005; Hsieh et al., 2012; Zhang et al., 2016), and *Drosophila melanogaster* (Wilfling et al., 2013; Thul et al., 2017). Based on these studies, different subpopulations of LDs were postulated to perform different functions. In plants, differential functionality of LD subpopulations has not been explored. Nonetheless, an attractive possibility is that LDs with different functions are established during seed germination and that PUX10 is involved in the required remodeling of the LD proteome to generate such functional subpopulations. The isolation of PUX10-enriched LDs followed by proteomic characterization might shed light on this possibility.

#### The Abundance of LD Proteins Appears to Influence the Size of LDs

The size of LDs is, at least partially, known to be determined by their protein-to-oil ratio. Ting et al. (1996) analyzed two different maize strains, which differ in their oil content but not in oleosin expression, and revealed differences in LD size and shape. In the high-oil strain, LDs were larger than in the strain strongly depleted of oil. Likewise, knockout mutants of *OLE1* and *OLE2* in Arabidopsis have enlarged LDs, especially in the case of *OLE1* (Siloto et al., 2006), apparently because oleosins prevent LD fusion, or perhaps due to their role in LD synthesis.

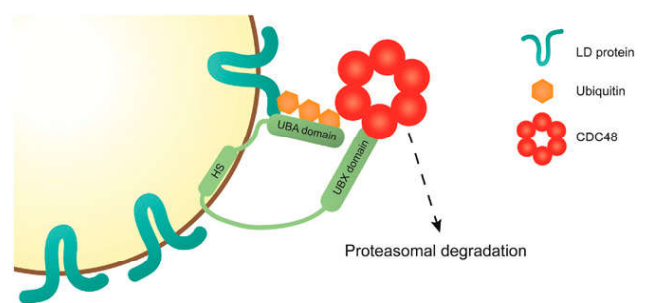
During germination and postgerminative growth in Arabidopsis, oleosins are degraded in parallel to, or even faster than, TAG (Deruyffelaere et al., 2015; Figure 7; Supplemental Figure 19). There are hints that oleosin degradation is a prerequisite to oil breakdown as it might shield the LDs from TAG lipases (Matsui et al., 1999). Thus, the increase of LD size that is observed in Arabidopsis seedlings (Figure 6) could be caused by the degradation of LD proteins leading to the fusion of LDs if neutral lipids are not degraded at the same pace. Consistent with this conclusion, a delay in the breakdown of LD proteins in the *qrt pux10-1* and *pux10-3* mutant seedlings (Figure 7) could account for the reduction in the size of LDs, in comparison to the wild type (Figure 6). Alternatively, PUX10 could be involved directly or indirectly in LD fusion, but this remains speculative, as the proteins involved in LD fusion in plants, if the process even takes place, are so far not known and it is not clear how PUX10 could be involved in such a process.

#### Degradation of LDs and LD Proteins

The degradation of LD components has been the subject of extensive research. To date, the catabolism of the stored oil in LDs, particularly TAG, is far better understood than the breakdown of the proteins associated with the organelle. The lipases responsible for the majority of TAG degradation during post-germinative growth, SDP1 and SDP1-like, have been identified (Eastmond, 2006; Kelly et al., 2011). However, degradation of the LD proteins that “coat” the surface of the organelle is far less understood.

Recently, it was suggested that oleosins are substrates for the ubiquitin-proteasome system (UPS) of the cell. Deruyffelaere et al. (2015) proposed several ubiquitination sites in different oleosins and showed that they were polyubiquitinated. We could add more ubiquitination sites in OLE4 and HSD1, and additionally detected enhanced ubiquitination of LD proteins from *qrt pux10-1* mutant seedlings (Figure 9B). However, due to the methods used, it is not possible to determine whether K48-linked ubiquitin peptides originate from oleosins. Still, our findings support the hypothesis that oleosins, and potentially other LD proteins, are substrates of the UPS and that PUX10 and CDC48 are involved in their degradation (Figure 10).

As PUX10 is also expressed during embryo development (Figures 5B and 5C), it might also be involved in protein turnover during this period. Furthermore, in the *pux10* mutants, LD protein levels were already higher in seeds imbibed for 45 min than in those from their respective backgrounds (Figure 7; Supplemental Figure 18). While the half-life of LD proteins during embryo development is unknown, an average value of 3.5 d has been measured in leaves (Ishihara et al., 2015), a time frame much shorter than embryo development (~3 weeks; Baud et al., 2002). Overall, the turnover of LD proteins during embryogenesis appears possible and our data supports the idea that PUX10 is involved in this process.



**Figure 10.** Proposed Model for the Function of PUX10 on the LD.

LD proteins destined for degradation are marked with a polyubiquitin motif by K48-linked ubiquitins. This motif is recognized by the UBA domain of PUX10. PUX10 is an LD protein, attached to the organelle via its hydrophobic stretch (HS), and an amphipathic helix N-terminal of the HS. Via its UBX domain, PUX10 interacts with CDC48, recruiting this protein to the LD. CDC48 helps to channel the polyubiquitinated substrate to the proteasome. Whether CDC48 interacts directly with the proteasome, helps to extract the substrate from the LD, or recruits additional factors remains to be determined.

The characterization of the *pux10* mutants also led to the conclusion that the action of PUX10 is by no means an exclusive mechanism related to LD protein turnover, since LD protein breakdown is only slowed, but was not abolished, in the *pux10* mutants (Figure 7). This indicates alternative pathways or at least proteins functionally redundant to PUX10 that are able to allow normal postgerminative growth.

So far, none of the other PUX proteins tested by us and others could be found to localize to LDs (Supplemental Figure 18) or, alternatively, to the ER (Park et al., 2007; Gallois et al., 2013). In other eukaryotic systems, homologs of PUX10, such as UBX2 in yeasts and UBXD8 in mammals, can shuttle between the ER and LDs (Suzuki et al., 2012; Wang and Lee, 2012). These proteins also contain a hydrophobic domain important for their proper intracellular localization and have long been identified as key components of the ERAD in their respective systems. While the plant homologs for other ERAD components have been identified, the ERAD PUX protein remains elusive (Liu and Li, 2014). This protein would be, considering the evidence from studies with yeast and mammals, a good candidate to potentially replace the function of PUX10. As this candidate remains elusive, the possibility exists that PUX10 serves this role. When expressed in tobacco pollen tubes, NtPUX10 localizes to both LDs and the ER (Figures 2A and 2B). Due to the close structural and functional relationship between the ER and LDs, it is also possible that factors that were originally identified as components of ERAD could operate at LDs. As mentioned above, Hs-UBXD8 relocates from the ER to LDs upon LD formation (Olzmann et al., 2013). Similarly, PUX10 could be LD-localized in LD-rich tissues and remain ER-bound in others. As for the ERAD pathway, whose importance is more pronounced during stress and disease (Liu et al., 2011; Guerriero and Brodsky, 2012), a functioning protein degradation machinery on the LD might also be particularly important under stress conditions, like temperature stress or pathogen infection.

The essential role of CDC48 has been studied in plants as well as other eukaryotic systems. Park et al. (2008) demonstrated the importance of At-CDC48A during various cellular processes including cytokinesis and cellular differentiation. All T-DNA insertion lines investigated in their study displayed defects in pollen tube growth and embryo development and were seedling lethal. In another study, the same group established a functional connection between At-CDC48A and another PUX protein, At-PUX1, which regulates CDC48A's hexameric structure and ATPase activity (Rancour et al., 2004). Loss of At-PUX1 resulted in accelerated growth of the mutant plants, an observation we could not confirm for *pux10* knockout lines, suggesting that while both proteins interact with CDC48A, their physiological functions differ. Nevertheless, even if evidence increases that LD proteins are substrate of the UPS system, many questions remain. As mentioned earlier, it is unclear how much of the necessary machinery is LD-specific as opposed to "borrowed" from the ER or the cytosol. The identification of potential players, like the E3 ligase responsible for the ubiquitination of oleosins, would help shed light on the mechanism.

Alternatively, the degradation of LDs via autophagy has been shown to be pivotal in animal and fungal cells, but the importance of the process in plants remains to be determined (recently reviewed in Elander et al., 2018). Studies in algae suggest auto-

phagy is involved in LD degradation (Zhao et al., 2014; Schwarz et al., 2017). Additionally, in Arabidopsis, the uptake of LDs into the vacuole for degradation, potentially achieved by autophagy and mediated via caleosins, has been previously proposed (Poxleitner et al., 2006). However, the focus of many studies on this topic is aimed more so at the degradation of the lipid components of LDs with the purpose of energy conversion, rather than the degradation of LD protein components. In general, the degradation of specific proteins by the UPS is a more flexible mechanism of spatial and temporal protein homeostasis.

In summary, we propose a mechanism for the degradation of LD proteins by the UPS. In our model (Figure 10), LD proteins are marked as substrate for the UPS by the addition of K48-linked polyubiquitin chains. This motif is recognized by the LD-bound PUX10 protein via its UBA domain. With its UBX domain, PUX10 interacts with the ubiquitous cytosolic AAA-type ATPase CDC48, which then channels the bound substrate to the proteasome. Whether CDC48 is responsible for removing LD proteins from the LD and/or whether it interacts directly with the proteasome or via additional factors remains to be elucidated. Furthermore, it remains to be determined whether PUX10 is involved in the degradation of ER proteins. This might be possible as PUX10 can also localize to the ER (Figure 2B), and as it has been speculated elsewhere that LDs stay connected to the ER throughout their life cycle (Goodman, 2009).

## METHODS

### Plant Growth

Tobacco (*Nicotiana tabacum*) plants were grown in the greenhouse as described (Rotsch et al., 2017). Pollen was harvested from freshly opened anther buds (six flowers of two plants per construct) when used for pollen transformation. For large-scale cultivation of pollen tubes for LD isolation, green anther buds close to dehiscence were collected and dried for 2 to 4 d.

*Arabidopsis thaliana* plants (ecotypes Col-0 and Ws-4) were grown on soil in a climate chamber (York) in 60% relative humidity, a constant temperature of 23°C and under 16-h-light/8-h-dark cycle with daytime lighting at 150  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  (the climate chamber was equipped with LuxLine Plus F36W 830 Warm White de Luxe fluorescent tubes; Osram Silvanial).

For all experiments conducted with the *qrt pux10-1* and the *pux10-2* mutant and their respective control lines, seeds were derived from one seed batch each originating from 10 mother plants grown side-by-side. In the case of *pux10-3* and its control line, seed batches of different individual mother plants were used for each biological replicate.

For seed germination on half-strength Murashige and Skoog (MS) medium without any supplemented sucrose, seeds were first surface sterilized in sodium-chlorate solution (6% [w/v]  $\text{NaClO}_3$  and 1% [v/v] Tween 20) for 10 min, washed five times with sterile water, and resuspended in 0.1% (w/v) agar. Seeds were imbibed for a given time at 4°C in the dark and grown under 22°C 16-h-light/8-h-dark cycle with daytime lighting at 150  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ . In all cases, the seeds of mutant lines and their respective controls were derived from plants grown side-by-side under identical conditions.

### Isolation of LD-Enriched Fractions of Tobacco Pollen Tubes

For each biological replicate, pollen was collected from ~40 plants for a week and the experiment repeated in subsequent weeks for each

biological replicate using the same plants. Two grams of pollen per replicate was resuspended in 40 mL of pollen tube medium (Read et al., 1993) and spread on a layer of 6- $\mu$ m Nylon mesh in four 20-cm Petri dishes. Pollen tubes were cultivated for 2 h at 22°C. All of the following steps were performed on ice or at 4°C. The tubes were collected by centrifugation at 1000g for 1 min and the pellet manually ground with sand in 10 mL grinding buffer (50 mM Tris-HCl, pH 7.5, 10 mM KCl, and 200  $\mu$ M proteinase inhibitor PMSF). An aliquot of the total extract was taken after a centrifugation at 1000g for 1 min and the rest of the extract was centrifuged at 100,000g and for 30 min in a swing-out rotor. The floating pad including the LDs was taken off with a spatula and resuspended in 1 mL grinding buffer. From the phase between the fat pad and the pellet, an aliquot was taken as a cytosolic fraction. The LD fraction was washed three times with grinding buffer using centrifugations of 20,000g for 5 min each in a tabletop centrifuge. During the isolation procedure, harsh washing steps with detergents or high salt concentrations were omitted, to avoid the detachment of loosely bound proteins.

### Isolation of Total Protein of Arabidopsis Seedlings

For total extract samples, 10 to 40 mg (~500–2000) dry seeds per biological replicate were sterilized and imbibed for 3 d at 4°C in the dark. Then, the seedlings were grown on half strength MS media for a 24-, 48-, and 72-h growth period. For 0 d after imbibition samples, seeds were soaked in water for 45 min. Seedlings were taken up in 1.6 mL grinding buffer (62.5 mM Tris-HCl, pH 7.5, 5% [w/v] SDS, 10% [v/v] glycerol, 1 mM DTT, and 200  $\mu$ M PMSF) and ground with sand in a mortar for 3 min. Then, the homogenate was incubated for 5 min at 95°C, transferred to a 15-mL reaction tube and 96% ethanol was added to a total volume of 15 mL. All steps were performed on ice or at 4°C.

### Isolation of LD-Enriched Fractions from Arabidopsis Seedlings

Two hundred milligrams (~10,000) of Arabidopsis seeds per replicate were sterilized and stratified for 3 d in the dark at 4°C. Then, seedlings were grown for 38 h (Supplemental Data Sets 40 to 43) or 1 and 2 d (Supplemental Data Sets 44 to 46) in 16-h-light/8-h-dark cycles at 22°C on half-strength MS media plates without sucrose. Seedlings were taken up in 3 mL grinding buffer (10 mM sodium phosphate buffer, pH 7.4, 0.5 mM Lohman's reagent, and 10 mM *N*-ethylmaleimide), ground with sand for 3 min with a mortar and pestle. The resulting homogenate was spun for 10 s at 100g and a 200- $\mu$ L aliquot of the supernatant ("total extract" sample) was taken. The remaining homogenate was centrifuged at 20,000g for 20 min at 4°C and the resulting fat pad was washed twice with grinding buffer, using the same centrifugation step. All steps were performed on ice or at 4°C.

### Proteomics Sample Preparation and LC/MS Analysis of Peptides

All protein fractions were precipitated and defatted with a final concentration of 80% ethanol and washed with 96% ethanol. The protein pellets were redissolved in 6 M urea and 5% (w/v) SDS. Protein concentrations were determined with a Pierce BCA protein assay kit (Thermo Fisher Scientific). Ten to twenty micrograms of protein was run on a SDS-PAGE gel until they entered the separation gel. A single gel piece per sample containing all proteins was excised, subjected to tryptic digestion, and derivatized as described (Shevchenko et al., 2006). Peptides were desalted over an Empore Octadecyl C18 47-mm extraction disks 2215 (Supelco) according to Rappsilber et al. (2007), dried, and dissolved in a 20  $\mu$ L 0.1% formic acid. Then, the peptides were subjected to LC-MS/MS (Schmitt et al., 2017). In detail, nano-flow liquid chromatography was done with the RSLCnano Ultimate 3000 (Thermo Fisher Scientific) system. Peptides of 1 to 3  $\mu$ L sample solution were loaded with 0.07% trifluoroacetic acid

on an Acclaim PepMap 100 precolumn (100  $\mu$ m  $\times$  2 cm, C18, 3  $\mu$ m, 100 Å; Thermo Fisher Scientific) at a flow rate of 20  $\mu$ L/min for 3 to 6 min. Analytical peptide separation by reverse phase chromatography was performed at a flow rate of 300 nL/min on an Acclaim PepMap RSLC column (75  $\mu$ m  $\times$  50 cm, C18, 3  $\mu$ m, 100 Å; Thermo Fisher Scientific). A gradient from 98% solvent A (0.1% formic acid) and 2% solvent B (80% acetonitrile, 0.1% formic acid) to 32% B was applied within 94 min and to 65% B within the next 26 min. Optima LC-MS solvents and acids were purchased from Thermo Fisher Scientific.

Nano-electrospray ionization mass spectrometry was done with the Orbitrap Velos Pro hybrid mass spectrometer (Thermo Fisher Scientific). Chromatographically eluting peptides were online ionized by nano-electrospray ionization using the Nanospray Flex Ion Source (Thermo Fisher Scientific) at 1.5 kV (liquid junction) and continuously transferred into the mass spectrometer. Full scans within the mass range of 300 to 1850 *m/z* were taken with the Orbitrap-FT analyzer at a resolution of 30,000 followed by data-dependent top 15 CID fragmentation (dynamic exclusion enabled) within the ion trap Velos Pro analyzer. LC-MS method programming and data acquisition was performed with XCalibur 2.2 software (Thermo Fisher Scientific).

### Computation Analysis of MS/MS2 Raw Data

MS/MS2 raw data processing for peptide analysis, protein identification, and quantification were performed with MaxQuant 1.5.6.0 or MaxQuant 1.6.1.0 (the ladder was used to analyze proteomic data related to the *pux10-3* mutant line; Cox and Mann, 2008).

Protein abundance was quantified using an LFQ label-free quantification algorithm implemented in MaxQuant software (Schaab et al., 2012) using the settings as described in the meta data files (Supplemental Data Sets 47 to 52). Default settings were used. Additionally, in group-specific parameter setting, label-free quantification was enabled, and the intensity determination was set to total sum. Within the global parameters, the PSM false discovery rate and protein false discovery rate were set to 0.02. Match between runs, dependent peptides, and iBAQ were enabled. After processing, the values were calculated as per mille of all values in one sample.

The iBAQ algorithm rather than the LFQ algorithm was used to calculate enrichment of proteins in subcellular fractions, as it is better suited to detect small abundance levels, and as the LFQ algorithm gave unrealistically high enrichment levels.

For tobacco, the *N. tabacum* Uniprot database was used (as of August 16, 2017) and for Arabidopsis the TAIR10 database.

### Molecular Cloning

Molecular cloning into the vectors pUC-LAT52-mVenusC, pUC-LAT52-mVenusN, and pUC-LAT52-mCherryC (Mähs et al., 2013; Steinhorst et al., 2015) as well as pLatMVC-GW, pLatMVN-GW (Müller et al., 2017), and pGWB604 (Nakamura et al., 2010) was performed using classical restriction sites or Gateway cloning as described (Müller et al., 2017). For a complete list of primers and strategies used, see Supplemental Data Set 17. All constructs used in this study were subject to full sequencing of the gene and its attachment sites to the promoter and the fluorophore. Sanger sequencing was performed by GATC Biotech (now Eurofins). The ERD-CFP construct for pollen tube expression has been previously described (Müller et al., 2017).

### Particle Bombardment and Microscopy of Tobacco Pollen Tubes

The transformation of pollen grains by particle bombardment, the cultivation of the pollen tubes on microscope slides, Nile Red and Bodipy 505/515 staining as well as confocal microscopy were performed as described



(Müller et al., 2017). Monodansylpentane (Abgent) was excited at 405 nm using a HFT 405/514/633 as major beam splitter and detected at a wavelength of 422 to 486 nm.

### Identification of T-DNA Insertion Lines and Generation of Transgenic Lines

The Arabidopsis mutant line described here as *qrt pux10-1* was obtained as SAIL\_1187B06 from the SAIL collection (Sessions et al., 2002) and is in the *qrt PUX10* background (Preuss et al., 1994). This line and the control line used (*qrt PUX10*) derive from the same line heterozygous for the *pux10-1* mutant allele. A second mutant line, *pux10-2*, was obtained as SALK\_139056 from the SALK collection (Alonso et al., 2003). A third independent mutant line, *pux10-3*, was obtained as FST EATTV209T3 from INRA in the Ws-4 background (Samson et al., 2002).

Homozygous plants were identified by performing PCR on genomic DNA using REDTaq ReadyMix (Sigma-Aldrich) according to the manufacturer's instructions.

Arabidopsis plants were transformed by the floral dip method (Clough and Bent, 1998). Arabidopsis Col-0 plants were transformed with *Agrobacterium tumefaciens* strain EHA105 containing pGWB604-At-*PUX10<sub>pro</sub>*::gAt-*PUX10-eGFP* construct. For complementation, those lines were crossed with *qrt pux10-1* plants and the F2-progeny was selected by PCR for lines homozygous for *pux10-1*.

### PUX10 Expression Analysis via RT-qPCR

Ten milligrams (~500) of dry seeds per biological replicate and line was frozen in liquid nitrogen and homogenized with the help of a retch mill. RNA was isolated with the Spectrum Plant Total RNA Kit (Sigma-Aldrich). The RNA was treated with DNaseI (Thermo Fisher Scientific) according to the manufacturer's instructions. For cDNA synthesis, 1 µg of DNaseI-treated RNA was reverse transcribed (Revert Aid Minus Reverse Transcriptase; Thermo Fisher Scientific). The reaction product was diluted 1:10 in double-distilled water before the RT-qPCR, and 2.5 µL of diluted cDNA per sample and Takyon No Rox SYBR MasterMix dTTP Blue (Eurogentec) were used for the RT-qPCR reaction. The following PCR program was applied in an iQ5 qPCR cyclor (Bio-Rad Laboratories): 95°C for 1 min 20 s (95°C for 20 s, 58°C for 20 s, 72°C for 40 s) × 39, 72°C 4 min (for primers, see Supplemental Data Set 17). At4g12590 was used as a reference gene (Dekkers et al., 2012).

### Microscopy of Arabidopsis Embryos, Seeds, Seedlings, and Pollen Tubes

Embryos were obtained from plants grown on soil and seeds were germinated on half-strength MS medium. Arabidopsis embryos and seedlings were gently squeezed out of the seed coat by pressure between a microscope slide and a cover slip. They were then stained with either Nile Red or Bodipy 505/515 (Thermo Fisher Scientific) for at least 5 min in a reaction cup before they were retransferred to a microscope slide.

Confocal microscopy was performed with a Zeiss LSM 510 confocal microscope (Carl Zeiss). eGFP was excited at 488 nm using a 488-nm major beam splitter and detected at a wavelength of 496 to 530 nm; Bodipy 505/515 was excited at 514 nm using a HFT 405/514/633 as major beam splitter and detected at a wavelength of 530 to 600 nm. Diameters of the LDs contained in those cells were quantified with ImageJ (Schneider et al., 2012). Epifluorescence was observed on a epifluorescence microscope (BX51; Olympus) supplied with an U-MF2 filter cube for mVenus fluorescence and recorded with an ORCA-Flash 4.0 V2 camera (Hamamatsu Photonics).

Pollen of Arabidopsis Col-0 At-*PUX10<sub>pro</sub>*::At-*PUX10-eGFP* plants were grown in vitro by gently brushing open 10 flowers over a microscope

slide and covering them with Arabidopsis pollen tube medium (0.01% [w/v] boric acid, 5 mM CaCl<sub>2</sub>, 5 mM KCl, 1 mM MgSO<sub>4</sub>, and 10% [w/v] sucrose). The microscope slides were incubated for 3 h in the dark. Pollen tubes were stained with 0.5% (w/v) Nile Red and fixed with 5% (v/v) formaldehyde both dissolved in Arabidopsis pollen tube medium. Pollen tubes were imaged by confocal microscopy as described above.

### Quantification of Fatty Acids

Arabidopsis seeds were sterilized and stratified at 4°C in the dark for 24 h and subsequently germinated on half-strength MS medium in a long-day chamber (16 h light/8 h dark). Seeds and seedlings (10 per biological replicate) were collected for analysis every 24 h and used directly for the generation of fatty acid methyl esters. C15:0 TAG (Sigma-Aldrich) was added as an internal standard. Fatty acid methyl esters were quantified by gas chromatography analysis with flame ionization detection as described (Hornung et al., 2002).

### Immunoblotting

Twenty micrograms of protein, derived from proteins extracted from 200 mg seedlings per biological replicate, was dissolved in 6 M urea and 5% (w/v) SDS. Then, the samples were run on a 12% or 13% SDS gel (LD and total samples, respectively). Proteins were transferred from the gel on a Nitrocellulose membrane (Roti-NC, 0.2 µm; Carl Roth) in 1× transfer buffer (25 mM Tris, 192 mM glycine, and 20% ethanol). Before blocking, the membrane was stained with Ponceau S (5% [v/v] acetic acid and 0.1% [w/v] Ponceau S; Serva) for 30 min, destained with double-distilled water, and documented. Then, the membrane was blocked in 1× TBS-T buffer (50 mM Tris, 0.15 M NaCl, and 0.05% [v/v] Tween 20, pH to 7.6 with HCl) supplemented with 3% (w/v) milk powder (Carl Roth) for 1 h at room temperature under agitation. The membrane was incubated with the first antibody, a rabbit antibody recognizing an epitope of Arabidopsis UBIQUITIN 11 (At4g05050, AS08 307 A by Agrisera), overnight at 4°C under agitation. The antibody was diluted 1:5000 in 1× TBS-T supplemented with 3% milk powder. After that incubation, the membrane was washed four times for 10 min per wash with 1× TBS-T supplemented with 3% milk powder and incubated with the secondary antibody for 1 h 30 min under agitation at room temperature. The goat anti rabbit IgG alkaline phosphates antibody (catalog no. A3678, lot no. 036K6031; Sigma-Aldrich) was diluted 1:30,000 in TBS-T supplemented with 3% milk powder. Then, the membrane was washed twice 10 min with 1× TBS-T and twice for 5 min with 1× TBS (50 mM Tris, and 0.15 M NaCl, pH to 7.6 with HCl). For revelation, the membrane was first equilibrated for 10 min with AP buffer (100 mM Tris, 100 mM NaCl, and 5 mM MgCl<sub>2</sub>, pH 9.5) and then incubated in 10 mL AP buffer supplemented with 33 µL BCIP (50 mg/mL in DMF) and 66 µL NBT (50 mg/mL in DMF) in the dark until bands were visible. The reaction was stopped with double-distilled water and the membrane documented.

### Bioinformatics

Phylogenetic trees were created with MEGA version X software using MUSCLE alignment with gap penalties set to -9 for gap open and to -3 for gap extension (Edgar, 2004; Kumar et al., 2018). The aligned protein sequences were used for phylogenetic tree construction using the maximum likelihood method based on the JTT matrix-based model (Jones et al., 1992). The phylogeny was tested with the Bootstrap method set for 1000 replicates (Felsenstein, 1985). No phylogeny testing was performed for the LDIP phylogenetic tree.

Hydropathy analyses of proteins were performed on the TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>) (Krogh et al., 2001). All statistical analysis was performed in and all box plots created

with RStudio 0.98.1060. Tobacco proteins were blasted against the Arabidopsis Uniprot database (8/3/2018) using the BLAST 2.8.0 algorithm (Altschul et al., 1997). Helical wheel projection of amino acids was performed with HeliQuest (Gautier et al., 2008).

#### Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under the following accession numbers: At-PUX10, AT4G10790; At-OLE1, AT4G25140; At-OLE2, AT5G40420; At-OLE3, AT5G51210; At-OLE4, AT3G27660; At-OLE5, AT3G01570; At-OLE6, AT1G48990; At-OLE7, AT2G25890; At-OLE8, AT3G18570; At-HSD1, AT5G50600/AT5G50700; At-HSD5, AT4G10020, At-CLO1, AT4G26740; At-CLO2, AT5G55240; At-LDAP1, AT1G67360; At-LDAP2, AT2G47780; At-LDAP3, AT3G05500; At-UBQ5, AT3G62250; At-PUX5, AT4G15410; At-PUX8, AT4G11740; At-PUX9, AT4G00752; At-PUX11, AT2G43210; At-PUX13, AT4G23040; At-CAS1, AT2G07050; At-CDC48A, AT3G09840; Nt-CLO1a, LOC107783728; Nt-OLE6b, LOC107824536; Nt-CLO1b, LOC107817909; Nt-OLE6a, LOC107780677; Nt-LDAP1a, LOC107765167; Nt-CASb, LOC107762593; Nt-PUX10c, LOC107769345; Nt-LDAP1b, LOC107804268; Nt-LDAP3a, LOC107827582; Nt-CASa, LOC107826198; Nt-SMT1c, LOC107813194; Nt-PTLD1b, LOC107815104; Nt-OBL1, LOC107788962; Nt-PUX10a, LOC107799924; Nt-LDIPa, LOC107763626; Nt-PTLD2a, LOC107796437; Nt-FEY1, LOC107800303; Nt-LDIPb, LOC107831283; Nt-ADCP, LOC107827250; Nt-PDCP, LOC107793109; Nt-RAB18, LOC107762524; Nt-PTLD2b, LOC107778489; Nt-CCD, LOC107830214; and Nt-RAB2c, LOC107794510. All proteomic data, including raw data files, MaxQuant search files, as well as protein group and peptide search results created by MaxQuant are available on ProteomeXchange/PRIDE (Vizcaino et al., 2014) under the identifiers PXD009184, PXD009186, PXD009207, PXD009397, PXD009248, and PXD009247.

#### Supplemental Data

**Supplemental Figure 1.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* oleosin proteins.

**Supplemental Figure 2.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* caleosin proteins.

**Supplemental Figure 3.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* steroleosin proteins.

**Supplemental Figure 4.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* plant UBX domain-containing proteins 10 (PUX10).

**Supplemental Figure 5.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* oil body lipases.

**Supplemental Figure 6.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* lipid droplet-associated proteins.

**Supplemental Figure 7.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* lipid droplet-associated protein-interacting proteins.

**Supplemental Figure 8.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* cycloartenol synthase proteins.

**Supplemental Figure 9.** Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* sterol methyltransferase proteins.

**Supplemental Figure 10.** Phylogenetic tree of “pollen tube lipid droplet protein” from *Nicotiana tabacum*.

**Supplemental Figure 11.** The cycloartenol synthase At-CAS1, Arabidopsis homolog to the LD-localized Nt-CAS1b, localizes to LDs.

**Supplemental Figure 12.** *Arabidopsis thaliana* plant UBX domain-containing (PUX) protein gene family.

**Supplemental Figure 13.** Hydropathy profiles of the Arabidopsis PUX proteins and NtPUX10c.

**Supplemental Figure 14.** Localization of Nt-PUX10c and truncated variants (magnified images).

**Supplemental Figure 15.** Uneven distribution of At-PUX10-eGFP on LD in germinating seeds remains after isolation.

**Supplemental Figure 16.** Relative expression of *PUX10* in mutants is strongly reduced compared with the mutant backgrounds.

**Supplemental Figure 17.** Relative LFQ intensity of PUX10 protein on LDs of the *qrt pux10-1* mutant and controls.

**Supplemental Figure 18.** LD proteins accumulate in the *pux10* mutant seedlings.

**Supplemental Figure 19.** Postgerminative fatty acid degradation in *qrt pux10-1* and the nonmutated background *qrt PUX10*.

**Supplemental Figure 20.** Antiubiquitin immunoblots.

**Supplemental Figure 21.** Annotated fragmentation spectrum of identified ubiquitination site of ubiquitin presented in Figure 9B.

**Supplemental Figure 22.** Annotated fragmentation spectrum of identified ubiquitination site of At-OLE4 presented in Figure 9B.

**Supplemental Figure 23.** Annotated fragmentation spectra of identified ubiquitination site of At-OLE4 presented in Figure 9B.

**Supplemental Figure 24.** Annotated fragmentation spectrum of identified ubiquitination site of At-HSD1 presented in Figure 9B.

**Supplemental Figure 25.** Hydropathy profiles of At-CAS1 and Nt-LDAP1a.

**Supplemental Figure 26.** Helical wheel projection of amino acids 93 to 111 of At-PUX10.

**Supplemental Figure 27.** Hydropathy profiles of the *N. tabacum* PTLD proteins.

**Supplemental Data Set 1.** Proteins found in tobacco pollen tubes (raw iBAQ and LFQ values).

**Supplemental Data Set 2.** Proteins found in tobacco pollen tubes (normalized and sorted data).

**Supplemental Data Set 3.** Lipid droplet proteins found in tobacco pollen tubes.

**Supplemental Data Set 4.** Organellar marker proteins in lipid droplet fractions of tobacco pollen tubes.

**Supplemental Data Set 5.** Proteins found in tobacco seeds (raw iBAQ and LFQ values).

**Supplemental Data Set 6.** Proteins found in tobacco seeds (normalized and sorted data).

**Supplemental Data Set 7.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* oleosin proteins.

**Supplemental Data Set 8.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* caleosin proteins.

**Supplemental Data Set 9.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* steroleosin proteins.

**Supplemental Data Set 10.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* plant UBX domain-containing proteins 10 (PUX10).

**Supplemental Data Set 11.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* oil body lipases.

**Supplemental Data Set 12.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* lipid droplet-associated proteins.

**Supplemental Data Set 13.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* lipid droplet-associated protein-interacting proteins.

**Supplemental Data Set 14.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* cycloartenol synthase proteins.

**Supplemental Data Set 15.** Sequence alignment of *Arabidopsis thaliana* and *Nicotiana tabacum* sterol methyltransferase proteins.

**Supplemental Data Set 16.** Sequence alignment of “pollen tube lipid droplet protein” from *Nicotiana tabacum*.

**Supplemental Data Set 17.** Cloning strategies used and list of primers.

**Supplemental Data Set 18.** Protein composition of rehydrated seeds of *pux10-1* and controls (raw iBAQ and LFQ values of total extracts).

**Supplemental Data Set 19.** Protein composition of seedlings 1 d after imbibition of *pux10-1* and controls (raw iBAQ and LFQ values of total extracts).

**Supplemental Data Set 20.** Protein composition of seedlings 2 d after imbibition of *pux10-1* and controls (raw iBAQ and LFQ values of total extracts).

**Supplemental Data Set 21.** Protein composition of seedlings 3 d after imbibition of *pux10-1* and controls (raw iBAQ and LFQ values of total extracts).

**Supplemental Data Set 22.** Protein composition of rehydrated seeds of *pux10-1* and controls (normalized iBAQ and LFQ values of total extracts).

**Supplemental Data Set 23.** Protein composition of seedlings 1 d after imbibition of *pux10-1* and controls (normalized iBAQ and LFQ values of total extracts).

**Supplemental Data Set 24.** Protein composition of seedlings 2 d after imbibition of *pux10-1* and controls (normalized iBAQ and LFQ values of total extracts).

**Supplemental Data Set 25.** Protein composition of seedlings 3 d after imbibition of *pux10-1* and controls (normalized iBAQ and LFQ values of total extracts).

**Supplemental Data Set 26.** Protein composition of rehydrated seeds and seedlings 1 to 3 d after imbibition of *pux10-1* and controls (averages of LFQ values of total extracts).

**Supplemental Data Set 27.** Protein composition of rehydrated seeds and seedlings 1 to 3 d after imbibition of *pux10-1* and controls (averages of iBAQ values of total extracts).

**Supplemental Data Set 28.** Abundance of lipid droplet proteins of rehydrated seeds and seedlings 1 to 3 d after imbibition of *pux10-1* and controls (averages of LFQ values of total extracts).

**Supplemental Data Set 29.** Protein composition of rehydrated seeds of Ws-4 and *pux10-3* - raw iBAQ and LFQ values of total extracts.

**Supplemental Data Set 30.** Protein composition of seedlings 1 d after imbibition of Ws-4 and *pux10-3* (raw iBAQ and LFQ values of total extracts).

**Supplemental Data Set 31.** Protein composition of seedlings 2 d after imbibition of Ws-4 and *pux10-3* (raw iBAQ and LFQ values of total extracts).

**Supplemental Data Set 32.** Protein composition of seedlings 3 d after imbibition of Ws-4 and *pux10-3* (raw iBAQ and LFQ values of total extracts).

**Supplemental Data Set 33.** Protein composition of rehydrated seeds of Ws-4 and *pux10-3* (normalized iBAQ and LFQ values of total extracts).

**Supplemental Data Set 34.** Protein composition of seedlings 1 d after imbibition of Ws-4 and *pux10-3* (normalized iBAQ and LFQ values of total extracts).

**Supplemental Data Set 35.** Protein composition of seedlings 2 d after imbibition of Ws-4 and *pux10-3* (normalized iBAQ and LFQ values of total extracts).

**Supplemental Data Set 36.** Protein composition of seedlings 3 d after imbibition of Ws-4 and *pux10-3* (normalized iBAQ and LFQ values of total extracts).

**Supplemental Data Set 37.** Protein composition of rehydrated seeds and seedlings 1 to 3 d after imbibition of Ws-4 and *pux10-3* (averages of LFQ values of total extracts).

**Supplemental Data Set 38.** Protein composition of rehydrated seeds and seedlings 1 to 3 d after imbibition of Ws-4 and *pux10-3* (averages of iBAQ values of total extracts).

**Supplemental Data Set 39.** Abundance of lipid droplet proteins of rehydrated seeds and seedlings 1 to 3 d after imbibition of Ws-4 and *pux10-3* (averages of LFQ intensity values of total extracts).

**Supplemental Data Set 40.** Proteins found in seedlings 38 h after imbibition of *pux10-1* and controls (raw iBAQ and LFQ values).

**Supplemental Data Set 42.** Proteins found in seedlings 38 h after imbibition of *pux10-1* and controls (normalized iBAQ and LFQ values).

**Supplemental Data Set 42.** Marker proteins found in seedlings 38 h after imbibition of *pux10-1* and controls (normalized iBAQ and LFQ values).

**Supplemental Data Set 43.** Ubiquitinated peptides found in seedlings 38 h after imbibition of *pux10-1* and controls.

**Supplemental Data Set 44.** Proteins found in seedlings 1 and 2 d after imbibition of *pux10-1* and controls (raw iBAQ and LFQ values).

**Supplemental Data Set 45.** Proteins found in seedlings 1 and 2 d after imbibition of *pux10-1* and controls (normalized iBAQ and LFQ values).

**Supplemental Data Set 46.** Ubiquitinated peptides found in seedlings 1 and 2 d after imbibition of *pux10-1* and controls.

**Supplemental Data Set 47.** Metadata: *Nicotiana tabacum* pollen tube lipid droplet proteome.

**Supplemental Data Set 48.** Metadata: *Nicotiana tabacum* seed lipid droplet proteome.

**Supplemental Data Set 49.** Metadata: Seedling proteome of *Arabidopsis thaliana* *qrt pux10-1* mutant versus the wild type and complemented lines.

**Supplemental Data Set 50.** Metadata: Seedling proteome of *Arabidopsis thaliana* *pux10-3* mutant versus the wild type.

**Supplemental Data Set 51.** Metadata: Proteome of *Arabidopsis thaliana* *qrt pux10-1* mutant versus wild type and complemented line lipid droplet and total fractions 38 h after imbibition.

**Supplemental Data Set 52.** Metadata: Proteome of *Arabidopsis thaliana* *qrt pux10-1* mutant lipid droplets versus the wild type 24 and 48 h after imbibition.

## ACKNOWLEDGMENTS

We thank Ivo Feussner for support and many helpful discussions and Christiane Gatz and Alexander Stein for valuable advice. We thank Kent Chapman, John Dyer, and Robert Mullen for discussing our work with us and for critically reading the manuscript. We also thank Jörg Großhans and Steven A. Johnson for granting access to their confocal microscopes as well as Florian Wegwitz and Johannes Sattmann for their assistance.

We thank Tsuyoshi Nakagawa (Shimane University) for providing the vector pGWB604 that contains the *bar* gene, which was identified by Meiji Seika Kaisha; and Leonie Steinhart and Jörg Kudla (University of Münster) for generating and providing plasmids. We thank Kaneschka Yaqubi, Anna-Lena Gippert, Jörg Heinz, Maurice Hädrich, Lara Baumbühl, and Milena Lewandowska for help in the lab and Benedikt Ni for proofreading. This work was supported by funding from the Deutsche Forschungsgemeinschaft (IS 273/2-2 to T.I. and GR1945/3-1, SFB937/TP10, and INST1525/16-1 FUGG to Jörg Großhans) and the Studienstiftung des Deutschen Volkes (to F.K.K.).

## AUTHOR CONTRIBUTIONS

F.K.K., G.H.B., and T.I. designed the work. F.K.K., L.A.M., A.O.M., K.S., K.F.B., O.V., and T.I. performed research. F.K.K., L.A.M., K.S., O.V., G.H.B., and T.I. analyzed data. F.K.K. and T.I. wrote the manuscript. All authors critically read and revised the manuscript and approved the final version.

Received April 5, 2018; revised June 12, 2018; accepted July 31, 2018; published August 7, 2018.

## REFERENCES

- Alonso, J.M., et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301**: 653–657.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389–3402.
- 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. *J. Amino Acids* **2013**: 183421.
- Baud, S., et al. (2009). Regulation of HSD1 in seeds of *Arabidopsis thaliana*. *Plant Cell Physiol.* **50**: 1463–1478.
- Baud, S., Boutin, J.-P., Miquel, M., Lepiniec, L., and Rochat, C. (2002). An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. *Plant Physiol. Biochem.* **40**: 151–160.
- Blée, E., Boachon, B., Burcklen, M., Le Guédard, M., Hanano, A., Heintz, D., Ehling, 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 Physiol.* **166**: 109–124.
- Bouvier-Navé, P., Berna, A., Noirié, A., Compagnon, V., Carlsson, A.S., Banas, A., Stymne, S., and Schaller, H. (2010). Involvement of the phospholipid sterol acyltransferase1 in plant sterol homeostasis and leaf senescence. *Plant Physiol.* **152**: 107–119.
- Brocard, L., Immel, F., Coulon, D., Esnay, N., Tophile, 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. *Front. Plant Sci.* **8**: 894.
- 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. *Plant Cell* **27**: 2616–2636.
- Cai, Y., McClinchie, E., Price, A., Nguyen, T.N., Gidda, S.K., Watt, S.C., Yurchenko, O., Park, S., Sturtevant, D., and Mullen, R.T. (2017). Mouse fat storage-inducing transmembrane protein 2 (FIT2) promotes lipid droplet accumulation in plants. *Plant Biotechnol. J.* **15**: 824–836.
- Chapman, K.D., Dyer, J.M., and Mullen, R.T. (2012). Biogenesis and functions of lipid droplets in plants. *J. Lipid Res.* **53**: 215–226.
- Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**: 735–743.
- Copeland, C., Woloshen, V., Huang, Y., and Li, X. (2016). AtCDC48A is involved in the turnover of an NLR immune receptor. *Plant J.* **88**: 294–305.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**: 1367–1372.
- d'Andréa, S., Canonge, M., Beopoulos, A., Jolivet, P., Hartmann, M.A., 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**: 222–229.
- Dekkers, B.J., Willems, L., Bassel, G.W., van Bolderen-Veldkamp, R.P., Ligterink, W., Hilhorst, H.W., and Bentsink, L. (2012). Identification of reference genes for RT-qPCR expression analysis in *Arabidopsis* and tomato seeds. *Plant Cell Physiol.* **53**: 28–37.
- 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 Cell Physiol.* **56**: 1374–1387.
- Eastmond, P.J. (2004). Cloning and characterization of the acid lipase from castor beans. *J. Biol. Chem.* **279**: 45540–45545.
- Eastmond, P.J. (2006). SUGAR-DEPENDENT1 encodes a patatin domain triacylglycerol lipase that initiates storage oil breakdown in germinating *Arabidopsis* seeds. *Plant Cell* **18**: 665–675.
- Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 113.
- Eisenberg-Bord, M., et al. (2018). Identification of seipin-linked factors that act as determinants of a lipid droplet subpopulation. *J. Cell Biol.* **217**: 269–282.
- Elander, P.H., Minina, E.A., and Bozhkov, P.V. (2018). Autophagy in turnover of lipid stores: trans-kingdom comparison. *J. Exp. Bot.* **69**: 1301–1311.
- Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.
- Gallois, J.L., Drouaud, J., Lécureuil, A., Guyon-Debast, A., Bonhomme, S., and Guerche, P. (2013). Functional characterization of the plant ubiquitin regulatory X (UBX) domain-containing protein AtPUX7 in *Arabidopsis thaliana*. *Gene* **526**: 299–308.
- Gautier, R., Douguet, D., Antonny, B., and Drin, G. (2008). HELIQUEST: a web server to screen sequences with specific alpha-helical properties. *Bioinformatics* **24**: 2101–2102.
- Gidda, S.K., et al. (2013). Lipid droplet-associated proteins (LDAPs) are involved in the compartmentalization of lipophilic compounds in plant cells. *Plant Signal. Behav.* **8**: e27141.
- 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 Physiol.* **170**: 2052–2071.
- Goodman, J.M. (2009). Demonstrated and inferred metabolism associated with cytosolic lipid droplets. *J. Lipid Res.* **50**: 2148–2156.
- Guerriero, C.J., and Brodsky, J.L. (2012). The delicate balance between secreted protein folding and endoplasmic reticulum-associated degradation in human physiology. *Physiol. Rev.* **92**: 537–576.

- Hanano, A., Burcklen, M., Flenet, M., Ivancich, A., Louwagie, M., Garin, J., and Blée, E. (2006). Plant seed peroxxygenase is an original heme-oxygenase with an EF-hand calcium binding motif. *J. Biol. Chem.* **281**: 33140–33151.
- 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 Physiol.* **162**: 1926–1936.
- Hornung, E., Pernstich, C., and Feussner, I. (2002). Formation of conjugated Delta11Delta13-double bonds by Delta12-linoleic acid (1,4)-acyl-lipid-desaturase in pomegranate seeds. *Eur. J. Biochem.* **269**: 4852–4859.
- Hsiao, E.S.L., and Tzen, J.T.C. (2011). Ubiquitination of oleosin-H and caleosin in sesame oil bodies after seed germination. *Plant Physiol. Biochem.* **49**: 77–81.
- Hsieh, K., Lee, Y.K., Londos, C., Raaka, B.M., Dalen, K.T., and Kimmel, A.R. (2012). Perilipin family members preferentially sequester to either triacylglycerol-specific or cholesteryl-ester-specific intracellular lipid storage droplets. *J. Cell Sci.* **125**: 4067–4076.
- Huang, A.H.C. (1996). Oleosins and oil bodies in seeds and other organs. *Plant Physiol.* **110**: 1055–1061.
- Huang, A.H.C. (2018). Plant lipid droplets and their associated oleosin and other proteins: potential for rapid advances. *Plant Physiol.* **176**: 1894–1918.
- Ischebeck, T. (2016). Lipids in pollen—They are different. *Biochim. Biophys. Acta* **1861**: 1315–1328.
- Ishihara, H., Obata, T., Sulpice, R., Fernie, A.R., and Stitt, M. (2015). Quantifying protein synthesis and degradation in Arabidopsis by dynamic <sup>13</sup>CO<sub>2</sub> labeling and analysis of enrichment in individual amino acids in their free pools and in protein. *Plant Physiol.* **168**: 74–93.
- 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 Cell Physiol.* **48**: 812–821.
- 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 Physiol. Biochem.* **42**: 501–509.
- 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. *J. Proteomics* **169**: 87–98.
- Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**: 275–282.
- 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 Physiol.* **157**: 866–875.
- Kim, E.Y., Park, K.Y., Seo, Y.S., and Kim, W.T. (2016). Arabidopsis small rubber particle protein homolog SRPs play dual roles as positive factors for tissue growth and development and in drought stress responses. *Plant Physiol.* **170**: 2494–2510.
- Kim, H.U., Hsieh, K., Ratnayake, C., and Huang, A.H. (2002). A novel group of oleosins is present inside the pollen of *Arabidopsis*. *J. Biol. Chem.* **277**: 22677–22684.
- 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. *J. Mol. Biol.* **305**: 567–580.
- Kumar, S., Stecher, G., Li, M., Niyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **35**: 1547–1549.
- Lee, K., Bih, F.Y., Learn, G.H., Ting, J.T.L., Sellers, C., and Huang, A.H.C. (1994). Oleosins in the gametophytes of Pinus and Brassica and their phylogenetic relationship with those in the sporophytes of various species. *Planta* **193**: 461–469.
- Li, F., Asami, T., Wu, X., Tsang, E.W., and Cutler, A.J. (2007). A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. *Plant Physiol.* **145**: 87–97.
- Liu, Y., and Li, J. (2014). Endoplasmic reticulum-mediated protein quality control in Arabidopsis. *Front. Plant Sci.* **5**: 162.
- Liu, H., Wang, C., Chen, F., and Shen, S. (2015). Proteomic analysis of oil bodies in mature *Jatropha curcas* seeds with different lipid content. *J. Proteomics* **113**: 403–414.
- Liu, L., Cui, F., Li, Q., Yin, B., Zhang, H., Lin, B., Wu, Y., Xia, R., Tang, S., and Xie, Q. (2011). The endoplasmic reticulum-associated degradation is necessary for plant salt tolerance. *Cell Res.* **21**: 957–969.
- Mähs, A., Steinhilber, L., Han, J.P., Shen, L.K., Wang, Y., and Kudla, J. (2013). The calcineurin B-like Ca<sup>2+</sup> sensors CBL1 and CBL9 function in pollen germination and pollen tube growth in Arabidopsis. *Mol. Plant* **6**: 1149–1162.
- Matsui, K., Hijiya, K., Tabuchi, Y., and Kajiwar, T. (1999). Cucumber cotyledon lipoxygenase during postgerminative growth. Its expression and action on lipid bodies. *Plant Physiol.* **119**: 1279–1288.
- Meinken, J., and Min, J. (2012). Computational prediction of protein subcellular locations in eukaryotes: an experience report. *Computational Molecular Biology* **2**: 1–7.
- Mellema, S., Eichenberger, W., Rawlyer, A., Suter, M., Tadege, M., and Kuhlemeier, C. (2002). The ethanolic fermentation pathway supports respiration and lipid biosynthesis in tobacco pollen. *Plant J.* **30**: 329–336.
- Meusser, B., Hirsch, C., Jarosch, E., and Sommer, T. (2005). ERAD: the long road to destruction. *Nat. Cell Biol.* **7**: 766–772.
- Müller, A.O., and Ischebeck, T. (2018). Characterization of the enzymatic activity and physiological function of the lipid droplet-associated triacylglycerol lipase AtOBL1. *New Phytol.* **217**: 1062–1076.
- Müller, A.O., Biersch, 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. *Plant J.* **89**: 1055–1064.
- 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. *Biosci. Biotechnol. Biochem.* **74**: 1315–1319.
- Olzmann, J.A., Richter, C.M., and Kopito, R.R. (2013). Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. *Proc. Natl. Acad. Sci. USA* **110**: 1345–1350.
- 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. *J. Biol. Chem.* **282**: 5217–5224.
- Park, S., Rancour, D.M., and Bednarek, S.Y. (2008). In planta analysis of the cell cycle-dependent localization of AtCDC48A and its critical roles in cell division, expansion, and differentiation. *Plant Physiol.* **148**: 246–258.
- Peters, J.-M., Franke, W.W., and Kleinschmidt, J.A. (1994). Distinct 19 S and 20 S subcomplexes of the 26 S proteasome and their distribution in the nucleus and the cytoplasm. *J. Biol. Chem.* **269**: 7709–7718.
- Poxleitner, M., Rogers, S.W., Lacey Samuels, A., Browse, J., and Rogers, J.C. (2006). A role for caleosin in degradation of oil-body storage lipid during seed germination. *Plant J.* **47**: 917–933.
- Preuss, D., Rhee, S.Y., and Davis, R.W. (1994). Tetrad analysis possible in Arabidopsis with mutation of the QUARTET (QRT) genes. *Science* **264**: 1458–1460.
- Pyc, M., Cai, Y., Greer, M.S., Yurchenko, O., Chapman, K.D., Dyer, J.M., and Mullen, R.T. (2017a). Turning over a new leaf in lipid droplet biology. *Trends Plant Sci.* **22**: 596–609.

- 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. (2017b). Arabidopsis lipid droplet-associated protein (LDAP) - interacting protein (LDIP) influences lipid droplet size and neutral lipid homeostasis in both leaves and seeds. *Plant J.* **92**: 1182–1201.
- Quettier, A.-L., and Eastmond, P.J. (2009). Storage oil hydrolysis during early seedling growth. *Plant Physiol. Biochem.* **47**: 485–490.
- Rancour, D.M., Park, S., Knight, S.D., and Bednarek, S.Y. (2004). Plant UBX domain-containing protein 1, PUX1, regulates the oligomeric structure and activity of arabidopsis CDC48. *J. Biol. Chem.* **279**: 54264–54274.
- Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**: 1896–1906.
- Read, S., Clarke, A., and Bacic, A. (1993). Stimulation of growth of cultured *Nicotiana tabacum* W 38 pollen tubes by poly (ethylene glycol) and Cu (II) salts. *Protoplasma* **177**: 1–14.
- 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. *Plant J.* **92**: 129–146.
- 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. *J. Exp. Bot.* **62**: 749–760.
- Samson, F., Brunaud, V., Balzergue, S., Dubreucq, B., Lepiniec, L., Pelletier, G., Caboche, M., and Lecharny, A. (2002). FLAGdb/FST: a database of mapped flanking insertion sites (FSTs) of *Arabidopsis thaliana* T-DNA transformants. *Nucleic Acids Res.* **30**: 94–97.
- Schaab, C., Geiger, T., Stoeck, G., Cox, J., and Mann, M. (2012). Analysis of high accuracy, quantitative proteomics data in the MaxQB database. *Mol. Cell Proteomics* **11**: M111.014068.
- Schmitt, K., Smolinski, N., Neumann, P., Schmaul, S., Hofer-Pretz, V., Braus, G.H., and Valerius, O. (2017). Asc1p/RACK1 connects ribosomes to eukaryotic phosphosignaling. *Mol. Cell. Biol.* **37**: 1–23.
- Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**: 671–675.
- Schuberth, C., and Buchberger, A. (2008). UBX domain proteins: major regulators of the AAA ATPase Cdc48/p97. *Cell. Mol. Life Sci.* **65**: 2360–2371.
- Schwahnäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. *Nature* **473**: 337–342.
- Schwarz, V., Andosch, A., Geretschlager, A., Affenzeller, M., and Lütz-Meindl, U. (2017). Carbon starvation induces lipid degradation via autophagy in the model alga *Micrasterias*. *J. Plant Physiol.* **208**: 115–127.
- Sessions, A., et al. (2002). A high-throughput Arabidopsis reverse genetics system. *Plant Cell* **14**: 2985–2994.
- Sharma, B., Joshi, D., Yadav, P.K., Gupta, A.K., and Bhatt, T.K. (2016). Role of ubiquitin-mediated degradation system in plant biology. *Front. Plant Sci.* **7**: 806.
- Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**: 2856–2860.
- Shimada, T.L., and Hara-Nishimura, I. (2010). Oil-body-membrane proteins and their physiological functions in plants. *Biol. Pharm. Bull.* **33**: 360–363.
- Shin, J.-B., et al. (2013). Molecular architecture of the chick vestibular hair bundle. *Nat. Neurosci.* **16**: 365–374.
- Siloto, R.M.P., 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 oilbodies in *Arabidopsis*. *Plant Cell* **18**: 1961–1974.
- Sonawane, P.D., et al. (2016). Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism. *Nat. Plants* **3**: 16205.
- Steinhorst, L., Mähs, A., Ischebeck, T., Zhang, C., Zhang, X., Arendt, S., Schüttke, S., Heilmann, I., and Kudla, J. (2015). Vacuolar CBL-CIPK12 Ca(2+)-sensor-kinase complexes are required for polarized pollen tube growth. *Curr. Biol.* **25**: 1475–1482.
- Suzuki, M., Otsuka, T., Ohsaki, Y., Cheng, J., Taniguchi, T., Hashimoto, H., Taniguchi, H., and Fujimoto, T. (2012). Derlin-1 and UBXD8 are engaged in dislocation and degradation of lipidated ApoB-100 at lipid droplets. *Mol. Biol. Cell* **23**: 800–810.
- 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 Physiol.* **176**: 1531–1546.
- Thul, P.J., Tschapalda, K., Kolkhof, P., Thiam, A.R., Oberer, M., and Beller, M. (2017). Targeting of the *Drosophila* protein CG2254/Ldsdh1 to a subset of lipid droplets. *J. Cell Sci.* **130**: 3141–3157.
- Ting, J.T.L., Lee, K., Ratnayake, C., Platt, K.A., Balsamo, R.A., and Huang, A.H.C. (1996). Oleosin genes in maize kernels having diverse oil contents are constitutively expressed independent of oil contents. Size and shape of intracellular oil bodies are determined by the oleosins/oils ratio. *Planta* **199**: 158–165.
- Tzen, J.T.C. (2012). Integral proteins in plant oil bodies. *Int. Sch. Res. Notices Botany* **2012**: 1–16.
- Tzen, J.T.C., Lai, Y.K., Chan, K.L., and Huang, A.H.C. (1990). Oleosin isoforms of high and low molecular weights are present in the oil bodies of diverse seed species. *Plant Physiol.* **94**: 1282–1289.
- 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 Environ.* **40**: 2292–2306.
- van Wijk, K.J., and Kessler, F. (2017). Plastoglobuli: plastid microcompartments with integrated functions in metabolism, plastid developmental transitions, and environmental adaptation. *Annu. Rev. Plant Biol.* **68**: 253–289.
- Villette, C., Berna, A., Compagnon, V., and Schaller, H. (2015). Plant sterol diversity in pollen from angiosperms. *Lipids* **50**: 749–760.
- Vizcaino, J.A., et al. (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **32**: 223–226.
- Wang, C.-W., and Lee, S.-C. (2012). The ubiquitin-like (UBX)-domain-containing protein Ubx2/Ubx8 regulates lipid droplet homeostasis. *J. Cell Sci.* **125**: 2930–2939.
- Wang, P., Mugume, Y., and Bassham, D.C. (2018). New advances in autophagy in plants: Regulation, selectivity and function. *Semin. Cell Dev. Biol.* **80**: 113–122.
- Wang, T.W., Balsamo, R.A., Ratnayake, C., Platt, K.A., Ting, J.T.L., and Huang, A.H.C. (1997). Identification, subcellular localization, and developmental studies of oleosins in the anther of *Brassica napus*. *Plant J.* **11**: 475–487.
- Wilfling, F., et al. (2013). Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocating from the ER to lipid droplets. *Dev. Cell* **24**: 384–399.

- Wolins, N.E., Quaynor, B.K., Skinner, J.R., Schoenfish, M.J., Tzekov, A., and Bickel, P.E.** (2005). S3-12, Adipophilin, and TIP47 package lipid in adipocytes. *J. Biol. Chem.* **280**: 19146–19155.
- 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. *Plant Cell* **21**: 3885–3901.
- 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. *Sci. Rep.* **6**: 29539.
- Zhao, L., Dai, J., and Wu, Q.** (2014). Autophagy-like processes are involved in lipid droplet degradation in *Auxenochlorella protothecoides* during the heterotrophy-autotrophy transition. *Front. Plant Sci.* **5**: 400.
- Zhi, Y., et al.** (2017). Comparative lipidomics and proteomics of lipid droplets in the mesocarp and seed tissues of Chinese tallow (*Triadica sebifera*). *Front. Plant Sci.* **8**: 1339.
- 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. *J. Exp. Bot.* **64**: 293–302.

**PUX10 Is a Lipid Droplet-Localized Scaffold Protein That Interacts with CELL DIVISION  
CYCLE48 and Is Involved in the Degradation of Lipid Droplet Proteins**

Franziska K. Kretzschmar, Laura A. Mengel, Anna O. Müller, Kerstin Schmitt, Katharina F. Blersch,  
Oliver Valerius, Gerhard H. Braus and Till Ischebeck

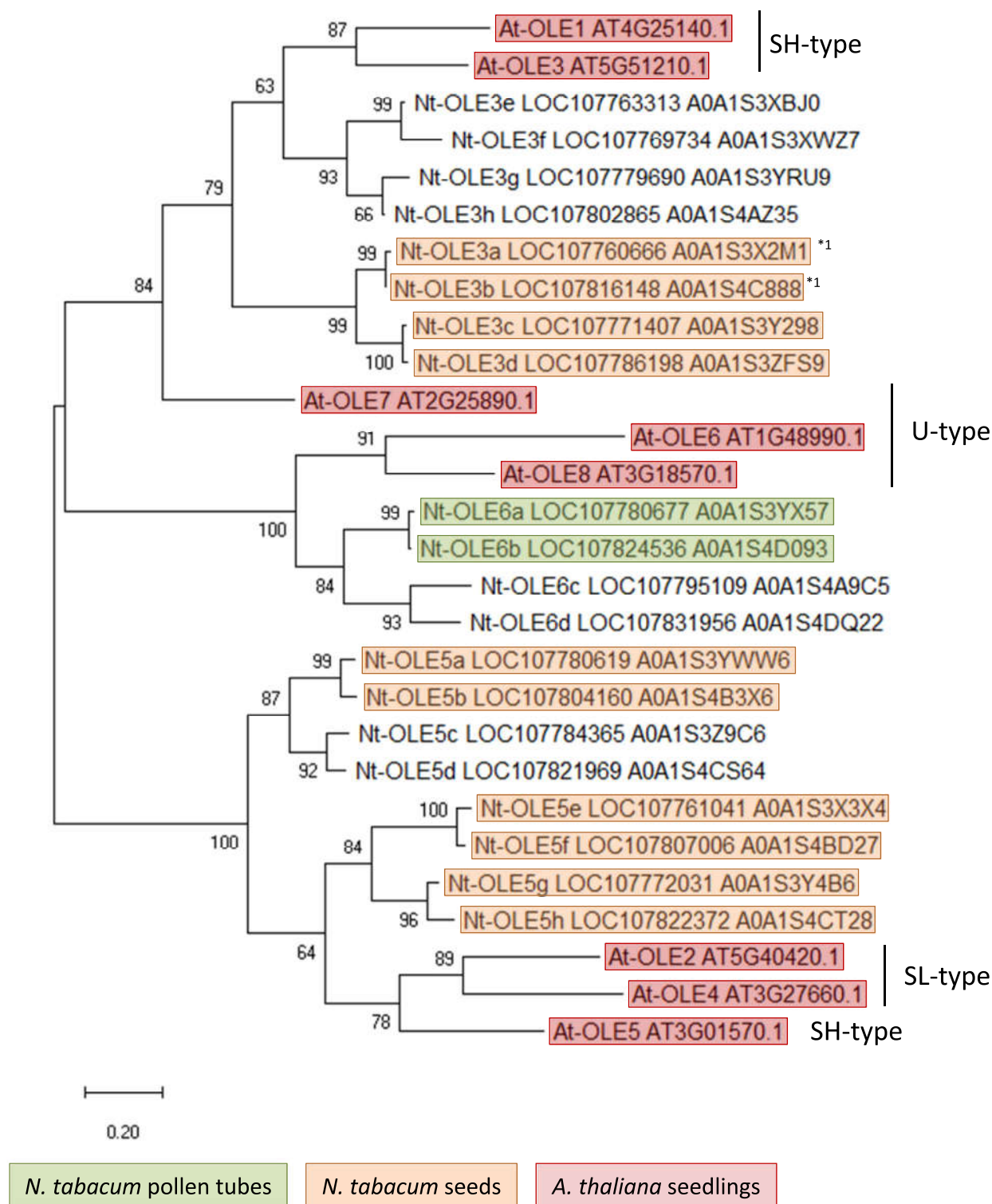
*Plant Cell* 2018;30;2137-2160; originally published online August 7, 2018;  
DOI 10.1105/tpc.18.00276

This information is current as of October 8, 2018

<b>Supplemental Data</b>	<a href="/content/suppl/2018/08/07/tpc.18.00276.DC1.html">/content/suppl/2018/08/07/tpc.18.00276.DC1.html</a>
<b>Permissions</b>	<a href="https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X">https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X</a>
<b>eTOCs</b>	Sign up for eTOCs at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a>
<b>CiteTrack Alerts</b>	Sign up for CiteTrack Alerts at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a>
<b>Subscription Information</b>	Subscription Information for <i>The Plant Cell</i> and <i>Plant Physiology</i> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a>



## Oleosins

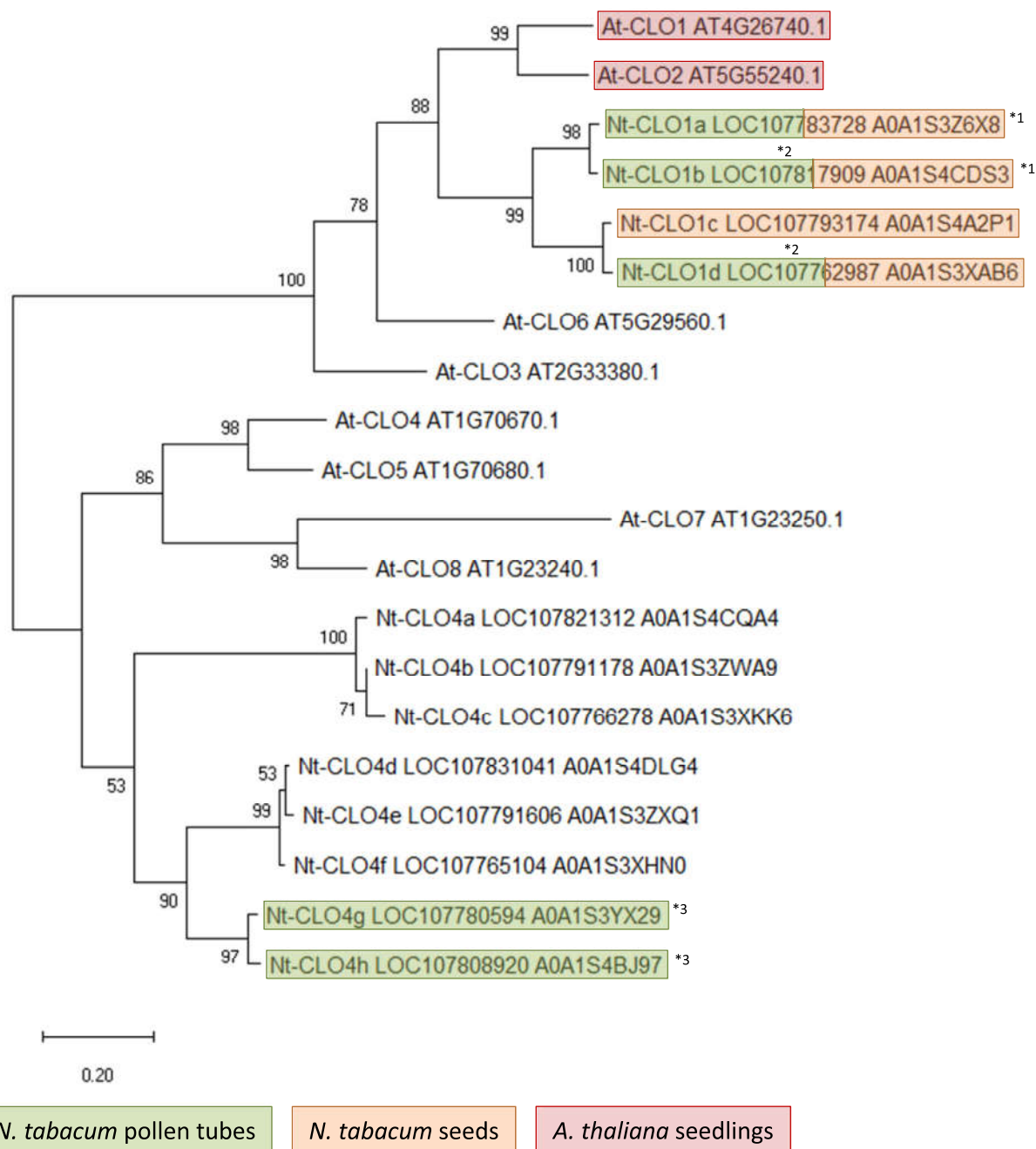


**Supplemental Figure 1. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* oleosin proteins.** Supporting Figure 1A. All *Arabidopsis* oleosins are detected at least at one time point during seed germination. In Tobacco, oleosins are differentially expressed in seeds and pollen tubes. LD localization of Nt-OLE6b is presented in Figure 1A.

Proteins marked by asterisk were not identified unambiguously.

Tree generated in MEGA X. Bar corresponds to substitutions per site.

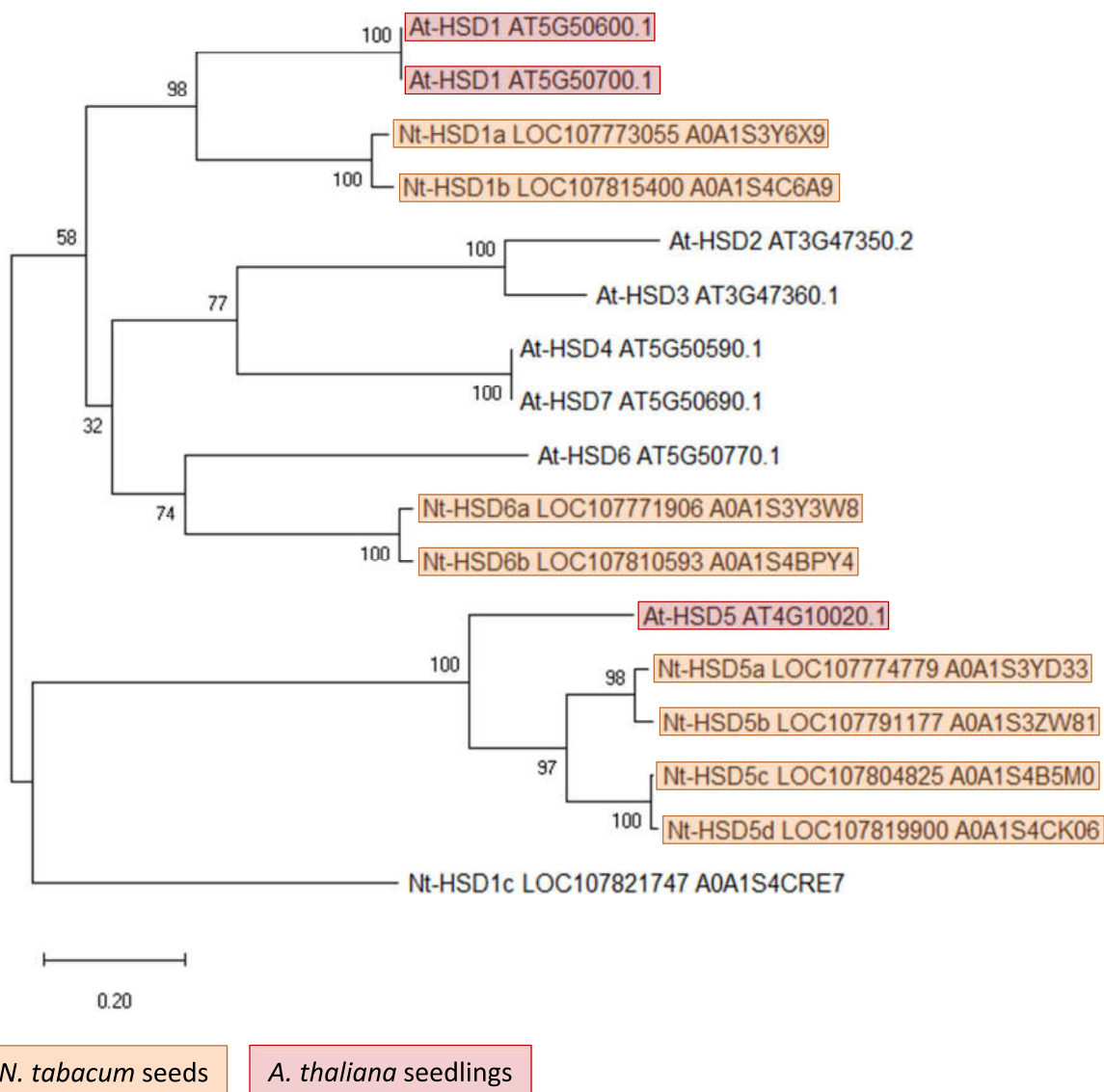
## Caleosins



**Supplemental Figure 2. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* caleosin proteins.** Supporting Figure 1B. In *Arabidopsis*, two caleosins are detected in germinating seeds. In Tobacco, expression of caleosins in seeds and pollen tubes partially overlaps, with one homolog being only seed-specific, two homologs being only pollen tube-specific and three homologs being expressed in both tissues. LD localization of Nt-Clo1a is presented in Figure 1B.

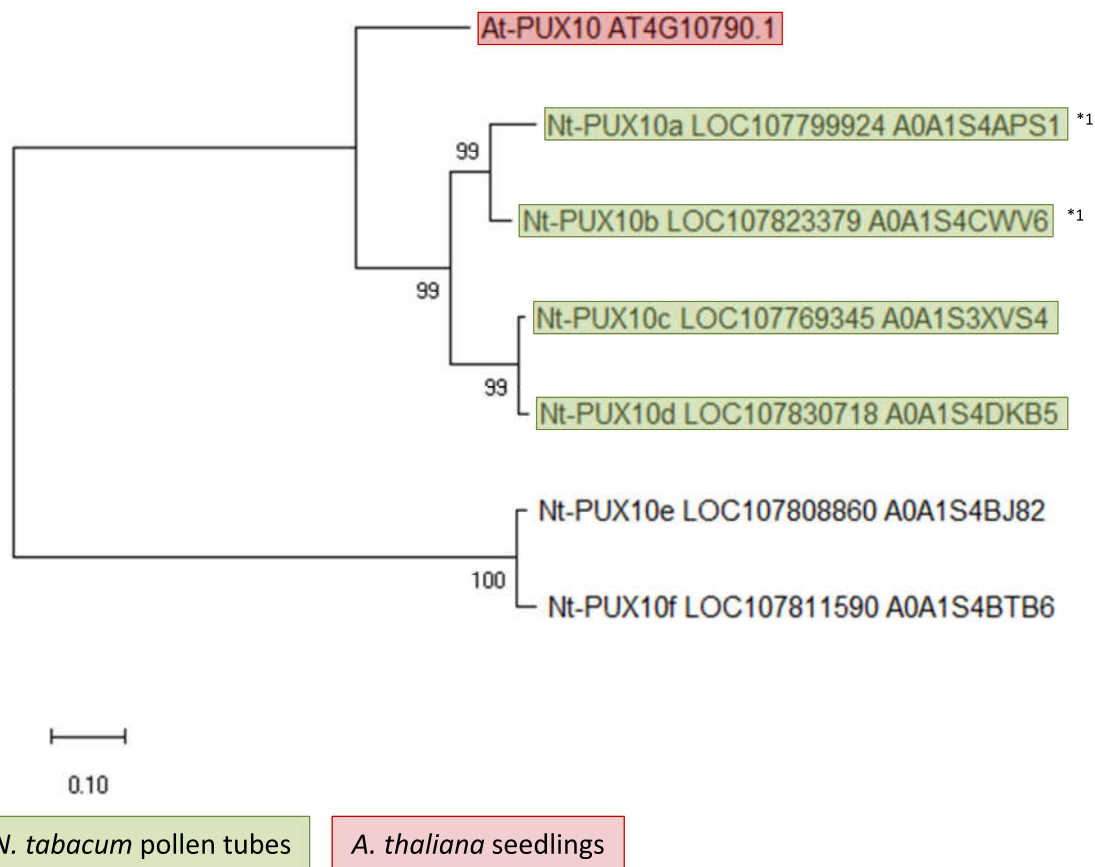
Proteins marked by asterisks and the same number were not identified unambiguously. Tree generated in MEGA X. Bar corresponds to substitutions per site.

## Steroleosins (hydroxysteroid dehydrogenase)



**Supplemental Figure 3. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* steroleosin proteins.** Supporting Figure 7. In *Arabidopsis*, three homologs of the protein have been detected in germinating seeds. Close homologs to the seed-specific *Arabidopsis* steroleosins have also been detected in tobacco seeds. Additionally, two more homologs were found in tobacco seeds. No steroleosin homologs were detected in tobacco pollen tubes. Tree generated in MEGA X. Bar corresponds to substitutions per site.

## Plant UBX domain-containing proteins 10

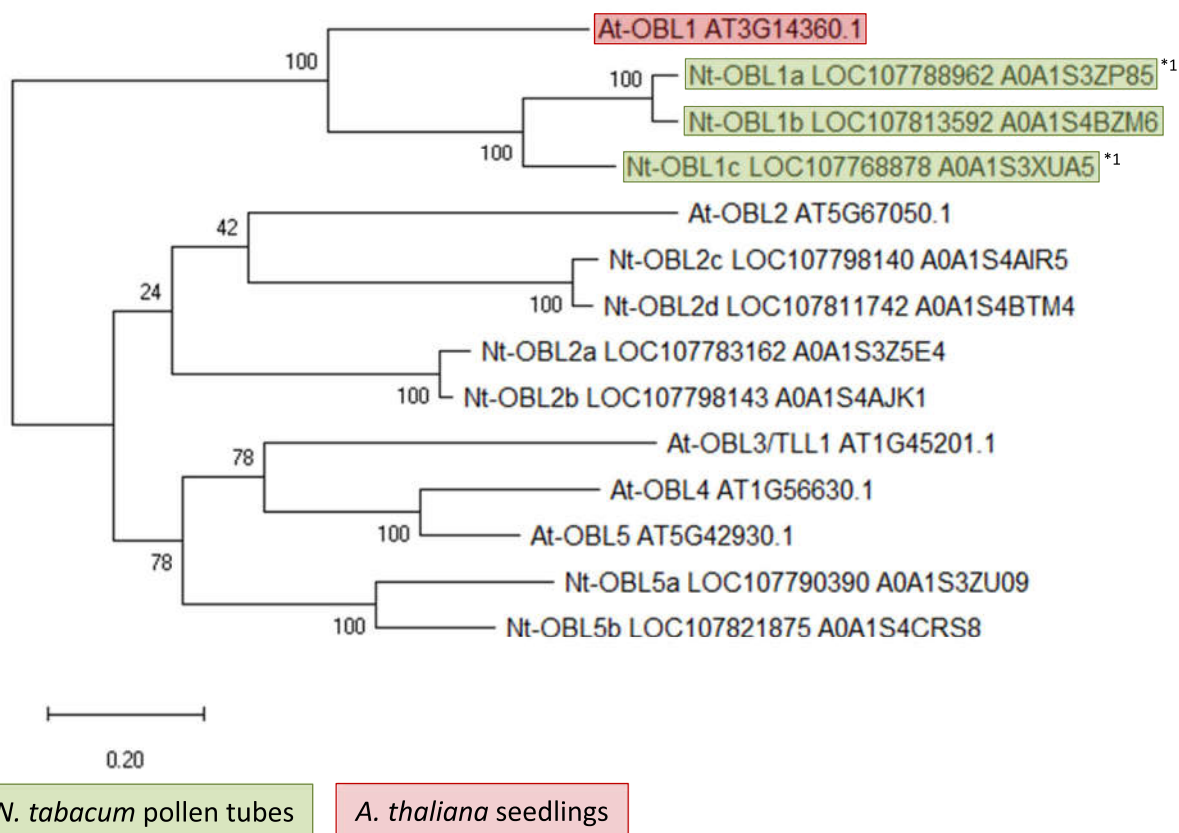


**Supplemental Figure 4. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* Plant UBX domain-containing proteins 10 (PUX10).** Supporting Figures 2A and 5. In *Arabidopsis*, the only PUX10 isoform was detected in germinating seeds. Four homologs were detected in tobacco pollen tubes. LD localization of Nt-PUX10c is presented in Figure 2A, and LD localization of At-PUX10 is shown in Figure 5.

Proteins marked by asterisk and the same number were not identified unambiguously. Tree generated in MEGA X. Bar corresponds to substitutions per site.

Please note that the family of PUX proteins contains 15 more members in *Arabidopsis* and 22 more members in tobacco.

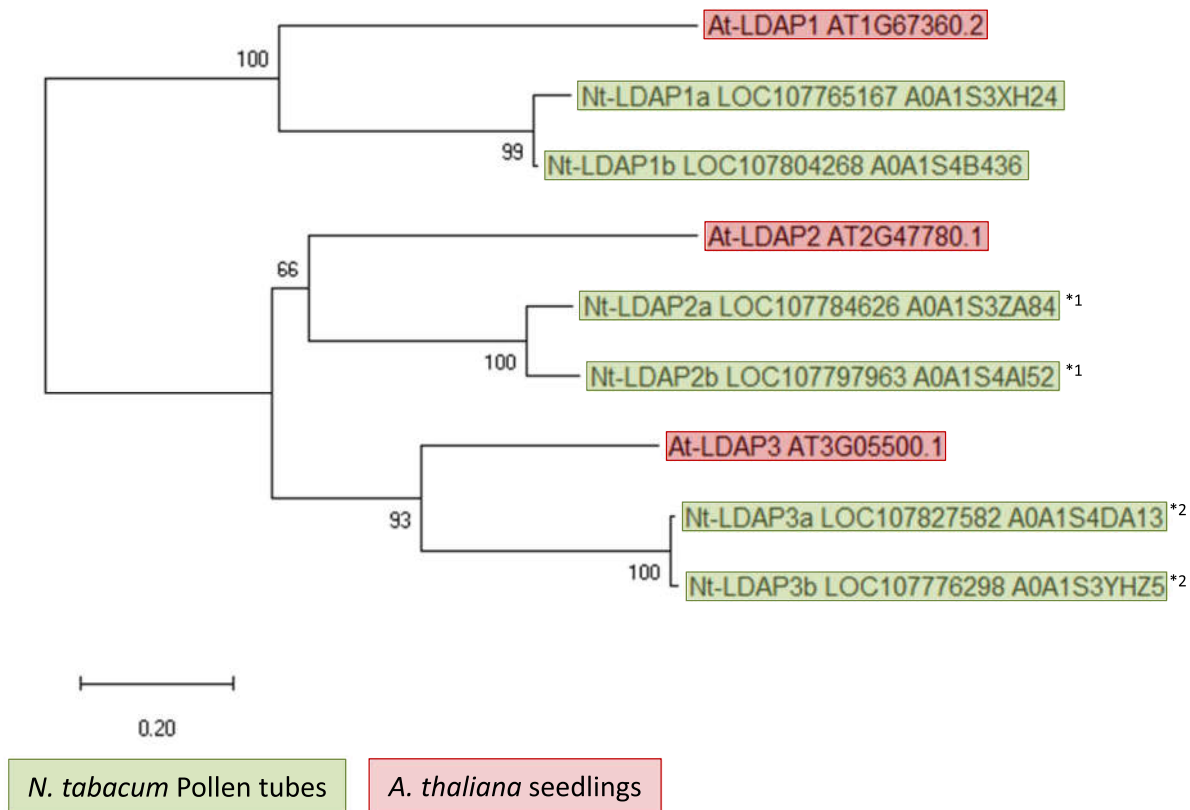
## Oil body lipases



**Supplemental Figure 5. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* Oil body lipases.** Supporting Table 1. In *Arabidopsis*, five homologs of the protein are known with one of them being expressed in germinating seeds. In Tobacco, ten homologs have been identified, and three of them were detected in pollen tubes.

Proteins marked by asterisk and the same number were not identified unambiguously. Tree generated in MEGA X. Bar corresponds to substitutions per site.

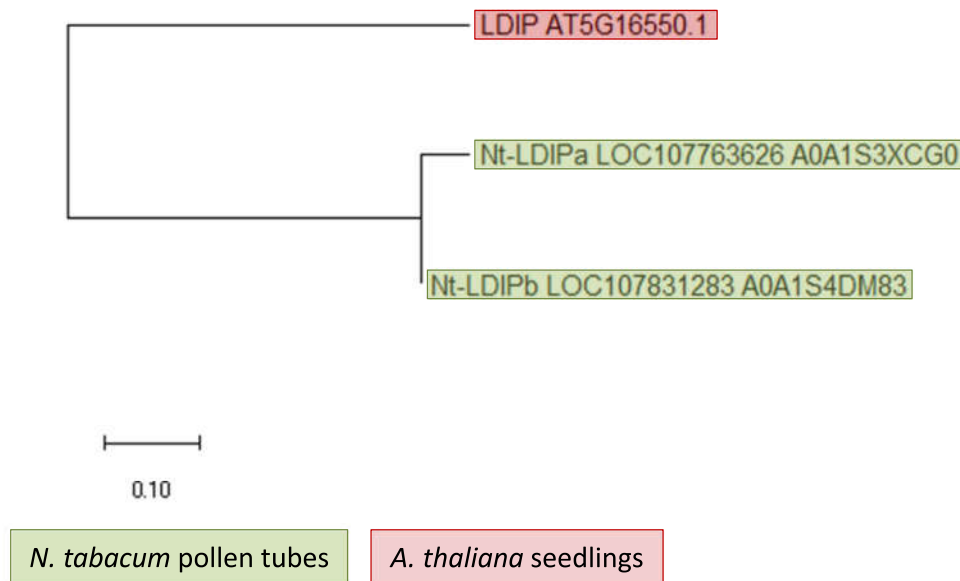
## Lipid droplet-associated protein



**Supplemental Figure 6. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* Lipid droplet-associated proteins.** Supporting Figure 1C. In *Arabidopsis*, all three known homologs of the protein were detected in *Arabidopsis* seedlings. All corresponding tobacco homologs were detected in tobacco pollen tubes. Proteins marked by asterisks and the same number were not identified unambiguously. Tree generated in MEGA X. Bar corresponds to substitutions per site.



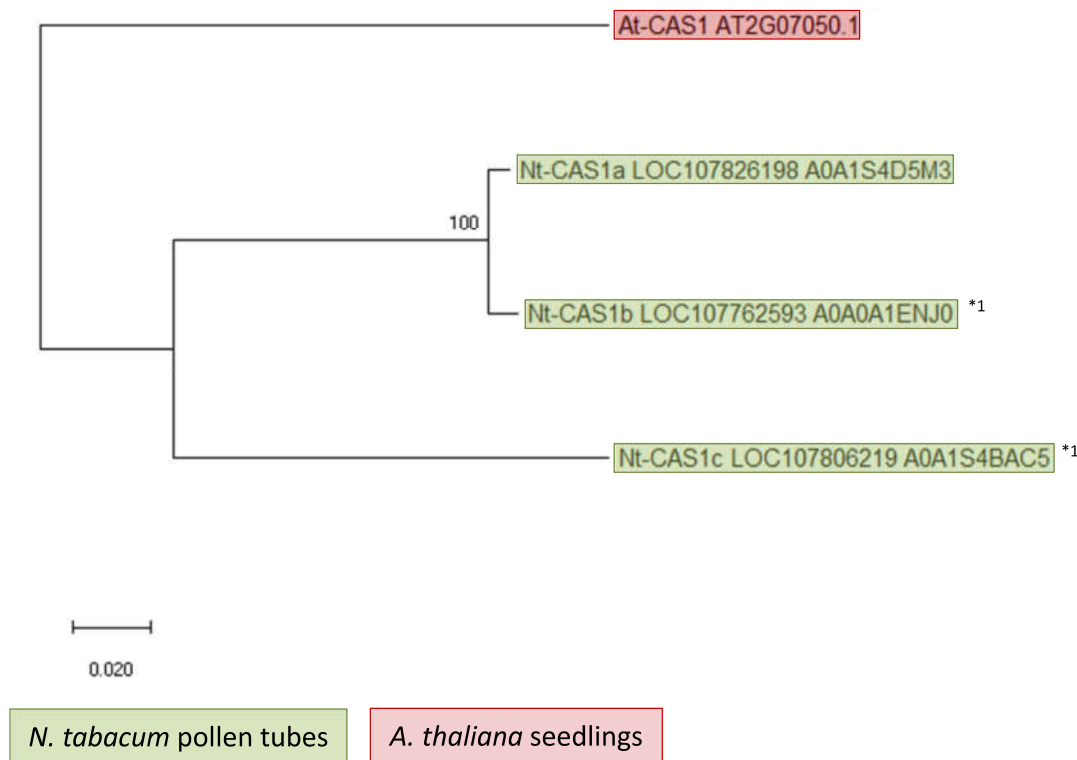
## Lipid droplet-associated protein-interacting protein



**Supplemental Figure 7. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* Lipid droplet-associated protein-interacting proteins.** Supporting Figures 1D and 1E. The one known *Arabidopsis* homolog was detected in germinating seeds. Both tobacco homologs were detected in tobacco pollen tubes.

Tree generated in MEGA X. Bar corresponds to substitutions per site.

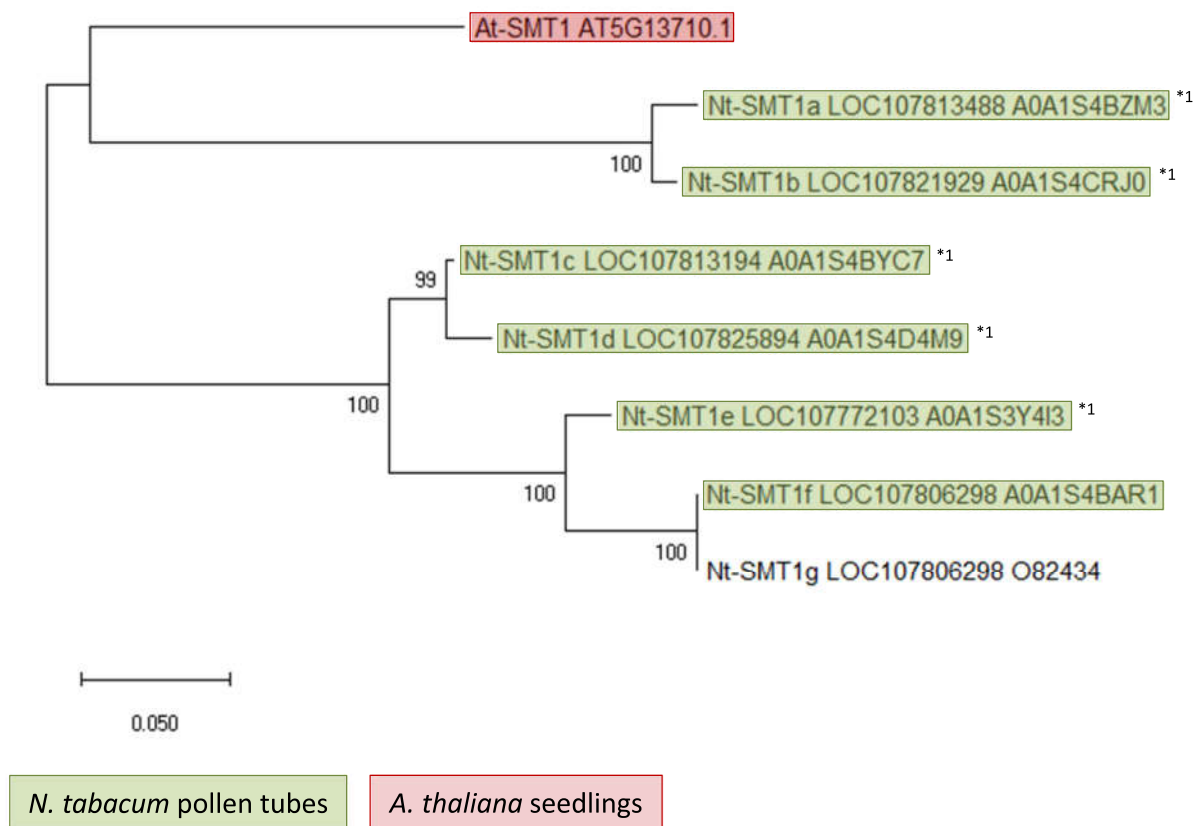
## Cycloartenol synthases



**Supplemental Figure 8. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* cycloartenol synthase proteins.** Supporting Figure 1G. In *Arabidopsis*, the homolog was detected in germinating seeds. In Tobacco pollen tubes, all three cycloartenol synthase isoforms were identified. LD localization At-CAS1 is shown in Supporting Figure 11. Proteins marked by asterisk and the same number were not identified unambiguously. Tree generated in MEGA X. Bar corresponds to substitutions per site.

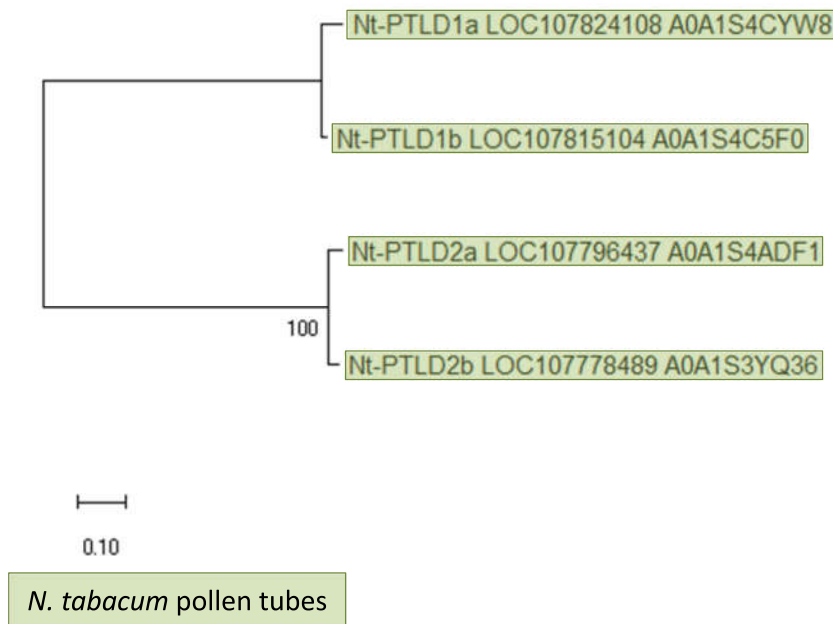


## Sterol methyltransferases

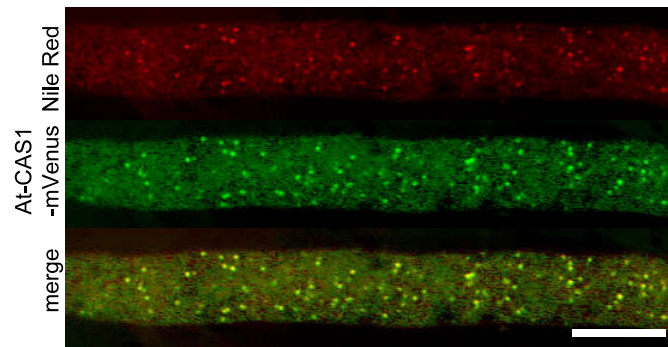


**Supplemental Figure 9. Phylogenetic tree of *Arabidopsis thaliana* and *Nicotiana tabacum* sterol methyltransferase proteins.** Supporting Figures 1H and 1I. In *Arabidopsis*, the one known homolog is expressed in germinating seeds. In *Tobacco*, seven homologs have been found, all but one were identified in pollen tubes. Proteins marked by asterisks and the same number were not identified unambiguously. Tree generated in MEGA X. Bar corresponds to substitutions per site.

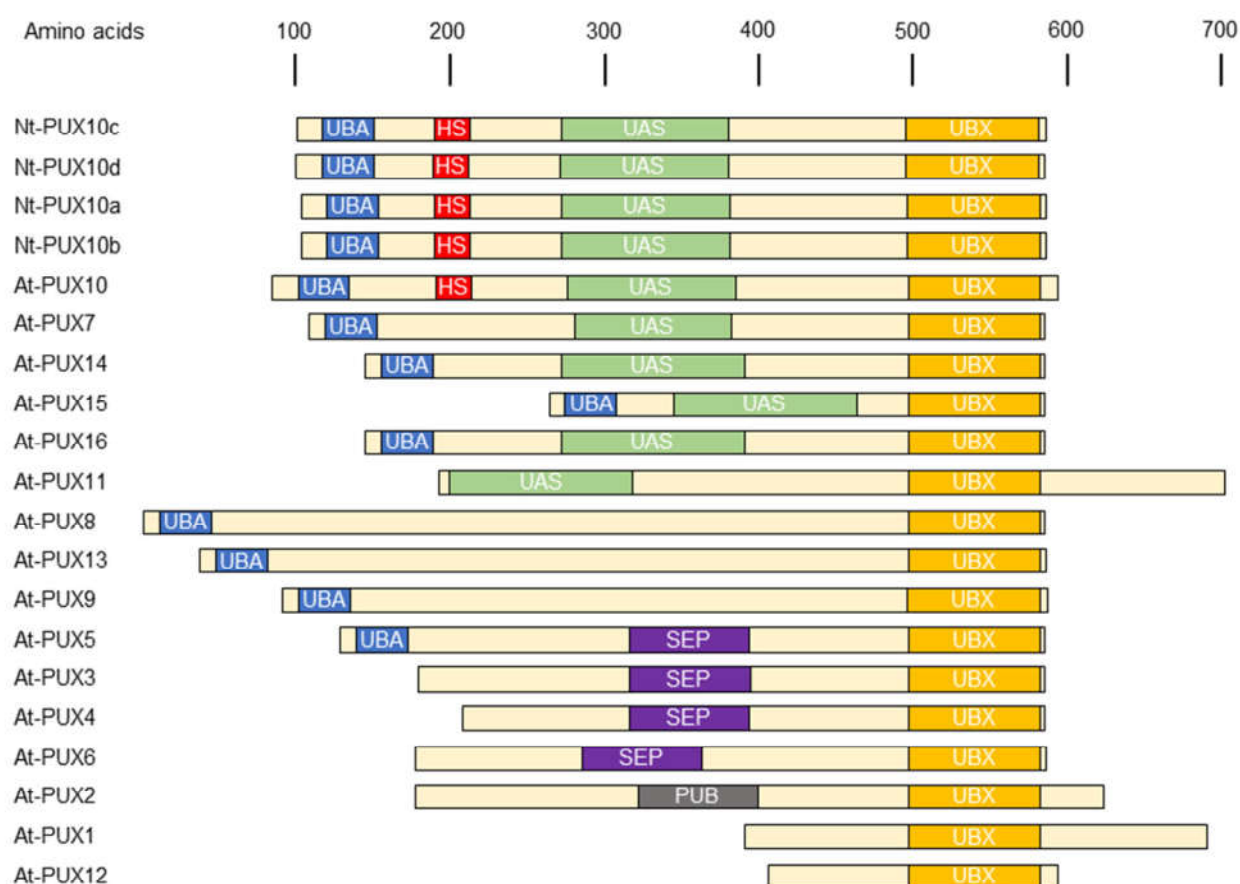
## Pollen tube lipid droplet protein



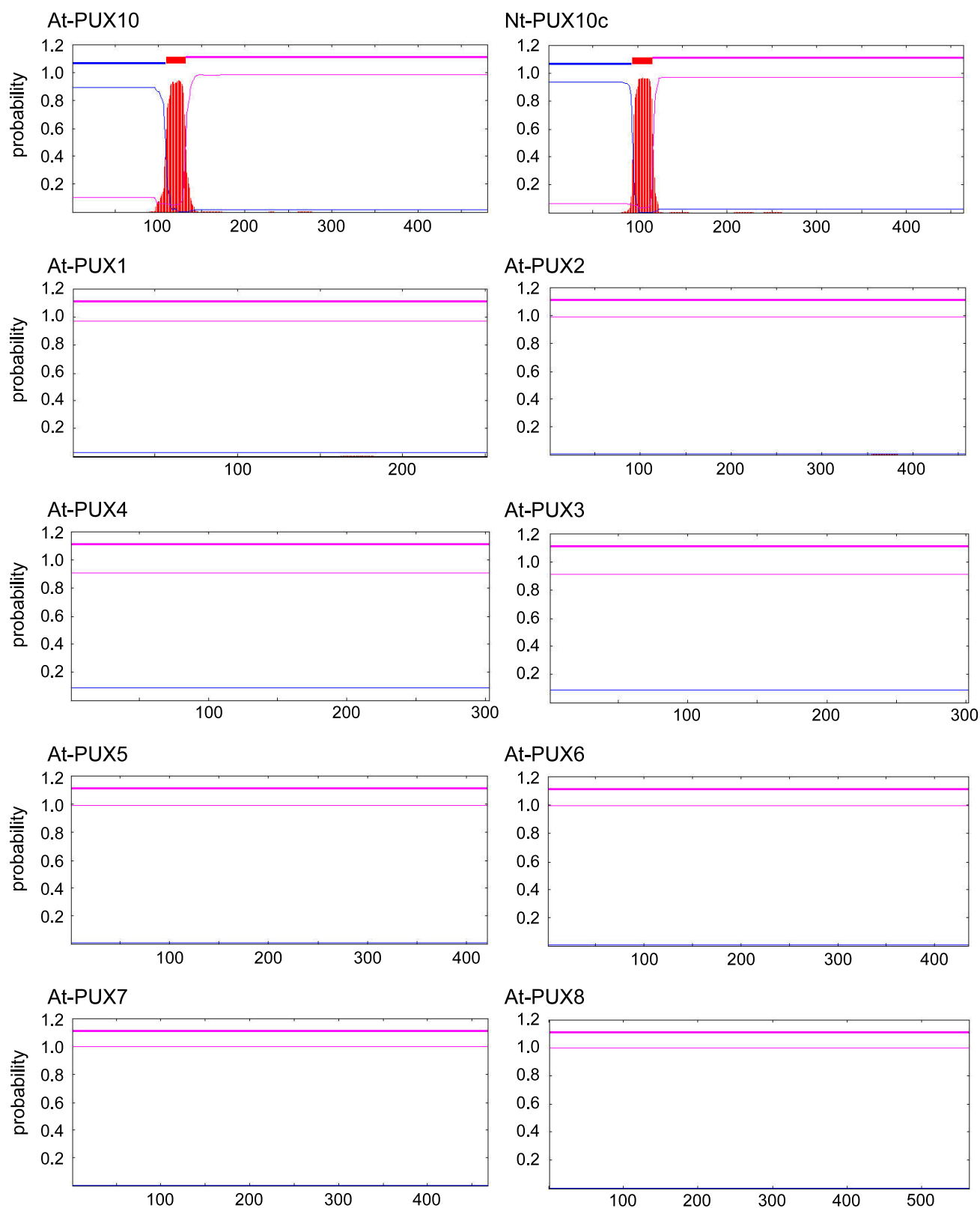
**Supplemental Figure 10. Phylogenetic tree of “Pollen tube lipid droplet protein” from *Nicotiana tabacum*.** Supporting Figure 1F. In Tobacco, four homologs have been found, and all were identified in pollen tubes. No homologs were found in *Brassicaceae*. The protein was previously annotated as unknown. Tree generated in MEGA X. Bar corresponds to substitutions per site.



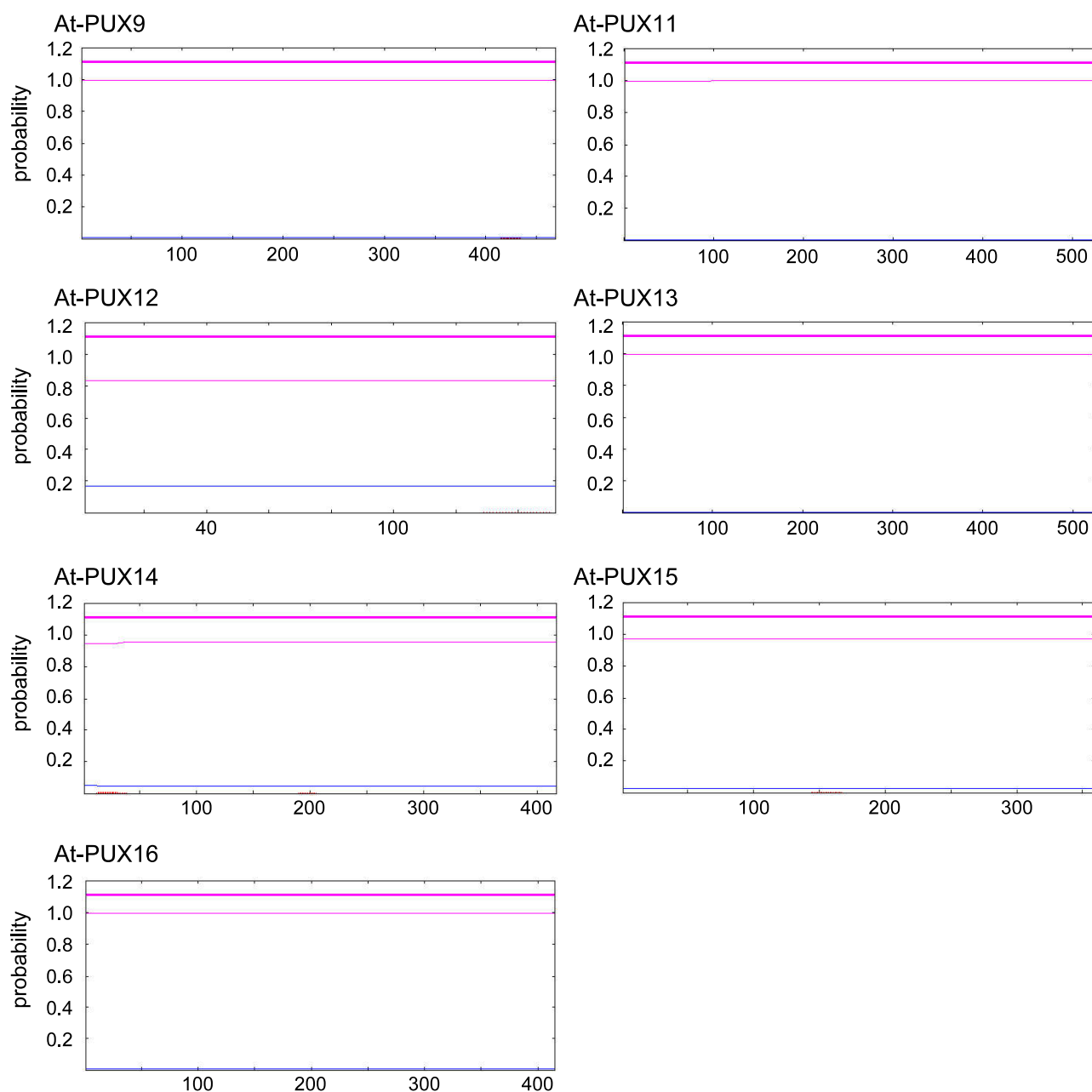
**Supplemental Figure 11. The cycloartenol synthase At-CAS1, Arabidopsis homolog to the LD-localized Nt-CAS1b localizes to LDs.** Supporting Figure 1G. The protein was transiently expressed in tobacco pollen tubes as an N-terminal fusion to the fluorescent protein mVenus. The tubes were cultivated for 5–8 h, fixed with formaldehyde, and stained with Nile Red. Then, they were monitored by confocal microscopy. Images are representative for 12 pollen tubes. Bar, 10  $\mu$ m.



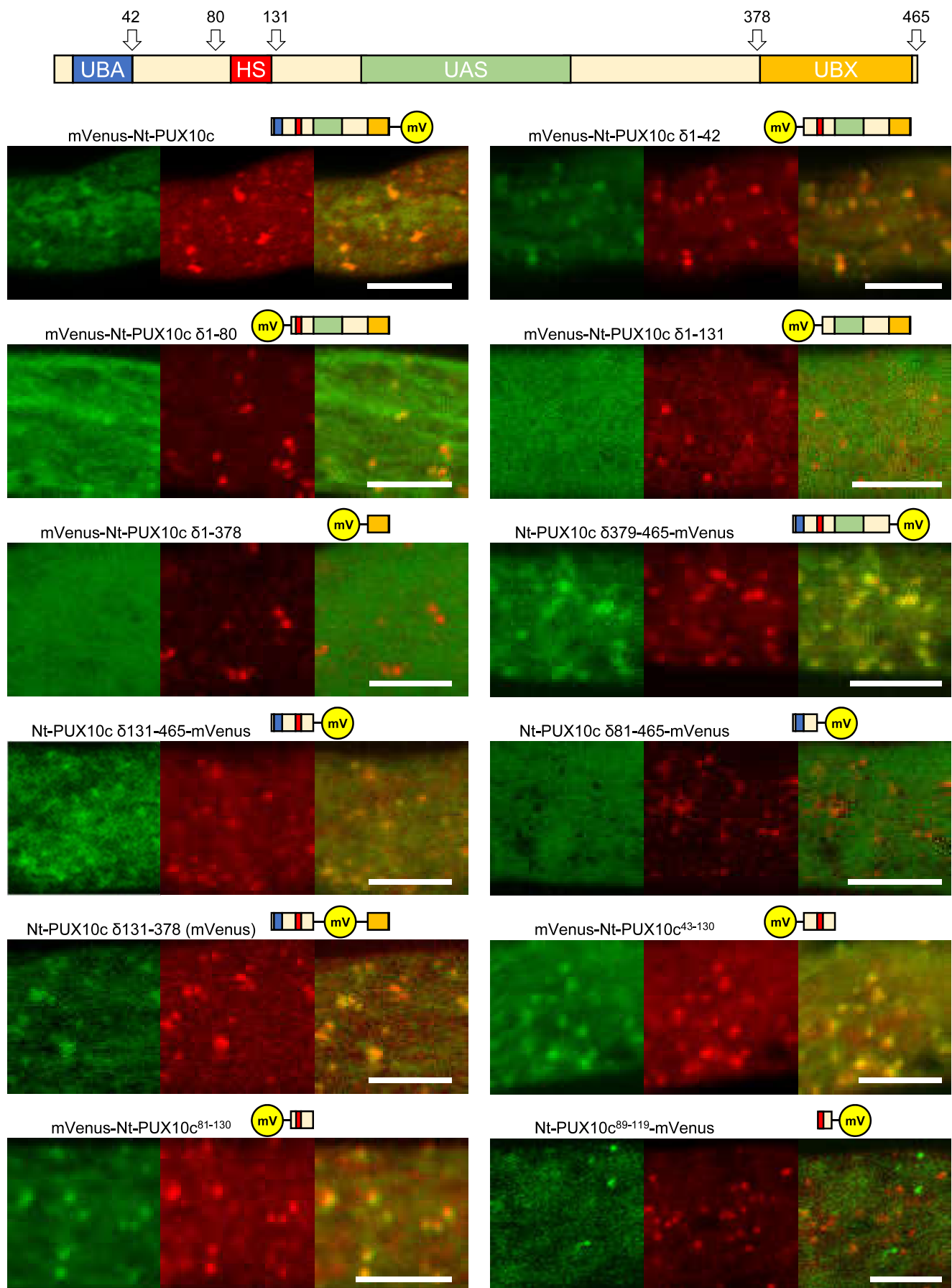
**Supplemental Figure 12. *Arabidopsis thaliana* Plant UBX domain-containing (PUX) protein gene family.** Supporting Figure 2D. All proteins of this family have the eponymous conserved C-terminal UBX domain. The UBX domain is described to interact with the AAA-type ATPase CDC48. The N-terminal UBA domain is a putative ubiquitin-interacting domain and conserved in nine proteins of the *Arabidopsis* gene family. UAS and SEP domains are domains of unknown function frequently found in UBX domain-containing proteins. For At-PUX10, there are six homologs identified in *N. tabacum*. The domain predictions presented here were applied in the design of the truncated versions of Nt-PUX10c presented in Figure 3D.



**Supplemental Figure 13. Hydropathy profiles of the Arabidopsis PUX proteins and Nt-PUX10c.** Supporting Figure 2D. The profiles, based on the protein sequences (position of amino acids is given on the x-axis), were created by the TMHMM server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>; Krogh et al., 2001). At-PUX10 and its homolog Nt-PUX10c are the only proteins of the Arabidopsis PUX gene family containing a predicted transmembrane domain (indicated as red bars). These predictions were applied in the design of the truncated versions of Nt-PUX10c presented in Figure 2D. Figure is continued on the next page.



**Supplemental Figure 13 (continued). Hydropathy profiles of the Arabidopsis PUX proteins and Nt-PUX10c.** Supporting Figure 2D. The profiles, based on the protein sequences (position of amino acids is given on the x-axis), were created by the TMHMM server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>; Krogh et al., 2001). At-PUX10 and its homolog Nt-PUX10 are the only proteins of the Arabidopsis PUX gene family containing a predicted transmembrane domain (indicated as red bars). These predictions were applied in the design of the truncated versions of Nt-PUX10c presented in Figure 2D.



**Supplemental Figure 14. The hydrophobic stretch is both essential and sufficient for LD association (magnified images).**

Supporting Figure 2. The legend is continued on the next page.

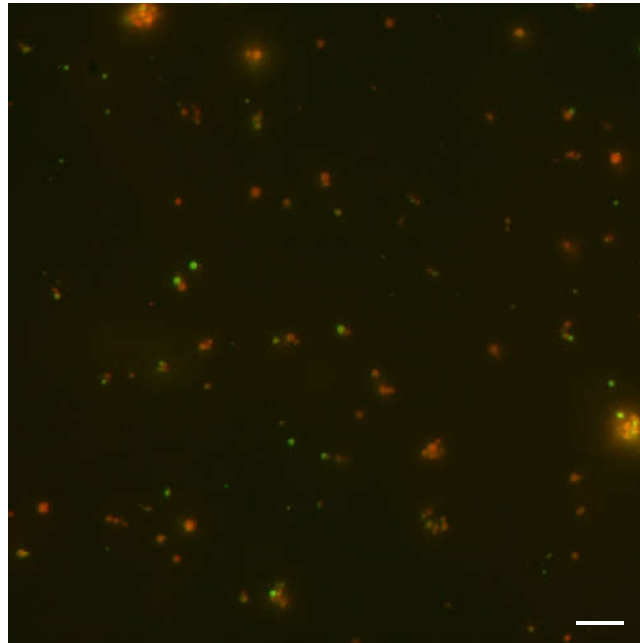
**Supplemental Figure 14 (continued). The hydrophobic stretch is both essential and sufficient for LD association (magnified images).** Supporting Figure 2.

The full-length Nt-PUX10c protein with a C-terminal mVenus fusion localizes to LDs and to structures similar to the ER. (Domain prediction made with CD-Search, hydrophobic region predicted with TMHMM, Supplemental Figure 13). All variants including the region comprising amino acids (AA) 81-131 colocalized at least in part with LDs, while all constructs excluding this region did not. A construct missing AA 1-88 and 120-465 localized to punctate structures not identical with LDs.

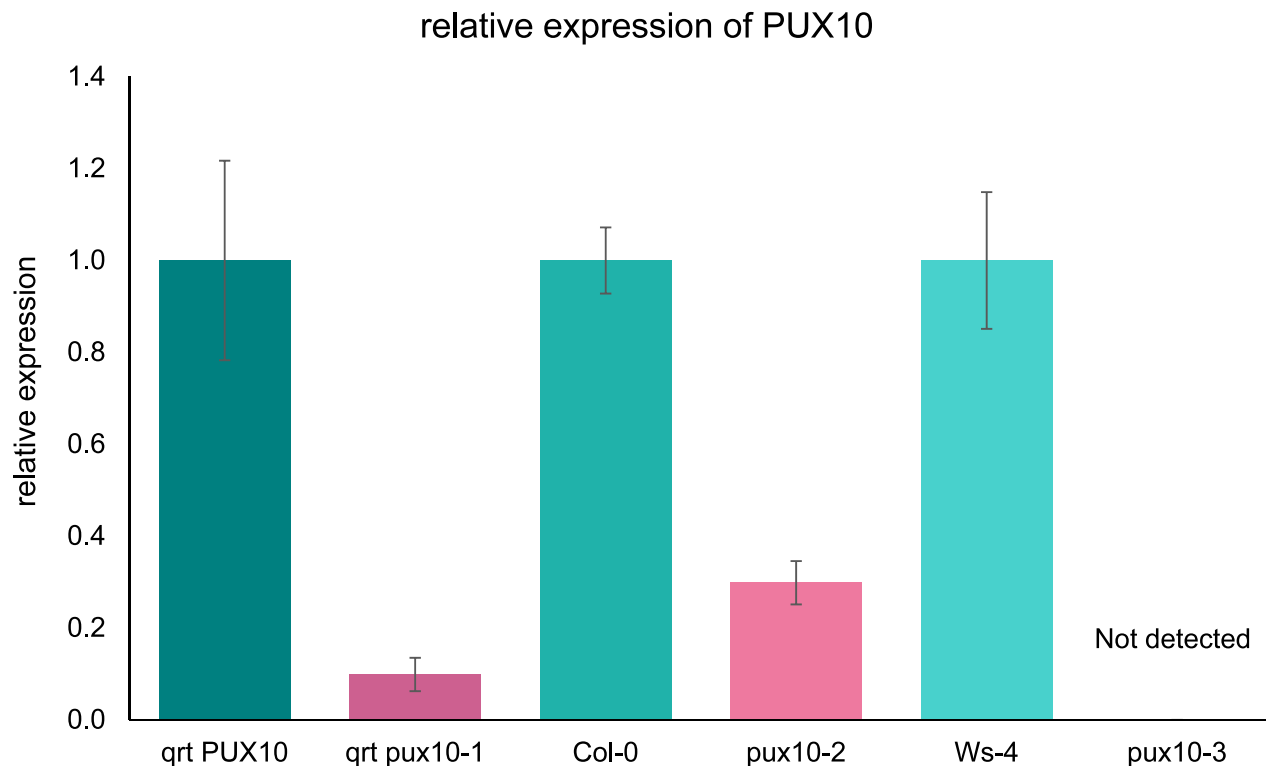
The proteins were transiently expressed in tobacco pollen tubes, as C- or N-terminal fusions to the fluorescent protein mVenus (green channel). The tubes were cultivated for 5-8 h, fixed with formaldehyde, and stained with Nile Red (red channel). Colocalization was observed in n = 6-11 independent pollen tubes per construct. Bars, 10  $\mu$ m. HS – hydrophobic stretch, UBA – ubiquitin associated domain, UBX – ubiquitin regulatory X.

The images represent magnifications of Figures 2A and 2D.



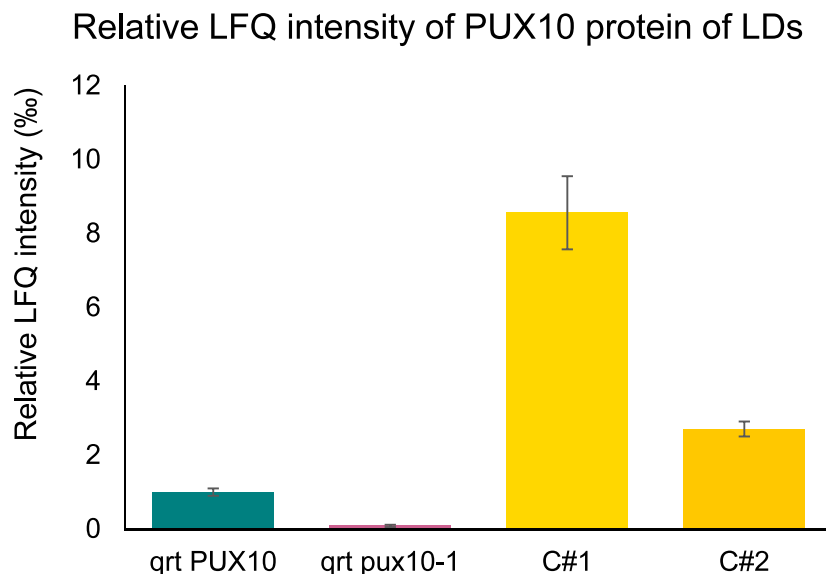


**Supplemental Figure 15. Uneven distribution of PUX10-eGFP on LDs in germinating seeds remains after isolation.** Supporting Figure 5K. LDs were isolated from Col-0 *At-PUX10<sub>pro</sub>:gAt-PUX10-eGFP* seedlings at 1 day after imbibition. After two washings steps, the LDs were resuspended in grinding buffer, stained with Nile Red, and observed by fluorescence microscopy. As observed for the intact seedlings (Figure 5K), the distribution of PUX10-eGFP on LDs is not homogenous. Bar, 10  $\mu$ m



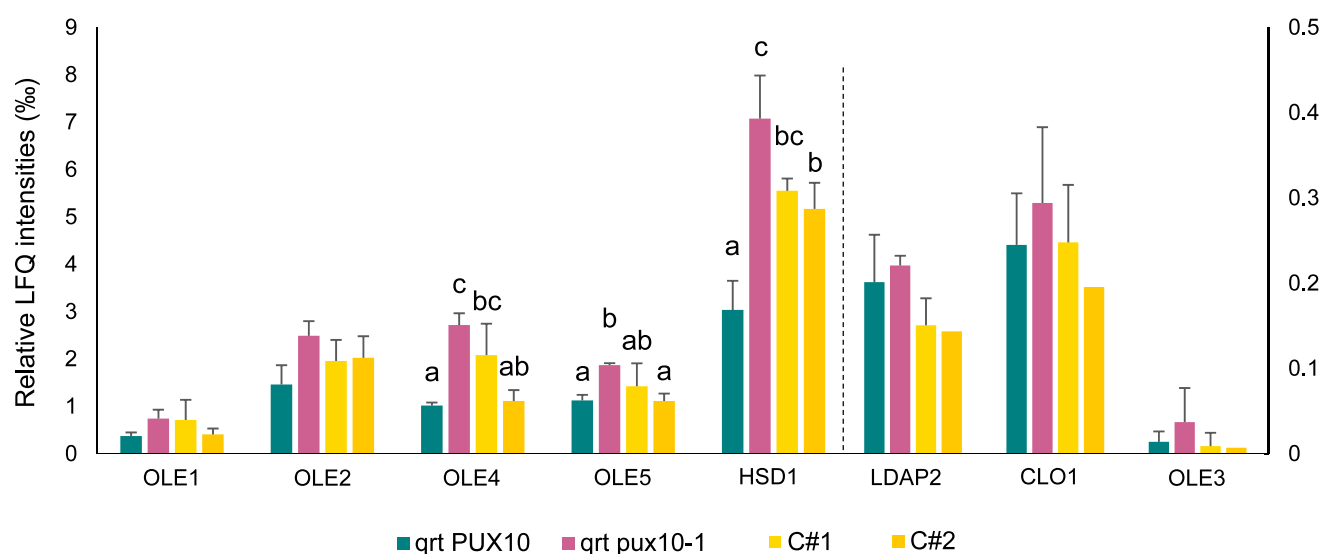
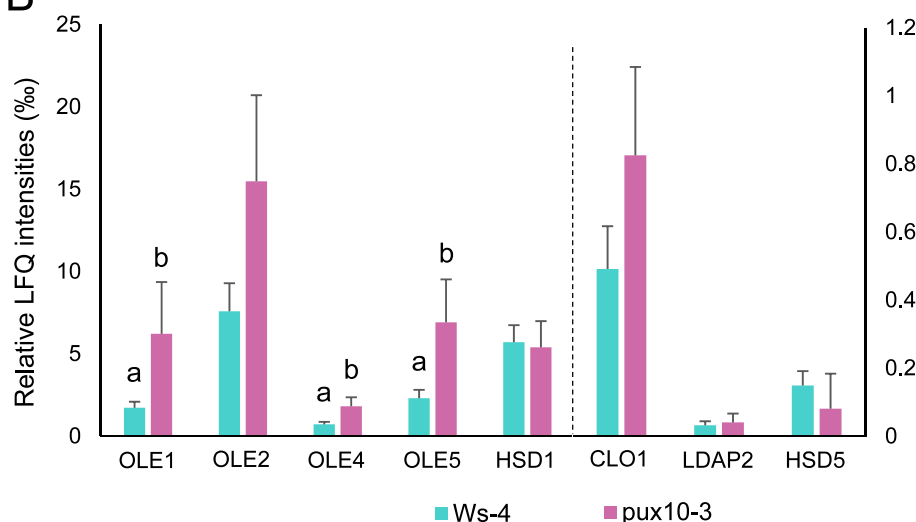
**Supplemental Figure 16. Relative expression of *PUX10* in mutants is strongly reduced compared to the mutant backgrounds.** Supporting Figure 6. cDNA was produced from RNA isolated from 10 mg dry seeds per biological replicate. No *PUX10* transcript was detected in the *pux10-3* mutant, the *PUX10* transcript in *qrt pux10-1* was reduced to 10 % compared to *qrt PUX10* non-mutated background, and the *PUX10* transcript in the *pux10-2* mutant was reduced to 30 % of Col-0 level. A graphic representation of the localization of the T-DNA insertions in the *PUX10* gene is presented in Figure 6A.

Average of three biological replicates. *PUX10* expression in the non-mutated background was set to 1. Error bars correspond to standard error. At4g12590 was used as a reference gene (Dekkers et al., 2012).



**Supplemental Figure 17. Relative LFQ intensity of PUX10 protein on LDs *qrt pux10-1* mutant and controls.** Supporting Figure 6. LDs were isolated from *qrt PUX10*, *qrt pux10-1* and the two complemented lines C#1 and C#2 at 38h after imbibition. Proteins were tryptically digested and peptides were analyzed by LC-MS/MS. Raw data were processed with the MaxQuant label-free quantification (LFQ) software (See Supplemental Tables 30 – 33 for raw data). PUX10 peptides were detected in all LD samples of the controls, and in one of three biological replicates of *qrt pux10-1*.

Average of three biological replicates. Relative LFQ intensity of PUX10 in the non-mutated background was set to 1. Error bars correspond to standard deviation.

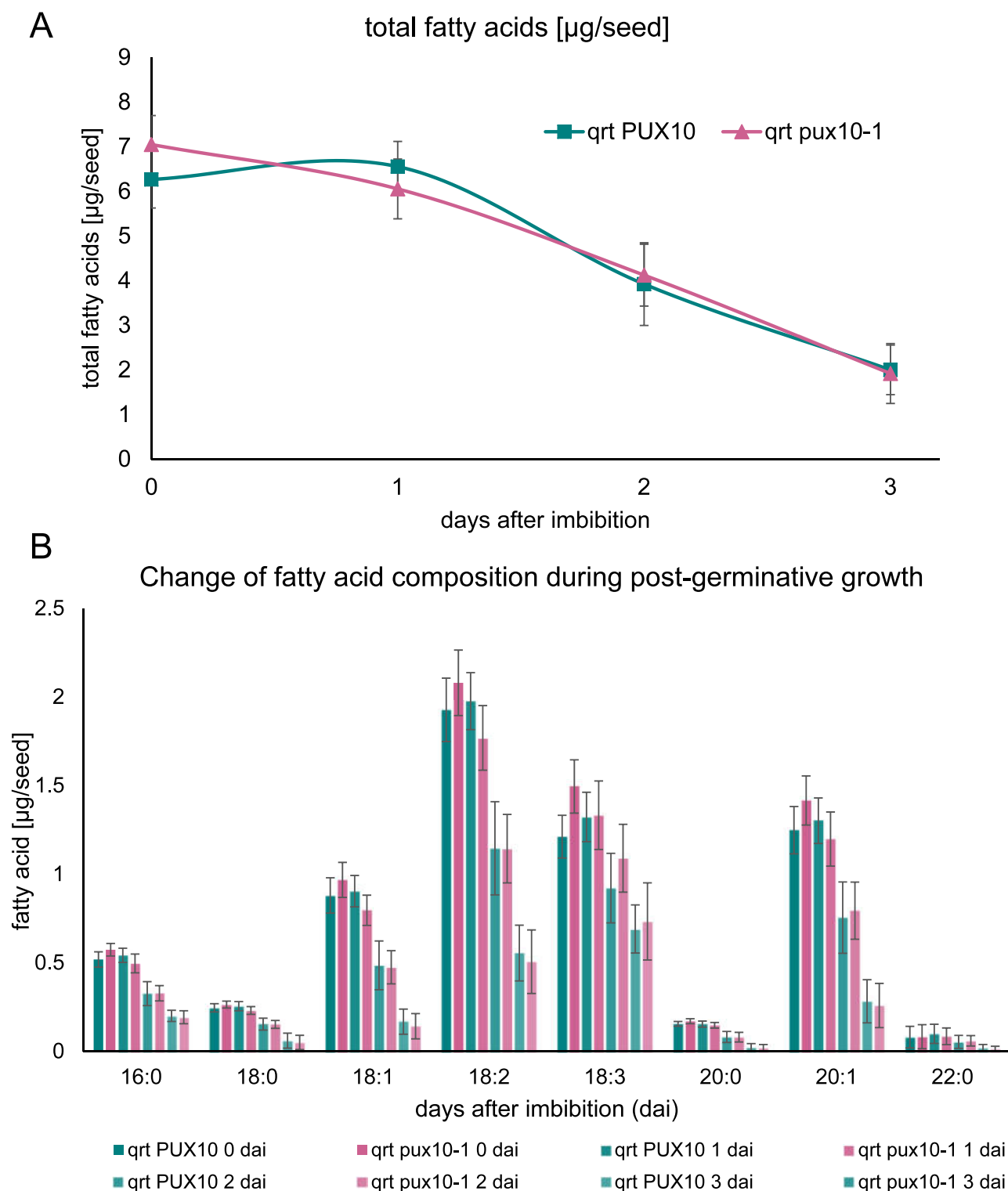
**A****B****Supplemental Figure 18. LD proteins accumulate in the *pux10* mutant seedlings.**

Separate representation of proteins added up in Figure 7A.

**A:** In the *qrt pux10-1* line, the LD proteins OLE4, OLE5 and HSD1 accumulate significantly at 2 days after imbibition (dai) when compared to the *qrt PUX10* line. This effect is partially rescued by the complemented lines. Proteins were isolated from total seedling homogenate at 2 dai.

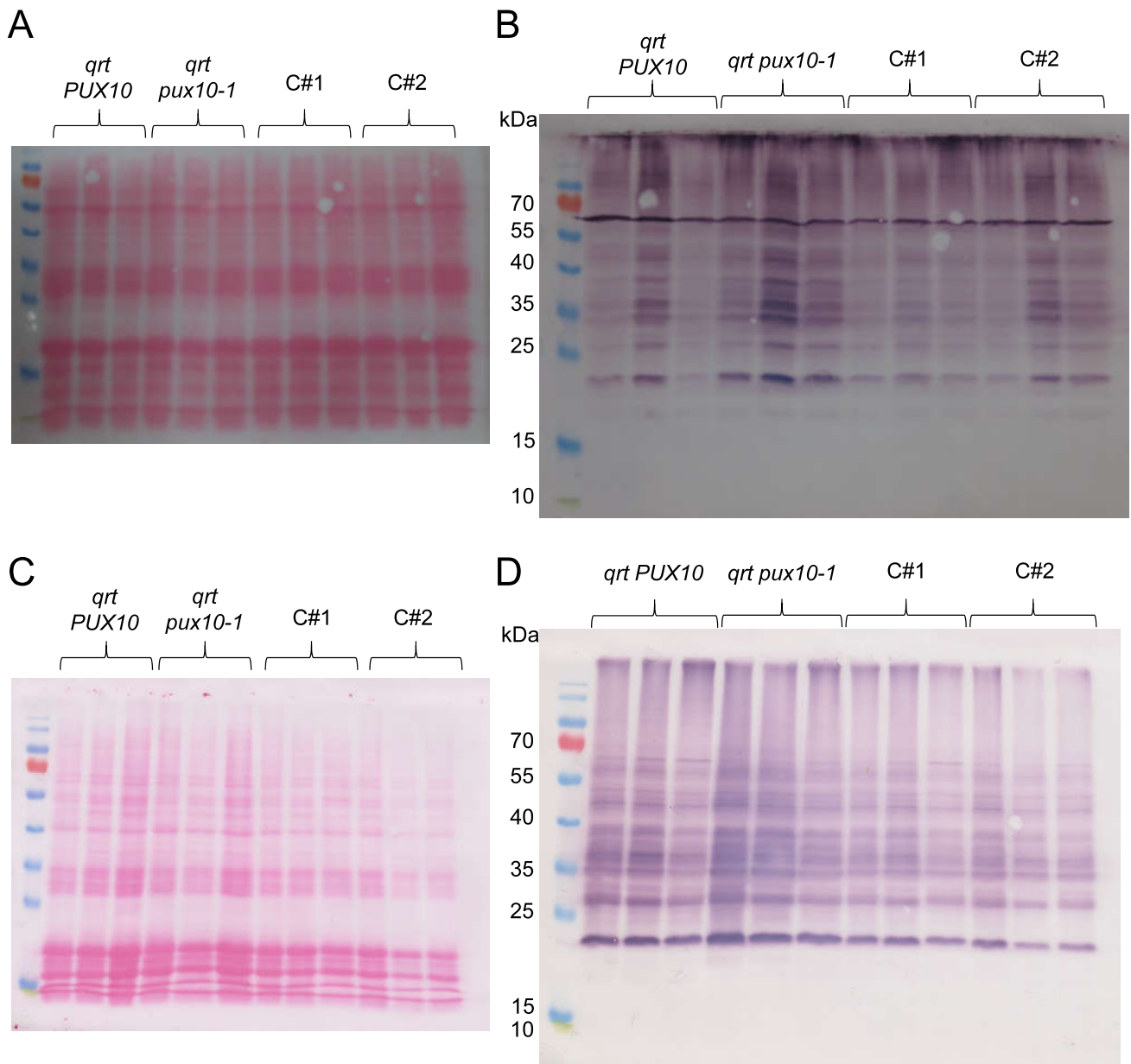
**B:** In the *pux10-3* line, the LD proteins OLE1, OLE4 and OLE5 accumulate significantly at 1 dai when compared to the Ws-4 wildtype background. Proteins were isolated from total seedling homogenate at 1 dai.

LC-MS/MS data were analyzed with the MaxQuant software label free quantification (LFQ) algorithm. Values are given as % of total LFQ intensities. Values are average of three biological replicates. Error bars correspond to standard deviation. Statistical analysis was performed as one-factorial ANOVA followed by Tukey's post-hoc test in R. Lowercase letters indicate significant differences ( $p < 0.05$ ).

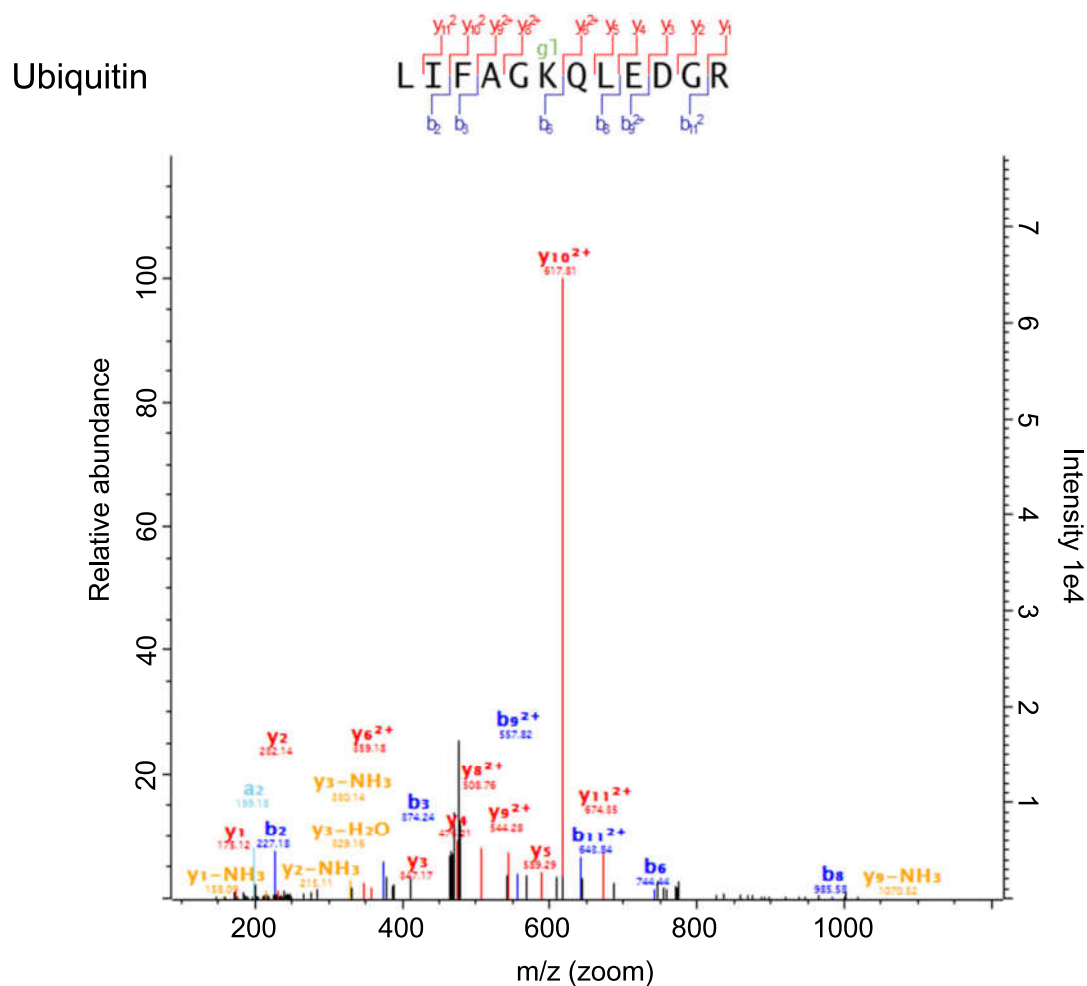


**Supplemental Figure 19. Post-germinative fatty acid degradation in *qrt pux10-1* and the non-mutated background *qrt PUX10*.** Supporting Figure 6. In contrast to the observed delay in protein degradation (Figure 7), no difference in fatty acid degradation between the two lines could be detected, neither when investigating total fatty acids (**A**), nor individual fatty acids including the marker fatty acid 20:1 (**B**).

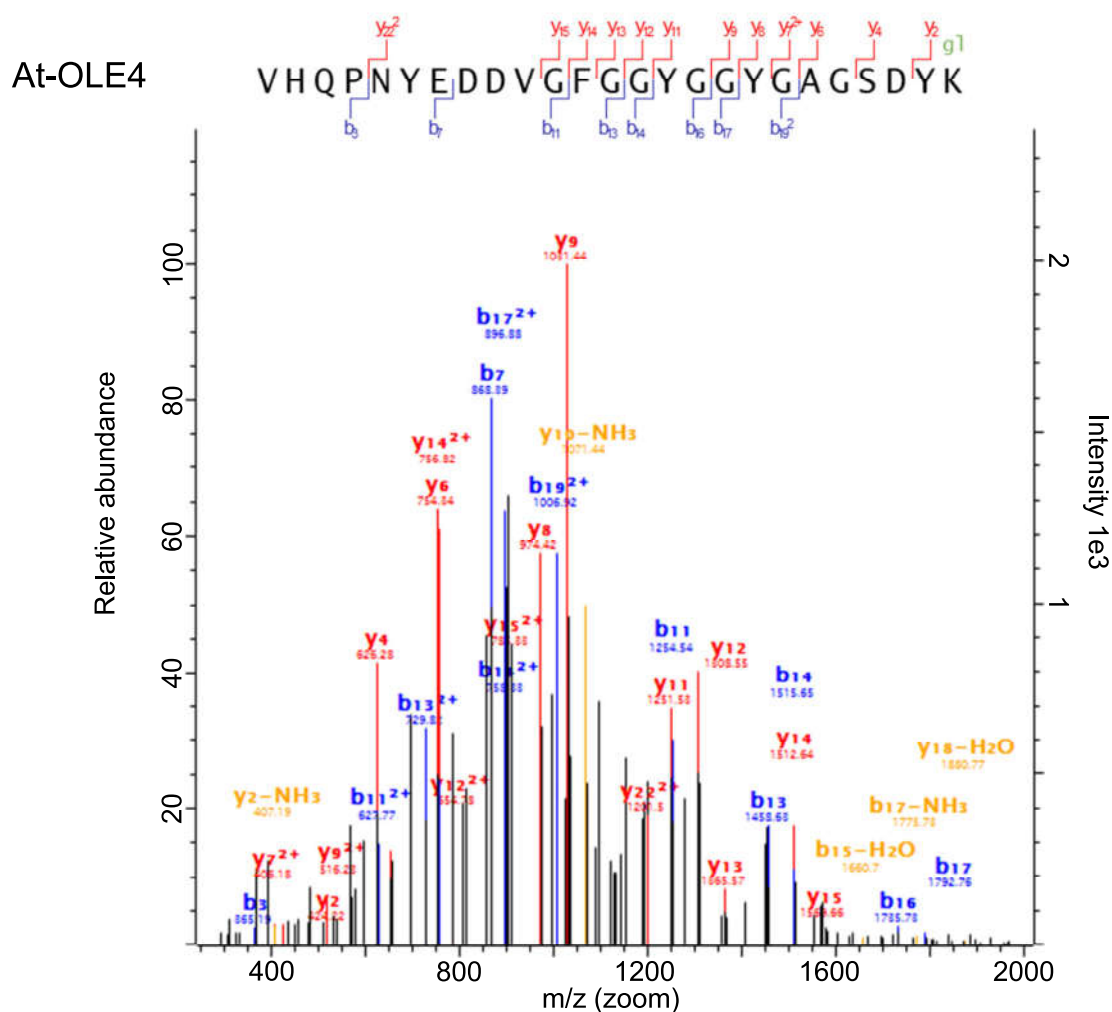
Fatty acid methyl esters created directly from rehydrated seed (0 days after imbibition, dai), and seedlings at one to three dai were quantified by gas chromatography with flame ion detection. C15:0 TAG was used as an internal standard. n=5 times 10 seeds for each of the time points.



**Supplemental Figure 20. Anti-ubiquitin immunoblots.** Supporting Figure 9A. The relative abundance of Ubiquitin increases in the *qrt pux10-1* mutant. Biological replicates from total protein fraction (A, B) or LD fraction (C, D) of *qrt PUX10*, *qrt pux10-1* and the two complementation lines C#1 and C#2 of seedlings 38 h after imbibition were loaded on the same gel. Samples correspond to the same presented in Figure 9a after proteomic analysis. Before blocking, membranes were stained with Ponceau S (A, C) to compare protein concentration. Membranes were treated with a Ubiquitin-specific antibody (B, D; UBQ11, Agrisera Antibodies).

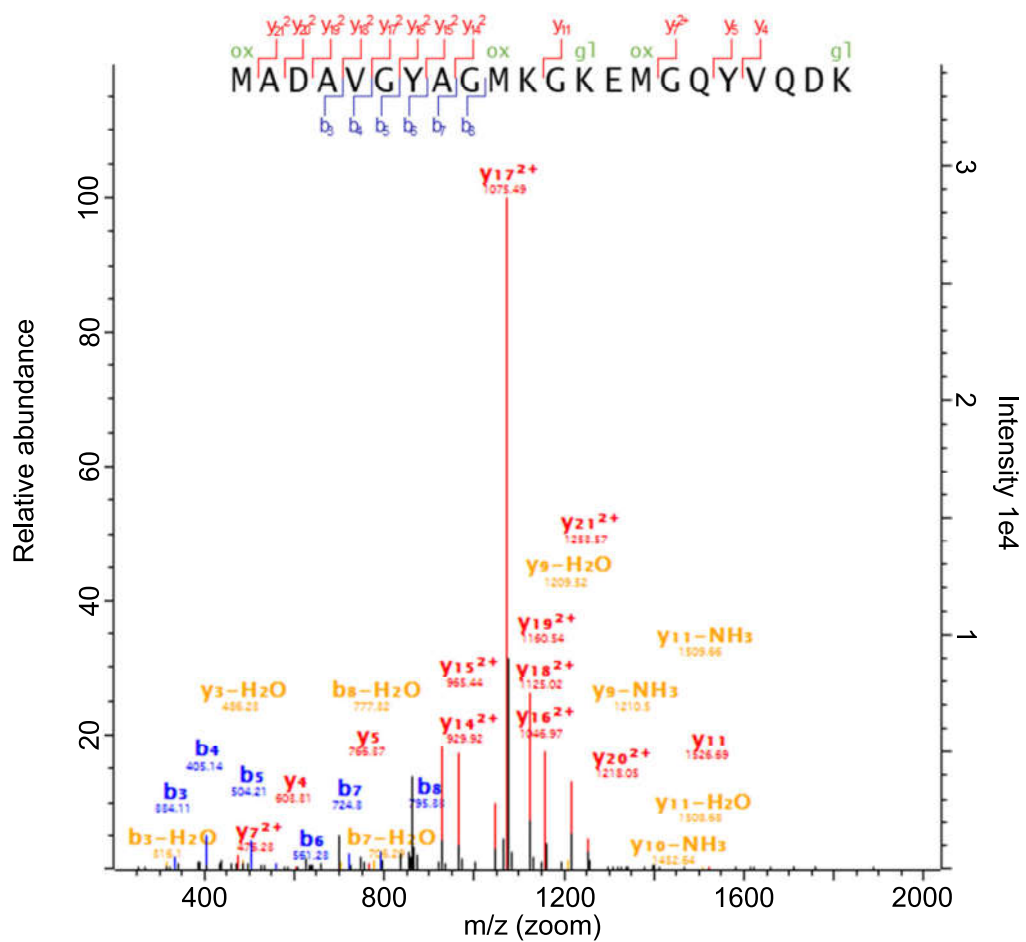
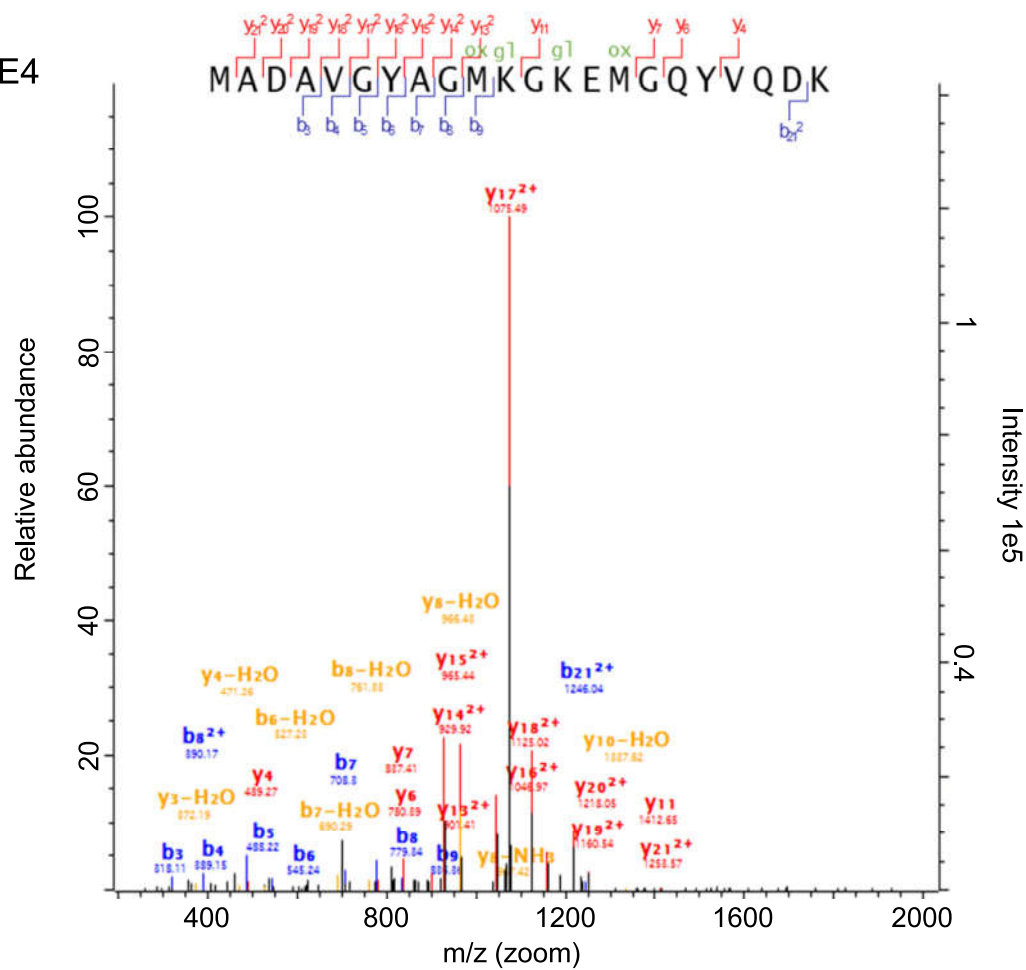


**Supplemental Figure 21. Annotated fragmentation spectrum of identified ubiquitination site of Ubiquitin presented in Figure 9B.** Spectrum was extracted from the MaxQuant software. Annotated fragmentation spectrum of the peptide LIFAGKQLEDGR (AA43 to AA54) of ubiquitin, ubiquitination position in protein on K48. AA – amino acid, K – Lysine.



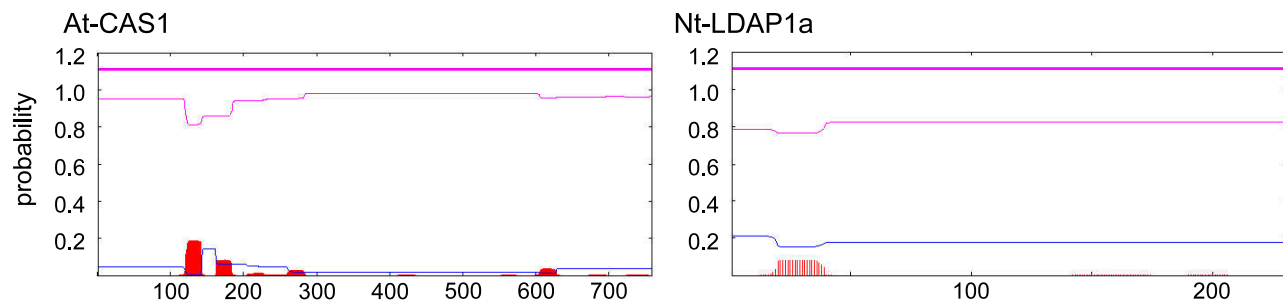
**Supplemental Figure 22. Annotated fragmentation spectrum of identified ubiquitination site of At-OLE4 presented in Figure 9B.** Spectrum was extracted from the MaxQuant software. Annotated fragmentation spectrum of the peptide VHQP NYE D D V G F G G Y G G Y G A G S D Y K (AA19 to AA43) of OLE4, ubiquitination position in protein K43. AA – amino acid, K – Lysine.





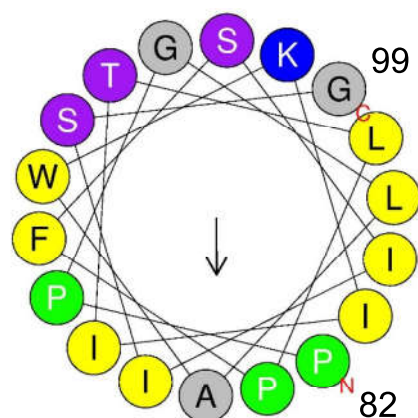
**Supplemental Figure 23. Annotated fragmentation spectra of identified ubiquitination site of At-OLE4 presented in Figure 9B.** Spectra were extracted from the MaxQuant software. Annotated fragmentation spectra of the peptide MADAVGYAGMKGKEMGQYVQDK (AA147 to AA168) of OLE4, ubiquitination positions in protein on K157 and K159 (upper panel), and K159 and K168 (lower panel). AA – amino acid, K – Lysine.

71

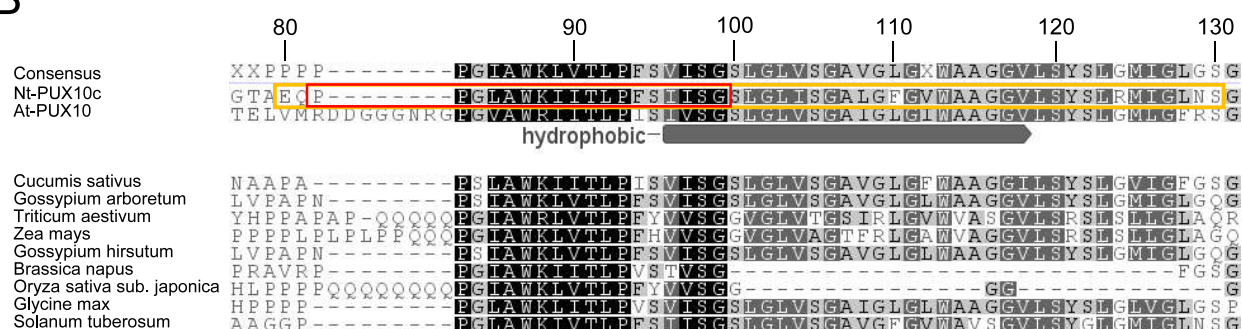


**Supplemental Figure 25. Hydropathy profiles of At-CAS1 and Nt-LDAP1a.** Supporting Figure 1. The profiles, based on protein sequences (position of amino acids is given on the x-axis), were created in TMHMM server 2.0 (Krogh *et al.*, 2001). Neither At-CAS1 (Supplemental Figure 11) nor the LDAP protein (Figure 1C) contain a predicted transmembrane (probability indicated as red bars) domain potentially involved in their LD targeting.

A



B

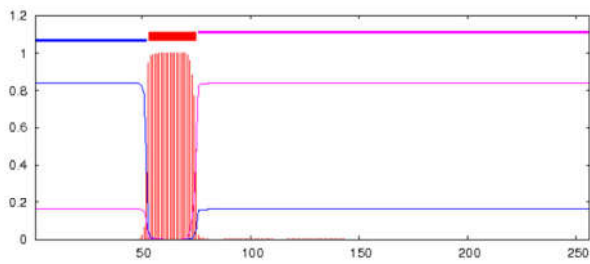


### Supplemental Figure 26. Helical wheel projection of amino acids 93 to 111 of At-PUX10.

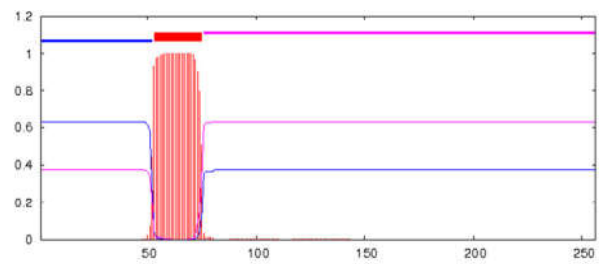
Supporting Figure 2. **A:** Via *in silico* analysis (HeliQuest), this part of the protein, located N-terminal of the hydrophobic sequence, is predicted to form an amphipathic  $\alpha$ -helix. Truncation experiments (Figure 3D) indicate that the combination of this amphipathic helix and the hydrophobic domain is required for the LD localization of a fluorescence protein. **B:** Comparison of the sequences of PUX10 homologs from different species indicate that both the amphipathic helix and the hydrophobic sequence are conserved among those homologs. The red box indicates the residues used for *in silico* helical wheel projection. The orange box indicates the residues needed for targeting a fluorophore to the LD. Alignment created in Geneious 8.18.

Protein identifiers: *Cucumis sativus* A0A0A0KKB6, *Gossypium arboreum* A0A0B0PRY7, *Triticum aestivum* A0A1D5RV37, *Zea mays* A0A1D6K6U6, *Gossypium hirsutum* A0A1U8NDX9, *Oryza sativa sub. japonica* A3C6J4, *Glycine max* 1LBY6, *Solanum tuberosum* M1A4Z9.

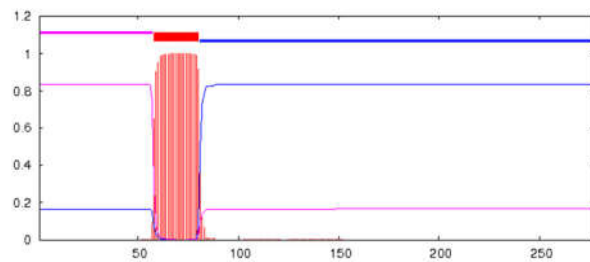
PTLD1a



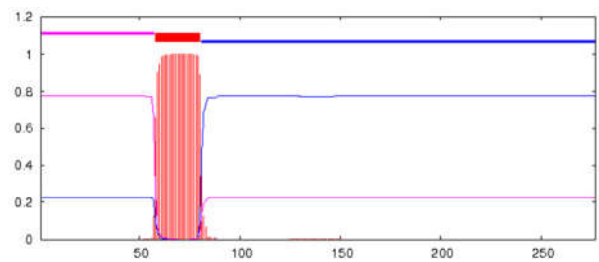
PTLD1b



PTLD2a



PTLD2b



**Supplemental Figure 27. Hydropathy profiles of the *Nicotiana tabacum* PTLD proteins.** Supporting figure 1. The profiles, based on the protein sequences (position of amino acids is given on the x-axis), were created by the TMHMM server v. 2.0 (Krogh et al., 2001). All proteins of this gene family carry a hydrophobic sequence predicted to be a transmembrane domain (indicated as red bars) in the N-terminal part of their peptide sequence. This hydrophobic sequence could be responsible for LD localization.

### 3. Manuscript II: Calculation of enrichment factors identifies low abundant lipid droplet-associated proteins from *Arabidopsis thaliana* seeds and seedlings

This is the status of the manuscript pre-submission. Supplemental figures are attached to the main part. The final, published manuscript including supplemental datasets and tables containing processed proteomics data are available online in the publication via <https://doi.org/10.1104/pp.19.01255>.

#### **Author contribution:**

Franziska K. Kretzschmar planned and performed the proteomics analysis with exception of the LC-MS/MS measurements. She processed and analyzed the data. She cloned, expressed and performed microscopy on the candidates proteins presented in Figure 9 A and B. She wrote this first draft of the manuscript.

# **Calculation of enrichment factors identifies low abundant lipid droplet-associated proteins from *Arabidopsis thaliana* seeds and seedlings**

Franziska K. Kretzschmar<sup>1</sup>, Nathan M. Doner<sup>2</sup>, Hannah Elisa Krawczyk<sup>1</sup>, Patricia Scholz<sup>1</sup>, Kerstin Schmitt<sup>3</sup>, Oliver Valerius<sup>3</sup>, Gerhard H. Braus<sup>3</sup>, Robert T. Mullen<sup>2</sup>, Till Ischebeck<sup>1,\*</sup>

<sup>1</sup>University of Goettingen, Albrecht-von-Haller-Institute for Plant Sciences, Department of Plant Biochemistry, Goettingen, Germany

<sup>2</sup>Department of Molecular and Cellular Biology, University of Guelph, Guelph, ON N1G 2W1, Canada

<sup>3</sup>University of Goettingen, Institute for Microbiology and Genetics, Department for Molecular Microbiology and Genetics, Goettingen, Germany

\*Author for correspondence: Till Ischebeck, Tel: +49 551 39 14430, Email: [tischeb@uni-goettingen.de](mailto:tischeb@uni-goettingen.de)

## **Abstract**

The developmental program of seed formation and seedling development requires not only tight regulation of cell division and metabolism but also the adaption of organelles in structure and function. Therefore, changes in organellar protein composition is one crucial factor in development. Of particular interest in plants is the switch to photoautotrophic growth, for which biosynthesis and degradation of lipid droplets (LDs) play a critical role. We present here a bottom-up proteomics study analyzing eight different developmental phases during silique development, seed germination and seedling establishment. We investigated both total protein fractions and LD-enriched fractions for each time point. The overall changes in the seed and seedling proteome during germination and seedling establishment monitored in this study present a rich resource for researchers interested in different questions of early seedling biology. The analysis of the proteome of LDs using LD-enrichment factors allowed the identification of four LD-associated protein families, which were subsequently confirmed by a cell biological approach. In addition to protein discovery, our dataset allows for the study of the dynamics of LD proteins throughout the developmental phases analyzed. We



found that the relative levels of oleosin stay stable, while many other proteins accumulate on LDs at later stages of seedling establishment. The methodology described here is shown to be well suited for describing a comprehensive and quantitative view of the Arabidopsis proteome across time, with a particular focus on proteins associated with LDs.

## Introduction

While the sporophyte of Arabidopsis as a whole is photoautotrophic during most of its life cycle, it is largely heterotrophic during its initial formation including embryo development and seedling establishment. During seed formation, the growing embryo is still protected by the mother plant, which provides it with nutrients for embryo growth and the accumulation of storage compounds crucial for seed germination and seedling establishment. Usually, three different phases of seed development are distinguished: embryogenesis, seed maturation including seed filling, and desiccation (Baud et al., 2002). During seed filling in Arabidopsis, a combination of storage proteins and lipids accumulate mostly in the embryo, with a minor proportion being deposited in the endosperm (Penfield et al., 2005). During late maturation, the seeds are prepared for a quiescent desiccated phase. Germination itself can again be defined as a three-step process: water intake (also called imbibition), reinitiation of metabolic processes, and radicle emergence that concludes germination *sensu stricto* (Nonogaki et al., 2010). Successful germination and seedling establishment is dependent on multiple factors, one of the most important ones being the ability to mobilize storage compounds. In Arabidopsis, the mobilization of storage lipids is important for successful germination. Arabidopsis dry seeds contain up to 40 % of their dry weight in lipids (Baud et al., 2002). Most of the lipids are triacylglycerols (TAGs) that are stored in lipid droplets (LDs). Numerous LDs with sizes between 0.5 to 2  $\mu\text{m}$  fill up to 60 % of the volume of a mature embryonic cell making these the most abundant organelles by volume together with the storage vacuoles (Kretzschmar et al., 2018; Mansfield and Briarty, 1992; Tzen et al., 1993). The degradation of the TAGs in the Arabidopsis seed during germination has been subject of extensive research, leading to a rather comprehensive picture of which components contribute how to germination success. In summary, TAGs are mobilized by the action of SUGAR-DEPENDENT LIPASE (SDP1 (Eastmond, 2004)) and SDP1-LIKE (Kelly et al., 2011). SDP1 localizes to the peroxisomal surface,

which is in close interaction with LDs during germination, possibly through peroxisomal extensions termed “peroxules” (Cui et al., 2016; Thazar-Poulot et al., 2015). The resulting fatty acids are metabolized in glyoxysomes, which are specialized peroxisomes and perform  $\beta$ -oxidation and the glyoxylate cycle in oil seeds during germination. Another step limiting germination success is the import of FAs into the glyoxysome. Several independent mutants of the ABC transporter in the glyoxysomal membrane have been identified (Footitt et al., 2002; Hayashi et al., 2002; Zolman et al., 2001). Products of the glyoxylate cycle are fed into gluconeogenesis which is essential for seedling growth (Penfield et al., 2005).

While the major function of LDs in seeds is energy storage, which is well understood, their roles in vegetative tissue and after seedling establishment are less explored. In recent years, through the discovery of new LD-associated proteins, the picture of the functional roles of LDs has increased in complexity (Pyc et al., 2017a). The first family of LD proteins discovered were the oleosins (Qu et al., 1986; Vance and Huang, 1987). They are a class of proteins predominantly found in higher plants, and more specifically, in tissues exposed to desiccation (Huang, 2017). Multiple studies have shown the importance of oleosins in LD formation and maintenance of seed vigor. The double knockout mutant *ole1 ole2* has strongly enlarged LDs and displays germination defects (Shimada et al., 2008; Siloto et al., 2006). The single mutants are more sensitive to freezing stress, which promotes unwanted LD coalescence, suggesting that oleosins enhance the integrity of LDs in desiccated seeds during winter. Steroleosins (also referred to as HYDROXYSTEROID DEHYDROGENASE, HSD) are homologs of metazoan sterol dehydrogenases (Lin et al., 2002). These enzymes mediate the homeostasis of the biologically active and less active forms of steroid-derived hormones (Chapman et al., 2012). While the exact substrate and products of steroleosins are still to be determined, plant mutants display a similar phenotype like plants impaired in brassinosteroid metabolism (Baud et al., 2009; Li et al., 2007). Caleosins are calcium-binding and heme-containing peroxygenases (Hanano et al., 2006; Næsted et al., 2000). They have been shown to produce hydroxylated fatty acids, which are important precursors for a large variety of secondary metabolites (Shimada et al., 2014). Eight caleosins genes are known in Arabidopsis, two of which have been shown to be seed specific (AtCLO1 and AtCLO2), and with two different isoforms being expressed in leaves (AtCLO3 and AtCLO4), the activity of caleosins seems to be of more systemic importance than oleosins (Chapman et al., 2012;

Hanano et al., 2006). Caleosins appear to work in close coordination with another LD protein:  $\alpha$ -dioxygenases ( $\alpha$ -DOX). These LD-proteins produce together the compound 2-hydroxy-octadecatrienoic acid that can act as a phytoalexine suggesting LDs as a production site for antimicrobial compounds within the plant cell (Shimada et al., 2014).

The discovery of the LD-associated protein (LDAP) family allowed the first insights into the function of LDs in vegetative tissues. Originally identified in the mesocarp of Avocado (Horn et al., 2013), LDAPs were later shown to be ubiquitously expressed in Arabidopsis tissues (Gidda et al., 2016). In leaves, LDAPs influence LD abundance, as was shown through overexpression and mutant studies. LDAP overexpression also confers resistance to drought stress (Kim et al., 2016). Since LD formation is induced under abiotic stress (Gidda et al., 2016; Pyc et al., 2017a; VanBuren et al., 2017) and the major structural protein oleosin has so far not been found in true leaves, LDAPs might play a role in LD formation and size. Through yeast-two-hybrid analysis, the LDAP-INTERACTING PROTEIN (LDIP) was discovered (Pyc et al., 2017b). While the physiological role of this protein remains to be determined, its ubiquitous expression, LD localization, and mutant phenotype suggest that it is important in LD biology. Recently, we investigated how the degradation of LD-associated proteins might be regulated. We discovered a member of the Plant UBX domain-containing (PUX) protein family found to be localized to tobacco pollen tube and Arabidopsis seed LDs (Deruyffelaere et al., 2018; Kretzschmar et al., 2018). Mutants of PUX10 showed a delay in LD protein breakdown during germination. This was in accordance with earlier findings that oleosins and caleosins are ubiquitinated (Hsiao and Tzen, 2011) and that oleosins are substrate for the Ubiquitin-Proteasome Pathway (Deruyffelaere et al., 2015). Thus, during germination, not only the lipid components of LDs are degraded but also the proteins associated with the organelle. How this protein degradation is regulated exactly in the fine regulatory network that mediates seed germination and seedling establishment is still to be determined.

Since the first complete genome sequence of Arabidopsis was released in 2000 (Arabidopsis Genome Initiative, 2000), seed formation and germination have been the topic of different proteomics studies. Often, two or more conditions are compared to each other, for example the influence of different hormones (Chibani et al., 2006; Li et al., 2016; Yin et al., 2015) or environmental conditions like salt or cold stress (Fercha et al., 2016; Lee et al., 2015; Xu et al., 2017). In addition, the influence of dormancy, seed ripening and seed aging were studied (Chibani et al., 2006; Gallardo et al., 2001;

Kubala et al., 2015; Nguyen et al., 2015; Yin et al., 2015). Other studies focus on specific post-translational modifications important during germination (Han, Wang, & Yang, 2014; Han, Yang, Sakata, & Komatsu, 2014; (Aguilar-Hernández et al., 2017; Aryal et al., 2015). In *Oryza sativa*, both seed formation and germination have been studied in extensive time course experiments (Han et al., 2014c; Lee and Koh, 2011; Li et al., 2016; Yang et al., 2007). Large studies of proteomic changes during seed formation and germination have also been performed for peanut (*Arachis hypogaea*, (Wang et al., 2016)), identifying more than 5500 proteins. Seed formation has been investigated in soybean (*Glycine max*; (Hajduch et al., 2005)), castor bean (*Ricinus communis*, (Houston et al., 2009)), and *Brassica napus* (Hajduch et al., 2006). In *Arabidopsis thaliana*, studies about the proteome in a collection of different tissues (Baerenfaller et al., 2008; Giavalisco et al., 2005) identify a large quantity of proteins in seeds and siliques. More detailed time course studies from the recent years focused on the seed development (Hajduch et al., 2010; Lorenz et al., 2018), and on different time points or conditions during germination (Durand et al., 2019; Quan et al., 2013), or after seedling establishment (Kim et al., 2013; De Michele et al., 2016). However, those studies are not as extensive as the ones existing for seed development. None of the studies offers a comprehensive proteome of any one specific organelle.

We present here an extensive time course proteomic study of two time points during silique development, and six time points during seed germination and seedling establishment, using a state-of-the-art proteomics platform. Additionally, we provide a proteome of a LD-enriched fraction for each of those time points and present four new protein families confirmed to be associated with LDs by fluorescence microscopy. The dynamics of LD protein composition during silique development and germination was monitored for 30 proteins from 14 protein families. Our proteomic data on LD-enriched proteins serve to enhance our understanding of the LD proteome in plants. By assessing different stages of seed development and germination, we provide quantitative temporal data that will inform further research aimed at better understanding these processes, and the roles that are played by LDs.

## Results

### **A large proteomics data set spanning seed development to seedling establishment was generated**

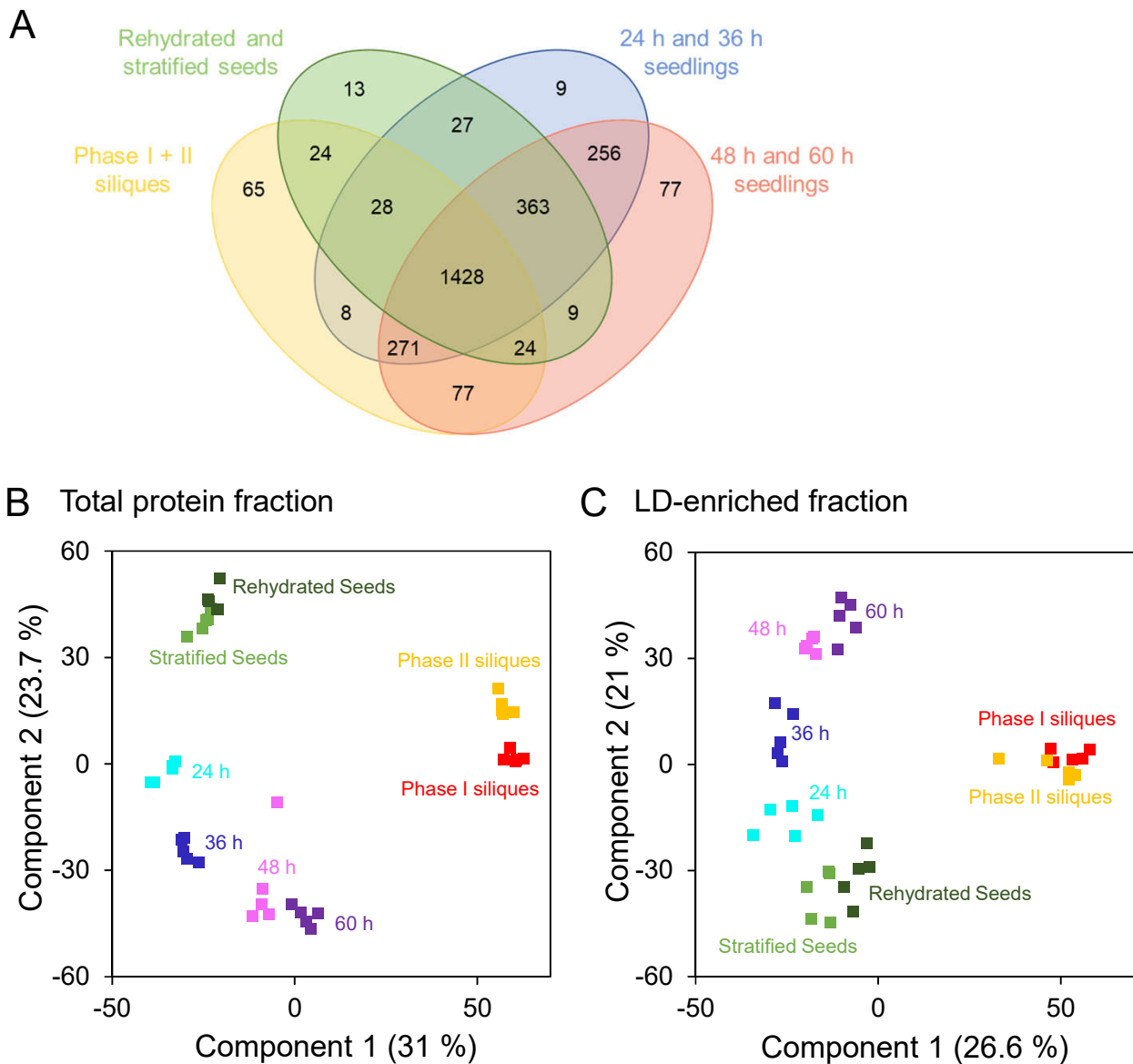
We have used label-free Orbitrap mass spectrometry to investigate *Arabidopsis thaliana* silique development, seed germination, and seedling establishment.

For this, proteins were isolated from two phases of silique growth, corresponding roughly to seed maturation and desiccation as defined in (Baud et al., 2002). However, no yellowing siliques were used in the experiment. Additionally, seed germination and seedling establishment was monitored over six time points: 30 min after seed rehydration, directly after 74 h stratification at 4°C in the dark, and at four time points during post-germinative growth under long-day conditions. After 24 h in long-day condition at 22 °C, seeds had completed germination *sensu stricto*, defined by radicle emergence. Within the next 24 h, cotyledons emerged and opened, and after 60 h, the cotyledons had grown and were dark green (Suppl. Figure 1).

For each time point, total protein and LD-enriched fractions were generated in five biological replicates. The LD-enriched fractions were generated from homogenized biological material without the addition of harsh chemical conditions. This approach was employed to avoid the removal of proteins loosely interacting with the LDs. Proteins were tryptically digested in-gel and analyzed on a Thermo Fisher Scientific Q Exactive HF Orbitrap mass spectrometer.

The mass spectrometral raw data files were processed in MaxQuant (for more information see Methods section, Suppl. Datasets 1 – 4, and Suppl. Table 1). The label-free quantification package MaxLFQ was used, and both the LFQ and iBAQ intensities determined. iBAQ values were calculated as ‰ of all intensities in each sample (relative iBAQ or riBAQ, relative LFQ or rLFQ), unless stated otherwise. When comparing and relating LD-enriched fractions to total protein fractions, riBAQ values were used, as this algorithm picks up smaller values more often. For quantitative comparison of the total proteome of different time points, rLFQ values are shown as these display smaller variations between replicates in our hands.

In total, we were able to detect 2696 protein groups that were identified by at least two peptides with the iBAQ data processing algorithm (Table 1). The number of protein groups in the total proteome is lowest in older siliques and rehydrated seeds, and



**Figure 1. Graphical representation of proteomic data derived from siliques, seeds and seedlings.** Proteins of two developmental stages of silique development (phase I, 7-14 d after fertilization and phase II, 14-21 d after fertilization) and six stages of seedling establishment (rehydrated and stratified seeds, and seedlings 24 h to 60 h after stratification). Both a total protein fraction and an LD-enriched fraction were obtained. Peptides were examined by LC-MS/MS after tryptic digest of the proteins. Data was processed by MaxQuant.  $n=5$  per stage and fraction. **A:** Venn diagram of the distribution of the all detected proteins across different time points. Many proteins are shared between all samples, while others are unique for the only one condition. Proteins of both total and LD-enriched fraction identified with at least two peptides with the iBAQ algorithm were grouped into 4 groups as depicted. Overlapping proteins of the groups were identified via InteractiVenn (Heberle et al., 2015). **B:** Principal component analysis (PCA) comparing the distribution of biological replicates within the total protein fraction of all time points. LFQ processed data was analyzed in Perseus. After initial filtering for potential contaminants, reverse peptides, and site-based identification, the data was log2-transformed and filtered for 5 valid values in at least one of the time points. After imputation of missing values by normal distribution (width: 0.9, down shift: 1.8), PCA was performed with Perseus default settings. Projections were exported and the graph for Component 1 vs. Component 2 was recreated in Excel. **C:** Principal component analysis comparing the distribution of biological replicates within the LD-enriched fraction of all time points. LFQ processed data of the LD-enriched fraction was analyzed like for the total protein fraction, with adapted imputation parameters (width: 0.9, down shift: 2.1).

highest during the late stages of seedling establishment. The complexity of the LD-enriched fractions behaved in a similar manner. Within the LD-enriched fraction, known LD-associated protein families (and families first described in this study) make up between 17 % (in the silique samples) and more than 30 % (in seeds and 60 h seedlings; Table 1, Suppl. Dataset 19), indicating contaminations by proteins that do not localize to LDs.

More than half of the proteins detected were shared between all conditions and samples (Figure 1A; Heberle et al., 2015). Other proteins were unique for one developmental phase or another.

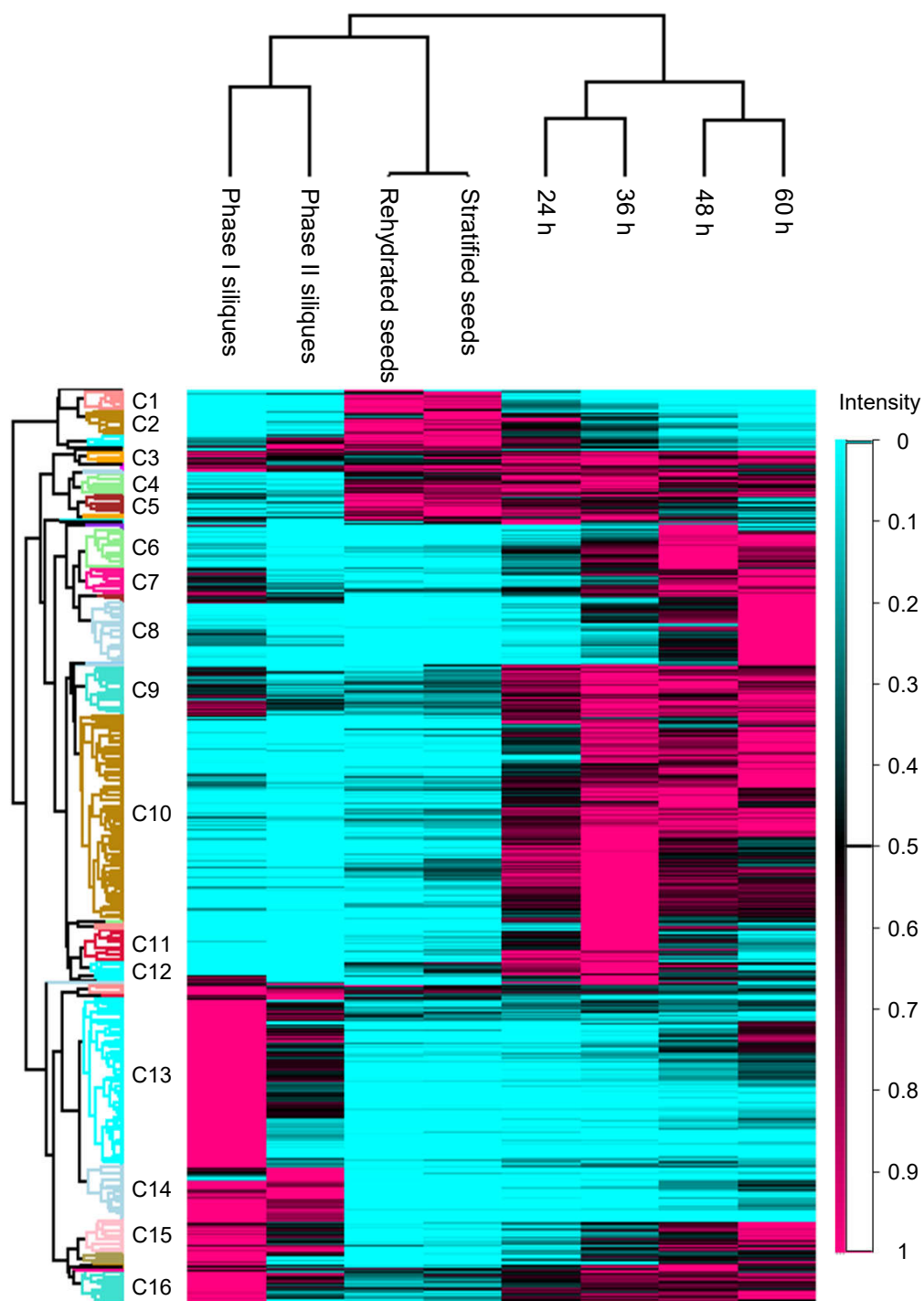
The five biological replicates of each time point reproducibly cluster together in the principal component analysis (PCA, Figure 1B, C, Suppl. Datasets 5 – 12). For this analysis, only proteins were taken into account that were found in all replicates of at least one of the stages, and the data is based on rLFQ values. The PCA plots also underline the unique proteome of siliques compared to all the other samples, as these are clearly separated by component 1. This observation can be made for both the total protein fractions and the LD-enriched fractions. However, the distribution along component 2 indicates in both cases that the younger siliques are more similar to the greenest seedlings (60 h seedlings), while the older siliques are more similar to seeds (rehydrated or stratified seeds). Along component 2, the seeds and seedlings are distributed in a time-dependent manner.

### **Seed and seedling establishment-specific proteins were identified in the total proteome**

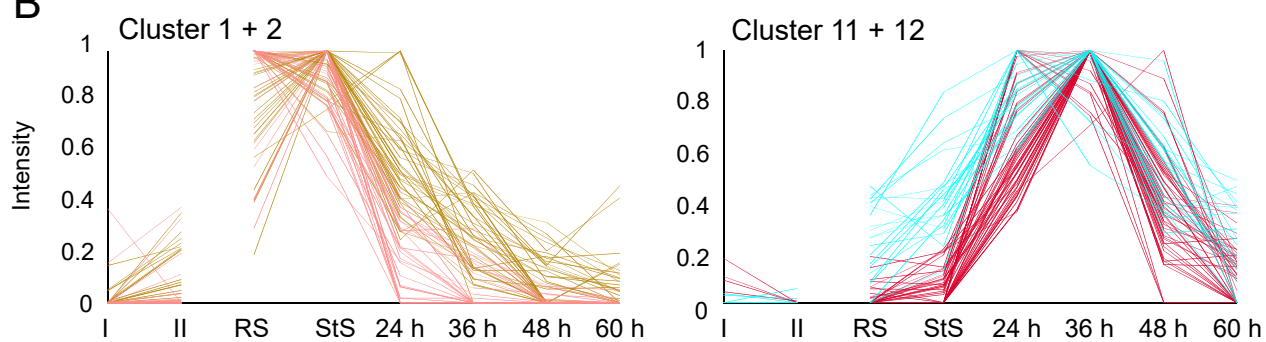
One objective for the creation of an extensive proteomics dataset for *Arabidopsis* germination and post-germinative growth was to identify growth stage-specific proteins for those critical periods in the plant's life cycle such as the desiccation phase and the switch from hetero- to photoautotrophic growth. Therefore, we analyzed our total proteome data for proteins that were increased in abundance during those developmental phases.

LFQ-processed total protein data were first filtered for four valid values in at least one of the eight different time points investigated. Then, the maximum intensity for each protein was set to one, and the intensities during the other time points were calculated as a proportion of this maximum value (Suppl. Datasets 11 and 12). This data matrix

A



B





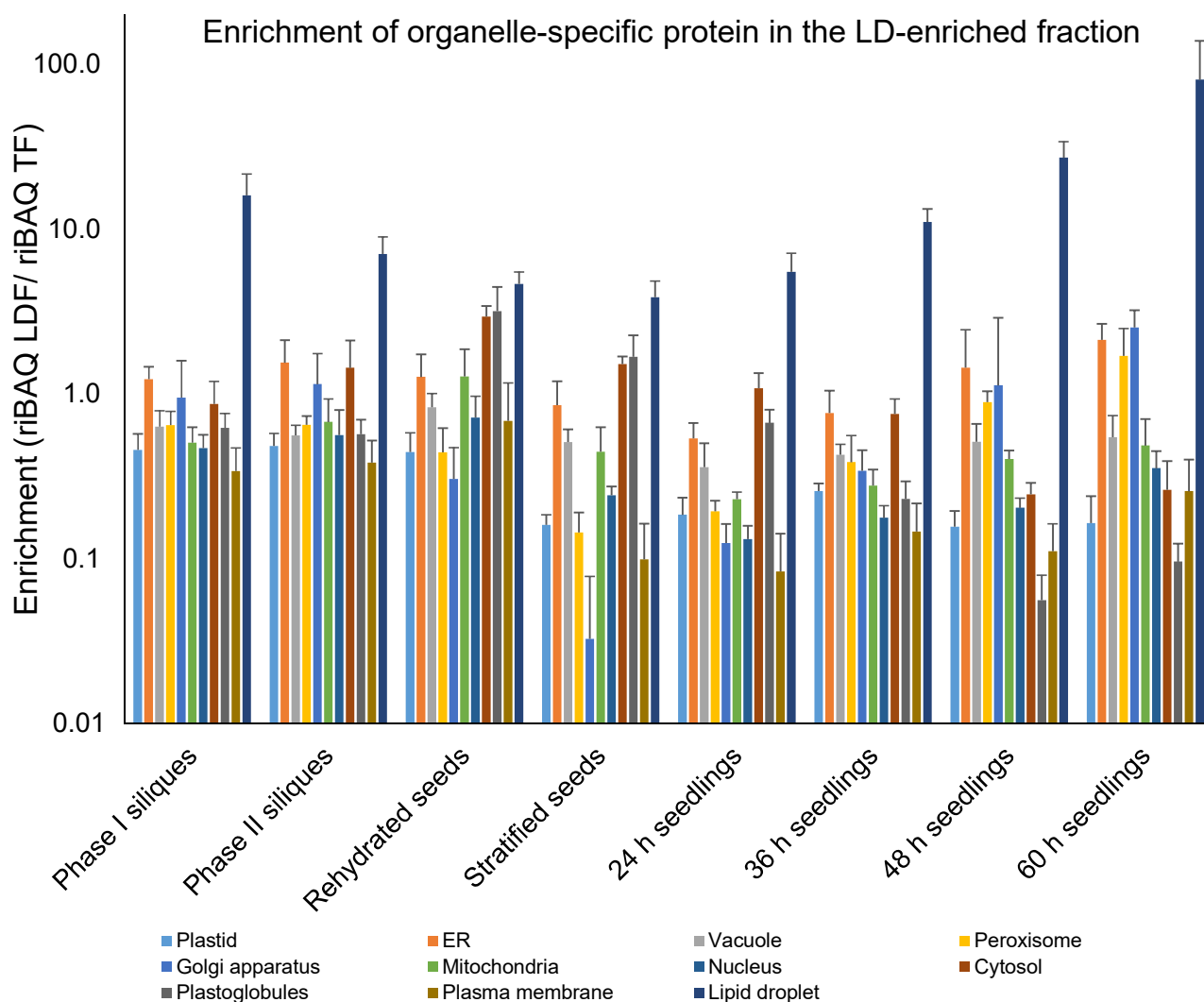
**Figure 2. Identification of protein expression clusters in total protein fractions. A:** Hierarchical clustering analysis of changes in protein abundance over the time points analyzed. LFQ-processed data was filtered for at least 4 valid values in at least one of the eight time points. For each protein, the maximum value was set to one, and the remaining time points were calculated as fractions of that value. The resulting data matrix was used for hierarchical clustering in Perseus 1.6.2.2 with Euclidean Distances and preprocessing with k-means. In the following, 40 row clusters were defined. Clusters harbouring more than 20 proteins are marked in the figure with labels. Several clusters of particular interest for this study could be identified. For example, in Clusters 1 and 2, seed-specific proteins are found. Clusters 11 and 12 contain proteins present predominantly during early stages of seedling establishment. **B:** Graphs showing the expression changes of all proteins collected in Clusters 1 and 2 (left graph) and Clusters 11 and 12 (right graph). Defining for the clusters is the maximum of protein expression achieved during seed stages, or after 24 h and 36 h of germination, in Clusters 1 and 2, and Cluster 11 and 12, respectively. I – Phase I siliques; II – Phase II siliques, RS – rehydrated seeds, StS – stratified seeds, 24 h to 60 h – seedlings 24h to 60 h after stratification.

(Suppl. Dataset 12) was used for hierarchical clustering in Perseus 1.6.2.2 using Euclidean distances and preprocessing with k-means. In the resulting heat map, 40 clusters were defined. Clusters containing more than 20 proteins are labeled in Figure 2A. Four of these clusters contain proteins that have highest intensities either during seed stages (Cluster 1 and Cluster 2, see Suppl. Dataset 12 for all the proteins sorted by cluster) or during early seedling establishment (Cluster 11 and Cluster 12). Cluster 1 and 2 together contain 71 proteins, Cluster 11 and 12, 70 proteins. In Cluster 1 and 2, we find late embryogenesis abundant proteins, two cell wall modifying enzymes and proteins of unknown function. While seed storage proteins like CRUCIFERIN2 and CRUCIFERIN3 also have their highest intensities during these stages, their degradation is slower than that of the proteins presented in Clusters 1 and 2. In Clusters 11 and 12, we find proteins involved in  $\beta$ -oxidation like LACS7 (LONG-CHAIN ACYL-COA SYNTHETASE 7), ACX1 (ACYL-COA OXIDASE 1) and the enoyl-CoA hydratase AIM1 (ABNORMAL INFLORESCENCE MERISTEM). Furthermore, other enzymes of lipid metabolism like CYCLOARTENOL SYNTHASE and OIL BODY LIPASE1 cluster with these enzymes. Nucleotide synthesis enzymes and proteins involved in vesicle transport are also found. Proteins of the glyoxylate cycle, like the peroxisomal NAD-MALATE DEHYDROGENASE 1 or MALATE SYNTHASE are grouped in Cluster 10. They too, have their highest intensity after 36 h of germination but remain present during later stages of seedling establishment.

We also employed GO Term analysis on our complete data set (Figure 3, Suppl. Datasets 13 – 15), by assigning all the proteins found in our study to all the GO terms they are associated with. Then, rLFQ intensities of all proteins assigned to each GO Term were added up. In the following, changes over the course of development could be identified. For example, fatty acid biosynthesis and degradation proteins have high abundance during silique development and seed germination, respectively. Proteins associated with the GO Term “nutrient reservoir activity” are most abundant in the seed stages, correlating the mobilization of protein storage molecules during germination. Proteins involved in processes that have been shown to be involved in seedling establishment are indeed upregulated during this phase, such as  $\beta$ -oxidation, glyoxylate cycle, and proteasomal degradation. As the abundance of photosynthesis-related proteins (e.g. RuBisCo subunits) strongly increases between the 48 h and 60 h seedlings, it is possible that these proteins are still increasing at later time points.

GO term selection		n TF+LDF	n TF	n total	Phase I silique		Phase II siliques	Rehydrated seeds		Stratified seeds	24 h seedlings	36 h seedlings	48 h seedlings	60 h seedlings
fatty acid biosynthetic process	GO:0006633	37	33	94	4.67	2.63	2.25	1.91	1.94	2.46	2.71	2.82		
	fatty acid biosynthetic process													
	GO:0006635	20	18	40	0.89	0.38	2.28	2.09	7.76	8.85	8.21	6.17		
	fatty acid beta-oxidation													
lipid transport	GO:0006869	23	22	147	6.98	9.41	4.57	4.74	2.03	1.13	1.38	0.35		
	lipid transport													
	GO:0006886	70	51	244	4.71	2.16	0.70	0.90	3.57	4.34	2.35	3.18		
	intracellular protein transport													
nutrient reservoir activity	GO:0045735	25	25	68	96.69	302.20	520.79	503.28	370.20	233.73	127.35	65.68		
	nutrient reservoir activity													
	GO:0006097	6	6	7	1.51	0.87	2.88	5.86	36.75	49.58	31.66	19.99		
	glyoxylate cycle													
gluconeogenesis	GO:0006094	19	19	22	65.23	37.48	15.45	16.38	24.63	33.06	40.03	41.06		
	gluconeogenesis													
	GO:0019253	15	15	19	137.83	81.53	6.26	5.92	15.80	65.41	166.64	207.54		
	reductive pentose-phosphate cycle													
protein catabolic process	GO:0030163	29	27	96	9.06	7.70	5.13	7.21	7.92	7.14	5.58	6.76		
	protein catabolic process													
	GO:0022626	149	146	194	133.20	75.73	18.00	24.54	81.50	157.63	263.70	287.38		
	cytosolic ribosome													
glycolytic process	GO:0006096	46	46	69	97.64	56.33	19.98	21.01	28.51	39.49	55.39	60.71		
	glycolytic process													
	GO:0000502	46	45	81	4.25	1.87	1.01	3.16	6.06	7.29	4.94	4.84		
	proteasome complex													
lipid storage	GO:0019915	10	9	27	4.65	12.10	10.67	12.54	7.29	4.62	2.54	1.24		
	lipid storage													
	GO:0006897	10	5	56	0	0	0.04	0.13	1.11	1.42	0.40	0.21		
	endocytosis													
tricarboxylic acid cycle	GO:0006099	34	34	54	6.57	3.92	7.50	10.61	47.45	63.50	44.26	31.56		
	tricarboxylic acid cycle													
	GO:0015995	30	30	43	1.44	0.43	0.10	0.11	2.24	8.79	14.64	16.43		
	chlorophyll biosynthetic process													
mRNA processing	GO:0006397	26	18	174	1.11	0.42	0.26	0.34	1.53	2.46	2.22	4.06		
	mRNA processing													
	GO:0043161	36	35	97	3.73	1.78	0.82	2.36	4.25	4.87	3.26	3.23		
	proteasome-mediated ubiquitin-dependent protein catabolic process													
cell wall organization	GO:0071555	29	24	288	3.44	1.69	1.32	1.48	3.00	4.85	3.22	2.26		
	cell wall organization													
	GO:0015979	66	65	124	130.62	87.24	1.04	1.09	4.46	14.17	36.39	67.20		
	photosynthesis													
starch biosynthetic process	GO:0019252	14	12	33	5.25	2.39	0.23	0.42	0.75	1.32	1.77	2.67		
	starch biosynthetic process													
	GO:0000278	15	12	85	1.97	0.41	0.43	0.28	5.20	9.70	9.82	10.10		
	mitotic cell cycle													
Cul4-RING E3 ubiquitin ligase complex	GO:0080008	10	7	124	0.03	0	0.03	0.02	0.60	0.81	0.36	0.44		
	Cul4-RING E3 ubiquitin ligase complex													

**Figure 3. Changes in protein intensity of functional groups in the total protein fraction.** LFQ-processed data were added in a GO Term-dependent manner. GO Term annotations were acquired from TAIR. “n total” corresponds to the complete number of genes assigned to one group, “n TF” to the amount of proteins within this group that were detected in this study in the total protein fraction, “n TF+LDF” to the corresponding number of proteins detected in both analyzed fractions of the study. Darker coloring represents higher intensities compared to the other phases. The figure displays a selection of GO Terms. The complete list and respective standard deviations can be found in Supplemental Datasets 15-17. TF – total protein fraction, LDF – lipid droplet-enriched fraction.



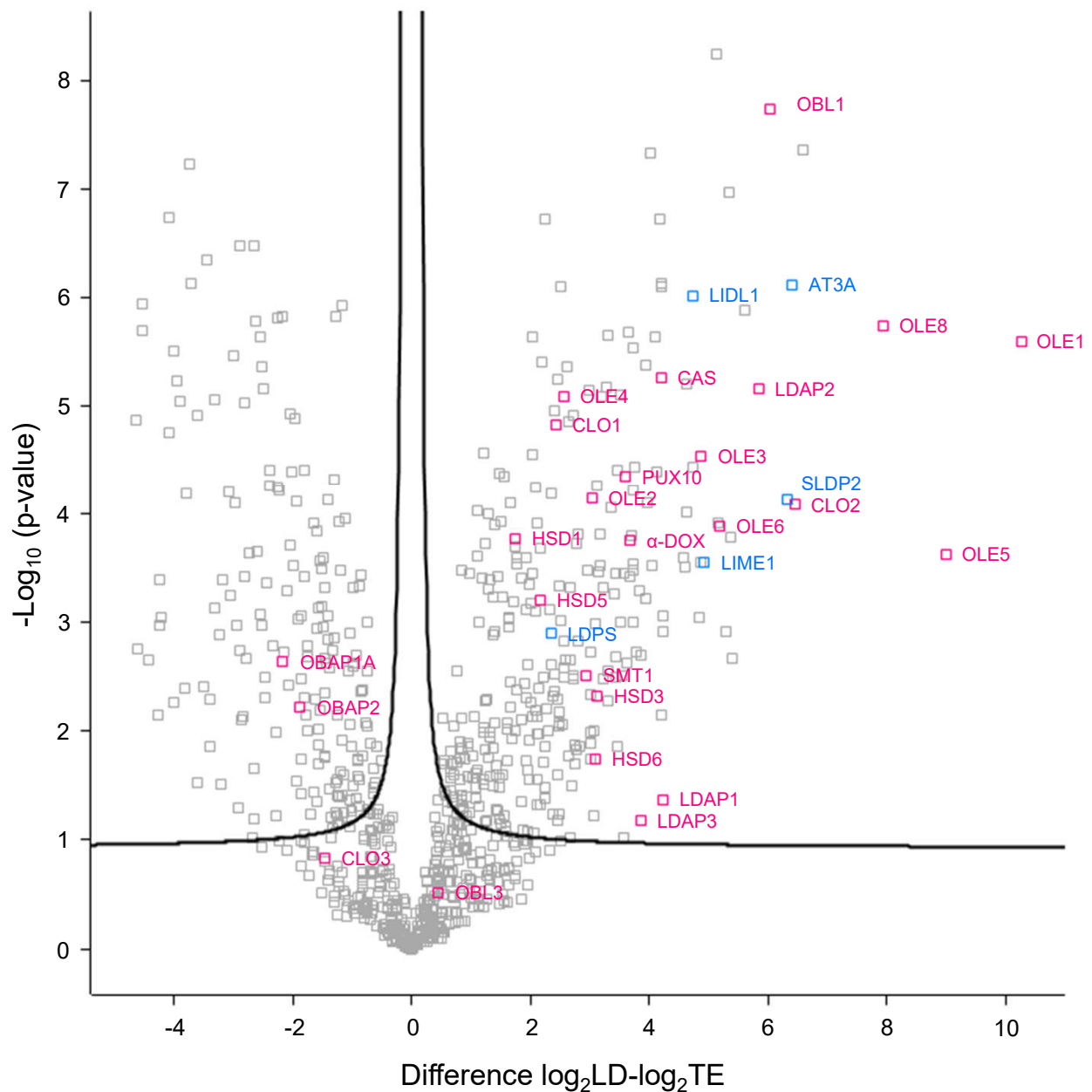
**Figure 4. Enrichment of organelle-specific proteins in the LD-enriched fraction.** Proteins were combined into eleven subcellular localization groups. The riBAQ intensities of all the proteins in one group were added for each stage and subcellular fraction. These sums were then used to calculate enrichment factors between the LD-enriched fraction (LDF) and the total protein fraction (TF) for each stage. Annotations of protein localization were obtained from the Plant Proteome Database (Sun et al., 2009). LD proteins were annotated manually taking into account the results of this work.

Furthermore, we analyzed the subcellular origin of the proteins. For this, we used the proteomics-confirmed annotation data set of the Plant Proteome Database (Sun et al., 2009). This dataset contains 78 different subcellular annotations that we summed up to ten different origins: plastid, ER, vacuole, peroxisome, golgi apparatus, mitochondria, nucleus, cytoplasm, plasma membrane, and plastoglobules. LD annotation was performed by us based on previous studies and taking into account LD-associated proteins identified in this study. The strongest changes in abundance can be observed for the proteins annotated as plastidial (Table 2, Suppl. Datasets 16 – 18). Their abundance changes tenfold between siliques and seeds, and seeds and 60 h seedlings. This curve shows the opposite trend compared with the changes in LD protein abundance, which is highest during the seed stages. The abundance of LD proteins within the total protein fraction correlates with the formation and degradation of LDs during seed development and germination, respectively. It doubles from younger to older siliques, is highest in rehydrated and stratified seeds and decreases during the next time points, approximately by a factor of 2 every twelve hours (Table 2, Suppl. Figure 2).

### **The calculation of enrichment factors allows the identification of LD-associated proteins**

The second objective of this work was to identify new protein families associated with LDs. The future in-depth study of these proteins could aid further understanding of LD functions, biogenesis, and breakdown.

Our approach of LD enrichment avoids the use of harsh chemicals or extensive washing of the LDs. While we hope to preserve loose interactions on the LDs that might yield interesting new discoveries, we also generate a large data set that likely still contains many contaminants. As shown in Table 1, more than 1000 proteins were identified in each of our LD-enriched fractions, and only between 16 and 34 percent of the protein intensities in those fractions originated from LD-associated proteins. Nevertheless, LD-associated proteins should be of higher abundance in the LD-enriched fraction in comparison to the total protein fraction, but we had to make sure that no other subcellular compartment was co-purified with the LDs. Therefore, we calculated the enrichment factor of proteins from different subcellular compartments at each time point (Figure 4, Suppl. Dataset 19). This indicates that while some organelles seem to be co-purified with LDs, the enrichment factor was always highest



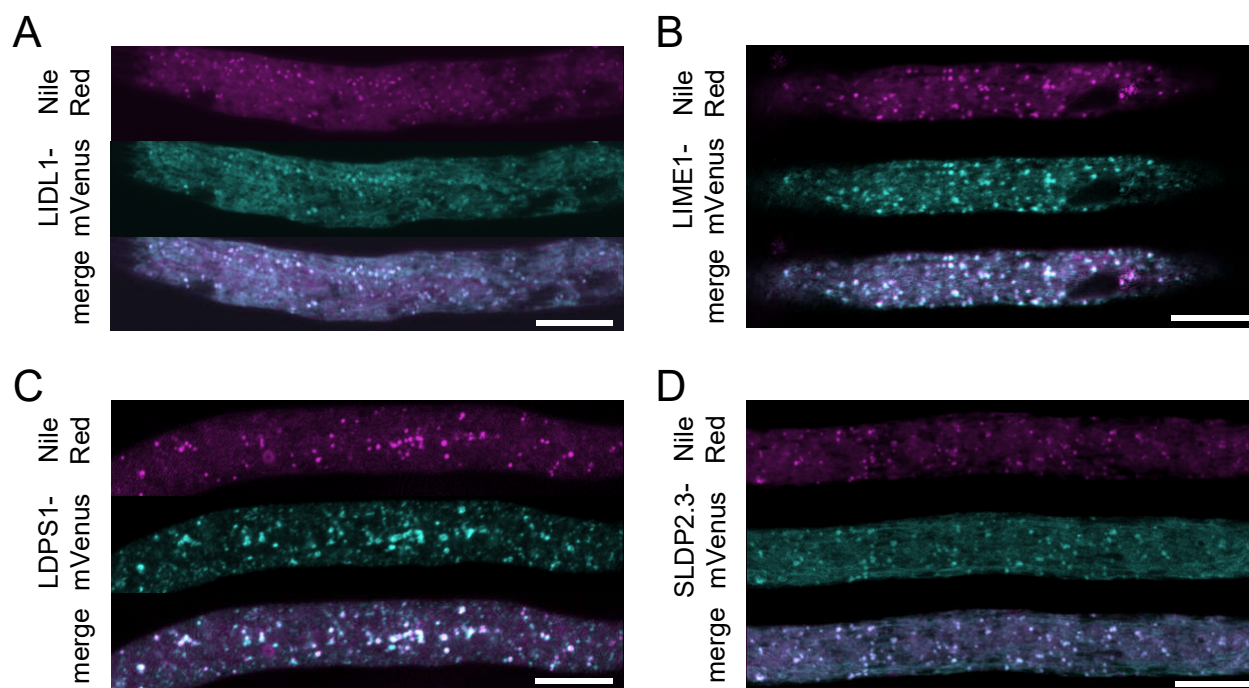
**Figure 5. Enrichment analysis of proteins in the LD-enriched fraction.** For each protein, the time point with the highest riBAQ intensity in the LD-enriched fraction (LDF) was chosen for this analysis, if the average riBAQ intensity at that stage was higher than 0.1. Then, the data was  $\log_2$ -transformed and a Vulcano plot created in Perseus with default parameters. For this, the fold change of a protein in the LDF compared to the TF was plotted against the respective p-value. Significantly depleted proteins are grouped in the left side of the plot, significantly enriched proteins on the right side. Proteins chosen for analysis of subcellular localization are marked in blue, previously known LD proteins are marked in magenta. LDF – Lipid droplet-enriched fraction, TF – total protein fraction.

for LDs (ranging from 3.9 to 80.8), while other subcellular compartments were only in some cases co-purified with LDs. The highest enrichment of any other organelle was seen in the LD-enriched fraction of rehydrated seeds, where plastoglobular and cytosolic proteins reached factors of 3.2 and 3.0, respectively.

To identify new LD protein candidates within the extensive proteome of the LD-enriched fraction, we used again enrichment factors and tested the enrichment of each protein for statistical significance. For each protein, the stage where its abundance was highest in the LD-enriched fraction was chosen for further analysis. We also only considered proteins that were identified in at least 4 of 5 replicates in this stage. The volcano plot in Figure 5 graphically represents the results of this analysis. On the right side of the plot, all proteins significantly enriched in the LD-enriched fraction are found, a total of 291 proteins. Most previously identified LD proteins are found in this area (marked in magenta). We studied the subcellular localization of a selection of promising candidates (marked in blue). Candidates were partially chosen because of their strong enrichment and high p-value, and partially due to promising features like the presence of hydrophobic regions in the peptide sequence. Other criteria for selection included interesting annotations either as unknown proteins, like for SEED LIPID DROPLET-ASSOCIATED PROTEINs (SLDPs), or as proteins potentially involved in either metabolism (like LIPID DROPLET-ASSOCIATED METHYLTRANSFERASE (LIME) and LIPID DROPLET-ASSOCIATED LIPASE (LIDL).

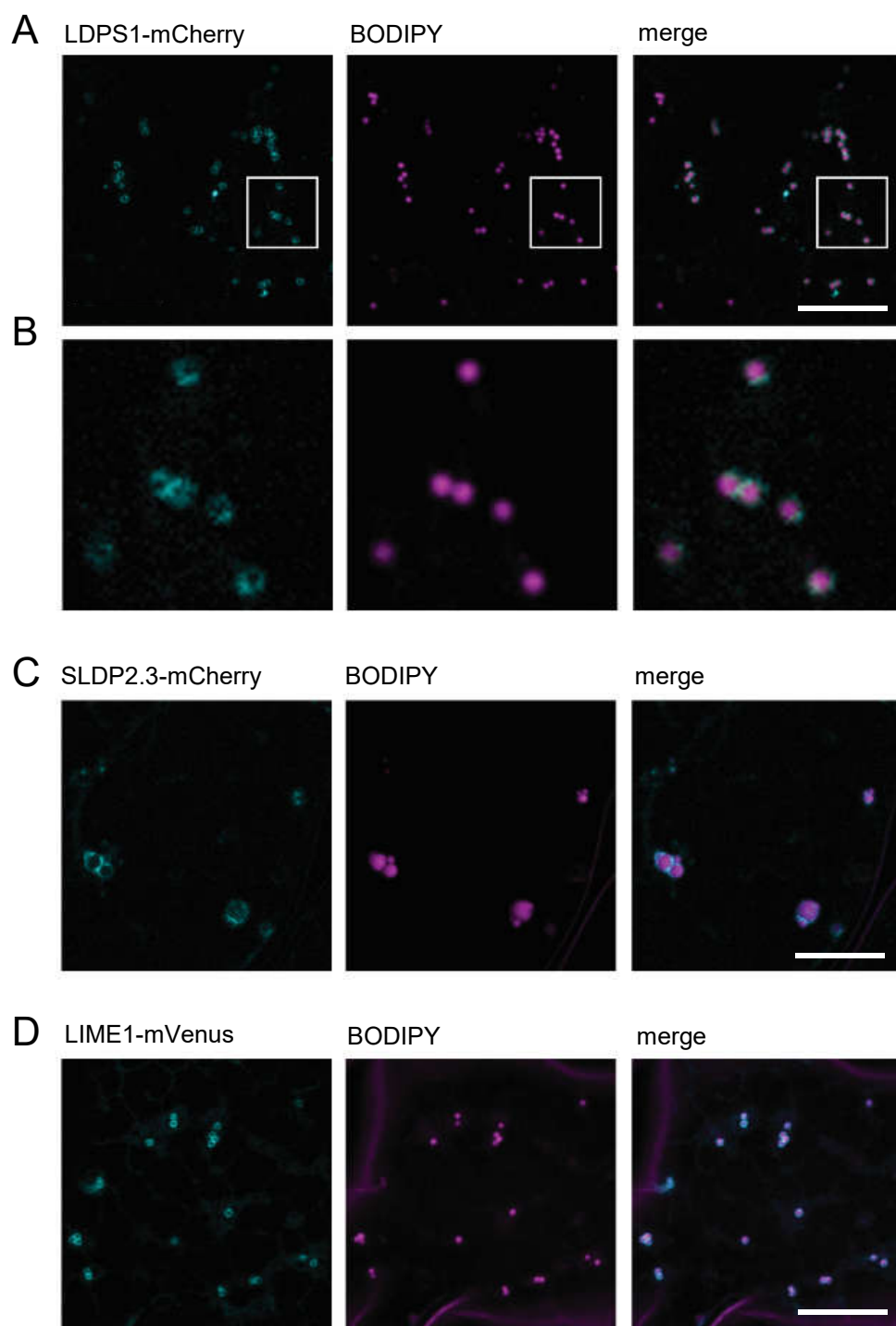
The candidates were then studied in regard to their subcellular localization using two independent cell biological systems, *Nicotiana tabacum* pollen tubes transformed by particle bombardment and *Nicotiana benthamiana* leaves transformed by *Agrobacterium tumefaciens* infiltration. For both cases, the genes were cloned as fluorescent fusion proteins, transiently expressed, and studied by confocal laser scanning microscopy. LD localization was confirmed by staining of LDs with Nile Red or BODIPY 493/503. ER localization was investigated with co-expression of ERD2-CFP.

The LD localization of four proteins could be confirmed with both transient expression systems (Figure 6 and Figure 7 for pollen tubes and leaves, respectively). We termed these proteins SLDP2, LIME1, LIDL1, and LDPS1. Another protein, though promising candidate marked by a highly significant, strong enrichment, did not localize to LD in the cell biological assays. For ATS3A, a punctuate localization could be observed in

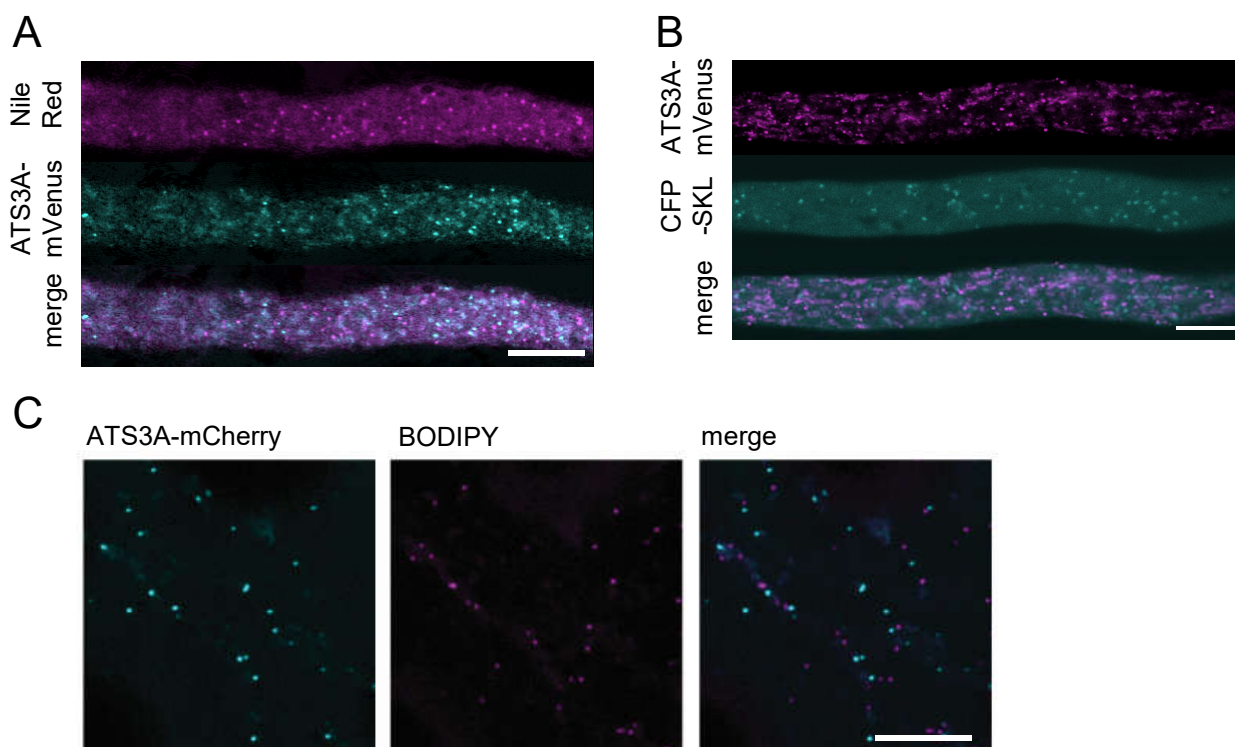


**Figure 6. Subcellular localization of LD protein candidates in *N. tabacum* pollen tubes.** Candidate proteins N-terminally fused to mVenus were transiently expressed in *Nicotiana tabacum* pollen tubes (cyan channel). LD were stained with Nile Red (magenta channel). In the merge channel, colocalization events appear white. Bar corresponds to 10  $\mu\text{m}$ .





**Figure 7. Subcellular localization of LD protein candidates in *N. benthamiana* leaves.** Candidate proteins N-terminally fused to mCherry were transiently expressed in *Nicotiana benthamiana* leaves (cyan channel). LD were stained with Bodipy (magenta channel). In the merge channel, the ring-like fluorescence of the protein around the BODIPY stain can be observed. B is a magnified version of A. Bar corresponds to 20  $\mu\text{m}$ .



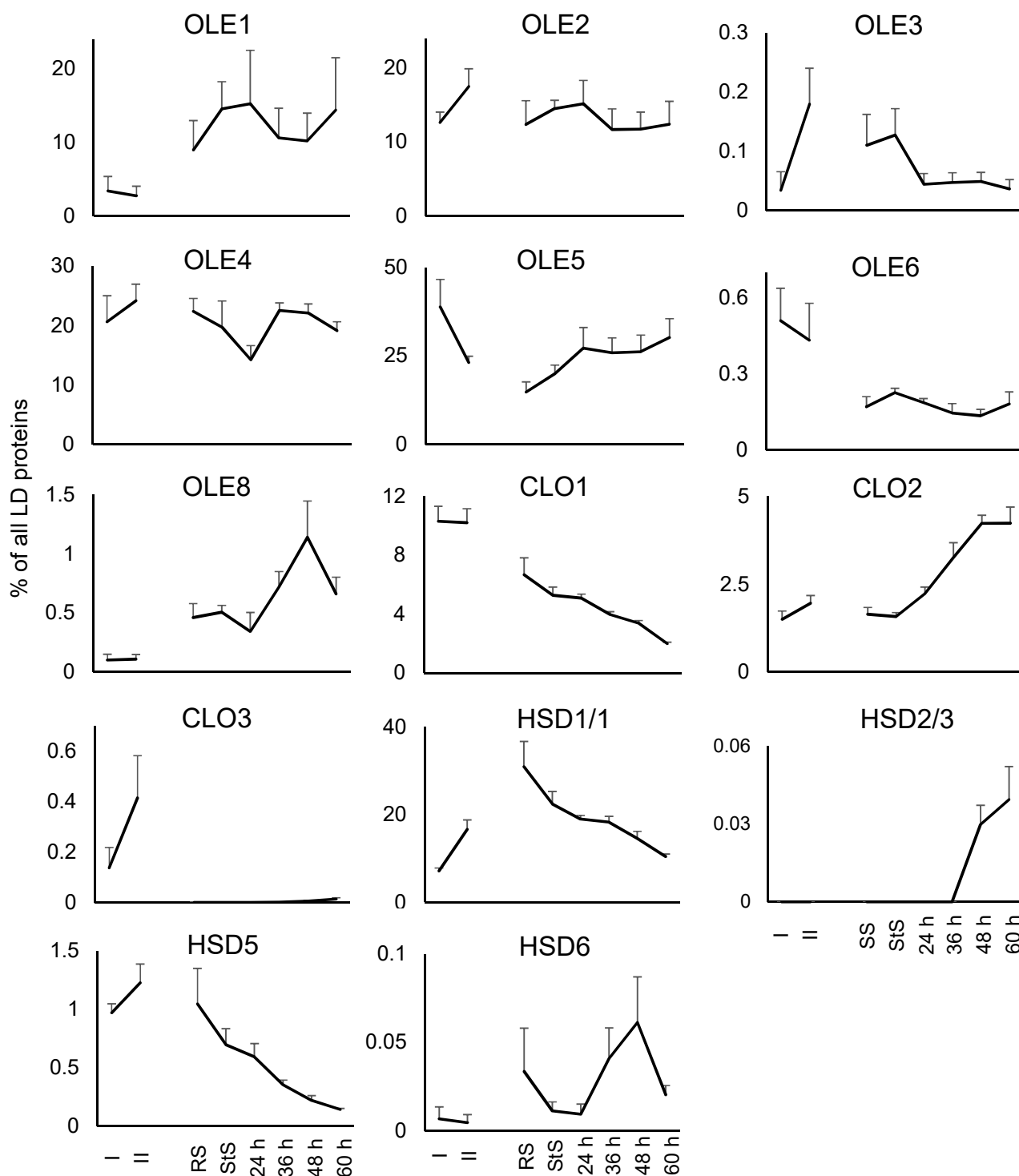
**Figure 8. Subcellular localization of LD protein candidate AT3A.** AT3A was N-terminally fused to mVenus (A, B) or mCherry (C) and transiently expressed in *Nicotiana tabacum* pollen tubes (A, B), or *Nicotiana benthamiana* leaves (C) (cyan channel). LD were stained with Nile Red (A) or Bodipy (C, magenta channel). As no colocalization was observed in either system, we coexpressed AT3A with a peroxisomal marker (B, in cyan), but no colocalization was observed in this case, either. Bar corresponds to 10  $\mu$ m (A, B) or 20  $\mu$ m (C).

both pollen tubes and leaves after transient expression (Figure 9). However, this punctuate structures did not colocalize with LD stainings, and also not with a peroxisomal marker.

A protein family, called oil body-associated proteins, have been described to be LD localized, using a maize isoform (Lopez-Ribera et al., 2014). However, in Arabidopsis, we continuously observed a depletion of these proteins in our LD fraction during different proteomics experiments (Kretzschmar et al., 2018; Pyc et al., 2017b). Also in this study, two proteins of the OBAP family are only slightly enriched in the LD-enriched fraction during silique stage, and strongly depleted during the other phases investigated (Suppl. Figure 3B and C). When transiently expressed as either C- or N-terminally mVenus fusion in tobacco pollen tubes, the protein OBAP1A shows a homogenous fluorescence in the pollen tube, indicating no specific membrane or organellar association (Suppl. Figure 3A).

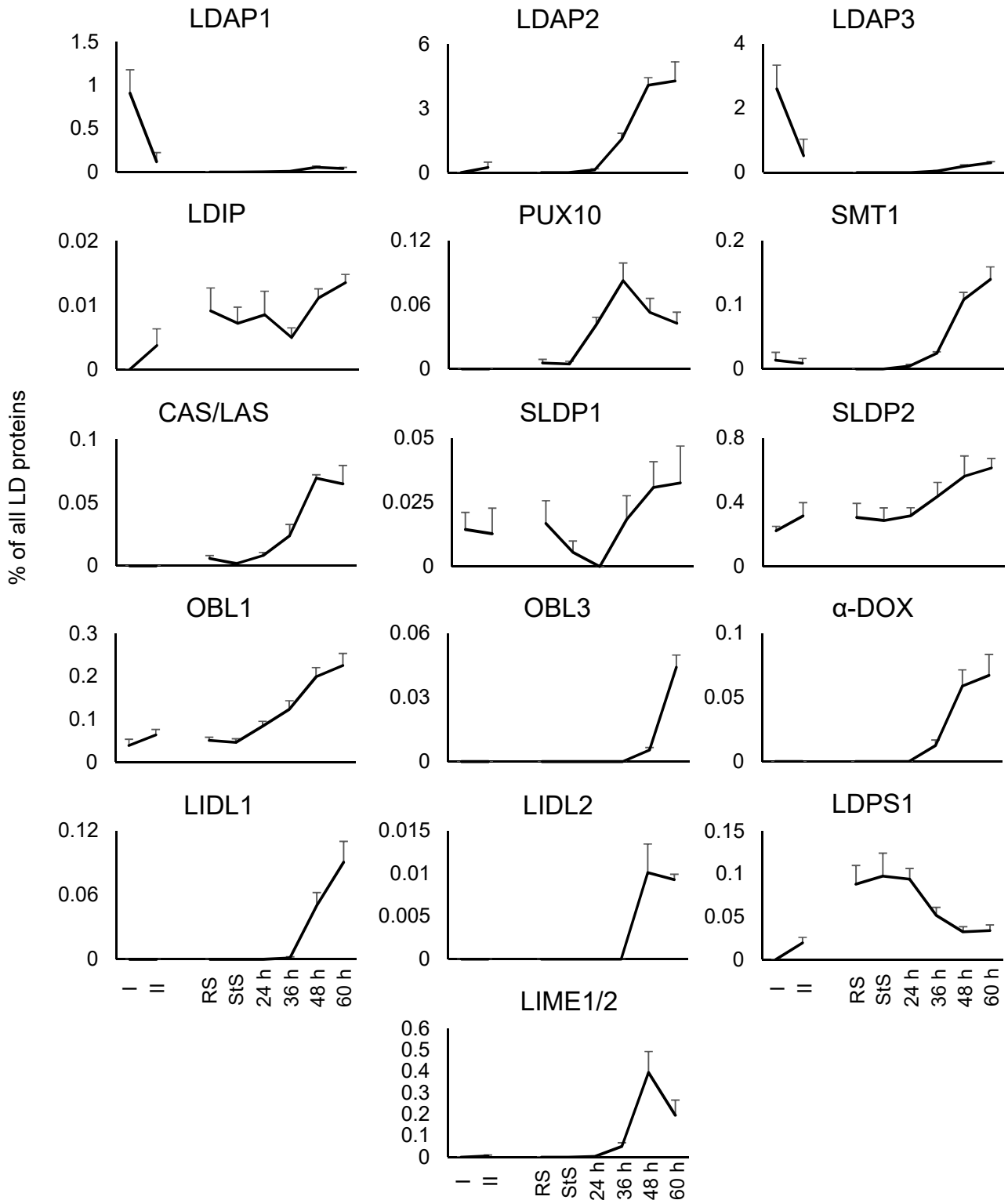
### **The lipid droplet proteome changes during seedling establishment**

In addition to the discovery of potential new LD proteins, the extensive LD proteome data collected highlighted the dynamics of the LD protein composition. As presented in Table 1, the abundance of LD proteins in the total proteome decreases after germination. In order to assess the changes of the composition of the LD proteome during development, the abundance of known LD-associated proteins was added in each stage, and the individual fraction for each protein was calculated (Suppl. Dataset 20). Oleosins make up a rather constant proportion of the LD proteome throughout seedling establishment both on an individual and total level (Figure 9, Suppl. Figure 4). The abundance of the major caleosin, CLO1, and the major steroleosin, HSD1, on the other hand, continuously decreases during the same time, while CLO2 and HSD2/3, increase. In general, more than 85 % of the LD proteome during all time points investigated is made up of only six proteins: OLE1, OLE2, OLE4, OLE5, CLO1, and HSD1. Very diverse dynamics can be observed for less abundant LD proteins (Figure 10). While some of these proteins can be detected during all time points investigated, namely LDIP, SLDP2, LDPS, and OBL1, they undergo strong changes in their levels. The contribution of OBL1, for example, increases over the course of seedling establishment, while that of LDPS decreases. In general, a few LD proteins that were identified only recently or even in this work contribute only to the seedling proteome, but not to that of seeds or siliques. This is the case for example for



**Figure 9. Dynamics of the LD coat protein families.** The riBAQ intensities of all isoforms of the LD coat protein families oleosin (OLE), caleosin (CLO), steroleosin (HSD) detected in the LD-enriched fraction were calculated as % of all known LD-associated proteins. This way, the contribution of each protein to the complete LD proteome and the dynamic changes in the abundance of the LD proteins can be investigated. 6 isoforms of the major LD protein families, namely OLE1, OLE2, OLE4, OLE5, CLO1, and HSD1 contribute more than 85 % to the LD proteome at all investigated time points.

I – Phase I siliques; II – Phase II siliques, RS – rehydrated seeds, StS – stratified seeds, 24 h to 60 h – seedlings 24h to 60 h after stratification.



**Figure 10. Dynamics of the less abundant LD protein families.** The riBAQ intensities of all isoforms of the less abundant LD protein families and LD proteins LD-ASSOCIATED PROTEINS (LDAP), LDAP-INTERACTING PROTEIN (LDIP), Plant UBX DOMAIN-CONTAINING PROTEIN 10 (PUX10), STEROLMETHYLTRANSFERASE 1 (SMT1), CYCLOARTENOL/LANOSTEROL SYNTHASE (CAS/LAS), SEED LD-ASSOCIATED PROTEINS (SLDP), OIL BODY LIPASEs (OBL),  $\alpha$ -DIOXYGENASE1 ( $\alpha$ -DOX), LD-ASSOCIATED LIPASE (LIDL), LD-ASSOCIATED PROTEIN OF SEEDS (LDPS), and LD-ASSOCIATED METHYLTRANSFERASE (LIME) detected in the LD-enriched fraction were calculated as % of all known LD proteins. This way, the contribution of each protein to the complete LD proteome and the dynamic changes in the abundance of the LD proteins can be investigated.

I – Phase I siliques; II – Phase II siliques, RS – rehydrated seeds, StS – stratified seeds, 24 h to 60 h – seedlings 24h to 60 h after stratification.

the LD lipases LIDL1 and LIDL2, or LIME1. The LDAPs are either not detected in seeds (LDAP2 and LDAP3) or contribute only very minor amounts to the complete proteome. Instead, they are most abundant in the siliques (LDAP1 and LDAP3) or in seedlings (LDAP2). Interestingly, also  $\alpha$ -DIOXYGENASE1 that was previously described to be LD-enriched under senescence (Brocard et al., 2017) and pathogen attack (Shimada et al., 2014) was already present after 36 h.

### **Phosphorylation and Ubiquitination of LD proteins**

Posttranslational modifications can influence the activity, localization and fate of a protein. Two of such modifications, phosphorylations and ubiquitinations have been identified on the LD protein families oleosin, caleosin and steroleosin (Deruyffelaere et al., 2015; Hsiao and Tzen, 2011; Kretzschmar et al., 2018). We analyzed our proteomic data for peptides with such modifications. Modifications were detected on 182 proteins, and 5 LD proteins (Suppl. Datasets 21 – 30). More than 80 of the proteins contain both a phosphorylation and a ubiquitination. In OLE4, we could identify a total of four ubiquitination sites and four phosphorylation sites; though not all modifications were found at all time points (Table 4). Of the four phosphorylations identified in OLE4, three occurred on a tyrosine, and one on a serine. A single ubiquitination site during one time point could be identified on OLE2, as well as a single Phosphorylation during a different time point. Modifications were also identified on HSD1, LDAP2 and CLO1. On UBIQUITIN itself, four different ubiquitination sites were identified including K48 and K63. These indicate ubiquitin chains most commonly associated with degradation, and these were found during seedling establishment (24 h – 48 h seedlings).

## **Discussion**

### **The seedling proteome dataset fills a gap in seed and seedling proteomics studies**

We present here a proteomic dataset of *Arabidopsis thaliana* seed germination and seedling establishment generated with an Orbitrap mass spectrometer X. Because of the depth and width of the dataset, it will become a helpful resource for researchers interested in different questions of early seedling biology.

For a long time, the use of 2D electrophoresis (2DE) combined with mass spectrometry was used to identify proteins that were differently expressed during two conditions. The limitation of that approach is that many of the protein spots remained unidentified. The amount of identified proteins in these studies is thus usually in the range of hundreds. Nevertheless, also bottom-up proteomics as presented here is likely far from covering the whole proteome of the sample investigated given that more than 10,000 transcripts have been described in seeds alone (Narsai et al., 2011). When compared to 2DE approaches, the total number of proteins identified is significantly larger in the bottom-up proteomic approach. For example, Gallardo et al., while detecting more than 1200 proteins from dry seeds via 2DE (Gallardo et al., 2001), they later on only identified a small amount of all detected proteins.

The 2DE-MS approach was for example combined with the [<sup>35</sup>S]-methionine labelling to identify newly synthesized proteins. Galland and colleagues monitored the protein synthesis during germination *sensu stricto* (Galland et al., 2013); the time points investigated are roughly comparable to the stages of rehydrated seeds, stratified seeds, and 24 h seedlings presented in this study. Remarkably, many of the proteins they found to be newly synthesized were not only detected in our study but show similar intensity patterns like they observed. For example, they already identified two of the late embryogenesis-abundant proteins that we also found in Cluster 1 and 2, defined by maximum intensities during seed stages (Figure 2B). They also detect a very early synthesis of proteins in central metabolism, like peroxisomal ketoacyl-CoA (AT2G33150) and isocitrate lyase (AT3G21720). We also observe an early increase in intensity of those proteins that continues to increase until later seedling establishment (Suppl. Dataset 12).

Our data is consistent with published proteomes, but offers coverage of a larger number of proteins and more developmental stages including seedling establishment.

While post-translational modification can be detected in bottom-up proteomics approaches without previous enrichment for specific modifications, the depth of analysis we achieved here is not comparable to what is possible after purification. For both ubiquitination and phosphorylation, more extensive datasets have been generated in *Arabidopsis thaliana* seedlings (Aguilar-Hernández et al., 2017; Aryal et al., 2015; Kim et al., 2013), but our data provides insights into changes of protein

modification during development. Interesting changes were observed for example for CRUCIFERINA or UBIQUITIN (Table 4; Suppl. Dataset 29).

### **The combination of a proteomics with a cell biological approach allows the identification of LD protein candidates**

With the rise of mass spectrometric analysis techniques in recent years, the number of studies investigating the protein composition of LDs has increased. Such studies were performed on LDs isolated from senescent leaves (Brocard et al., 2017) or a variety of algae species (Lupette et al., 2019; Nguyen et al., 2011; Siegler et al., 2017; Wang et al., 2009). Bottom-up proteomics allows the detection of hundreds or even thousands of proteins in a fraction. However, one should not assume that all proteins identified in LD-enriched fraction, no matter how stringently prepared, are LD proteins. In our approach, we therefore combine two different measures of quality control to ensure the discovery of true LD-associated proteins. In a first step, in addition to sampling the LD-enriched fraction, we take a sample of the total protein before LD enrichment. This way, we can calculate the enrichment or depletion of every single detected protein between the two fractions (Figure 5, Suppl. Dataset 19). Secondly, after the identification of potential candidates, we perform a cell biological study to confirm the LD localization of the candidate protein. Here, we present a combination of two cell biological approaches (Figures 6 – 8) that have been previously used to study the subcellular localization of LD proteins.

Our work shows that while our approach is suitable for the identification of LD-associated proteins, there are still obstacles to overcome. For example, as seen in Figure 5, the number of significantly enriched proteins is too great to quickly analyze and present them in this study. On the other hand, we applied stringent criteria for a protein to be considered for the analysis presented in Figure 5. Other true LD proteins may remain undiscovered because their abundance was too low, and they were therefore not detected in enough biological replicates to be considered for the analysis.

We employed two cell biological approaches to verify LD localization. We have shown before that *Nicotiana tabacum* pollen tubes are a suitable tool for this experiment, as they are easy to transform, allow quick analysis, and are a naturally LD-rich organ (Kretzschmar et al., 2018; Müller et al., 2017; Siegler et al., 2017). The more common tool of transient transformation of *Nicotiana benthamiana* leaves are more laborious, and have the drawback not containing many LDs. Through co-expression of the



transcription factor LEC2, LD biogenesis is stimulated and can facilitate LD-localization experiments. We used this system to strengthen the results obtained for the pollen tube system.

While we are able to present four new LD-localized protein families in this study (Figures 6 and 7), we also present the subcellular localization of a protein that did not localize to LDs despite being excellent candidate (Figure 8), based on enrichment in the LD-enriched fraction. This emphasizes that not every protein in the LD-enriched fraction is a LD protein. Localization experiments are required in order to distinguish true LD proteins from contaminating proteins. We also provide evidence that proteins previously annotated as LD-localized by homology to other species might in fact not be LD-associated in Arabidopsis. In our hands, we did not observe an enrichment of OIL BODY-ASSOCIATED PROTEIN1 (OBAP1A) in the LD fractions of seed and seedlings, nor did the transiently expressed protein localize to LDs in tobacco pollen tubes (Suppl. Figure 3). These observations contradict the localization presented by López-Ribera, though these studies were performed with the maize OBAP1 which shares about 64 % sequence identity with AtOBAP1A (Lopez-Ribera et al., 2014). However, the LD-localization of ZmOBAP1 seemed to be mediated via the hydrophilic C-terminus, a rather unusual observation for a LD protein (Kory et al., 2016). López-Ribera discusses that this might indicate an indirect localization to the LD. We cannot exclude the possibility that this mediator protein is missing in tobacco pollen tubes, therefore leading to a cytosolic localization of AtOBAP1A, and that the interaction between OBAP and the mediator is very weak, leading to its removal from the LD during preparation.

### **The lipid droplet proteome is varied across species and developmental stages**

One objective of this work was the creation of a time course data set to monitor the changes in the LD proteome during and after germination. The total number of LD proteins decreased during seedling establishment, and since LDs are degraded during seedling establishment to supply the growing seedling with energy and molecular building blocks, this observation was expected. In recent years, insight into the mechanism of protein breakdown on the LDs has been gained. It has been shown that oleosins and steroleosin are ubiquitinated and are removed from the LD with the involvement of the LD-localized protein PUX10 (Deruyffelaere et al., 2015, 2018;

Kretzschmar et al., 2018). PUX10 helps to mediate the transfer of ubiquitinated oleosins to the proteasome.

Oleosins have been described as proteins specific to pollen, the tapetum and seeds (Fang et al., 2014). However, in our data set, we did not observe the replacement of oleosins in favor of other LD coat proteins. On the contrary, the contribution of oleosins to the complete LD proteome remains very constant (Suppl. Figure 4). Therefore, it is worth investigating if this replacement takes place at all or if the LDs found in vegetative tissue do contain oleosins. Proteomic analysis of LDs from senescent *Arabidopsis* leaves did not identify oleosins, only CLO3 and  $\alpha$ -DOX, which have been previously shown to be involved in pathogen defense in leaves (Brocard et al., 2017; Shimada et al., 2014). If oleosins do indeed vanish from vegetative LD, the question is when the transition from oleosin-bearing seed LD to oleosin-free vegetative LD takes place. According to microarray data, the oleosin transcripts are already reduced in dry seeds compared to developing seeds, indicating that no new oleosin synthesis takes place after germination (Winter et al., 2007).

Despite the continuous presence of oleosins, the LD proteome diversifies over the course of seedling establishment (Figure 10). This diversification could be representing a transition from seed-type LDs to vegetative-type LDs. Further research on vegetative-type LDs will show which proteins we show here to be expressed during later stages of seedling establishment are present at LDs from other tissues, as shown for the LDAPs, LDIP and  $\alpha$ -DOX (Gidda et al., 2016; Pyc et al., 2017b; Shimada et al., 2014)

The different tissues analyzed in this work do not only vary in comparison with each other, but also compared to pollen tubes. In pollen tubes from *N. tabacum*, no HSDs were identified (Kretzschmar et al., 2018). Proteins first described in this study, LIME, LIDL, LDPS and SLDP, were not detected in the pollen tube proteome. On a quantitative level, LDAP proteins were more abundant in pollen tubes than seedling LDs. Oleosins and caleosins, on the other hand, were detected, though with varying intensities, in both tobacco pollen tubes and all tissues of this study. One could hypothesize that in fact, the core plant LD proteome might be rather stable evolutionary, as most LD proteins identified are also found in lower plants like mosses (Huang et al., 2009). This not the case however for algae. In algal models, proteomics studies have indicated a unique set of LD localized proteins with low conservation to

plants or animals (Moellering and Benning, 2010; Nguyen et al., 2011; Siegler et al., 2017; Vieler et al., 2012; Yoneda et al., 2016).

### **New LD proteins add new potential functions of LDs**

In this study, we successfully identified four protein families so far not described to be LD localized in plants: a putative lipase (LIDL), and a potential methyl transferase (LIME1), as well as two proteins exclusively expressed in seeds (SLDP and LDPS1).

Both LDPS and SLDP are so far annotated as unknown proteins. We identified two SLDP isoforms in *Arabidopsis thaliana*. Homology analysis revealed that SLDP is specific for higher plants, while LDPS is also found in mosses. However, no homologs were discovered in animals or fungi. Other recently discovered LD proteins are plant specific, like the LDAPs and LDIP. Because there is low homology between animal or yeast LD proteomes and plant LD proteomes, it is likely that there are unique functions of plant LDs not found in the animal or fungal kingdoms. On the other hand, non-homologous proteins might play similar roles in different species and have emerged by convergent evolution.

LIME and LIDL were named according to their putative annotation as a putative coclaurine N-methyltransferase and a lipase family protein, respectively. In this study, we could detect two distinct LIDL isoforms, while the two *Arabidopsis* LIME isoforms were not distinguishable from the proteomic results. In contrast to the above-mentioned proteins, LIDL and LIME homologs were detected not only in plants and mosses but also algae, fungi and animals. Based on their current annotation, both LIME and LIDL are putative enzymes. So far, a few enzymes related to the lipid metabolism have been described at LDs. Members of a family of oil body lipases have been shown to be LD localized and to hydrolyze TAG, diacylglycerol, and monoacylglycerol to free fatty acids and glycerol (Eastmond, 2004; Müller and Ischebeck, 2018). Also, lipoxygenase and phospholipase A activity has been detected on the surface of LDs (Feussner and Kindl, 1992; Noll et al., 2000; Zienkiewicz et al., 2013). LD store other hydrophobic compounds in addition to triacylglycerol. Another storage molecule are sterol esters, which are found in only minor amounts in *Arabidopsis* seed LDs but can make up a much higher proportion of the neutral lipid weight in other organisms and tissues (Onal et al., 2017; Rotsch et al., 2017). In plants, there is so far no sterol esterase described that localizes to the LDs. A homolog of LIDL in *Saccharomyces cerevisiae* has been shown to have sterol esterase activity, and

loss-of-function mutants accumulate sterol esters (Athenstaedt et al., 1999; Jandrositz et al., 2005).

Coclaurine is an intermediate from the morphine biosynthetic pathway. Morphine is a hydrophobic compound found in the latex of the opium poppy. Latex particles are a class of lipid droplets that instead of triacylglycerols and sterol esters, contain polyisoprenoids (Nawamawat et al., 2011). However, neither coclaurine nor the pathway leading to its synthesis exist in *Arabidopsis*. Similar pathways to those described for coclaurine can be imagined with other secondary metabolites, which are stored within *Arabidopsis* LDs due to their high hydrophobicity.

In conclusion, the discovery of new LD proteins opens up the possibility for new perspectives on plant LD functions, biogenesis, and turnover. The existence of several proteins conserved across kingdoms implies that much of LD biology is conserved, yet the existence of plant-specific proteins suggests that plant LDs play additional roles not found in other organisms. Insight into such functions will be gained by identifying and confirming new LD proteins.

## Experimental procedures

### Plant growth

*Nicotiana tabacum* (tobacco) plants and *Arabidopsis thaliana* ecotype Col-0 plants were grown as described before (Kretzschmar et al., 2018).

For the two silique development phases, one freshly bloomed flower of 5 weeks old Col-0 *Arabidopsis* plants was removed from each inflorescence. Exactly one week later, all developing flowers of every inflorescence were removed to abolish any further fertilization. This way, another week later, two phases of seed development could be distinguished: one between two and three weeks after fertilization below the first marking, and a second one between one and two weeks after fertilization above the marking. Complete siliques were harvested, and 2 g of the older siliques, or 3 g per of the younger siliques were pooled for one biological replicate.

For rehydrated seeds (RS), stratified seeds (StS), and seedlings 24 h, and 36 h post stratification, 100 mg of dry seed material per biological replicate was used. For 48 h and 60 h, dry seed starting material was increased to 160 mg per biological replicate.

RS were incubated in water for 30 minutes without surface-sterilization. For all other germination time points, surface-sterilized seeds were spread on half-strength Murashige and Skoog media and incubated in the dark at 4 °C for 74 h. Then, StS samples were harvested, and all other plants transferred into a 22 °C 16-h-light/8-h-dark cycle growth chamber with 150  $\mu\text{mol photons m}^{-2}\text{s}^{-1}$  daytime light strength. The time points (24 h, 36 h, 48 h, 60 h) indicate the time spent by the seeds in this condition.

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

After harvest, each sample was mixed with appropriate amounts (2 mL for RS, 3 mL for IS – 48 h samples, 3.5 mL for 60 h, 15 mL for younger siliques, 20 mL for older siliques) of grinding buffer (10 mM sodium phosphate buffer pH 7.4, 200  $\mu\text{M}$  PMSF, 0.5 mM Lohman's reagent) and ground with sand to homogeneity with a pre-cooled mortar and pestle. For the total protein sample, the homogenates were centrifuged for 1 min at 100 x g to precipitate sand and larger plant debris, and 100  $\mu\text{L}$  of the homogenate was precipitated in 900  $\mu\text{L}$  96% ethanol. For enrichment of LDs, the homogenate was then subjected to three consecutive 20,000xg centrifugations for 15 min at 4 °C. After each centrifugation step, the resulting fat pad was taken off the aqueous phase and transferred into a fresh aliquot of grinding buffer, where it was resuspended. After the third centrifugation step, the fat pad was resuspended in 1 mL 96% ethanol.

### **Proteomic Sample preparation and LC-MS/MS Analysis of Peptides**

Proteins were isolated, protein concentration determined and in-gel tryptically digested as described before (Kretzschmar et al., 2018), but with twenty micrograms of protein digested per replicate. Peptides were then subjected to LC-MS/MS analysis modified from what was previously described (Schmitt et al., 2017). First, 2  $\mu\text{L}$  peptide samples were separated by nano-flow LC on a RSLCnano Ultimate 3000 system (Thermo Fisher Scientific). The peptides were loaded with 0.07% trifluoroacetic acid on an Acclaim PepMap 100 precolumn (100  $\mu\text{m}$  x 2 cm, C18, 3  $\mu\text{m}$ , 100 Å; Thermo Fisher Scientific) with a flow rate of 20  $\mu\text{L}/\text{min}$  for 3 minutes. Then, peptides were separated by reverse phase chromatography on an Acclaim PepMan PSLC column (75  $\mu\text{m}$  x 50 cm, C18, 3  $\mu\text{m}$ , 100 Å; Thermo Fisher Scientific) with a flow rate of 300

nL/min. The solvent composition was gradually changed within a time period of 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 minutes, to 30 % solvent B within the following 58 min, to 45% solvent B within the next 22 min, and to 90 % solvent B within the following 12 min. All solvents and acids were of Optima LC/MS quality and purchased at Thermo Fisher Scientific. Eluted peptides were ionized on-line by nano-electrospray ionization (nESI) with a Nanspray Flex Ion Source (Thermo Fisher Scientific) at 1.5 kV (liquid junction) and analyzed with a Q Exactive HF mass spectrometer (Thermo Fisher Scientific). Full scans were recorded in a mass range of 300 to 1650 m/z at a resolution of 30,000 followed by data dependent top 10 HCD fragmentation (dynamic exclusion enabled). LC-MS method programming and data acquisition was performed with the XCalibur 4.0 software (Thermo Fisher Scientific).

### **Mass spectrometral data processing**

MS and MS/MS raw data was processed for feature detection, peptide identification, and protein group assembly with MaxLFQ algorithm in the MaxQuant software version 1.6.2.10 (Cox and Mann, 2008; Cox et al., 2014). Settings are specified in the metadata file (Suppl. Tables 1/2). In brief, default settings were used. Additionally, label-free quantification was enabled in group-specific parameter settings. In global parameter settings, match between runs and iBAQ were enabled. The TAIR10 protein database was used as reference. Data analysis was performed in Perseus 1.6.2.2 and in Excel as described in the corresponding figures or supplementary datasets.

### **Particle Bombardment and Microscopy of *Nicotiana tabacum* Pollen Tubes**

Pollen grains were transformed by particle bombardment, in vitro cultivated on microscope slides, stained with Nile Red and analyzed by confocal laser scanning microscopy as described (Müller et al., 2017) or with a Zeiss LSM780 (Carl Zeiss, Jena, Germany).

### **Molecular Cloning etc.**

LDPS, LIDL, SLDP2 and ATS3a were amplified from Arabidopsis mature (dry) seed cDNA and LIME1 was amplified from 7 d old seedling cDNA using Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific), as per the manufacturer's protocol. LDPS1 and ATS3A PCR products were inserted into the pDONR/zeo entry vector using Gateway technology (Curtis and Grossniklaus, 2003) with the Gateway BP clonase enzyme mix (Invitrogen). Entry vector genes were subsequently subcloned into the appropriate binary expression vectors (e.g., pMDC32-ChC) using Gateway LR clonase enzyme mix (Invitrogen). Molecular cloning into the Gateway vector pUC-LAT52-mVenusC-GW and pUC-LAT52-mVenusN-GW was performed as described in (Müller et al., 2017). The use of CFP-SKL for peroxisomal marking has been previously described (Müller et al., 2017).

The pMDC32-ChC plant expression binary vector, encoding the monomeric Cherry fluorescent protein open reading frame (ORF) adjacent to a 5' recombination site that allows for mCherry to be translational fused to the C terminus a protein of interest, was constructed in the following manner. First the mCherry ORF was amplified using the mCherry-FP-*PacI* and mCherry-RP-*SacI* primers and pRTL2-Cherry (Gidda et al., 2011) as template DNA. The resulting PCR products were then digested with *PacI* and *SacI* and inserted into similarly-digested pMDC32 (Curtis and Grossniklaus, 2003), yielding pMDC32-ChC1. Thereafter, the *Cm<sup>r</sup>/ccdB* region of pMDC32/ChC1 was amplified using the primers *ccdB-FP-KpnI* and *ccdB-RP-PacI*, which resulted in the removal of a stop codon upstream of the Cherry ORF, and reinserted into *KpnI-PacI*-digested pMDC32-ChC1, yielding pMDC32-ChC.

Custom oligonucleotide primers were synthesized by Sigma-Aldrich; a complete list of all primers is given in Suppl. Table 3. All plasmids constructed in this study, including their promoter and cloning regions and any fusion protein ORFs, were verified by automated sequencing performed at Microsynth AG (Balgach, Switzerland) or the Advanced Analysis Centre Genomics Facility (Guelph, Canada).

### **Transient transformation of *Nicotiana benthamiana* leaves**

*Nicotiana benthamiana* plants were grown in soil with a 16-h/8-h day/night cycle at 22 °C. Leaves of ~4-week-old plants were infiltrated with *Agrobacterium tumefaciens* (strain LBA4404) harboring the selected expression vector as described in Pyc et al., 2017b. All infiltrations were performed with pORE04-35S:p19 containing the tomato bushy stunt virus gene P19 in order to enhance transgene expression (Petrie et al., 2010).

## **Microscopy**

*A. tumefaciens*-infiltrated tobacco leaves were processed for CLSM imaging, including staining of LDs with BODIPY 493/503 (Invitrogen), as previously described (Gidda et al., 2016). Micrographs of tobacco leaves were acquired using a Leica SP5 CLSM (Leica Microsystems). Excitations and emission signals for fluorescent proteins and BODIPY were collected sequentially as single optical sections in double-labelling experiments as those described in Gidda et al. (2016); single-labelling experiments showed no detectable crossover at the settings used for data collection. All fluorescent images of cells shown are representative of at least two separate experiments, including at least three separate transformations of tobacco leaf cells.

## **Accession numbers**

AtLIDL1 - AT1G18460; AtLIDL2 - AT1G73920; AtLDPS1 - AT3G19920; AtLIME1 - AT4G33110; AtLIME2 - AT4G33120; AtSLDP1 - AT5G36100; AtSLDP2 - AT1G65090; AtOBAP1A - AT1G05510; AtOLE1 - AT4G25140; AtOLE2 - AT5G40420; AtOLE3 - AT5G51210; AtOLE4 - AT3G27660; AtOLE5 - AT3G01570; AtOLE6 - AT1G48990; AtOLE7 - AT2G25890; AtOLE8 - AT3G18570; AtCLO1 - AT4G26740; AtCLO2 - AT5G55240; AtCLO3 - AT2G33380; AtCLO5 - AT1G70680; AtHSD1/1 - AT5G50600/AT5G50700; AtHSD2 - AT3G47350; AtHSD3 - AT3G47360; AtHSD5 - AT4G10020; AtHSD6 - AT5G50770; AtLDAP1 - AT1G67360; AtLDAP2 - AT2G47780; AtLDAP3 - AT3G05500; AtLDIP - AT5G16550; AtPUX10 - AT4G10790; AtSMT1 - AT5G13710; AtCAS - AT2G07050; AtLAS - AT3G45130; AtOBL1 - AT3G14360; AtOBL3 - AT1G45201; AtαDOX - AT3G01420.



## **Supplemental Data**

**Supplemental Figure 1. Post-germinative growth of *Arabidopsis thaliana* Col-0.**

**Supplemental Figure 2. Abundance of LD proteins within the total protein fraction.**

**Supplemental Figure 3. Characterization of the AtOBAP protein family.**

**Supplemental Figure 4. Changes in contribution of all detected oleosins to the complete LD proteome.**

**Supplemental Dataset 1. Proteins found in siliques and seedlings - raw LFQ values.**

**Supplemental Dataset 2. Proteins found in siliques and seedlings - normalized and sorted LFQ.**

**Supplemental Dataset 3. Proteins found in siliques and seedlings - raw iBAQ values.**

**Supplemental Dataset 4. Proteins found in siliques and seedlings - normalized and sorted iBAQ.**

**Supplemental Dataset 5. Proteins found in siliques and seedlings - imputed log<sub>2</sub> transformed LFQ values of Total protein fraction.**

**Supplemental Dataset 6. Loadings of PCA plot created with Supplemental Dataset 5.**

**Supplemental Dataset 7. Projections of PCA plot created with Supplemental Dataset 5.**

**Supplemental Dataset 8. Proteins found in siliques and seedlings - imputed log<sub>2</sub> transformed LFQ values of lipid droplet-enriched fractions.**

**Supplemental Dataset 9. Loadings of PCA plot created with Supplemental Dataset 8.**

**Supplemental Dataset 10. Projections of PCA plot created with Supplemental Dataset 8.**

**Supplemental Dataset 11. Proteins found in the total protein fraction of siliques and seedlings - normalized and sorted LFQ with at least four valid values in at least one condition.**

**Supplemental Dataset 12. Data set used to create heat map presented in Figure 2.**

**Supplemental Dataset 13. Sums of rLFQ values of proteins associated with the same GO ID.**

**Supplemental Dataset 14. Phase-dependent averages of sums of rLFQ values of proteins associated with the same GO ID.**

**Supplemental Dataset 15. Selected GO ID sums.**

**Supplemental Dataset 16. Subcellular localization of proteins**

**Supplemental Dataset 17. Sums of riBAQ values of proteins with the same localization.**

**Supplemental Dataset 18. List of curated localization acquired from the <http://ppdb.tc.cornell.edu/>.**

**Supplemental Dataset 19. Results Matrix from enrichment analysis.**

**Supplemental Dataset 20: Contribution of LD proteins to the total LD proteome.**

**Supplemental Dataset 21: Modified proteins identified in Phase I silique samples.**

**Supplemental Dataset 22: Modified proteins identified in Phase II silique samples.**

**Supplemental Dataset 23: Modified proteins identified in rehydrated seeds samples.**

**Supplemental Dataset 24: Modified proteins identified in stratified seeds samples.**

**Supplemental Dataset 25: Modified proteins identified in 24 h seedling samples.**

**Supplemental Dataset 26: Modified proteins identified in 36 h seedling samples.**

**Supplemental Dataset 27: Modified proteins identified in 48 h seedling samples.**

**Supplemental Dataset 28: Modified proteins identified in 60 h seedling samples.**

**Supplemental Dataset 29: All modified proteins identified across samples including their modified sites**

**Supplemental Dataset 30: All modified LD proteins identified across samples including their modified sites.**

**Supplemental Table 1: Metadata file for LC-MS/MS data processing with MaxQuant.**

**Supplemental Table 2: Metadata file for LC-MS/MS data processing with MaxQuant for posttranslationally modified peptides.**

**Supplemental Table 3: Primers used for molecular cloning via GateWay cassette, and sequencing.**

## **Author contribution**

F.K.K., G.H.B., R.T.M. and T.I. designed the work, F.K.K., N.D., H.E.K., P.S., K.S., and T.I. performed research, F.K.K., H.E.K., P.S., K.S., G.H.B. and T.I. analyzed data, and F.K.K. and T.I. wrote the manuscript. All authors critically read and revised the manuscript and approved the final version.

## **Acknowledgements**

We thank Ivo Feussner for all his support and many helpful discussions and Prof. Gatz and Dr. Alexander Stein for valuable advice. Thanks goes to Prof. Kent Chapman, Dr. John Dyer for discussing our work with us.

We are also grateful to Prof. Großhans and Prof. Johnsen for granting access to their confocal microscopes as well as Dr. Florian Wegwitz and Johannes Sattmann for their assistance. We also thank Dr. Leonie Steinhorst and Prof. Jörg Kudla (University of Münster) for generating and providing plasmids. Thanks for help in the lab to Siqi Sun and Antony Grüness.

This work was supported by funding from the DFG (IS 273/2-2 to T.I., GR1945/3-1, SFB937/TP10 and INST1525/16-1 FUGG to Prof. Großhans) and the Studienstiftung des deutschen Volkes (to F.K.K. and P.S.).

## Tables

**Table 1:** Summary of proteins identified across samples and the percentage of LD proteins within all samples. iBAQ processed proteomic data of both total protein fraction and LD-enriched fraction was filtered for at least 2 peptides per protein group. Then, the relative abundance of LD-associated proteins in the LD-enriched fraction was calculated based on their iBAQ scores. LD proteins were chosen taking into account the results of this work.

	<b>Total protein fraction</b>	<b>LD-enriched fraction</b>	<b>% of LD-associated proteins in LD-enriched fraction</b>
<b>Stage I siliques</b>	1723	1266	16.8±4.4
<b>Stage II siliques</b>	1417	1337	17.5±1.1
<b>Rehydrated seeds</b>	1425	1353	32±1.4
<b>Stratified seeds</b>	1511	1024	31.8±8
<b>24 h seedlings</b>	2004	1158	24.4±7.9
<b>36 h seedlings</b>	2197	1368	26±5.1
<b>48 h seedlings</b>	2218	1478	28.5±4.4
<b>60 h seedlings</b>	2198	1723	34.1±12.2

**Table 2:** Annotation of subcellular localization of proteins in the LD-enriched fraction. All proteins were annotated with 78 subcellular localization obtained from the Plant Proteome Database and combined into 11 groups. The riBAQ intensities (in %) of the proteins were added up for 11 different subcellular compartments.

Total protein fraction fractions								
	Phase I siliques	Phase II siliques	Rehydrated seeds	Stratified seeds	24 h seedlings	36 h seedlings	48 h seedlings	60 h seedlings
<b>Plastid</b>	571.3±24.6	392.3±59.4	44.1±8.8	40±5.9	52.4±10.3	154±26.2	334.3±42.7	477.5±33.2
<b>ER</b>	8±0.5	5.1±0.9	2.2±0.3	2.6±0.5	8.8±1.6	13.2±1.5	13.1±3.5	15.9±10.6
<b>Vacuole</b>	14.8±1.8	16.1±2.1	4.8±1	5.5±2.4	7.1±1.7	9.9±1.1	12.1±3.5	23.6±2.2
<b>Peroxisome</b>	15.1±0.2	9.8±1.1	9.6±2.4	10±2.5	22.7±3.6	26.6±2.6	26.4±2	21.2±13.7
<b>Golgi apparatus</b>	0.3±0.1	0.2±0.1	0.5±0.1	0.6±0.2	1.1±0.3	1.4±0.2	0.7±0.4	0.5±0.6
<b>Mitochondria</b>	22.8±1.6	13.5±1.6	9±1.3	10.7±1.7	32±3.8	39.5±1.3	33.9±3.5	34.5±4.4
<b>Nucleus</b>	29.1±2.5	18.7±3.1	17.3±4	19.3±2.3	29.2±2.9	34.3±1.4	35.8±6.5	26.8±1.8
<b>Cytosol</b>	67.8±5.4	28.3±4.7	40.3±1.9	48.3±4	148.1±6.1	195.7±2.3	205.9±10.7	183.5±15.9
<b>Plastoglobules</b>	12.3±0.9	14.1±3.2	0.8±0.2	0.8±0.2	0.9±0.1	2±0.2	5.1±1.3	6.4±0.1
<b>Plasma membrane</b>	15.4±2.	9.1±1.5	2.5±0.7	3±0.7	6.3±1.6	8.9±1.7	9.3±1.4	9.2±1.5
<b>LD</b>	12.3±8	25.5±7.7	57.4±7.8	69.9±8.9	38±5	19±1.1	8.3±1.8	4.8±4.5

**Table 3:** Proteins chosen as LD protein candidates. Enrichment factors and p-values were calculated from the given developmental stages at which the protein had the highest riBAQ intensity value in the LD-enriched fraction.

Protein name	AGI code	Description	Stage with highest abundance in LD fraction	Abundance at highest stage (riBAQ ‰)	Enrichment factor (log2 LD - log2 TE)	p-value (-log)	localization
LIDL1	AT1G18460.1	Lipase family protein	60 h seedlings	0.31±0.15	4.75	6.01	LD
LIME1	AT4G33110.1	Coclaurine methyltransferase	48 h seedlings	1.12±0.62	4.92	3.55	LD
SLDP2	AT1G65090.3	unknown	60 h seedlings	2.1±0.45	4.13	6.31	LD
LDPS1	AT3G19920.1	unknown	Imbibed seeds	0.31±0.19	2.89	2.36	LD
ATS3A	AT2G41475.1	Embryo-specific protein related	48 h seedlings	1.16±0.54	6.12	6.39	Punctate structures

**Table 4: Posttranslational modifications detected on LD proteins and Ubiquitin.** Sites marked with an asterisk were not identified unambiguously. For complete list of modified proteins and modification sites, see Supplemental Dataset 26-35.

	Ubiquitination						Phosphorylation					
	RS	StS	24	36	48	60	RS	StS	24	36	48	60
<b>OLE2</b>					K146		S18					
<b>OLE4</b>	K157*	K157*	K43	K157*	K157*	K157*	Y153	Y153	S40			
	K159*	K159*	K157*	K159*	K159*	K159*	Y164		Y42			
	K168		K159*	K168	K168	K168						
			K168									
<b>CLO1</b>		K4		K4								
<b>HSD1</b>				K289			S340	S348		T283		
				K295			T343					
							S348					
<b>LDAP2</b>				K108						T118		S128
				K125						S122		
<b>UBQ13</b>			K48	K6	K48				S246	T7	S246	
			K63	K11	K63				S247	T9	S247	
				K48						S246	T249	
				K63						S247		



## References

- Aguilar-Hernández, V., Kim, D.Y., Stankey, R.J., Scalf, M., Smith, L.M., and Vierstra, R.D. (2017). Mass spectrometric analyses reveal a central role for ubiquitylation in remodeling the *Arabidopsis* proteome during photomorphogenesis. *Mol. Plant* **10**, 846–865.
- Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815.
- Aryal, U.K., Ross, A.R.S., and Krochko, J.E. (2015). Enrichment and analysis of intact phosphoproteins in *Arabidopsis* seedlings. *PLoS One* **10**, e0130763.
- Athenstaedt, K., Zweytick, D., Jandrositz, A., Kohlwein, S.D., and Daum, G. (1999). Identification and characterization of major lipid particle proteins of the yeast *Saccharomyces cerevisiae*. *J. Bacteriol.* **181**, 6441–6448.
- Baerenfaller, K., Grossmann, J., Grobei, M. a, Hull, R., Hirsch-hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W., and Baginsky, S. (2008). Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics. *Science* (80-. ). **320**, 938–941.
- Baud, S., Boutin, J.P., Miquel, M., Lepiniec, L., and Rochat, C. (2002). An integrated overview of seed development in *Arabidopsis thaliana* ecotype WS. *Plant Physiol. Biochem.* **40**, 151–160.
- Baud, S., Dichow, N.R., Kelemen, Z., D'Andréa, S., To, A., Berger, N., Canonge, M., Kronenberger, J., Viterbo, D., Dubreucq, B., et al. (2009). Regulation of HSD1 in seeds of *Arabidopsis thaliana*. *Plant Cell Physiol.* **50**, 1463–1478.
- 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. *Front. Plant Sci.* **8**.
- Chapman, K.D., Dyer, J.M., and Mullen, R.T. (2012). Biogenesis and functions of lipid droplets in plants. *J. Lipid Res.* **53**, 215–226.
- Chibani, K., Ali-Rachedi, S., Job, C., Job, D., Jullien, M., and Grappin, P. (2006). Proteomic analysis of seed dormancy in *Arabidopsis*. *Plant Physiol.* **142**, 1493–1510.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372.
- Cox, J., Hein, M.Y., Lubner, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513–2526.
- Cui, S., Hayashi, Y., Otomo, M., Mano, S., Oikawa, K., Hayashi, M., and Nishimura, M. (2016). Sucrose production mediated by lipid metabolism suppresses the physical interaction of peroxisomes and oil bodies during germination of *arabidopsis thaliana*. *J. Biol. Chem.* **291**, 19734–19745.
- Curtis, M.D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* **113**, 462–469.

- 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 Cell Physiol.* **56**, 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*. *Plant Cell* **30**, 2116–2136.
- Durand, T.C., Cue, G., Godin, B., Valot, B., and Cl, G. (2019). Combined proteomic and metabolomic profiling of the *Arabidopsis thaliana* vps29 mutant reveals pleiotropic functions of the retromer in seed development. *Int. J. Mol. Sci.* **20**, 1–22.
- Eastmond, P.J. (2004). Cloning and characterization of the acid lipase from Castor beans. *J. Biol. Chem.* **279**, 45540–45545.
- Fang, Y., Zhu, R.L., and Mishler, B.D. (2014). Evolution of oleosin in land plants. *PLoS One* **9**, 1–10.
- Fercha, A., Capriotti, A.L., Caruso, G., Cavaliere, C., Stampachiacchiere, S., Zenezini Chiozzi, R., and Laganà, A. (2016). Shotgun proteomic analysis of soybean embryonic axes during germination under salt stress. *Proteomics* **16**, 1537–1546.
- Feussner, I., and Kindl, H. (1992). A lipoxygenase is the main lipid body protein in cucumber and soybean cotyledons during the stage of triglyceride mobilization. *FEBS Lett.* **298**, 223–225.
- Footitt, S., Slocombe, S.P., Lerner, V., Kurup, S., Wu, Y., Larson, T., Graham, I., Baker, A., and Holdsworth, M. (2002). Control of germination and lipid mobilization by COMATOSE, the *Arabidopsis* homologue of human ALDP. *EMBO J.* **21**, 2912–2922.
- Galland, M., Huguet, R., Arc, E., Cueff, G., Job, D., and Rajjou, L. (2013). Dynamic proteomics emphasizes the importance of selective mRNA translation and protein turnover during *Arabidopsis* seed germination. *Mol. Cell. Proteomics* **13**, 252–268.
- Gallardo, K., Job, C., Groot, S.P.C., Puype, M., Demol, H., Vandekerckhove, J., and Job, D. (2001). Proteomic analysis of *Arabidopsis* seed germination and priming. *Plant Physiol.* **126**, 835–848.
- Giavalisco, P., Nordhoff, E., Kreitler, T., Klöppel, K.-D., Lehrach, H., Klose, J., and Gobom, J. (2005). Proteome analysis of *Arabidopsis thaliana* by two-dimensional gel electrophoresis and matrix-assisted laser desorption/ionisation-time of flight mass spectrometry. *Proteomics* **5**, 1902–1913.
- Gidda, S.K., Shockey, J.M., Falcone, M., Kim, P.K., Rothstein, S.J., Andrews, D.W., Dyer, J.M., and Mullen, R.T. (2011). Hydrophobic-domain-dependent protein-protein interactions mediate the localization of GPAT enzymes to ER subdomains. *Traffic* **12**, 452–472.
- 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 Physiol.* **170**, 2052–2071.
- Hajduch, M., Ganapathy, A., Stein, J.W., and Thelen, J.J. (2005). A systematic proteomic study of seed filling in Soybean. Establishment of high-resolution two-

- dimensional reference maps, expression profiles, and an interactive proteome database. *Plant Physiol.* **137**, 1397–1419.
- Hajduch, M., Casteel, J.E., Hurrelmeyer, K.E., Song, Z., Agrawal, G.K., and Thelen, J.J. (2006). Proteomic analysis of seed filling in *Brassica napus*. Developmental characterization of metabolic isozymes using high-resolution two-dimensional gel electrophoresis. *Plant Physiol.* **141**, 32–46.
- Hajduch, M., Hearne, L.B., Miernyk, J.A., Casteel, J.E., Joshi, T., Agrawal, G.K., Song, Z., Zhou, M., Xu, D., and Thelen, J.J. (2010). Systems analysis of seed filling in *Arabidopsis*: Using general linear modeling to assess concordance of transcript and protein expression. *Plant Physiol.* **152**, 2078–2087.
- Han, C., Wang, K., and Yang, P. (2014a). Gel-based comparative phosphoproteomic analysis on rice embryo during germination. *Plant Cell Physiol.* **55**, 1376–1394.
- Han, C., Yang, P., Sakata, K., and Komatsu, S. (2014b). Quantitative proteomics reveals the role of protein phosphorylation in rice embryos during early stages of germination. *J. Proteome Res.* **13**, 1766–1782.
- Han, C., He, D., Li, M., and Yang, P. (2014c). In-depth proteomic analysis of rice embryo reveals its important roles in seed germination. *Plant Cell Physiol.* **55**, 1826–1847.
- Hanano, A., Burcklen, M., Flenet, M., Ivancich, A., Louwagie, M., Garin, J., and Blée, E. (2006). Plant seed peroxxygenase is an original heme-oxygenase with an EF-hand calcium binding motif. *J. Biol. Chem.* **281**, 33140–33151.
- Hayashi, M., Nito, K., Takei-Hoshi, R., Yagi, M., Kondo, M., Suenaga, A., Yamaya, T., and Nishimura, M. (2002). Ped3p is a peroxisomal ATP-binding cassette transporter that might supply substrates for fatty acid  $\beta$ -oxidation. *Plant Cell Physiol.* **43**, 1–11.
- 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**, 169.
- 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 Physiol.* **162**, 1926–1936.
- Houston, N.L., Hajduch, M., and Thelen, J.J. (2009). Quantitative proteomics of seed filling in *Castor*: Comparison with soybean and rapeseed reveals differences between photosynthetic and nonphotosynthetic seed metabolism. *Plant Physiol.* **151**, 857–868.
- Hsiao, E.S.L., and Tzen, J.T.C. (2011). Ubiquitination of oleosin-H and caleosin in sesame oil bodies after seed germination. *Plant Physiol. Biochem.* **49**, 77–81.
- Huang, A.H.C. (2017). Plant lipid droplets and their associated oleosin and other proteins: potential for rapid advances. *Plant Physiol.* pp.01677.2017.
- Huang, C.-Y., Chung, C.-I., Lin, Y.-C., Hsing, Y.-I.C., and Huang, A.H.C. (2009). Oil bodies and oleosins in *Physcomitrella* possess characteristics representative of early trends in evolution. *Plant Physiol.* **150**, 1192–1203.
- Jandrositz, A., Petschnigg, J., Zimmermann, R., Natter, K., Scholze, H., Hermetter,

- A., Kohlwein, S.D., and Leber, R. (2005). The lipid droplet enzyme Tgl1p hydrolyzes both steryl esters and triglycerides in the yeast, *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* 1735, 50–58.
- 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 Physiol.* 157, 866–875.
- Kim, D.-Y., Scalf, M., Smith, L.M., and Vierstra, R.D. (2013). Advanced proteomic analyses yield a deep catalog of ubiquitylation targets in *Arabidopsis*. *Plant Cell* 25, 1523–1540.
- Kim, E.Y., Park, K.Y., Seo, Y.S., and Kim, W.T. (2016). *Arabidopsis* small rubber particle protein homolog SRPs play dual roles as positive factors for tissue growth and development and in drought stress responses. *Plant Physiol.* 170, 2494–2510.
- Kory, N., Farese, R. V., and Walther, T.C. (2016). Targeting fat: Mechanisms of protein localization to lipid droplets. *Trends Cell Biol.* 26, 535–546.
- Kretschmar, F.K., Mengel, L.A., Müller, A.O., Schmitt, K., Bliersch, K.F., Valerius, O., Braus, G.H., and Ischebeck, T. (2018). PUX10 is a lipid droplet-localized scaffold protein that interacts with CELL DIVISION CYCLE48 and is involved in the degradation of lipid droplet proteins. *Plant Cell* 30, 2137–2160.
- Kubala, S., Garnczarska, M., Wojtyła, Ł., Clippe, A., Kosmala, A., Zmieńko, A., Lutts, S., and Quinet, M. (2015). Deciphering priming-induced improvement of rapeseed (*Brassica napus* L.) germination through an integrated transcriptomic and proteomic approach. *Plant Sci.* 231, 94–113.
- Lee, J., and Koh, H.J. (2011). A label-free quantitative shotgun proteomics analysis of rice grain development. *Proteome Sci.* 9, 61.
- Lee, J., Lee, W., and Kwon, S.W. (2015). A quantitative shotgun proteomics analysis of germinated rice embryos and coleoptiles under low-temperature conditions. *Proteome Sci.* 13, 1–11.
- Li, F., Asami, T., Wu, X., Tsang, E.W.T., and Cutler, A.J. (2007). A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. *Plant Physiol.* 145, 87–97.
- Li, Q.-F., Xiong, M., Xu, P., Huang, L.-C., Zhang, C.-Q., and Liu, Q.-Q. (2016). Dissection of brassinosteroid-regulated proteins in rice embryos during germination by quantitative proteomics. *Sci. Rep.* 6, 34583.
- Lin, L.-J., Tai, S.S.K., Peng, C.-C., and Tzen, J.T.C. (2002). Steroleosin, a sterol-binding dehydrogenase in seed oil bodies. *Plant Physiol.* 128, 1200–1211.
- Lopez-Ribera, I., La Paz, J.L., Repiso, C., Garcia, N., Miquel, M., Hernandez, M.L., Martinez-Rivas, J.M., and Vicient, C.M. (2014). The evolutionary conserved oil body associated protein OBAP1 participates in the regulation of oil body size. *Plant Physiol.* 164, 1237–1249.
- Lorenz, C., Brandt, S., Borisjuk, L., Rolletschek, H., Heinzl, N., Tohge, T., Fernie, A.R., Braun, H.-P., and Hildebrandt, T.M. (2018). The role of persulfide metabolism during *Arabidopsis* seed development under light and dark conditions. *Front. Plant Sci.* 9, 1–15.

- Lupette, J., Jaussaud, A., Seddiki, K., Morabito, C., Brugière, S., Schaller, H., Kuntz, M., Putaux, J.L., Jouneau, P.H., Rébeillé, F., et al. (2019). The architecture of lipid droplets in the diatom *Phaeodactylum tricornutum*. *Algal Res.* 38, 101415.
- Mansfield, S.G., and Briarty, L.G. (1992). Cotyledon cell development in *Arabidopsis thaliana* during reserve deposition. *Can. J. Bot.* 70, 151–164.
- De Michele, R., McFarlane, H.E., Parsons, H.T., Meents, M.J., Lao, J., González Fernández-Niño, S.M., Petzold, C.J., Frommer, W.B., Samuels, A.L., and Heazlewood, J.L. (2016). Free-flow electrophoresis of plasma membrane vesicles enriched by two-phase partitioning enhances the quality of the proteome from *Arabidopsis* seedlings. *J. Proteome Res.* 15, 900–913.
- Moellering, E.R., and Benning, C. (2010). RNA interference silencing of a major lipid droplet protein affects lipid droplet size in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* 9, 97–106.
- Müller, A.O., and Ischebeck, T. (2018). Characterization of the enzymatic activity and physiological function of the lipid droplet-associated triacylglycerol lipase AtOBL1. *New Phytol.* 217, 1062–1076.
- Müller, A.O., Bliersch, 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. *Plant J.* 89, 1055–1064.
- 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 Mol. Biol.* 44, 463–476.
- Narsai, R., Law, S.R., Carrie, C., Xu, L., and Whelan, J. (2011). In-depth temporal transcriptome profiling reveals a crucial developmental switch with roles for RNA processing and organelle metabolism that are essential for germination in *Arabidopsis*. *Plant Physiol.* 157, 1342–1362.
- Nawamawat, K., Sakdapipanich, J.T., Ho, C.C., Ma, Y., Song, J., and Vancso, J.G. (2011). Surface nanostructure of *Hevea brasiliensis* natural rubber latex particles. *Colloids Surfaces A Physicochem. Eng. Asp.* 390, 157–166.
- Nguyen, H.M., Baudet, M., Cuiné, S., Adriano, J.-M., Barthe, D., Billon, E., Bruley, C., Beisson, F., Peltier, G., Ferro, M., et al. (2011). Proteomic profiling of oil bodies isolated from the unicellular green microalga *Chlamydomonas reinhardtii*: With focus on proteins involved in lipid metabolism. *Proteomics* 11, 4266–4273.
- Nguyen, T.-P., Cueff, G., Hegedus, D.D., Rajjou, L., and Bentsink, L. (2015). A role for seed storage proteins in *Arabidopsis* seed longevity. *J. Exp. Bot.* 66, 6399–6413.
- Noll, F., May, C., and Kindl, H. (2000). Phospholipid monolayer of plant lipid bodies attacked by phospholipase A2 shows 80 nm holes analyzed by atomic force microscopy. *Biophys. Chem.* 86, 29–35.
- Nonogaki, H., Bassel, G.W., and Bewley, J.D. (2010). Germination—Still a mystery. *Plant Sci.* 179, 574–581.
- Onal, G., Kutlu, O., Gozuacik, D., and Dokmeci Emre, S. (2017). Lipid Droplets in Health and Disease. *Lipids Health Dis.* 16, 128.
- Penfield, S., Graham, S., and Graham, I.A. (2005). Storage reserve mobilization in

germinating oilseeds: Arabidopsis as a model system. *Biochem. Soc. Trans.* 33, 380–383.

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, 8.

Pyc, M., Cai, Y., Greer, M.S., Yurchenko, O., Chapman, K.D., Dyer, J.M., and Mullen, R.T. (2017a). Turning over a new leaf in lipid droplet biology. *Trends Plant Sci.* 22, 596–609.

Pyc, M., Cai, Y., Gidda, S.K., Yurchenko, O., Park, S., Kretschmar, F.K., Ischebeck, T., Valerius, O., Braus, G.H., Chapman, K.D., et al. (2017b). Arabidopsis LDAP-interacting protein (LDIP) influences lipid droplet size and neutral lipid homeostasis in both leaves and seeds. *Plant J.* 12, 3218–3221.

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.). *Biochem. J.* 235, 57–65.

Quan, S., Yang, P., Cassin-Ross, G., Kaur, N., Switzenberg, R., Aung, K., Li, J., and Hu, J. (2013). Proteome analysis of peroxisomes from etiolated Arabidopsis seedlings identifies a peroxisomal protease involved in  $\alpha$ -oxidation and development. *Plant Physiol.* 163, 1518–1538.

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. *Plant J.* 92, 129–146.

Schmitt, K., Smolinski, N., Neumann, P., Schmaul, S., Hofer-Pretz, V., Braus, G.H., and Valerius, O. (2017). Asc1p/RACK1 connects ribosomes to eukaryotic phosphosignaling. *Mol. Cell. Biol.* 37.

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*. *Plant J.* 55, 798–809.

Shimada, T.L., Takano, Y., Shimada, T., Fujiwara, M., Fukao, Y., Mori, M., Okazaki, Y., Saito, K., Sasaki, R., Aoki, K., et al. (2014). Leaf oil body functions as a subcellular factory for the production of a phytoalexin in Arabidopsis. *Plant Physiol.* 164, 105–118.

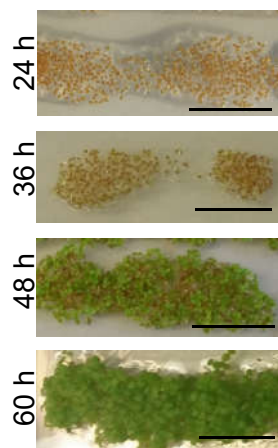
Siegler, H., Valerius, O., Ischebeck, T., Popko, J., Tourasse, N.J., Vallon, O., Khozin-Goldberg, I., Braus, G.H., and Feussner, I. (2017). Analysis of the lipid body proteome of the oleaginous alga *Lobosphaera incisa*. *BMC Plant Biol.* 17, 1–17.

Siloto, R.M.P., 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 oilbodies in Arabidopsis. *Plant Cell* 18, 1961–1974.

Sun, Q., Zybaylov, B., Majeran, W., Friso, G., Olinares, P.D.B., and van Wijk, K.J. (2009). PPDB, the plant proteomics database at Cornell. *Nucleic Acids Res.* 37, D969–D974.

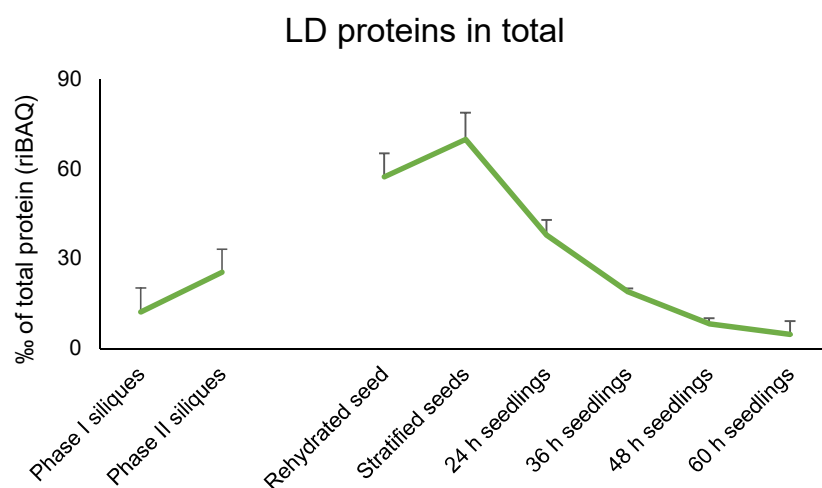
Thazar-Poulot, N., Miquel, M., Fobis-Loisy, I., and Gaude, T. (2015). Peroxisome extensions deliver the Arabidopsis SDP1 lipase to oil bodies. *Proc. Natl. Acad. Sci. U. S. A.* 112, 4158–4163.

- Tzen, J.T.C., Cao, Y., Laurent, P., Ratnayake, C., and Huang, A.H.C. (1993). Lipids, proteins, and structure of seed oil bodies from diverse species. *Plant Physiol.* **101**, 267–276.
- 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 Environ.* **40**, 2292–2306.
- Vance, V.B., and Huang, A.H. (1987). The major protein from lipid bodies of maize. Characterization and structure based on cDNA cloning. *J. Biol. Chem.* **262**, 11275–11279.
- 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 Physiol.* **158**, 1562–1569.
- Wang, Y., Ma, X., Zhang, X., He, X., Li, H., Cui, D., and Yin, D. (2016). ITRAQ-based proteomic analysis of the metabolic mechanisms behind lipid accumulation and degradation during peanut seed development and postgermination. *J. Proteome Res.* **15**, 4277–4289.
- Wang, Z.T., Ullrich, N., Joo, S., Waffenschmidt, S., and Goodenough, U. (2009). Algal lipid bodies: Stress induction, purification, and biochemical characterization in wild-type and starchless *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **8**, 1856–1868.
- 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**, e718.
- Xu, E., Chen, M., He, H., Zhan, C., Cheng, Y., Zhang, H., and Wang, Z. (2017). Proteomic analysis reveals proteins involved in seed imbibition under salt stress in rice. *Front. Plant Sci.* **7**, 1–11.
- Yang, P., Li, X., Wang, X., Chen, H., Chen, F., and Shen, S. (2007). Proteomic analysis of rice (*Oryza sativa*) seeds during germination. *Proteomics* **7**, 3358–3368.
- Yin, X., He, D., Gupta, R., and Yang, P. (2015). Physiological and proteomic analyses on artificially aged *Brassica napus* seed. *Front. Plant Sci.* **6**, 1–11.
- Yoneda, K., Yoshida, M., Suzuki, I., and Watanabe, M.M. (2016). Identification of a major lipid droplet protein in a marine diatom *Phaeodactylum tricornutum*. *Plant Cell Physiol.* **57**, 397–406.
- 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. *J. Exp. Bot.* **64**, 293–302.
- 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 Physiol.* **127**, 1266–1278.

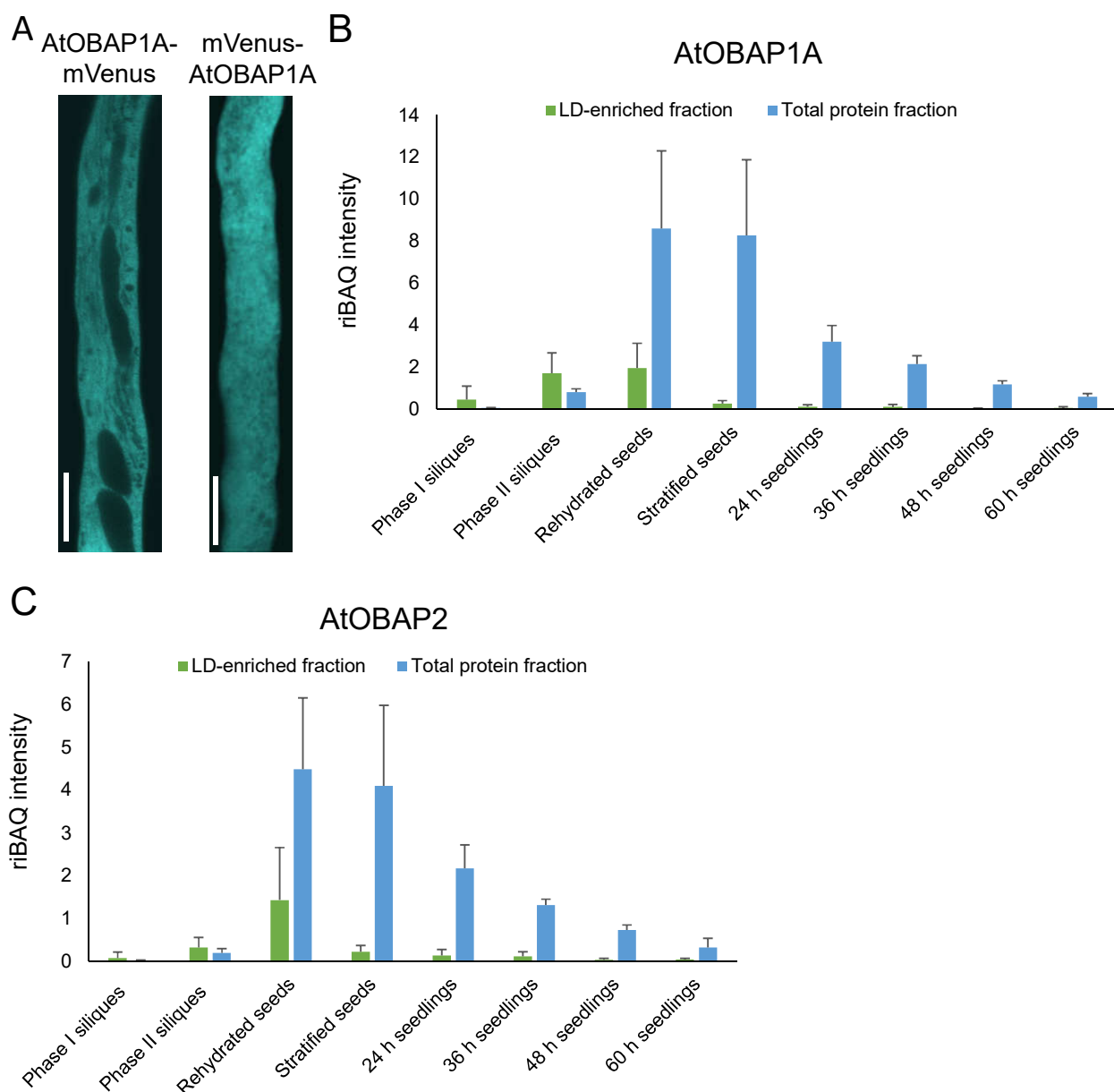


**Supplemental Figure 1. Post-germinative growth of *Arabidopsis thaliana* Col-0.** After a 74 h stratification period at 4 °C in the dark, seeds were incubated in long-day conditions (16 h light/8 h dark) for the given period of time. Germination *sensu stricto* was complete after 24 h. After 48 h, cotyledons had opened. Bar 1 cm

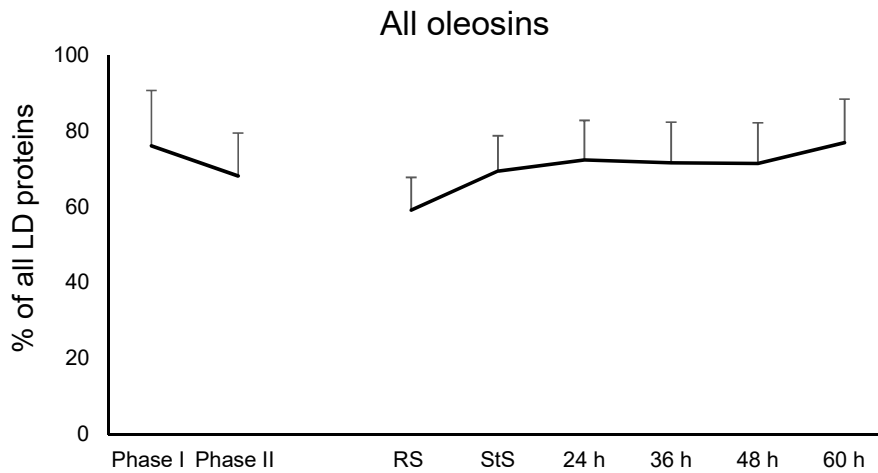




**Supplemental Figure 2. Abundance of LD proteins within the total protein fraction.** Abundance of LD protein within the total protein fraction increases during silique ripening. After germination *sensu stricto*, the abundance of LD proteins decreases constantly. All known LD proteins and LD protein identified in this study were taken into account.



**Supplemental Figure 3. Characterization of the AtOBAP protein family.** **A)** AtOBAP1A was cloned in both C- and N-terminal fusion with mVenus. The fluorescence of the fusion protein in transformed *Nicotiana tabacum* pollen tubes was homogenous, not indicating any membrane association. **B)** and **C)** The riBAQ intensity in the total and the LD-enriched fraction of AtOBAP1A and AtOBAP2. In the samples isolated from siliques, the riBAQ intensity is higher in the LD-enriched fraction than in the total protein fraction. During all other stages, the riBAQ intensities of both homologs is higher in the total protein fraction, and strongly depleted in the LD-enriched fraction. Bars 10  $\mu$ m,



**Supplemental Figure 4. Changes in contribution of all detected Oleosins to the complete LD proteome.** The contribution of each LD protein during each stage to the complete LD proteome (defined as presented in Suppl. Dataset 19) was calculated in % (Figures 9 and 10). Over the time points monitored, the contribution of the oleosins to the LD proteome is very stable.

#### 4. Article III: Arabidopsis lipid droplet-associated protein (LDAP) – interacting protein (LDIP) influences lipid droplet size and neutral lipid homeostasis in both leaves and seeds

This article was published online in the journal The Plant Journal in October 2017. The supplemental tables containing processed mass spectrometry data can be found online together with the full article:

<https://doi.org/10.1111/tpj.13754>

##### **Author contribution:**

Franziska K. Kretzschmar performed the LD isolation and proteomics experiment presented in Figure 1B and C with exception of the LC-MS/MS analysis. She processed and helped analyze the data and wrote the corresponding parts of the manuscript.

# Arabidopsis lipid droplet-associated protein (LDAP) – interacting protein (LDIP) influences lipid droplet size and neutral lipid homeostasis in both leaves and seeds

Michal Pyc<sup>1</sup>, Yingqi Cai<sup>2</sup>, Satinder K. Gidra<sup>1</sup>, Olga Yurchenko<sup>3</sup>, Sunjung Park<sup>3</sup>, Franziska K. Kretschmar<sup>4</sup>, Till Ischebeck<sup>4</sup>, Oliver Valerius<sup>5</sup>, Gerhard H. Braus<sup>5</sup>, Kent D. Chapman<sup>2</sup>, John M. Dyer<sup>3,\*</sup> and Robert T. Mullen<sup>1,\*</sup>

<sup>1</sup>Department of Molecular and Cellular Biology, University of Guelph, Guelph, ON N1G 2W1, Canada,

<sup>2</sup>Department of Biological Sciences, Center for Plant Lipid Research, University of North Texas, Denton, TX 76203, USA,

<sup>3</sup>US Department of Agriculture, Agricultural Research Service, US Arid-Land Agricultural Research Center, Maricopa, AZ 85138, USA,

<sup>4</sup>Department of Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences, University of Goettingen, Justus-von-Liebig-Weg 11, 37007 Goettingen, Germany, and

<sup>5</sup>Department of Molecular Microbiology and Genetics, Institute for Microbiology and Genetics, University of Goettingen, Grisebachstrasse 8, 37077 Goettingen, Germany

Received 10 August 2017; revised 25 September 2017; accepted 17 October 2017; published online 30 October 2017.

\*For correspondence (e-mails john.dyer@ars.usda.gov; rtmullen@uoguelph.ca).

## SUMMARY

Cytoplasmic lipid droplets (LDs) are found in all types of plant cells; they are derived from the endoplasmic reticulum and function as a repository for neutral lipids, as well as serving in lipid remodelling and signalling. However, the mechanisms underlying the formation, steady-state maintenance and turnover of plant LDs, particularly in non-seed tissues, are relatively unknown. Previously, we showed that the LD-associated proteins (LDAPs) are a family of plant-specific, LD surface-associated coat proteins that are required for proper biogenesis of LDs and neutral lipid homeostasis in vegetative tissues. Here, we screened a yeast two-hybrid library using the Arabidopsis LDAP3 isoform as 'bait' in an effort to identify other novel LD protein constituents. One of the candidate LDAP3-interacting proteins was Arabidopsis At5g16550, which is a plant-specific protein of unknown function that we termed LDIP (LDAP-interacting protein). Using a combination of biochemical and cellular approaches, we show that LDIP targets specifically to the LD surface, contains a discrete amphipathic  $\alpha$ -helical targeting sequence, and participates in both homotypic and heterotypic associations with itself and LDAP3, respectively. Analysis of *LDIP* T-DNA knockdown and knock-out mutants showed a decrease in LD abundance and an increase in variability of LD size in leaves, with concomitant increases in total neutral lipid content. Similar phenotypes were observed in plant seeds, which showed enlarged LDs and increases in total amounts of seed oil. Collectively, these data identify LDIP as a new player in LD biology that modulates both LD size and cellular neutral lipid homeostasis in both leaves and seeds.

**Keywords:** *Arabidopsis thaliana*, endoplasmic reticulum, LDAP, LDIP, lipid droplet, neutral lipids, organelle biogenesis, protein targeting.

## INTRODUCTION

Cytoplasmic lipid droplets (LDs) are evolutionarily conserved organelles that compartmentalize storage lipids such as sterol esters and triacylglycerols (TAGs). Structurally, LDs consist of a neutral lipid core uniquely enclosed by a single phospholipid monolayer and coated with a diverse array of 'coat' proteins that either bind the LD surface directly or embed themselves into the monolayer (Murphy, 2012). The formation of LDs occurs *de novo*

between the leaflets of the endoplasmic reticulum (ER) membrane in what is considered to be a complex, multi-step process (Thiam and Forêt, 2016; Barneda and Christian, 2017; Chen and Goodman, 2017). Briefly, the synthesis and accumulation of TAG within the ER bilayer serves to initiate the core of a nascent LD, which is thought to occur at specialized sites (subdomains) of the ER. Then, through the concerted action of various ER membrane and

soluble cytoplasmic proteins that are recruited to the growing LD, along with the continued addition of TAG to the LD core and coordinated changes in the phospholipid composition of the LD monolayer, the LD matures and eventually detaches from the ER into the cytoplasm; in some instances it can stay permanently connected with the ER. While the mechanistic details underlying these events have begun to be elucidated, albeit less so for plant LDs (Chapman *et al.*, 2012; Pyc *et al.*, 2017), the general process of LD biogenesis appears to be relatively well conserved among evolutionarily diverse species (Thiam and Beller, 2017). There is also a growing appreciation that LDs, long considered to be static fat depots that merely serve as energy reservoirs, are far more dynamic in nature and function in a multitude of cellular and physiological processes (Welte and Gould, 2017).

In plants, ER-derived cytoplasmic LDs have mostly been studied in pollen and oilseeds; in the latter they are responsible for providing the required carbon and energy for pre-photosynthetic development of the germinated seedling. In seeds, oleosins are the predominant LD coat proteins and play a role in both the formation and stabilization of LDs during seed desiccation (Huang, 1996; Purkrtova *et al.*, 2008; Laibach *et al.*, 2015; Pyc *et al.*, 2017). However, LDs are present in virtually all plant cell types, many of which are devoid of oleosins. This observation has recently led to increased efforts to identify LD proteins in non-seed tissues/organs (Gidda *et al.*, 2013; Horn *et al.*, 2013; Davidi *et al.*, 2015; Huang and Huang, 2016; Brocard *et al.*, 2017), as well as to probe the function of LDs in cell types that do not specialize in fat storage (Shimada *et al.*, 2014; Gidda *et al.*, 2016; McLachlan *et al.*, 2016; Zhang *et al.*, 2016).

Lipid droplet-associated proteins (LDAPs) are among the proteins implicated recently in LD biology in non-seed tissues (Horn *et al.*, 2013). These are considered to be functionally similar to the small rubber particle proteins (SRPPs) involved in the formation and stabilization of polyisoprenoid-containing LDs in rubber-producing plants (Berthelot *et al.*, 2016). In Arabidopsis, LDAPs constitute a three-member family (termed LDAP1–3) of ubiquitously expressed LD coat proteins that are critical for the proper maintenance and regulation of LDs during various developmental and stress-related processes (Gidda *et al.*, 2016). For instance, the abundance of LDs in leaves varies throughout the diurnal cycle, and this process is regulated, at least in part, by LDAP. Further, despite similarities of the LDAPs in terms of their ability to modulate LD abundance, exposure of plants to heat or cold stress results in the differential induction of LDAP genes, and analysis of loss-of-function mutants confirmed that specific LDAPs are required for the proliferation of LDs under different stress conditions (Gidda *et al.*, 2016). Consistent with these results, the LDAPs have also been implicated in drought

stress response as well as in overall plant growth and development (Kim *et al.*, 2010, 2016; Seo *et al.*, 2010; Brocard *et al.*, 2017), suggesting broader roles for LDAPs and LDs in plant physiological processes and stress adaptations.

Given that LDAPs are involved in multiple aspects of LD biology, and that LD coat proteins are known to function in general as part of highly regulated protein–protein interaction networks in other organisms (Tsai *et al.*, 2015; Kolkhof *et al.*, 2017), we employed Arabidopsis LDAP3 as ‘bait’ in a yeast two-hybrid (Y2H) assay screen to identify new proteins involved in biogenesis and/or function of plant LDs. We chose LDAP3 as the specific ‘bait’ since it is the most highly and ubiquitously expressed LDAP gene in Arabidopsis, including in seeds (Gidda *et al.*, 2016). One of the interacting proteins identified was Arabidopsis At5g16550, which, depending on the database, is annotated to be of unknown function or a putative voltage-dependent L-type calcium channel subunit, but has regions of similarity to the mycobacterial membrane protein large (MmpL) family of proteins and the promethin protein in mammals, both of which are known to be involved in various aspects of lipid metabolism (Yu *et al.*, 2004; Viljoen *et al.*, 2017). Here, we provide evidence that At5g16550, which we refer to hereafter as LDIP (LDAP-interacting protein), is a *bona fide* constituent of LDs in plant cells which physically interacts with LDAP3 on the LD surface; furthermore, it is also found on oleosin-containing LDs derived from germinated Arabidopsis seedlings. Analysis of two independent mutants with disruptions in LDIP gene expression revealed pronounced changes in LD abundance and morphology in plant leaves, with a reduction in the number of LDs and increases in LD size and variability, as well as increases in the total neutral lipid content in leaves. Similar changes in LD morphology and neutral lipid content were observed in LDIP mutant seeds. Together, these and other results identify LDIP as a new component of the cellular machinery involved in the modulation of LD abundance and size in plants; further, this proper compartmentalization is required for the maintenance of neutral lipid homeostasis in both leaves and seeds.

## RESULTS

### Identification of LDIP as a new LD protein in plant cells

The use of LDAP3 as ‘bait’ in a Y2H screen against a normalized Arabidopsis cDNA library resulted in identification of nine candidate interacting proteins (Table 1). To determine whether any of these proteins represented a new LD protein constituent, each was fused to the C-terminus of green fluorescent protein (GFP) then transiently expressed via *Agrobacterium tumefaciens* infiltration into *Nicotiana benthamiana* tobacco leaves and visualized using confocal laser-scanning microscopy (CLSM). As shown in Figure S1

**Table 1** List of candidate LDAP3-interacting proteins identified by yeast two-hybrid (Y2H) analysis

AGI no. <sup>a</sup>	Name	Y2H association <sup>b</sup>	Description <sup>c</sup>	SUBA localization <sup>d</sup>	GFP localization <sup>e</sup>
At2g39990	EIF2	Strong (1)	Translation initiation factor eIF2 p47 subunit homologue	Nucleus, cytoplasm	Cytoplasm
At3g21190	MSR1	Strong (2)	O-fucosyltransferase family protein	Golgi	Golgi
At4g08320	TPR8	Strong (1)	Tetratricopeptide repeat (TPR)-like superfamily protein	Nucleus	Cytoplasm
At1g55190	PRA7	Weak (1)	PRA1 (prenylated rab acceptor) family protein	Vacuole	ER
At1g79690	NUDT3	Weak (1)	Nudix hydrolase homologue 3	Cytoplasm	Cytoplasm
At5g13420	TRA2	Weak (2)	Aldolase-type triosephosphate isomerase (TIM barrel) family protein	Plastid	Cytoplasm
At5g16550	LDIP <sup>f</sup>	Weak (6)	Unknown protein; putative voltage-dependent L-type calcium channel subunit	ER	LD
At5g17920	ATMS1	Weak (1)	Cobalamin-independent synthase family protein	Cytoplasm	Cytoplasm
At5g24420	PGL5	Weak (1)	6-phosphogluconolactonase 5	Cytoplasm	Cytoplasm

<sup>a</sup>The Arabidopsis gene identifier (AGI) number represents the systematic designation given to each locus, gene and its corresponding protein product(s) by The Arabidopsis Information Resource (TAIR; <http://www.arabidopsis.org>).

<sup>b</sup>Strong and weak interacting 'prey' proteins identified in the Y2H screen based on the relative growth of yeast on selective media. The number of times that each 'prey' protein was identified in the screen is indicated in parentheses.

<sup>c</sup>Information on protein function and/or homology summarized from TAIR, Araport (<https://www.araport.org>) and SUBA (Subcellular Localization Database for Arabidopsis Proteins; <http://suba.live/>).

<sup>d</sup>Protein intracellular (consensus) localization based on SUBA.

<sup>e</sup>Intracellular localization of C-terminal GFP-tagged fusion protein in *Agrobacterium*-infiltrated tobacco leaf epidermal cells (refer to Figure S1).

<sup>f</sup>Named LDIP (lipid drop-associated protein-interacting protein) in this study.

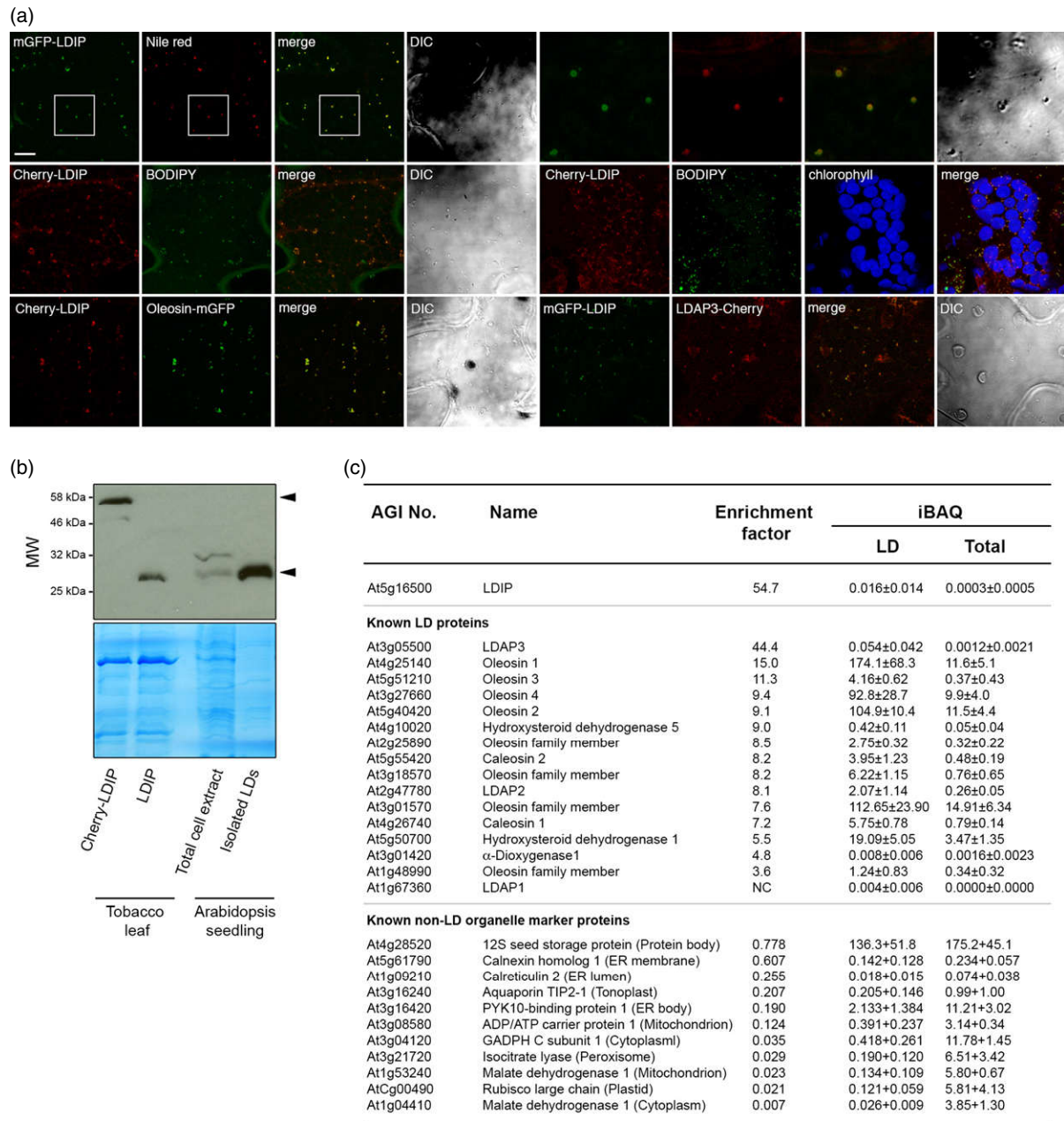
in the Supporting Information, GFP-At5g16550 (GFP-LDIP) was the only fusion protein that localized to Nile red-stained LDs. All other candidate proteins displayed exclusively cytoplasmic localizations, except for GFP-At1g55190 (i.e. PRA7) and GFP-At3g21190 (i.e. MSR1), which were localized to the ER and Golgi, respectively (Figure S1). As shown Figure 1(a), high-magnification images of LDIP appended to a monomeric version of GFP (mGFP-LDIP) in tobacco epidermal cells revealed distinct toroidal shapes that encircled the Nile red-stained LD cores, indicating that mGFP-LDIP is localized to the surface of the LD (Figure 1a). Localization to LDs was also observed in BODIPY-stained epidermal and mesophyll cells of tobacco leaves when LDIP was fused to the Cherry fluorescent protein (Cherry-LDIP) (Figure 1a). Similarly, Cherry- and mGFP-tagged LDIP co-localized with the LD marker proteins oleosin-mGFP and LDAP3-Cherry (Figure 1a).

To confirm that LDIP represents a *bona fide* LD protein in plants, 2-day-old, germinated Arabidopsis seedlings were harvested and homogenized and then LDs were purified by differential centrifugation. The resulting total cell and isolated LD fractions were then analysed using Western blotting or mass spectrometry (MS)-based proteomics. Antibodies were raised to a unique peptide sequence in Arabidopsis LDIP (see Experimental Procedures for details) and the quality of antibodies was verified by Western blot analysis of proteins extracted from tobacco leaves transiently expressing either Cherry-tagged or non-tagged LDIP. As shown in Figure 1(b), bands of the expected sizes

were observed for both Cherry-LDIP (about 55.4 kDa) and native LDIP (about 26.6 kDa) proteins. A band of the expected size for native LDIP was also observed in total protein extracts derived from Arabidopsis seedlings, and this protein was strongly enriched in the purified LD fraction (Figure 1b). Moreover, the presence of LDIP in these fractions was confirmed by label-free quantitative proteomics, which showed that while LDIP occurs at relatively low abundance it was clearly enriched in the purified LD fraction, along with several other known LD marker proteins (Figure 1c). Marker proteins for various other subcellular compartments, however, were not enriched in the LD fraction (Figure 1c) (refer to Tables S1 and S2 for all the proteomics data obtained). Taken together with the results above, these data indicate that LDIP represents a new LD protein in plants.

#### **LDIP is a plant-specific protein that is constitutively expressed in Arabidopsis**

To gain insight to the evolutionary history of LDIP and the distribution of homologues within the plant kingdom, phylogenetic analyses were performed comparing the deduced polypeptide sequence for Arabidopsis LDIP with other protein sequences encoded in extant genomic databases. The results indicated that LDIP is a plant-specific protein with closely related homologues found throughout most of the plant kingdom, including in bryophytes, gymnosperms, monocots and dicots (Figure 2a). However, no homologues were detected in algal species or in yeasts



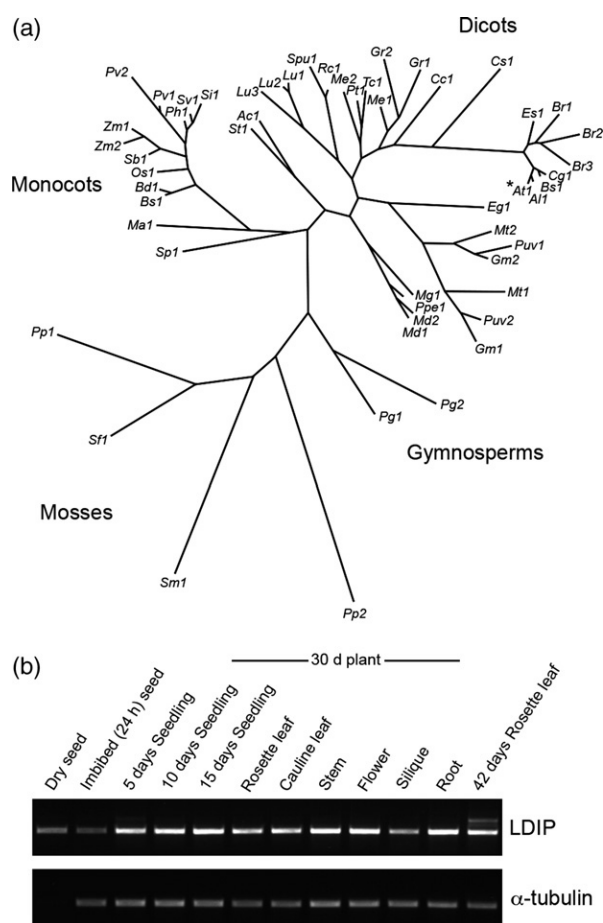
**Figure 1.** Localization of lipid drop-associated protein (LDAP)-interacting protein (LDIP) in plant cells.

(a) Representative confocal laser-scanning microscopy images of tobacco leaf cells transiently (co)expressing (as indicated by labels) mGFP- or Cherry-tagged LDIP and either oleosin-mGFP or LDAP3-Cherry, or stained with the neutral lipid-selective dye Nile red or BODIPY. For each set of images the corresponding merged and differential interference contrast (DIC) images or the corresponding image of endogenous chlorophyll autofluorescence of the mesophyll cells (middle row, right set of images) is shown; all other sets of images in (a) are of leaf epidermal cells. Also, images of the mesophyll cells are a z-stack series, whereas all other images shown are single (individual) optical sections. Boxes (top row) represent the portion of the cell shown at higher magnification to the right, revealing the localization of mGFP-At5g16500 (LDIP) to the surface of several, individual Nile red-stained lipid droplets (LDs). Bar = 20 µm.

(b) Western blot analysis of total cell protein extracts from tobacco leaves transiently transformed (via *Agrobacterium* infiltration) with Cherry-tagged LDIP or non-tagged LDIP and total cell protein and isolated LD protein extracts from 2-day-old Arabidopsis seedlings. The relative positions of expressed Cherry-LDIP or non-tagged LDIP in tobacco or native (endogenous) LDIP in Arabidopsis are indicated with arrowheads. Positions of molecular mass markers are also indicated. Also shown is the corresponding Coomassie-stained gel of the same protein extracts that were analysed by Western blotting with anti-LDIP IgGs.

(c) Identification of LDIP in the label-free LD proteome of Arabidopsis seedlings. The identical 2-day-old Arabidopsis seedling total cell protein and LD fractions used in (b), as well as two additional replicates ( $n = 3$ ), were subjected to LC-MS/MS-based shotgun proteomic analysis. Protein levels were calculated using the intensity-based quantification (iBAQ) label-free algorithm. The values were normalized, setting the sum of all abundances to 1000. Depicted are iBAQ values for LDIP and selected known LD and non-LD proteins (AGI numbers and protein names were obtained from the TAIR database) as the mean of the total values of three separate samples (±SD). Also depicted is the corresponding LD/total cell enrichment ratio as determined for each protein. See Table S1 and S2 for the iBAQ values and enrichment ratios for all proteins identified in all samples.





**Figure 2.** Phylogenetic analysis of lipid drop-associated protein-interacting proteins (LDIPs) and expression of *LDIP* in Arabidopsis.

(a) Phylogenetic tree depicting the evolutionary relationship of LDIP-like proteins across various plant species, including those found in distinct clades corresponding to mosses, gymnosperms, monocots and dicots. Each protein is labelled with the respective genus and species, and numerically, including those that exist as multiple homologues in some species. Arabidopsis LDIP (*At1*) is indicated with an asterisk.

(b) RT-PCR analysis of *LDIP* gene expression in various organs and developmental stages in Arabidopsis, as indicated by labels.  $\alpha$ -tubulin served as an endogenous control.

and metazoans. In Arabidopsis, LDIP is present as a single copy, whereas two or sometimes three homologues exist in some other plant species (Figure 2a), reflecting whole-genome duplication events known to have occurred during plant evolution (Paterson *et al.*, 2010).

To determine the expression pattern of *LDIP* in Arabidopsis, reverse transcription–polymerase chain reaction (RT-PCR) assays were used to assess the transcript levels in mRNA derived from various organs at different developmental stages. As shown in Figure 2(b), *LDIP* is expressed at comparable levels in all organs examined during plant growth and development, including in mature dry seeds. This expression pattern is similar to that of the Arabidopsis *LDAP* genes, especially *LDAP2* and *LDAP3*, which are also

constitutively expressed, including in seeds (Gidda *et al.*, 2016). However, the pattern is different from certain other LD-specific proteins, such as oleosins, which are expressed predominantly in seeds and pollen (Kim *et al.*, 2002; Ischebeck, 2016). Taken together, these observations suggest that LDIP is probably involved in LD-related processes throughout the entire plant life cycle, including in both leaves and seeds.

#### LDIP contains an amphipathic $\alpha$ -helix-type LD targeting signal

To further analyse the relationship of LDIP with LDs, we next investigated whether the protein contains a discrete LD targeting sequence. To help guide the design of these experiments, we analysed the Arabidopsis LDIP sequence along with several phylogenetically diverse homologues using both polypeptide sequence alignments and hydropathy profile analyses (Figure 3a, b). The results indicated that each of the proteins has three fairly distinct regions: (i) an N-terminal section (amino acid residues 1–106 in Arabidopsis LDIP) that possess relatively little sequence similarity/identity and is hydrophilic overall, with the exception of a mildly hydrophobic segment (residues 50–65 in Arabidopsis LDIP); (ii) a middle section (residues 107–211 in Arabidopsis LDIP) consisting of a more highly conserved and strongly hydrophobic sequence, including several predicted transmembrane-spanning domains (Figure 3b); and (iii) a C-terminal section (residues 212–249 in Arabidopsis LDIP) that, similar to the N-terminal region, is a less conserved and primarily hydrophilic sequence. Notably, BLASTP analysis with Arabidopsis LDIP or its homologues revealed little sequence conservation with any other proteins, although the middle hydrophobic region possesses some similarity to domains present in the MmpL family and the mammalian promethin protein (Yu *et al.*, 2004; Viljoen *et al.*, 2017).

Based on this information, we constructed a series of truncation mutants of Arabidopsis LDIP that were subsequently fused to the Cherry fluorescent protein, and then transiently expressed the proteins in tobacco epidermal cells to assess their subcellular localization. As shown in Figure 3(c), Cherry-LDIP<sup>1–106</sup>, which lacks the C-terminal and hydrophobic middle regions of the protein, localized to LDs in a manner similar to full-length Arabidopsis LDIP (Cherry-LDIP), indicating that the N-terminal region of the protein (i.e. residues 1–106) is sufficient for targeting LDs. Deletion of this region confirmed it was also necessary for LD targeting, since Cherry-LDIP<sup>211–249</sup>, consisting of the C-terminal section alone, and Cherry-LDIP<sup>107–249</sup>, consisting of the C-terminal and hydrophobic middle sections together, mislocalized to the cytoplasm and ER, respectively (Figures 3c and S2).

To gain additional insight into the nature of the N-terminal LD targeting signal in Arabidopsis LDIP, we analysed

the 1–106 polypeptide sequence for possible structural features including hydrophobic domains and amphipathic  $\alpha$ -helices, which are known to serve as LD targeting signals in mammals and yeast (Kory *et al.*, 2016; Bersuker and Olzmann, 2017). This analysis revealed a possible amphipathic  $\alpha$ -helix (Figure 3d) that corresponded to the moderately hydrophobic sequence present in the N-terminal region (residues 50–65, see Figure 3b). The amphipathic characteristics of this hydrophobic region were relatively well conserved in other LDIP homologues (Figures 3b and S3). Moreover, analysis of the full-length Arabidopsis LDIP protein sequence using the structural homology-modelling tool available at SWISS-MODEL showed that while there were no sequences that were sufficiently similar for modelling the entire 3D structure of the protein, a significant portion of region 1–106 (residues 32–115; italicized in Figure 3a) matched several known protein structures available in the database. Development of a homology model for this N-terminal region of Arabidopsis LDIP revealed five secondary structures (Figure 3e; see also Figure 3b), the second of which was an  $\alpha$ -helix that corresponded almost perfectly to the region predicted to form the amphipathic  $\alpha$ -helix (Figure 3d, e).

To help determine the importance of the predicted amphipathic  $\alpha$ -helix sequence for LDIP targeting to LDs, we generated two mutant versions of Cherry-LDIP<sup>1–106</sup>: (i) Cherry-LDIP<sup>1–106</sup> Mut1, whereby amino acids 50–65 in LDIP comprising the hydrophobic sequence and the majority of the amphipathic helix (Figure 3d), were deleted; and (ii) Cherry-LDIP<sup>1–106</sup> Mut2, whereby the two isoleucine residues at positions 55 and 59 located near the centre of the hydrophobic face of the amphipathic helix (refer to Figure 3d) were replaced with hydrophilic glutamic acids. As shown in Figure 3(f), both mutant fusion proteins, in contrast to LD-localized, wild-type Cherry-LDIP<sup>1–106</sup> (Figures 3c and S2), mislocalized to the cytoplasm in tobacco leaf epidermal cells, suggesting that LDIP contains an amphipathic-helix-type LD targeting signal similar to that found in other proteins that target to LDs via the cytoplasm.

#### LDIP interacts with itself and LDAP3 in yeast and plant cells

We next performed a more detailed assessment of the physical interactions between Arabidopsis LDIP and LDAP3. As shown in Figure 4(a), co-expression of full-length LDIP and LDAP3 as 'bait' and 'prey' fusion proteins, respectively, in the Y2H system resulted in cell growth under selective conditions, as expected (see Table 1). Similar results were observed when the orientation of proteins was switched between the bait and prey vectors (Figure S4), and also when LDIP was paired with itself (Figure 4a). Cell growth was not observed, however, when LDIP was co-expressed with a mutant version of LDAP3 lacking its C-terminal 100 amino acids (i.e. LDAP3 $\Delta$ C100)

(Figure 4a), which disrupts the association of LDAP3 with LDs in plant cells (Gidda *et al.*, 2016). Cell growth was also abolished when each protein was expressed with the corresponding empty vector control, as expected (Figures 4a and S4). Western blotting was used to confirm the presence of proteins in all yeast strains tested (Figure S5). Taken together, these data indicate that LDIP and LDAP3 interact in yeast cells, and that LDIP is capable of homotypic association.

To characterize the interaction of Arabidopsis LDIP and LDAP3 *in planta*, the proteins were fused to the C-terminal half of cyan fluorescent protein (cCFP) and the N-terminal half of the yellow fluorescent protein Venus, respectively, which by themselves are not fluorescent. The resulting fusion proteins were then transiently expressed in tobacco leaf epidermal cells, and protein–protein interactions assessed based on bimolecular fluorescence complementation (BiFC). Given that BiFC can result from false positive interactions (i.e. the two halves of the fluorescent proteins can assemble if the proteins being tested are merely in close proximity to one another and are not necessarily interacting (see Stefano *et al.* (2015) and Xing *et al.* (2016))), two experiments were performed in parallel: tobacco leaves were co-infiltrated with either cCFP-LDIP and nVenus-LDAP3 or cCFP-LDIP and nVenus-LDAP3 $\Delta$ C100, the latter serving as a negative control since LDAP3 $\Delta$ C100 does not target to LDs (Gidda *et al.*, 2016). In addition, cells were transformed with Cherry-PeroX, which is a marker protein that targets to peroxisomes (Ching *et al.*, 2012) and was used to identify transformed cells, regardless of whether a BiFC signal was present or not. As shown in Figure 4(b) and quantified in Figure 4(c), co-expression of cCFP-LDIP and nVenus-LDAP3 resulted in numerous BiFC puncta, and staining of these cells with monodansylpentane (MDH), a blue-fluorescent neutral lipid dye (Yang *et al.*, 2012), confirmed that the puncta were LDs (Figure 4d). Further, the LDs were aggregated in a manner similar to the LDs in cells transformed with GFP-LDIP (compare the images in Figures 4d and S1). By contrast, co-expression of cCFP-LDIP and nVenus-LDAP3 $\Delta$ C100 resulted in significantly fewer BiFC puncta compared with co-expression of full-length proteins (Figure 4b, c). The decrease in fluorescence of the negative control was not due to differences in transgene expression, since RT-PCR analysis confirmed similar levels of gene transcripts in both experiments (Figure S6).

Biochemical evidence in support of interaction of Arabidopsis LDIP and LDAP3 was obtained using a GFP-based affinity-capture method, whereby GFP-LDIP was first transiently expressed in tobacco leaves, then leaves were homogenized, proteins were solubilized in detergent and GFP-LDIP and interacting proteins were isolated using the GFP trap system (see Experimental Procedures for details). Co-purifying proteins were then identified using MS and peptide mass fingerprinting, which identified a number of

different LD proteins, including endogenous tobacco LDAPs and SEIPINs, as well as a tobacco homologue of Arabidopsis LDIP (Figure 4e). These results were consistent with Y2H analyses (Figure 4a), which showed both homotypic and heterotypic associations between LDIP and LDAP3. Furthermore, co-expression in tobacco leaves of GFP-LDIP with the Arabidopsis LEAFY COTYLEDON2 (LEC2) transcription factor, which induces genes for seed oil-like synthesis in leaves (Santos Mendoza *et al.*, 2005; Vanhercke *et al.*, 2017), resulted in co-purification of GFP-LDIP with several tobacco oleosins (Figure 4e), which are known to be induced by LEC2 when expressed in tobacco leaves (Kim *et al.*, 2013). None of these GFP-LDIP co-purifying proteins were identified when GFP was expressed on its own, with or without LEC2, in tobacco leaves (Figure 4e). Taken together, the data in Figure 4 provide strong evidence that LDIP and LDAPs interact at the surface of LDs, and furthermore that LDIP might have a role in conjunction with oleosins in plant seeds.

### Disruption of *LDIP* alters LD morphology and abundance and increases neutral lipid content in leaves

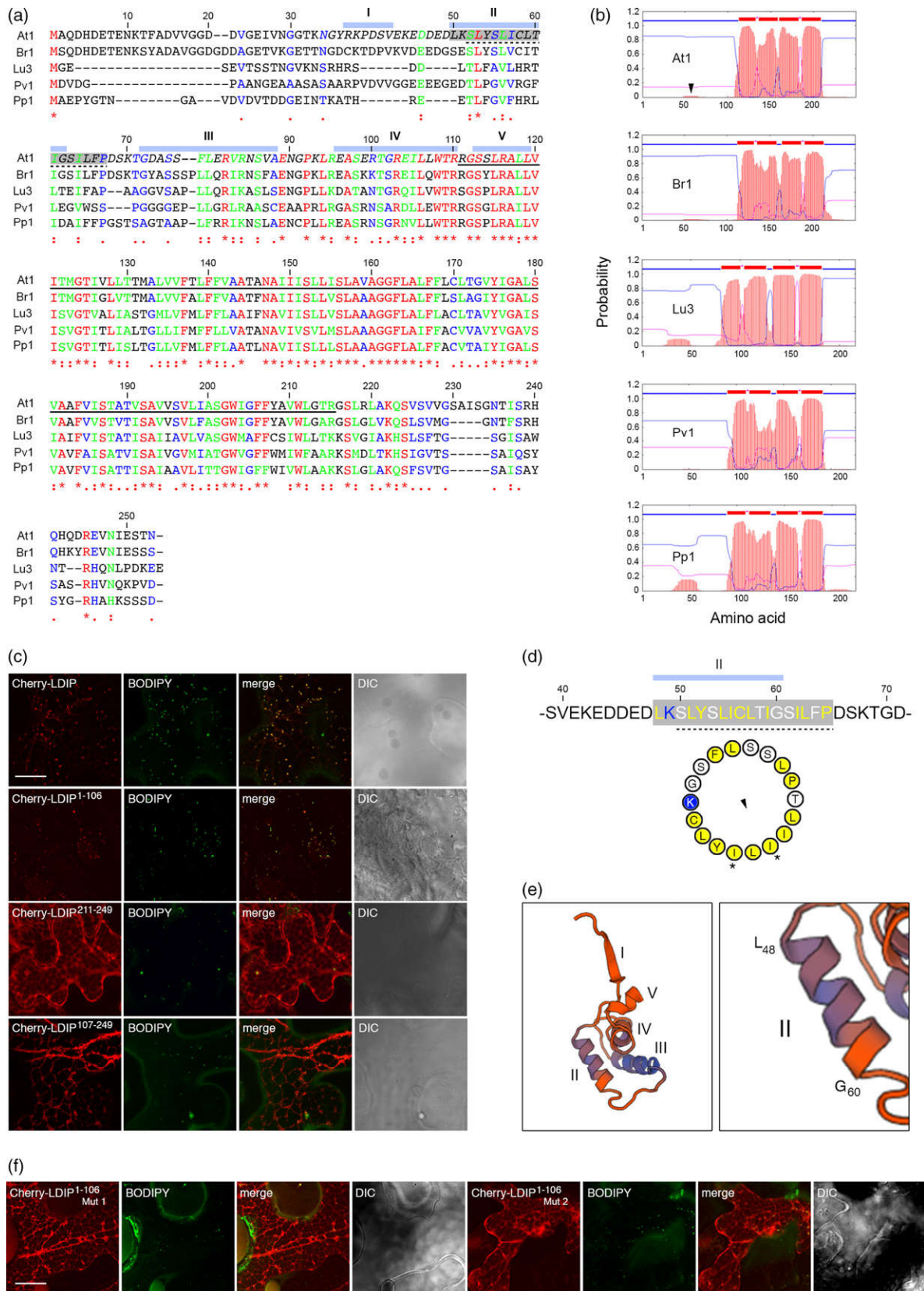
To assess the function of LDIP in plants, we obtained two Arabidopsis *LDIP* T-DNA mutants from the Arabidopsis Biological Resource Center: SALK\_084555, which contains a T-DNA inserted in the first intron, and SAIL\_335\_H11, which contains a tandem T-DNA inserted in the first exon (Figure S7). Each line was genotyped

and advanced to homozygosity, then expression of *LDIP* was determined using RT-PCR. The results revealed that SALK\_084555 was a knockdown (KD) mutant with reduced gene expression, while SAIL\_335\_H11 was a knockout (KO) mutant with no full-length transcripts detected (Figure S7).

Prior studies showed that the abundance of LDs in plant leaves changes throughout the diurnal cycle, with the highest number of LDs being present at the end of the dark period and the lowest numbers present at the end of the light period (Gidda *et al.*, 2016). To determine whether disruption of *LDIP* had any effects on LDs in leaves, wild-type (WT), KD and KO Arabidopsis lines were germinated on ½ MS plates, then 15-day-old seedlings were formaldehyde-fixed at either the end of the dark cycle or the end of the light cycle, followed by visualization of LDs with BODIPY staining and CLSM. As shown in Figure 5(a), disruption of *LDIP* resulted in progressive decreases in the abundance of LDs in both KD and KO lines compared with WT. This trend was observed at the end of both the dark and light periods (Figure 5a). Furthermore, disruption of *LDIP* led to progressively greater variation in LD size, with the appearance of several prominent 'supersized' LDs (i.e. LDs  $\geq 2$   $\mu\text{m}$  diameter) in the KO leaves (Figure 5a). Indeed, these larger LDs were even more obvious, and clearly external to chloroplasts, in three-dimensional (3D) projections of surface-rendered, highly magnified CLSM images of cells (Figure 5a). Supersized LDs were also observed in older (i.e. 28 days

**Figure 3.** Identification of a lipid drop (LD)-targeting signal in lipid drop-associated protein-interacting protein (LDIP).

- (a) Deduced polypeptide sequence alignment of Arabidopsis LDIP and selected homologues. Proteins are labelled with the respective genus and species and numerically, and correspond to the same labels as in Figure 2(a): *Arabidopsis thaliana* (At1), *Brassica rapa* (Br1), *Linum usitatissimum* (Lu3), *Panicum virgatum* (Pv1) and *Physcomitrella patens* (Pp1). Identical and similar amino acid residues in each protein are highlighted red and blue or green and indicated with asterisks and colons or periods, respectively. Numbers above each row of sequences represent specific amino acid residues. The Arabidopsis LDIP sequence (i.e. residues 32–115) that matches several known protein structures available in the SWISS-MODEL database is italicized and Roman numerals (I–V) and blue-shaded boxes correspond to the regions within this sequence that are predicted to form distinct secondary structures, as illustrated in (e). Also indicated for the Arabidopsis LDIP sequence is the region predicted to form an amphipathic  $\alpha$ -helix (highlighted with a grey background; refer also to (d)), the overlapping, moderately hydrophobic sequence (residues 50–65, dashed underline; refer also to (b)), and the middle, strongly hydrophobic section of the protein (residues 107–211, solid underline; refer also to (b)).
- (b) Hydropathy profiles of the deduced polypeptide sequence of Arabidopsis LDIP and selected homologues based on the TMHMM algorithm (Krogh *et al.*, 2001). Proteins shown are the same as those in (a). Note the strong hydrophobic sequence in the middle portion of each protein, as well as the relatively moderate hydrophobic sequence located in the vicinity of residue 50 in each protein, indicated with the arrowhead in Arabidopsis LDIP.
- (c) Truncation analysis of Arabidopsis LDIP targeting to LDs in tobacco leaf cells. Representative confocal laser-scanning microscopy (CLSM) images of leaf epidermal cells transiently expressing (as indicated by labels) full-length or truncated versions Cherry-tagged LDIP along with the corresponding BODIPY-stained LDs, as well as the corresponding merged and differential interference contrast (DIC) images for each set of images. The numbers in the name of each mutant construct denote the amino acid residues in LDIP that were fused to Cherry fluorescent protein. Bar = 20  $\mu\text{m}$ .
- (d) Helical wheel projection of amino acid residues 48–65 in Arabidopsis LDIP. Shown is a portion of the Arabidopsis LDIP sequence (numbers represent specific amino acid residues) that includes the region predicted by HeliQuest to form an amphipathic  $\alpha$ -helix (residues 48–65, highlighted with a grey background), which is also depicted in the  $\alpha$ -helical wheel projection. Hydrophobic amino acid residues are coloured yellow, hydrophilic and charged residues are white and red or blue, respectively. The direction of the arrowhead in the helical wheel indicates the position of the hydrophobic face along the axis of the helix. The sequence predicted, based on SWISS-MODEL, to form an  $\alpha$ -helix (residues 48–60, region II) is indicated by the blue-shaded box (refer also to (a) and (e)). The dashed line indicates the sequence (i.e. residues 50–65) that was deleted from Cherry-LDIP<sup>1–106 Mut1</sup>; asterisks indicate the two isoleucine residues i.e. I<sub>55</sub> and I<sub>59</sub>, that were replaced with glutamic acid residues in Cherry-LDIP<sup>1–106 Mut2</sup> (refer to (f)).
- (e) Structural modelling of Arabidopsis LDIP. BLAST alignments performed at the SWISS-MODEL website identified a region in Arabidopsis LDIP, i.e. residues 32–115 (refer to (a)) that was sufficiently similar to proteins with known 3D structures that allowed for homology-based modelling. Shown (left panel) are the relative positions of the five secondary structures (domains I–V) modelled in the 32–115 amino acid sequence of LDIP. Shown also (right panel) is a close-up image of the predicted  $\alpha$ -helix, beginning at leucine 48 and ending at glycine 60 (i.e. domain II; refer also to (d)).
- (f) Mutational analysis of the predicted amphipathic  $\alpha$ -helix in Arabidopsis LDIP and effects on targeting to LDs in tobacco leaf cells. Representative CLSM images of leaf epidermal cells transiently expressing (as indicated by labels) modified versions of Cherry-tagged LDIP<sup>1–106</sup> (see text for details), along with the corresponding BODIPY-stained LDs, as well as the corresponding merged and DIC images. Bar = 20  $\mu\text{m}$ .



old) leaves of KO plants (Figure S8), indicating that this LD phenotype persists in other stages of plant growth and development.

To determine whether the changes in LD abundance and size in *LDIP* mutants affected leaf neutral lipid content, total lipids were extracted from 15-day-old seedlings at the end of both dark and light periods then neutral lipids and polar lipids were isolated by solid phase extraction, and fatty acid content determined using gas chromatography with flame ionization detection (GC/FID). As shown in Figure 5(b), neutral lipids were progressively increased in the KD and KO lines compared with WT, and these increases were apparent at the end of the dark period and, in the KO line, also at the end of the light period. Analysis of fatty acid composition revealed that these increases were due primarily to increases in linoleic (18:2) and linolenic (18:3) acids (Figure 5b), which are the most abundant fatty acids in leaf tissues (Li-Beisson *et al.*, 2013). By contrast, analysis of polar lipids showed no obvious changes in content or composition between WT, KD and KO lines (Figure 5c), suggesting that the effects of *LDIP* disruption were specific to neutral lipid homeostasis.

Taken together, these data indicate that loss of *LDIP* results in a decrease in LD abundance in plant leaves, an increase in the variability of LD size, the appearance of drastically enlarged LDs and changes in neutral lipid metabolism that result in increases in total neutral lipid content. Additional evidence that *LDIP* is responsible for these phenotypes comes from experiments where the KO line was backcrossed with WT, which resulted in normal LDs in  $F_1$  plants, then a distribution of both normal and supersized LDs in a sorting  $F_2$  population that was consistent with the phenotype being associated with a single Mendelian locus (Figure S9). Further, transformation of the KO line with either non-tagged *LDIP* or Cherry-*LDIP* resulted in complementation of the supersized LD phenotype in  $T_2$  transgenic Arabidopsis lines, and Cherry-*LDIP* targeted to LDs in these plant lines, just as Cherry-*LDIP* did in WT plants when expressed under control of its native promoter (Figure S10).

#### Disruption of *LDIP* alters LD morphology and increases neutral lipid content in seeds

The observed expression of *LDIP* in seeds (Fig. 2), the co-enrichment of *LDIP* and oleosins in LDs isolated from germinated seedlings (Fig. 1c) and the co-immunoprecipitation of *LDIP* and oleosins from tobacco leaves transformed with LEC2 (Figure 4) prompted us to examine whether disruption of *LDIP* had any effects on LDs and/or neutral lipid content in seeds. As shown in Figure 6(a), CLSM analysis of mature dry seeds stained with BODIPY revealed the presence of noticeably enlarged LDs in cotyledons of KD and KO lines compared with WT. Analysis of lipids further revealed significant and progressive

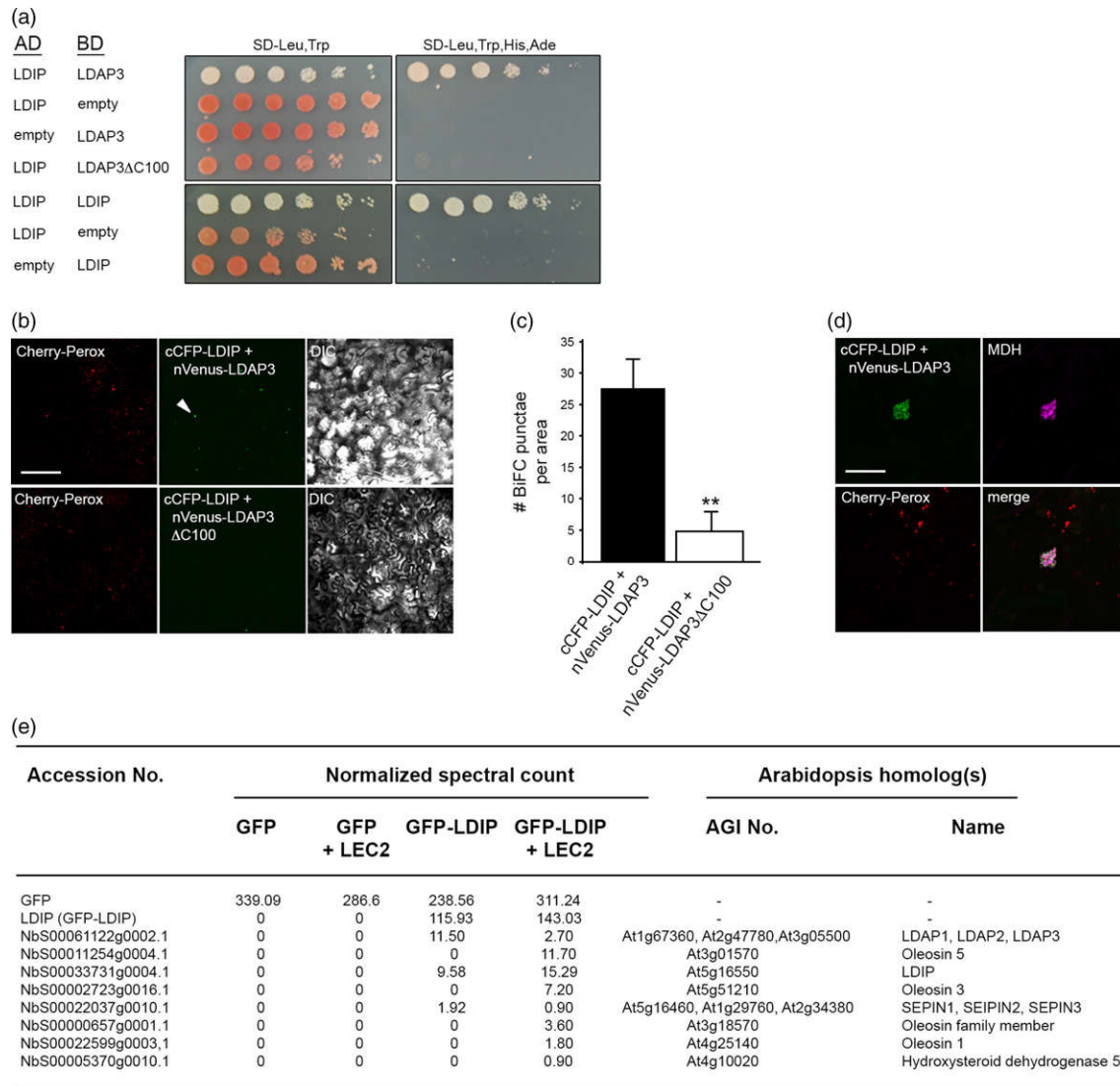
increases in the oil content of mature dry seeds in KD and KO lines relative to WT (Figure 6b), and these increases were due to increases in nearly all of the seed oil fatty acids (Figure 6c). The enlarged LDs and relatively higher neutral lipid content persisted throughout germination and seedling establishment, and notably more so in the KD line. While the reason for this more persistent phenotype in the KD line is currently unclear, RT-PCR analysis revealed that while the KO line did not produce any normal-length transcripts (Figure S7), it did possess transcripts that included a portion of the first exon of the *LDIP* sequence and the T-DNA insert sequence (Figure S11). Since translation of this latter RNA would generate a truncated version of the *LDIP* protein that also contains the LD targeting signal (Figure S11), perhaps this truncated protein contributed to changes in the LD proteome that affected the rate of breakdown of stored lipids during seedling establishment, different from that of the KD line. Regardless, by 4 days after the initiation of germination, the LDs and storage oil were largely degraded in both the KD and KO lines, as in the WT (Figure 6a–c). Taken together with the results presented above for leaves (Figure 5), these data clearly identify *LDIP* as a new and important player in neutral lipid metabolism and the modulation of LD abundance and size.

## DISCUSSION

### Identification of *LDIP* as an LD-localized, plant-specific protein

The identification and characterization of the Arabidopsis LDAPs has increased our understanding of plant LD biology as a whole and provided new insights into the roles of LDs, particularly in non-seed tissues (Gidda *et al.*, 2013, 2016; Horn *et al.*, 2013; Kim *et al.*, 2016). To identify additional proteins involved in LD biogenesis and/or function, we used LDAP3 as bait in a Y2H screen and identified nine candidate interacting proteins. Microscopic analysis of transiently expressed fluorescent protein-tagged versions of all nine proteins determined that only GFP-*LDIP* was targeted to LDs (Figure S1), although the non-LD localization of the others does not preclude them from having a role in LD-related processes (e.g. LDs might interact with other organelles or with proteins in the cytoplasm). *LDIP* was subsequently confirmed as a *bona fide* constituent of LDs using a variety of approaches, including co-localizations with oleosin or LDAP at LDs in tobacco leaves (Figure 1), isolation of LDs from germinating seedlings followed by Western blotting or proteomics, which showed clear enrichment of Arabidopsis *LDIP* in the LD fraction (Figure 1), and stable expression of Cherry-tagged *LDIP* in transgenic plants, which confirmed localization to LDs in Arabidopsis leaves and complemented the *LDIP* KO phenotype (Figure S10). Consistent with these data, *LDIP* was also identified as a





**Figure 4.** Interaction of lipid drop-associated protein (LDAP)-interacting protein (LDIP) and LDAP3 in yeast and plant cells.

(a) Interaction of LDIP with itself and LDAP3 in the yeast two-hybrid (Y2H) assay. Yeast strains were co-transformed with the indicated pairs of GAL4-activating domain (AD) and GAL4-binding domain (BD) fusion proteins or the corresponding 'empty' AD or BD plasmids, serving as negative controls. Serial dilutions of cells were spotted onto plates containing either low-stringency [synthetic dextrose media lacking leucine and tryptophan (SD-Leu, Trp)] or high-stringency (SD-Leu, Trp, His, Ade) selection, the latter of which requires protein-protein interactions for yeast growth.

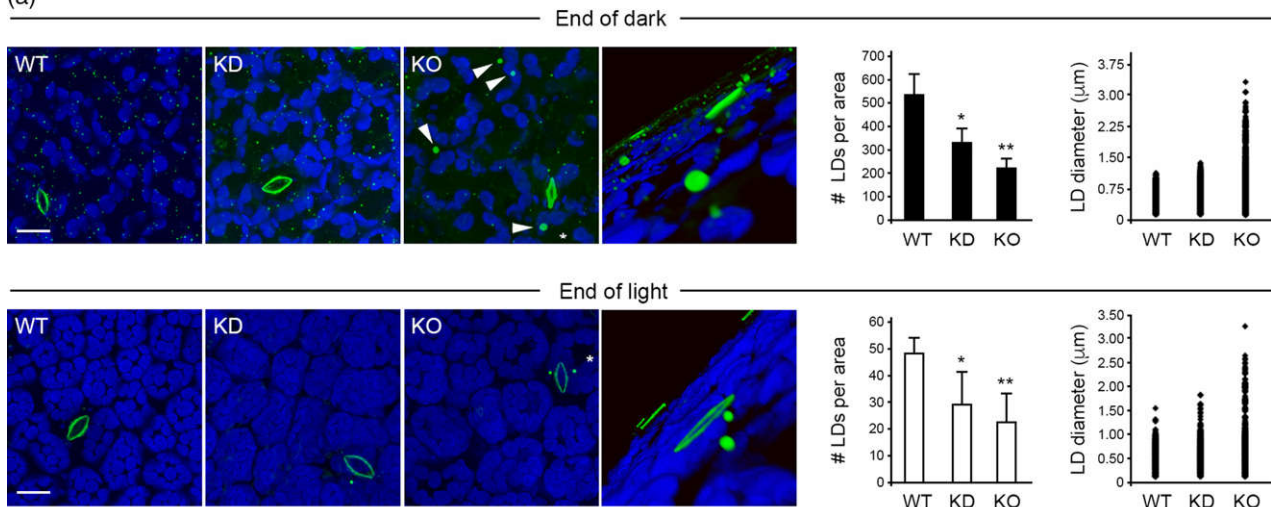
(b) Interaction of LDIP and LDAP3 in the bimolecular fluorescence complementation (BiFC) assay in tobacco leaf cells. Representative confocal laser-scanning microscopy (CLSM) images of leaf epidermal cells transiently (triple) transformed (via *Agrobacterium* infiltration) with Cherry-PeroX, which serves as a transformation control, cCFP-LDIP, and either nVenus-tagged full-length LDAP3 or LDAP3 lacking its C-terminal 100 amino acid residues (nVenus-LDAP3ΔC100). Also shown are the corresponding differential interference contrast images. Note the relative abundance of the BiFC puncta (arrowhead), which are aggregates of LDs (refer to (d)), in areas of cells transformed (based on Cherry-PeroX fluorescence) with cCFP-LDIP and nVenus-LDAP3, compared with cells transformed with cCFP-LDIP and nVenus-LDAP3ΔC100; refer also to (c). Bar = 20 μm.

(c) Quantification of BiFC assays with LDIP and LDAP3 in tobacco leaf cells. Results from at least 20 areas of transformed epidermal leaf cells, similar to those shown in (a), were analysed from three independent experiments, and the mean number of BiFC puncta per area (±SD) are plotted in the graph on the right. Asterisks indicate significant difference at  $P \leq 0.01$  determined by Student's *t*-test.

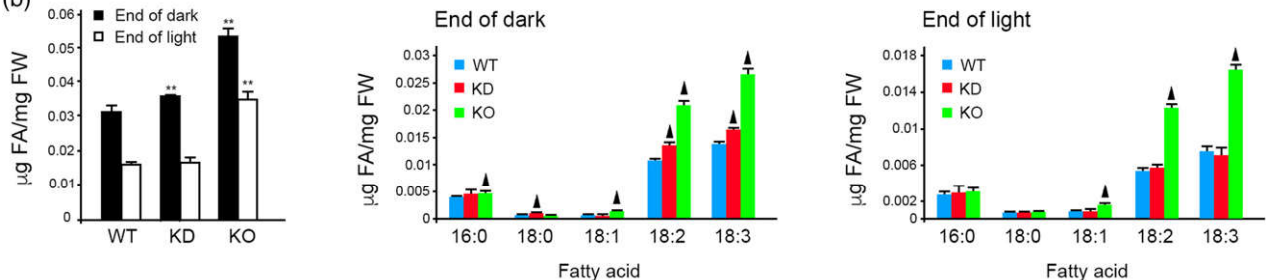
(d) LDIP and LDAP3 interact at LDs in the BiFC assay in tobacco leaf cells. Representative CLSM images of an individual BiFC puncta in a cCFP-LDIP and nVenus-LDAP3-transformed leaf epidermal cell (refer to the example shown at lower magnification and indicated with an arrowhead in (b)) and the corresponding staining of LDs with monodansylpentane (MDH; false coloured magenta). Also shown is the corresponding image of co-expressed Cherry-PeroX at peroxisomes, as well as the merged image. The aggregation of LDs is presumably due to organelle 'zippering' caused by the dimerization of the reconstituted fluorescent protein, similar to that observed in GFP-LDIP-transformed cells (Figure S1). Bar = 5 μm.

(e) Pull-down of LDAP and LDIP in tobacco leaf cells. Listed are selected MS-identified *N. benthamiana* LD proteins that co-immunoprecipitated with expressed GFP-LDIP or GFP alone, with or without co-expressed LEC2, in total protein extracts obtained from *Agrobacterium*-infiltrated tobacco leaves. Accession numbers of *N. benthamiana* proteins were obtained from the *N. benthamiana* genome, available at the Sol Genomics Network (<http://www.solgenomics.net>). AGI numbers and protein names of Arabidopsis homologues were obtained from TAIR database. The quantities (normalized spectra counts) of the co-purifying proteins were analysed using Scaffold software. Spectra counts of each protein were normalized to the average of the sums of all MS samples in the experiment.

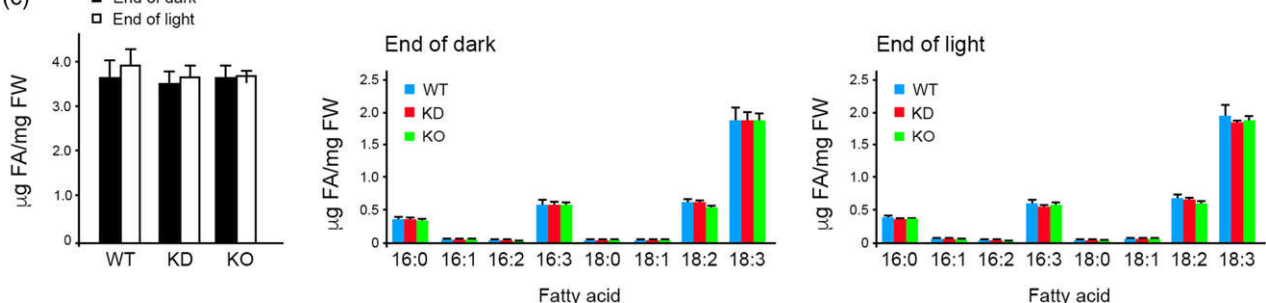
(a)



(b)



(c)



**Figure 5.** Effects of *LDIP* disruption on lipid droplets (LDs) and neutral lipid content in leaves.

(a) Abundance and size of LDs during the diurnal cycle in leaves of Arabidopsis *LDIP* mutant plants. Wild-type (WT) and *LDIP* knockout (KO) and knockdown (KD) mutant lines (Figure S7) were grown on ½ MS plates for 15 days in a 16-h/8-h day/night cycle, then leaves were harvested at the end of the night or end of the day, stained with BODIPY, and then the LDs examined by confocal laser-scanning microscopy. Representative images of BODIPY-stained LDs in leaves from each line (as indicated by labels) are shown on the left and quantifications of LD number and LD diameter are indicated by the graphs on the right. The blue colour in the micrographs is attributable to chlorophyll autofluorescence. Micrographs on the far right in each row of images are 3D projections of surface-rendered, high-magnification (zoom in) Z-stack images of a selected region of the same cells shown in micrographs of the KO leaves; asterisks represent the point of view for the 3D images and arrowheads indicate obvious examples of supersized LDs. Bars = 20 μm. Values of LD number are averages (±SD) from three biological replicates, each replicate consisting of eight leaf samples per line. LD diameter was calculated using the same data set (i.e. micrographs). Single and double asterisks in graphs represent statistically significant differences (relative to WT) at  $P \leq 0.05$  and  $P \leq 0.01$ , respectively, as determined by Student's *t*-test.

(b) Neutral lipid content and composition of Arabidopsis leaves in 15-day-old seedlings at the end of the dark or end of the light period (as indicated in graph legends) in WT and *LDIP* KD or KO mutant lines. FA, fatty acids; FW, fresh weight. Asterisks and arrowheads represent statistically significant differences at  $P \leq 0.01$  and  $P \leq 0.05$ , respectively, as determined by Student's *t*-test.

(c) Polar lipid content and composition of Arabidopsis leaves in 15-day-old seedlings at the end of the dark or end of the light in (as indicated in graph legends) WT and *LDIP* KD or KO mutant lines.

constituent of LDs in a recent proteomics study of LDs isolated from Arabidopsis senescing leaves (Brocard *et al.*, 2017).

We also showed that *LDIP* is present as a single-copy gene in Arabidopsis and, similar to the Arabidopsis *LDAP* genes (Gidda *et al.*, 2016; Kim *et al.*, 2016), is expressed in

a variety of organs and developmental stages, including seeds (Figure 2). Like the LDAPs (Horn *et al.*, 2013), as well as the oleosins (Huang and Huang, 2015), *LDIP* is strongly conserved amongst plant species (Figure 2), with no homologues detected outside the plant kingdom, indicating it is a plant-specific, LD-localized protein. Further, these general evolutionary similarities between LDIPs and LDAPs, taken together with their physical interaction (Figure 4), imply that they may have co-evolved in terms of a functional relationship.

### LDIP is class II-type LD coat protein

LD proteins can target the LD surface either by binding the lipids of the phospholipid monolayer and/or neutral lipid core or by associating with other coat proteins (Kory *et al.*, 2016; Bersuker and Olzmann, 2017). In general, they are divided into two classes: class I proteins, including the oleosins, which target to the LD surface via the ER (Huang and Huang, 2017), and class II proteins, which target directly from the cytoplasm. When transiently over-expressed in tobacco leaf cells, Cherry-LDIP localized to LDs and the cytoplasm, but was never observed in association with the ER, even after extended periods of transient expression in tobacco (Figure S12) or when stably (ectopically) expressed in Arabidopsis plants (Figure S10), suggesting that LDIP is a class II-type LD protein.

Class II LD proteins typically possess one or more structural motifs that allow them to associate specifically with LDs, such as an amphipathic  $\alpha$ -helix or hydrophobic domains (Kory *et al.*, 2016). Hydropathy analyses and 3D structural modelling revealed that amino acid residues 48–60 within Arabidopsis LDIP formed a potential amphipathic  $\alpha$ -helix, followed by a downstream, strongly hydrophobic region (amino acids 107–211) (Figure 3) that had limited sequence similarity to the promethin proteins of mammals or MmpL proteins of mycobacteria. Subsequent mutational analyses showed that the N-terminal region of LDIP (i.e. LDIP<sup>1–106</sup>), which contained the amphipathic  $\alpha$ -helix, was both necessary and sufficient for LD targeting, while a mutant version of the protein (i.e. LDIP<sup>107–249</sup>) containing the promethin/MmpL-like domain, but lacking the N-terminal region, targeted to the ER (Figure 3). This latter result might imply that the ER is important for some aspect(s) of LDIP targeting and/or function, as evidenced by LDIP being previously identified in the microsomal fraction of avocado mesocarp, while avocado LDAPs were enriched in purified LDs (Horn *et al.*, 2013). LDs are also known to maintain intimate associations with the ER (Chapman *et al.*, 2012; Gao and Goodman, 2015), and, as discussed in more detail below, perhaps LDIP is localized to ER–LD junctions prior to its localization to the LD surface. However, we also cannot rule out the possibility that the localization of

the LDIP<sup>107–249</sup> mutant to the ER simply reflects an artefact of protein (mis)folding or a default targeting pathway, since hydrophobic proteins lacking their normal targeting information are often misdirected to the ER (Walter and Johnson, 1994).

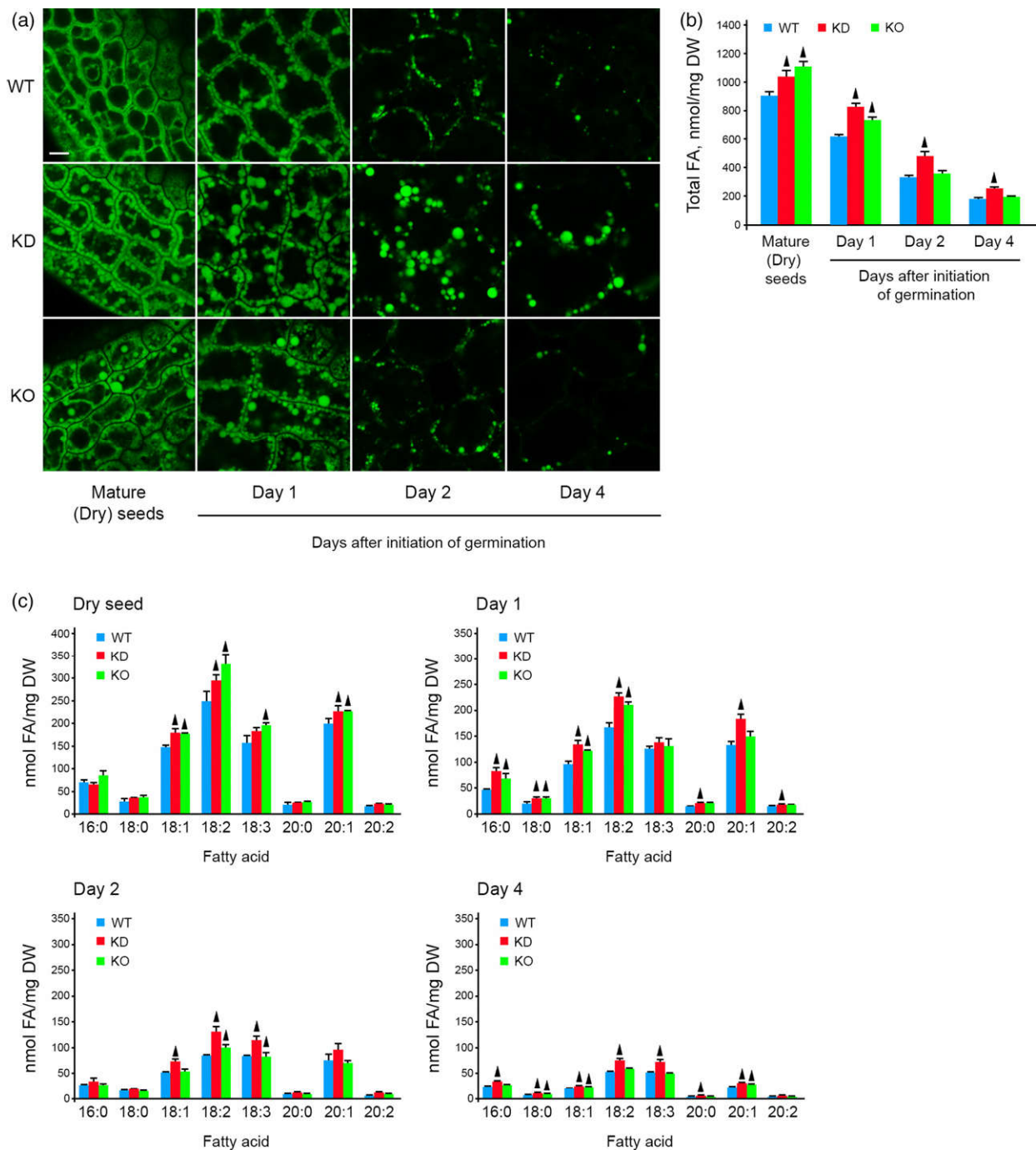
### LDIP–LDAP3 interactions and potential mechanisms of LDIP activity

LDIP was initially identified in a Y2H screen using LDAP3 as ‘bait’, and the physical association of LDIP and LDAP3 on LDs was confirmed using BiFC (Figure 4). Additional Y2H studies suggested that LDIP can form homotypic associations with itself (Figure 4) as well as bind to Arabidopsis LDAP1 and LDAP2, although relatively less so with LDAP1 (Figure S13). Nonetheless, all of these interactions were further supported by affinity-capture experiments, wherein GFP-tagged LDIP was transiently expressed in tobacco leaves then precipitated using affinity-capture methods, with co-purifying proteins, most notably the tobacco LDIP and LDAPs, being identified by MS and peptide mass fingerprinting (Figure 4).

In considering the functional significance of the LDIP–LDAP protein interaction, one hypothesis is that LDIP serves as an anchor for targeting LDAP to the LD surface. Prior studies have shown that recombinantly expressed and purified LDAP (LDAP3) binds weakly and non-specifically to synthetic liposomes containing various phospholipid compositions (Gidda *et al.*, 2016), indicating that other factors are probably involved in determining its LD-specific association. The LDAPs are also generally hydrophilic proteins, with no predicted membrane-spanning regions or other obvious regions of strong hydrophobicity (Gidda *et al.*, 2016). As such, they probably need to interact with other proteins for recruitment to and association with the LD surface, and given the strong hydrophobic domain within LDIP (Figure 4) it is possible that it serves as an anchor for this process. Alternatively, it is possible that the previously reported *in vitro* liposome-binding assays (Gidda *et al.*, 2016) did not faithfully recapitulate some needed, but as yet unknown, aspect of the targeting conditions, and thus LDAP might target specifically to LDs in a protein-independent manner *in vivo*, or through interaction with protein(s) other than LDIP.

Additional support for a functional connection between LDIP and LDAP comes from analyses of knockout mutants of both genes. We showed previously that disruption of any of the three *LDAP* genes in Arabidopsis decreased the number of LDs in 15-day-old leaves (Gidda *et al.*, 2016), and while we did not observe any apparent changes in the size of LDs, a recent study of a *LDAP1* knockout in Arabidopsis senescing (42–49 days old) leaves, which contain more neutral lipid than younger leaves, did show both a decrease in the number of LDs and an increase in





**Figure 6.** Effects of *LDIP* disruption on seed development and post-germinative growth.

(a) Representative confocal laser-scanning microscopy images of BODIPY-stained lipid droplets (LDs) in mature, dry seeds and seedlings 1, 2 or 4 days after the onset of germination of wild type (WT) and *LDIP* knockdown (KD) or knockout (KO) mutant lines. Note the presence of larger LDs in the KD and KO seeds and seedlings. Bar = 5  $\mu$ m.

(b) Total lipid content on a fatty acid basis in WT and *LDIP* KD and KO mutant mature seeds and in seedlings during post-germinative growth. DW, dry weight; FA, fatty acids. Results shown (and in (c)) represent averages ( $\pm$ SD) of three biological replicates, and arrowheads represent statistically significant differences at  $P \leq 0.05$  determined by Student's *t*-test.

(c) Fatty acid composition in WT and *LDIP* KD and KO mutant mature seeds and in seedlings during post-germinative growth.

variability of LD size, with the appearance of some noticeably larger LDs (Brocard *et al.*, 2017). This phenotype is generally similar to the *LDIP* mutant phenotype described

here, which included a reduction in the number of LDs in leaves and an increase in the variability of LD size, with the appearance of some considerably larger LDs (i.e.

supersized LDs) (Figure 5). One potential model for explaining these observations is that, as mentioned above, LDIP is required for the association of LDAP with LDs, which are subsequently important for stabilizing the LDs and preventing LD–LD fusion, as previously proposed (Gidda *et al.*, 2016). In the absence of LDIP, LDAPs would not associate with LDs as effectively, leading to a reduction of LDAP proteins on the LD surface and a concomitant increase in the propensity for LDs to fuse in the cytoplasm. In this model, fusion of LDs would not only reduce LD abundance but also lead to increases in the size of LDs. The increase in LD size might also account for the observed increase in neutral lipid content in *LDIP* mutant plants, since the lower surface-to-volume ratio of larger LDs would potentially reduce access of their TAG contents to lipase enzymes. There is, however, an important difference in the phenotypes of *LDIP* and *LDAP* mutant plants: disruption of *LDAPs* resulted in no change in the total neutral lipid content of leaves (Gidda *et al.*, 2016; Kim *et al.*, 2016), while disruption of *LDIP* nearly doubled the amount of leaf oil (Figure 5). This suggests that (i) LDIP has functions beyond just the recruitment and association of LDAP to LDs and/or (ii) the model for interaction of LDIP and LDAP is more complex, and the similarities in the LD phenotypes of the *LDIP* and *LDAP* mutants arise from distinct underlying molecular mechanisms.

Disruption of *LDIP* in seeds also resulted in the appearance of supersized LDs in the seed embryo and significant increases in total seed oil content (Figure 6). At the cell biology level, this enlarged LD phenotype is reminiscent of the phenotype observed in *oleosin* KO mutants, where a reduction of oleosin results in fusion of LDs in developing seeds (Siloto *et al.*, 2006; Schmidt and Herman, 2008; Shimada *et al.*, 2008; Miquel *et al.*, 2014). At the biochemical level, however, seed oil content is reduced in the *oleosin* mutant background (Siloto *et al.*, 2006) while it is increased in the *LDIP* mutants. Indeed, in two different organ types (leaves and seeds) and two different developmental stages (mature seeds and 15-day-old leaves), disruption of *LDIP* not only affects LD size but also increases total neutral lipid content (Figures 5 and 6). On the other hand, this cellular and biochemical phenotype is more similar to the phenotype observed when the *Arabidopsis SEIPIN1* gene is over-expressed in transgenic plants (Cai *et al.*, 2015). SEIPIN proteins are conserved ER membrane proteins that localize specifically to ER–LD junctions and promote LD formation (Chen and Goodman, 2017). The SEIPINs are also known to form large multimeric complexes that recruit and interact with multiple protein partners, thereby serving as important organizing centres for LD production. Although SEIPIN has no known enzymatic functions, SEIPINs in yeast and mammals interact with several different enzymes of the Kennedy pathway (Sim *et al.*, 2013; Talukder *et al.*, 2015; Pagac *et al.*, 2016), which couples LD

formation with localized synthesis of neutral lipids, including metabolites such as diacylglycerol and phosphatidic acid that are known to be essential for LD formation (Pol *et al.*, 2014). Also key to the growth of nascent LDs is the coordinated synthesis and enrichment of certain phospholipids within the LD monolayer. As such, the composition of the phospholipid monolayer is an important determinant of the size of LDs, as well as their propensity for subsequent LD–LD fusion (Thiam *et al.*, 2013; M'barek *et al.*, 2017). Indeed, the majority of genes identified in screens for altered LD sizes in insect and yeast cells are involved in phospholipid metabolism (Guo *et al.*, 2008; Fei *et al.*, 2011; Li *et al.*, 2016; Fan *et al.*, 2017).

Given that the size of LDs is affected in *LDIP* mutants and there is an increase in neutral lipid content (Figures 5 and 6), and that SEIPIN is one of the LD proteins that co-purified with LDIP in affinity-capture experiments in tobacco (Figure 4), one possibility is that LDIP functions at an earlier step in LD biogenesis that also involves SEIPIN activity. For instance, LDIP might serve at the cytoplasmic surface of ER–LD junctions to modulate the phospholipid content of a growing nascent LD, thereby affecting local glycerolipid metabolism, which is also influenced by SEIPIN activity. In the absence of LDIP, however, phospholipids might be altered or reduced, thereby promulgating larger LDs and/or the fusion of normal-sized LDs into supersized LDs, and redirecting glycerolipid intermediates from phospholipid metabolism to neutral lipid synthesis. In this scenario, LDIP would function as part of the machinery that helps coordinate the ordered formation of nascent LDs of a specific size and composition. Moreover, LDIP could still serve as an anchor for the recruitment of LDAPs to the growing LDs, although in seeds the LDAPs may be supplanted by the oleosin proteins.

Lastly, it is also possible that LDIP might serve at a later step in the LD life cycle by transferring neutral lipids from LDs to other organelles. The hydrophobic MmpL domains within mycobacterial MmpL proteins are known to associate to form complexes that transport various lipid compounds across the plasma membrane to the outer cell wall (Viljoen *et al.*, 2017). Based on its MmpL-like domain sequence (Figure 3) and its ability to self-associate (Figure 4), perhaps LDIP, in an analogous manner, is involved in the transfer of lipids from LDs in plant cells, such that disruption of *LDIP* results in the continued growth of LDs and a subsequent increase in cellular neutral lipid content. This model is somewhat difficult to reconcile in seeds, however, where oil production is known to include both rapid synthesis of oil as well as some turnover in the latter stages of seed development due to peroxisomal  $\beta$ -oxidation (Theodoulou and Eastmond, 2012). While a reduced transfer of lipids from LDs to peroxisomes in *LDIP* mutant seeds might account for the increase in seed oil content of mature seeds, there is apparently no defect in the

mobilization of seed oil during post-germination growth, i.e. the elevated seed oil content in mature *LDIP* mutants seeds was rapidly degraded and achieves nearly wild-type levels in 4-day-old seedlings (Figure 6). Furthermore, disruption of *LDIP* results in a decrease in LD abundance in leaves (Figure 5), which would not be expected if *LDIP* served primarily in the transfer of neutral lipids out of LDs. As such, if *LDIP* is indeed involved in lipid transfer, it is perhaps more likely that the activity is required for modulating the phospholipid content of the monolayer, either during LD formation, as discussed above, or during a later stage in the lifecycle of mature LDs. Current studies are now aimed at distinguishing between these various possibilities and also determining whether *LDIP* is required for the targeting of LDAPs to LDs.

## EXPERIMENTAL PROCEDURES

### Plant material, growth conditions and transformations

All *Arabidopsis* (*Arabidopsis thaliana*)-based experiments employed the WT Columbia-0 (Col-0) ecotype and derivatives thereof, including the T-DNA insertional mutant lines [i.e. *LDIP*-KD (SALK\_084555) and *LDIP*-KO (SAIL\_335\_H11)] obtained from the Arabidopsis Biological Resource Center (ABRC; <https://abrc.osu.edu>). Unless indicated otherwise, *Arabidopsis* plants were cultivated in soil in an environmental room with a 16-h/8-h day/night cycle at 22°C and 50  $\mu\text{E m}^{-2} \text{sec}^{-1}$  light intensity, or seeds were sterilized and plated on plates containing ½ MS media (Murashige and Skoog, 1962), then stratified for 3 days in the dark at 4°C before being moved into a growth chamber for the initiation of germination, with similar growth conditions to those described above. For the WT  $\times$  *ldip* KO cross, F<sub>1</sub> and F<sub>2</sub> seeds (progeny) were collected and plated on full-strength MS plates containing 50  $\mu\text{g } \mu\text{L}^{-1}$  Basta (phosphinothricin; Gold Biotechnology, <https://www.goldbio.com/>). BODIPY-stained LDs in all F<sub>1</sub> and F<sub>2</sub> seedlings that survived selection were analysed by CLSM and a chi-square test was used to determine the significance of the segregation pattern of the F<sub>2</sub> progeny. *LDIP*-KO plants were stably transformed (via *A. tumefaciens* infiltration, strain GV3101) with pMDC32/mCherry-*LDIP* using the floral dip method (Clough and Bent, 1998). The procedures for *A. tumefaciens* growth, transformation, infiltration and processing of leaf material for microscopy have been described elsewhere (McCartney et al., 2005; Cai et al., 2015).

*Nicotiana benthamiana* plants used for all *A. tumefaciens*-mediated transient expression experiments were grown in soil at 22°C with a 16-h/8-h day/night cycle and 50  $\mu\text{E m}^{-2} \text{sec}^{-1}$  light intensity. Leaves of 28-day-old *N. benthamiana* plants were infiltrated with *A. tumefaciens* (strain LBA4404 or, for co-immunoprecipitations, GV3101) carrying selected binary vectors. *Agrobacterium tumefaciens* transformed with the tomato bushy stunt virus gene *P19* was also included in all infiltrations to enhance transgene expression (Petrie et al., 2010).

### Gene cloning and plasmid construction

Molecular biology reagents were purchased from New England Biolabs (<https://www.neb.com/>), Thermo Fisher Scientific (<https://www.thermofisher.com/>) or Invitrogen (<http://www.invitrogen.com/>), and custom oligonucleotides were synthesized by Sigma-Aldrich (<http://www.sigmaaldrich.com/>). Sequence information for all primers used in gene cloning and plasmid construction is

available upon request. All DNA constructs were verified using automated sequencing performed at the University of Guelph Genomics Facility. Plasmids harbouring full-length open reading frames (ORFs) for each of the candidate proteins identified in the LDAP3 Y2H screen, including *LDIP*, were obtained from the ABRC, then ORFs were PCR-amplified and subcloned into the plant expression binary vector pMDC43 using Gateway technology (Curtis and Grossniklaus, 2003). pMDC43 contains the ORF of dimeric GFP, followed by a multiple cloning site (MCS), and the 35S cauliflower mosaic virus promoter [as do all other plant expression vectors used in this study, with the exception of *LDIP*::Cherry-MMPL (see below)]. pMDC43/mGFP-*LDIP*, consisting of a monomerized version of GFP (mGFP) linked to the N terminus of *LDIP*, was generated using PCR-based site-directed mutagenesis, whereby the leucine at position 221 in the GFP ORF in pMDC43/GFP-*LDIP* (serving as the template DNA) was replaced with a lysine (Zacharias et al., 2002). pMDC32/Cherry-*LDIP* was constructed by amplifying the full-length ORF of *LDIP* and cloning the resulting PCR products into pRTL2/Cherry, a plant expression vector containing the monomeric red fluorescent protein Cherry (Gidda et al., 2011). Thereafter, the coding region for the Cherry-*LDIP* fusion protein was subcloned into pMDC32 using Gateway technology. All the truncation and site-specific mutations of Cherry-*LDIP* used for analysing the putative LD targeting of *LDIP* signal were generated by PCR-based site-directed mutagenesis with pRTL2/Cherry-*LDIP* or pRTL2/Cherry-*LDIP*<sup>1–106</sup> as a template, followed by subcloning into pMDC32. pMDC32/GFP and pMDC32/Cherry, respectively encoding GFP and Cherry alone, and serving as cytoplasmic marker proteins, were constructed by amplifying the ORF of each fluorescent protein from pRTL2/GFP (Clark et al., 2009) or pRTL2/Cherry, then subcloning into pMDC32. pMDC99/*LDIP*::Cherry-MMPL, consisting of Cherry-*LDIP* driven by the native *LDIP* promoter sequence, was constructed by PCR amplifying (with genomic DNA as template) the 490-nucleotide sequence upstream of the *LDIP* ORF (and downstream of the ORF adjacent to *LDIP* in the *Arabidopsis* genome) then subcloning the resulting PCR products into pRTL2/Cherry-MMPL. Thereafter, the entire *LDIP*::Cherry-MMPL sequence was cloned into pMDC99 using Gateway technology.

Plasmids used for BiFC assays were generated based on the Gateway-compatible vectors pDEST-VYNE/nVenus and pDEST-SCYCE/cCFP, which encode the N-terminal and C-terminal halves of Venus and CFP, respectively (Gehl et al., 2009), and were obtained from the ABRC. The full-length ORF of *LDIP* and *LDAP3* or *LDAP3* $\Delta$ C100, whereby a premature stop codon was introduced into the *LDAP3* ORF, resulting in a 100-amino-acid long C-terminal truncation (Gidda et al., 2016), were PCR-amplified from their respective pRTL2-based plasmids and then subcloned into pDEST-VYNE/nVenus or pDEST-SCYCE/cCFP using Gateway technology.

Other plant expression binary vectors used in this study have been described elsewhere, including the following: pMDC32/*LDAP3*-Cherry and *OLEO1*-mGFP, encoding *Arabidopsis* *LDAP3* and oleosin isoform 1 appended to Cherry and mGFP, respectively (Horn et al., 2013; Gidda et al., 2016); ST-mRFP, encoding the N terminus of a rat *trans*-Golgi sialyl transferase appended to the monomeric red fluorescent protein (Boevink et al., 1998); pBIN/ER-GK and pBIN/ER-RK, encoding ER (lumen)-localized green and red fluorescent fusion proteins, and referred to in this study as GFP-ER and RFP-ER, respectively (Nelson et al., 2007) [obtained from the ABRC, clone numbers CD3-955 and CD3-959]; pMDC32/Cherry-PTS1, encoding Cherry linked to type 1 peroxisomal matrix targeting signal and referred to in this study as Cherry-PeroX (Ching et al., 2012); and pORE04/LEC2 and pORE04/P19, encoding

Arabidopsis LEC2, a regulator of seed development, and the tomato bushy stunt virus RNA-silencing suppressor p19 (Petrie *et al.*, 2010).

For Y2H library screening, the full-length ORF of Arabidopsis LDAP3 was PCR-amplified from pRTL2/LDAP3-Cherry (Gidda *et al.*, 2016) then subcloned into the 'bait' vector pGBKT7-DNA-BD (Clontech). Similarly, for directed Y2H assays, the LDAP3 ORF was subcloned from pRTL2/LDAP3-Cherry into the 'prey' vector pGADT7-AD, and the ORF of Arabidopsis LDIP was PCR-amplified from a plasmid encoding the full-length LDIP ORF obtained from the ABRC (see above) and cloned into pGADT7-AD and pGBKT7-DNA-BD. pGADT7-AD/LDAP3ΔC100 was generated using PCR-based site-directed mutagenesis with pGADT7-AD/LDAP3 serving as the template.

### LD isolations and intensity-based quantification (iBAQ) label-free proteomics

Two hundred and fifty milligrams of Arabidopsis seed (Col-0) was sterilized and stratified for 4 days in the dark at 4°C and then grown for 2 days in long-day conditions at 22°C on ½ MS medium plates without sucrose. Seedlings were ground in sand and 4 ml of grinding buffer [10 mM sodium phosphate buffer, pH 7.4, 0.5 mM Lohman's reagent and 10 mM *N*-ethylmaleimide (Sigma-Aldrich)] for 1 min with a mortar and pestle. The resulting homogenate was spun for 10 sec at 100 *g* and a 100 µl aliquot of the supernatant ('total cell protein' sample) was precipitated in 90% (v/v) ethanol at –20°C. The rest of the supernatant was centrifuged at 20 000 *g* for 20 min at 4°C and the resulting fat pad was washed twice with grinding buffer and finally delipidated in ethanol at –20°C. The protein pellet ('LD fraction' sample) was washed twice with ethanol, dried and resuspended in 6 M SDS and 5% (w/v) urea. Protein concentrations were determined with a Pierce BCA protein assay kit (Thermo Fisher Scientific). Ten micrograms of protein was run on a 12% SDS-PAGE gel until it entered into the separation gel. The Coomassie blue-stained protein bands were then excised and processed as described by Shevchenko *et al.* (1996). Peptides were desalted over a Supelco C18 column (Sigma-Aldrich) according to Rappsilber *et al.* (2007) and then subjected to LC-MS/MS (Schmidt *et al.*, 2017). Protein abundance was quantified using iBAQ label-free quantification implemented in MaxQuant software (Schaab *et al.*, 2012) and the values were calculated as a percentage of all values in one sample. For a detailed description of the data acquisition and processing see Methods S1. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository (Vizcaino *et al.*, 2016) with the dataset identifier PXD007192.

For Western blotting, protein extracts were separated by SDS-PAGE and electroblotted onto Hybond<sup>®</sup> nitrocellulose (GE LifeSciences, Mississauga, ON, Canada). Membranes were then incubated with anti-LDIP IgGs and immunoreactive proteins visualized using a Western Lightning<sup>®</sup> Plus-ECL kit and Blue XB film (PerkinElmer, Waltham, MA, USA). Rabbit anti-LDIP IgGs raised against a synthetic peptide corresponding to the LDIP amino acid sequence, -CNTISRHHQHDREVNIESTN- (residues 231–249), and purified using an LDIP peptide-Sepharose-linked column were generated by Cedarlane Labs (<https://www.cedarlanelabs.com/>).

### Bioinformatics

Construction of the LDIP phylogenetic tree was carried out using BioEdit (v.7.2.5) (Hall, 1999) with sequence alignments performed using the embedded ClustalW software, followed by manual

adjustment, and phylogenetic reconstruction of the aligned sequences performed using the embedded ProML Protein Maximum Likelihood program. The phylogram was generated using TreeView (v.1.6.6) (Page, 1996). Polypeptide sequences of various LDIP proteins were identified using the Protein Homologs tool at Phytozome (<http://www.phytozome.net>) (Goodstein *et al.*, 2011). Other protein sequence alignments were performed using the ClustalW algorithm at Prabi-Gerland (<https://npsa-prabi.ibcp.fr/>). Hydropathy analyses of amino acid sequences were carried out using the TMHMM server (<http://www.cbs.dtu.dk/services/TMHMM/>) and helical wheel projections were generated using HeliQuest (<http://heliquest.ipmc.cnrs.fr>) (Gautier *et al.*, 2008). Structural homology modelling was conducted using SWISS-MODEL (<https://swissmodel.expasy.org>) (Biasini *et al.*, 2014), whereby the Arabidopsis LDIP polypeptide sequence was used as a query to search for templates, then a structural homology model was built for the N-terminal region of LDIP (amino acid residues 32–115) using default parameters.

### RT-PCR and genotyping

Assessment of *LDIP* gene expression in leaves of 15-day-old WT Arabidopsis and T-DNA transgenic lines, as well as the expression of *LDIP* and *LDAP3* BiFC constructs in infiltrated *N. benthamiana* leaves, was carried out using RT-PCR based on procedures described by Cai *et al.* (2015) and Gidda *et al.* (2016). *LDIP* and *LDAP3* were amplified by 30 cycles of 94°C for 30 sec, 55°C for 30 sec and 72°C for 90 sec, while *TUBULIN*, *ACTIN* and *EF1α*, serving as control genes, were amplified by 30 cycles of 94°C for 30 sec, 55°C for 30 sec and 72°C for 1 min. All reactions contained 500 ng of total RNA. Specific forward and reverse primers used for RT-PCRs, as well as those used for genotyping of *LDIP* T-DNA insertional transgenic lines, are provided in Table S3.

### Microscopy

Wild-type and transgenic Arabidopsis seeds and leaves, as well as *A. tumefaciens*-infiltrated tobacco leaves, were processed for CLSM imaging, including staining of LDs either with BODIPY 493/503 (Invitrogen), Nile red (Sigma-Aldrich) or MDH (Abgent, <http://www.abgent.com/>), as previously described (Cai *et al.*, 2015; Gidda *et al.*, 2016). Micrographs of Arabidopsis and tobacco leaves were acquired using either a Leica DM RBE microscope equipped with a 63× Plan Apochromat oil-immersion objective and TCS SP2 scanning head, or a Leica SP5 CLSM equipped with a Radius 405-nm laser (Leica Microsystems, <https://www.leica-microsystems.com/>). Micrographs of dry seeds and germinated seedlings were acquired with a Zeiss LSM710 with a 63× water-immersion objective lens (Carl Zeiss Inc., <https://www.zeiss.com/>). Excitations and emission signals for fluorescent proteins, LD stains and/or chlorophyll autofluorescence collected sequentially as single optical sections or Z-series in double- or triple-labelling experiments are the same as those described in Gidda *et al.* (2016); single-labelling experiments showed no detectable cross-over at the settings used for data collection. 3D volume renderings were generated using Volocity imaging software (v.6.3) (PerkinElmer, <http://www.perkinelmer.com/>). The numbers and diameters of LDs in leaves of Arabidopsis seedlings were quantified according to Cai *et al.* (2015) using the Analyze Particles function in ImageJ (v.1.43; <https://imagej.net>), and the significance assessments of these data sets were performed using Student's *t*-test. All fluorescence images of cells shown in individual figures are representative of at least two separate experiments, including at least three separate transformations of tobacco leaf cells. Figure compositions

were generated using Adobe Photoshop CS (Adobe Systems, <http://www.adobe.com/>).

## Y2H and BiFC assays

Screening of a Y2H library, consisting of Arabidopsis cDNA from various plant tissues and cloned into the appropriate prey vector, using Arabidopsis LDAP3 (pGBKT7/LDAP3) as 'bait', was carried out with the Matchmaker Gold Y2H System (Clontech Laboratories, Inc., <http://www.clontech.com/>) as described by the manufacturer. All yeast strains that grew on low-selection [synthetic dextrose (SD) media lacking tryptophan and leucine, but containing X- $\alpha$ -Gal and Aureobasidin A] or high-selection (the same as low-selection media, but also lacking histidine and adenine) were designated as 'weak' or 'strong' interactors, respectively. Plasmids were extracted from yeast cells to determine the identity of encoded prey proteins using DNA sequencing. None of the prey plasmids autoactivated the Y2H reporter genes when retransformed into yeast cells with appropriate corresponding empty vectors. Directed Y2H assays were carried out as described previously (Richardson *et al.*, 2011), and, unlike for Y2H screening, plates used for high-stringency growth conditions consisted of SD media lacking tryptophan, leucine, histidine and adenine. The results of growth assays shown in figures are representative of those obtained from analysing three isolated yeast colonies from at least two independent transformations. In addition, all fusion proteins were confirmed to be expressed properly by Western blot analysis, as described above.

The BiFC assays in tobacco leaves were performed according to Stefano *et al.* (2015). Briefly, leaves were infiltrated with *Agrobacterium* containing plasmids encoding cCFP-LDIP and nVenus appended to either LDAP3 or LDAP3 $\Delta$ C100, the latter serving as a negative control based on guidelines described elsewhere for assessing protein interactions using the BiFC assay (Lee *et al.*, 2012). All infiltrations also included Cherry-Perox, serving as a transformation marker. Transformed cells in leaf areas were visualized (via CLSM) based on Cherry fluorescence, and both Cherry and reconstituted BiFC (cCFP/nVenus) fluorescence signals were collected with identical image acquisition settings for all samples analysed. ImageJ was used to quantify spectral counts from acquired micrographs of at least 20 leaf areas from three separate infiltrations. CLSM acquisition settings, amounts of *Agrobacterium* infiltrated and post-infiltration times were chosen based on preliminary optimization experiments aimed at minimizing the possibility of non-specific interactions based on guidelines for assessing protein-protein interactions using the BiFC assay described by Stefano *et al.* (2015).

## Analysis of lipids

For analysis of the content and fatty acid composition of neutral and polar lipids from Arabidopsis leaves, total lipids were extracted from 500 mg (fresh weight, FW) of 15-day-old seedlings grown on ½ MS medium, using a hexane/isopropanol method (Hara and Radin, 1978) with the addition of C17:0 TAG (Sigma-Aldrich) and C15:0 phosphatidylcholine (Avanti Polar Lipids, Inc., <https://avantilipids.com/>) as internal standards. Total lipid extracts in hexane were separated into neutral and polar lipids on solid-phase extraction cartridges (Supelco Discovery DSC-Si 6 ml, Sigma-Aldrich), as described (Gidda *et al.*, 2016). To prepare fatty acid methyl esters (FAMES), 0.5 ml of 0.5 N sodium methoxide solution in methanol was added to neutral or polar lipid extracts, and samples were incubated at room temperature (22°C) for 25 min. The reaction was quenched with 1 ml of saturated NaCl solution in water, and FAMES were extracted with 1 ml of hexane.

The FAME samples were analysed on an Agilent HP 6890 series GC system equipped with a 7683 series injector and autosampler (Agilent Technologies, <http://www.agilent.com/>) and a BPX70 (SGE Analytical Science, <http://www.sge.com/>) capillary column (10 m  $\times$  0.1 mm  $\times$  0.2 mm) with a constant pressure of 25 p.s.i., as described in Gidda *et al.* (2016). Compounds were identified by comparing with the GLC-10 FAME standard mix (Sigma-Aldrich). Analyses of the content and fatty acid composition of neutral and polar lipids from dry seeds and germinated seeds were performed as described in Gidda *et al.* (2016).

## Co-immunoprecipitation using GFP-Trap-A beads

Leaves of approximately 28-day-old *N. benthamiana* plants were transiently transformed (via *Agrobacterium* infiltration) with binary plasmids encoding GFP alone or GFP-LDIP and with and without LEC2. Expression of GFP and GFP-LDIP was verified by CLSM and transformed leaves were collected (3 days post-infiltration) for protein extraction. Briefly, approximately 1.5 g of leaf material was ground to a fine powder in liquid nitrogen using a mortar and pestle. Then 3 ml of extraction buffer [50 mM TRIS-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, 1% NP-40, 1 mM PMSF, Roche protease inhibitor cocktail (2 tablets/10 ml buffer)] (Roche Diagnostics, <http://www.roche.com/>) was added to the tissue powder; samples were transferred into 15-ml test tubes and incubated for 30 min on ice and vortexed every 10 min. Cell lysates were centrifuged at 2500 *g* at 4°C for 10 min to remove cell debris and the supernatants were further clarified by centrifugation at 16 000 *g* at 4°C for 20 min. Supernatants were then incubated with 30  $\mu$ l of GFP-Trap-A beads (ChromoTek, <https://www.chromotek.com/>) for 3 h at 4°C on a rotating shaker. After incubation, beads were washed four times with extraction buffer and proteins bound to the beads were then eluted with 100  $\mu$ l of 4 $\times$  SDS sample buffer and heated at 70°C for 10 min. To confirm the presence of the 'bait' proteins (i.e. free GFP or GFP-LDIP), an aliquot of each immunoprecipitated sample was separated by SDS-PAGE and detected by Western blotting with anti-GFP antibodies (Thermo Fisher Scientific). The remaining protein was concentrated at the top of an SDS-PAGE resolving gel and the Coomassie blue-stained protein bands were excised and submitted to the Michigan State University Proteomics Core Service (<https://rtsf.natsci.msu.edu/proteomics/>).

Processing of samples (gel bands) for MS analysis was carried out by dehydrating the samples using 100% acetonitrile and incubating with 10 mM dithiothreitol in 100 mM ammonium bicarbonate (pH ~8) at 56°C for 45 min, followed by dehydration and incubation in the dark with 50 mM iodoacetamide in 10 mM ammonium bicarbonate for 20 min. Gel bands were then washed with ammonium bicarbonate and dehydrated again. Sequencing-grade modified trypsin (0.01  $\mu$ g  $\mu$ l<sup>-1</sup>) in 50 mM ammonium bicarbonate was added to the gel band and incubated at 37°C overnight. Peptides were extracted by water bath sonication in a solution of 60% acetonitrile and 1% trichloroacetic acid, vacuum dried to about 2  $\mu$ l and then re-suspended in 2% acetonitrile/0.1% trifluoroacetic acid to 25  $\mu$ l. Five microlitres was automatically injected by a Thermo EASYnLC 1000 (Thermo Fisher Scientific) onto a Thermo Acclaim PepMap 0.1 mm  $\times$  20 mm C18 peptide trap and washed for about 5 min. Bound peptides were then eluted onto a Thermo Acclaim PepMap RSLC 0.075 mm  $\times$  250 mm C18 column over 65 min with a gradient of 5% to 28% buffer B (consisting of 99.9% acetonitrile and 0.1% formic acid) in 54 min, ramping to 100% buffer B at 55 min and held at 100% buffer B for the duration of the run at a constant flow rate of 0.3  $\mu$ l min<sup>-1</sup>. Eluted peptides were sprayed into a Thermo Fisher

Q-Exactive MS using a FlexSpray spray ion source. Survey scans were taken in the Orbitrap (70 000 resolution, determined at  $m/z$  200) and the top 10 ions in each survey scan were then subjected to automatic higher-energy collision-induced dissociation with fragment spectra acquired at 17 500 resolution. The resulting MS/MS spectra were converted to peak lists using Mascot Distiller (v.2.6.1; <http://www.matrixscience.com/>) and searched against a database of all *N. benthamiana* protein sequences available from the Sol Genomics Network (v.0.4.4; <http://www.solgenomics.net>) and appended with common laboratory contaminants (cRAP project; <http://www.thegpm.org>) using the Mascot searching algorithm (v.2.6.0). The Mascot output was then analysed using Scaffold (v.4.7.5; <http://www.proteomesoftware.com/>) to probabilistically validate protein identifications. Assignments validated using the Scaffold 1% false discovery rate (FDR) confidence filter are considered true. Mascot parameters for all databases were as follows: (i) allow up to two missed tryptic sites; (ii) fixed modification of carbamidomethyl cysteine, variable modification of oxidation of methionine, deamidation of glutamine and asparagine; (iii) peptide tolerance of  $\pm 10$  p.p.m.; (iv) MS/MS tolerance of 0.3 Da; and (v) FDR calculated using randomized database search.

### ACCESSION NUMBERS

Arabidopsis Information Resource numbers and/or GenBank accession numbers for the Arabidopsis proteins described in this study are as follows: LDIP (At5g16550, NP\_568333), LDAP3 (At3g05500, NP\_187201), EIF2 (At2g39990, NP\_181528), MSR1 (At3g21190, NP\_566677), TPR8 (At4g08320, NP\_001031594), PRA7 (At1g55190, NP\_564679), NUDT3 (At1g79690, NP\_565218), TRA2 (At5g13420, NP\_196846), ATMS1 (At5g17920, NP\_001078599), and PGL5 (At5g24420, NP\_197830), OLEO1 (At4g25140, NP\_194244),  $\alpha$ -TUBULIN (At5g44340, NM\_123801), and LEC2 (At1g28300, NP\_564304). Other protein accession numbers: tomato bushy stunt virus p19 (CAC01278) and *N. benthamiana* ACTIN (AY179605). Accession numbers for LDIP homologues from various plant species used in the construction of the phylogenetic tree (Figure 2) and the protein sequence alignment (Figure 3) are provided in Table S4.

### ACKNOWLEDGEMENTS

We are grateful to Dr Reynald Tremblay (University of Guelph) for assistance with the phylogenetic analysis of LDIPs, Dr Frederica Brandizzi (Michigan State University) for ST-mRFP, Dr Qing Liu (CSIRO Plant Industry) for pORE04/LEC2 and pREO4/P19 and Ms Vivien Truong (University of Guelph) for assistance with figure construction. We are also grateful to Drs Toby Kellogg and Michael McKain (Donald Danforth Plant Science Center) for their insights on the potential co-evolution of LDIPs and LDAPs, and Drs Jay Thelen and Veronika Lancikova (University of Missouri-Columbia) for their assistance with preliminary LD proteomics experiments. Proteomics analysis of GFP capture experiments was carried out by Douglas Whitten (Michigan State University Proteomics Core Facility). Research support was provided by the US Department of Energy (DOE) Office of Science, BES-Physical Biosciences program (DE-SC0016536), the Natural Sciences and Engineering Research Council of Canada (217291-2013), the German Research Foundation (DFG) and the Studienstiftung des Deutschen Volkes. Mention of trade names or commercial products in this article is solely for

the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture (USDA). The USDA is an equal opportunity provider and employer.

### CONFLICT OF INTEREST

All the authors declare that they have no conflict of interests.

### SUPPORTING INFORMATION

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

**Figure S1.** Localization of GFP-tagged candidate LDAP3-interacting proteins in tobacco leaf epidermal cells.

**Figure S2.** Localization of Cherry-tagged LDIP truncation mutants and modified variants thereof in tobacco leaf epidermal cells.

**Figure S3.** Helical wheel projections of N-terminal polypeptide sequences from selected LDIP homologues.

**Figure S4.** Yeast two-hybrid growth assays with Arabidopsis LDIP and LDAP3.

**Figure S5.** Western blot analysis of yeast two-hybrid fusion proteins.

**Figure S6.** Confirmation of expression of bimolecular fluorescence complementation fusion constructs in tobacco leaf epidermal cells using RT-PCR.

**Figure S7.** Characterization of Arabidopsis LDIP T-DNA insertional transgenic lines.

**Figure S8.** BODIPY-stained lipid droplets in leaves of 28-day-old Arabidopsis wild-type and LDIP knockout plants.

**Figure S9.** Progeny analysis and BODIPY-stained lipid droplets in leaves of Arabidopsis F<sub>1</sub> and F<sub>2</sub> plants from a wild type  $\times$  LDIP knockout mutant backcross.

**Figure S10.** BODIPY-stained lipid droplets in leaves of Arabidopsis wild type and various LDIP knockout mutant and transgenic plants, and the localization of Cherry-LDIP in a stably transformed LDIP knockout Arabidopsis plant.

**Figure S11.** RT-PCR analysis of modified LDIP transcripts in the LDIP knockout mutant.

**Figure S12.** Localization of over-expressed Cherry-LDIP in tobacco leaf epidermal cells.

**Figure S13.** Interaction of Arabidopsis LDIP and LDAP1 and LDAP2 in yeast cells.

**Table S1.** Quantitative proteomics data from 2-day-old Arabidopsis seedling total cell extracts and isolated lipid droplets.

**Table S2.** Qualitative proteomics data depicting identified peptides.

**Table S3.** List of primers used for RT-PCR and genotyping.

**Table S4.** Accession numbers of LDIP homologues from various plant species.

**Methods S1.** Detailed workflow of proteomics data acquisition and processing.

### REFERENCES

- Barneda, D. and Christian, M. (2017) Lipid droplet growth: regulation of a dynamic organelle. *Curr. Opin. Cell Biol.* **47**, 9–15.
- Bersuker, K. and Olzmann, J.A. (2017) Establishing the lipid droplet proteome: mechanisms of lipid droplet protein targeting and degradation. *Biochim. Biophys. Acta* **1862**, 1166–1177. <https://doi.org/10.1016/j.bbalip.2017.06.00>



- Berthelot, K., Peruch, F. and Lecomte, S. (2016) Highlights on *Hevea brasiliensis* (pro) hevein proteins. *Biochimie*, **127**, 258–270.
- Biasini, M., Bienert, S., Waterhouse, A. et al. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**(W1), W252–W258.
- Boevink, P., Oparka, K., Cruz, S.S., Martin, B., Betteridge, A. and Hawes, C. (1998) Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. *Plant J.*, **15**, 441–447.
- Brocard, L., Immel, F., Coulon, D., Esnay, N., Tophile, K., Pascal, S., Claverol, S., Bessoule, J.J. and Bréhélin, C. (2017) Proteomic analysis of lipid droplets from Arabidopsis aging leaves brings new insight into their biogenesis and functions. *Front. Plant Sci.* **8**, 894.
- 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. *Plant Cell*, **27**, 2616–2636.
- Chapman, K.D., Dyer, J.M. and Mullen, R.T. (2012) Biogenesis and functions of lipid droplets in plants Thematic Review Series: Lipid droplet synthesis and metabolism: from yeast to man. *J. Lipid Res.* **53**, 215–226.
- Chen, X. and Goodman, J.M. (2017) The collaborative work of droplet assembly. *Biochim. Biophys. Acta* **1862**, 1205–1211. <https://doi.org/10.1016/j.bbalip.2017.07.003>
- Ching, S.L., Gidda, S.K., Rochon, A., van Cauwenberghe, O.R., Shelp, B.J. and Mullen, R.T. (2012) Glyoxylate reductase isoform 1 is localized in the cytosol and not peroxisomes in plant cells. *J. Integr. Plant Biol.* **54**, 152–168.
- Clark, S.M., Di Leo, R., Dhanoa, P.K., Van Cauwenberghe, O.R., Mullen, R.T. and Shelp, B.J. (2009) Biochemical characterization, mitochondrial localization, expression, and potential functions for an Arabidopsis  $\gamma$ -aminobutyrate transaminase that utilizes both pyruvate and glyoxylate. *J. Exp. Bot.* **60**, 1743–1757.
- Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.
- Curtis, M.D. and Grossniklaus, U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* **133**, 462–469.
- Davidi, L., Levin, Y., Ben-Dor, S. and Pick, U. (2015) Proteome analysis of cytoplasmatic and plastidic  $\beta$ -carotene lipid droplets in *Dunaliella bardawil*. *Plant Physiol.* **167**, 60–79.
- Fan, W., Lam, S.M., Xin, J., Yang, X., Liu, Z., Liu, Y., Wang, Y., Shui, G. and Huang, X. (2017) Drosophila TRF2 and TAF9 regulate lipid droplet size and phospholipid fatty acid composition. *PLoS Genet.* **13**, e1006664.
- Fei, W., Shui, G., Zhang, Y. et al. (2011) A role for phosphatidic acid in the formation of “supersized” lipid droplets. *PLoS Genet.* **7**, e1002201.
- Gao, Q. and Goodman, J.M. (2015) The lipid droplet—a well-connected organelle. *Front. Cell Dev. Biol.* **3**, 49.
- Gautier, R., Douguet, D., Antony, B. and Drin, G. (2008) HELIQUEST: a web server to screen sequences with specific  $\alpha$ -helical properties. *Bioinformatics*, **24**, 2101–2102.
- Gehl, C., Waadt, R., Kudla, J., Mendel, R.R. and Hänsch, R. (2009) New GATEWAY vectors for high throughput analyses of protein–protein interactions by bimolecular fluorescence complementation. *Mol. Plant*, **2**, 1051–1058.
- Gidda, S.K., Shockey, J.M., Falcone, M., Kim, P.K., Rothstein, S.J., Andrews, D.W., Dyer, J.M. and Mullen, R.T. (2011) Hydrophobic-domain-dependent protein–protein interactions mediate the localization of GPAT enzymes to ER subdomains. *Traffic*, **12**, 452–472.
- Gidda, S.K., Watt, S.C., Collins-Silva, J. et al. (2013) Lipid droplet-associated proteins (LDAPs) are involved in the compartmentalization of lipophilic compounds in plant cells. *Plant Signal. Behav.* **8**, e27141.
- 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 Physiol.* **170**, 2052–2071.
- Goodstein, D.M., Shu, S., Howson, R. et al. (2011) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* **40**(D1), D1178–D1186.
- Guo, Y., Walther, T.C., Rao, M., Stuurman, N., Goshima, G., Terayama, K., Wong, J.S., Vale, R.D., Walter, P. and Farese, R.V. Jr (2008) Functional genomic screen reveals genes involved in lipid-droplet formation and utilization. *Nature*, **453**, 657–661.
- Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98.
- Hara, A. and Radin, N.S. (1978) Lipid extraction of tissues with a low-toxicity solvent. *Anal. Biochem.* **90**, 420–426.
- 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 Physiol.* **162**, 1926–1936.
- Huang, A.H. (1996) Oleosins and oil bodies in seeds and other organs. *Plant Physiol.* **110**, 1055–1061.
- Huang, M.D. and Huang, A.H. (2015) Bioinformatics reveal five lineages of oleosins and the mechanism of lineage evolution related to structure/function from green algae to seed plants. *Plant Physiol.* **169**, 453–470.
- Huang, M.D. and Huang, A.H. (2016) Subcellular lipid droplets in vanilla leaf epidermis and avocado mesocarp are coated with oleosins of distinct phylogenetic lineages. *Plant Physiol.* **171**, 1867–1878.
- Huang, C.Y. and Huang, A.H. (2017) Motifs in oleosin target the cytosolic side of endoplasmic reticulum and budding lipid droplet. *Plant Physiol.* **174**, 2248–2260. <https://doi.org/10.1104/pp.17.00366>
- Ischebeck, T. (2016) Lipids in pollen – They are different. *Biochim. Biophys. Acta* **1861**, 1315–1328.
- Kim, H.U., Hsieh, K., Ratnayake, C. and Huang, A.H. (2002) A novel group of oleosins is present inside the pollen of Arabidopsis. *J. Biol. Chem.* **277**, 22677–22684.
- Kim, E.Y., Seo, Y.S., Lee, H. and Kim, W.T. (2010) Constitutive expression of CaSRP1, a hot pepper small rubber particle protein homolog, resulted in fast growth and improved drought tolerance in transgenic Arabidopsis plants. *Planta*, **232**, 71–83.
- Kim, H.U., Jung, S.J., Lee, K.R., Kim, E.H., Lee, S.M., Roh, K.H. and Kim, J.B. (2013) Ectopic overexpression of castor bean LEAFY COTYLEDON2 (LEC2) in Arabidopsis triggers the expression of genes that encode regulators of seed maturation and oil body proteins in vegetative tissues. *FEBS Open Bio.* **4**, 25–32.
- Kim, E.Y., Park, K.Y., Seo, Y.S. and Kim, W.T. (2016) Arabidopsis small rubber particle protein homolog SRPs play dual roles as positive factors for tissue growth and development and in drought stress responses. *Plant Physiol.* **170**, 2494–2510.
- Kolkhof, P., Werthebach, M., van de Venn, A., Poschmann, G., Chen, L., Welte, M., Stühler, K. and Beller, M. (2017) A luciferase-fragment complementation assay to detect lipid droplet-associated protein–protein interactions. *Mol. Cell Proteomics*, **16**, 329–345.
- Kory, N., Farese, R.V. and Walther, T.C. (2016) Targeting fat: mechanisms of protein localization to lipid droplets. *Trends Cell Biol.* **26**, 535–546.
- 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. *J. Mol. Biol.* **305**, 567–580.
- Laibach, N., Post, J., Twyman, R.M., Gronover, C.S. and Prufer, D. (2015) The characteristics and potential applications of structural lipid droplet proteins in plants. *J. Biotechnol.* **201**, 15–27.
- Lee, L.Y., Wu, F.H., Hsu, C.T. et al. (2012) Screening a cDNA library for protein–protein interactions directly in planta. *Plant Cell*, **24**, 1746–1759.
- Li, S., Xu, S., Ma, Y. et al. (2016) A Genetic screen for mutants with super-sized lipid droplets in *Caenorhabditis elegans*. *G3* **6**, 2407–2419.
- Li-Beisson, Y., Shorrosh, B., Beisson, F. et al. (2013) Acyl-lipid metabolism. *Arabidopsis Book*, **11**, e0161.
- M'barek, K. B., Ajjaji, D., Chorlay, A., Vanni, S., Forêt, L. and Thiam, A.R. (2017) ER Membrane phospholipids and surface tension control cellular lipid droplet formation. *Dev. Cell*, **41**, 591–604.
- McCartney, A.W., Greenwood, J.S., Fabian, M.R., White, K.A. and Mullen, R.T. (2005) Localization of the tomato bushy stunt virus replication protein p33 reveals a peroxisome-to-endoplasmic reticulum sorting pathway. *Plant Cell*, **17**, 3513–3531.
- McLachlan, D.H., Lan, J., Geiffus, C.M. et al. (2016) The breakdown of stored triacylglycerols is required during light-induced stomatal opening. *Curr. Biol.* **26**, 707–712.
- 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 Physiol.* **164**, 1866–1878.

- Murashige, T. and Skoog, F. (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Plant Physiol.* **15**, 473–497.
- Murphy, D.J. (2012) The dynamic roles of intracellular lipid droplets: from archaea to mammals. *Protoplasma*, **249**, 541–585.
- Nelson, B.K., Cai, X. and Nebenführ, A. (2007) A multicolored set of *in vivo* organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J.* **51**, 1126–1136.
- Pagac, M., Cooper, D.E., Qi, Y. *et al.* (2016) SEIPIN regulates lipid droplet expansion and adipocyte development by modulating the activity of glycerol-3-phosphate acyltransferase. *Cell Rep.* **17**, 1546–1559.
- Page, R.D.M. (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357–358.
- Paterson, A.H., Freeling, M., Tang, H. and Wang, X. (2010) Insights from the comparison of plant genome sequences. *Annu. Rev. Plant Biol.* **61**, 349–372.
- 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**, 8.
- Pol, A., Gross, S.P. and Parton, R.G. (2014) Biogenesis of the multifunctional lipid droplet: lipids, proteins, and sites. *J. Cell Biol.* **204**, 635–646.
- Purkrtova, Z., Jolivet, P., Miquel, M. and Chardot, T. (2008) Structure and function of seed lipid-body-associated proteins. *Curr. Rev. Biol.* **331**, 746–754.
- Pyc, M., Cai, Y., Greer, M.S., Yurchenko, O., Chapman, K.D., Dyer, J.M. and Mullen, R.T. (2017) Turning over a new leaf in lipid droplet biology. *Trends Plant Sci.* **22**, 596–609.
- Rappsilber, J., Mann, M. and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906.
- Richardson, L.G., Howard, A.S., Khuu, N., Gidda, S.K., McCartney, A., Morphy, B.J. and Mullen, R.T. (2011) Protein–protein interaction network and subcellular localization of the *Arabidopsis thaliana* ESCRT machinery. *Front. Plant Sci.* **2**, 20.
- Santos Mendoza, M., Dubreucq, B., Miquel, M., Caboche, M. and Lepiniec, L. (2005) LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves. *FEBS Lett.* **579**, 4666–4670.
- Schaab, C., Geiger, T., Stoehr, G., Cox, J. and Mann, M. (2012) Analysis of high accuracy, quantitative proteomics data in the MaxQB database. *Mol. Cell Proteomics*, **11**, M111–014068.
- Schmidt, M.A. and Herman, E.M. (2008) Suppression of soybean oleosin produces micro-oil bodies that aggregate into oil body/ER complexes. *Mol. Plant*, **1**, 910–924.
- Schmidtt, K., Smolinski, N., Neumann, P., Schmaul, S., Hofer-Pretz, V., Braus, G.H. and Valerius, O. (2017) Asc1p/RACK1 connects ribosomes to eukaryotic phosphosignaling. *Mol. Cell Biol.* **37**, pii:e00279-16.
- Seo, S.G., Kim, J.S., Yang, Y.S., Jun, B.K., Kang, S.W., Lee, G.P., Kim, W., Kim, J.B., Lee, H.U. and Kim, S.H. (2010) Cloning and characterization of the new multiple stress responsible gene I (MuSI) from sweet potato. *Genes Genomics*, **32**, 544–552.
- Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858.
- 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*. *Plant J.* **55**, 798–809.
- Shimada, T.L., Takano, Y., Shimada, T. *et al.* (2014) Leaf oil body functions as a subcellular factory for the production of a phytoalexin in Arabidopsis. *Plant Physiol.* **164**, 105–118.
- 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 oilbodies in Arabidopsis. *Plant Cell*, **18**, 1961–1974.
- Sim, M.M., Dennis, R.J., Aubry, E.M., Ramanathan, N., Sembongi, H., Sauddek, V., Siniosoglou, S. and Rochford, J.J. (2013) The human lipodystrophy protein seipin is an ER membrane adaptor for the adipogenic PA phosphatase lipin 1. *Mol. Metab.* **2**, 38–46.
- Stefano, G., Renna, L. and Brandizzi, F. (2015) BiFC for protein–protein interactions and protein topology: discussing an integrative approach for an old technique. *Methods Mol. Biol.* **1242**, 173–182.
- Talukder, M.M.U., Sim, M.M., O’Rahilly, S., Edwardson, J.M. and Rochford, J.J. (2015) Seipin oligomers can interact directly with AGPAT2 and lipin 1, physically scaffolding critical regulators of adipogenesis. *Mol. Metab.* **4**, 199–209.
- Theodoulou, F.L. and Eastmond, P.J. (2012) Seed storage oil catabolism: a story of give and take. *Curr. Opin. Plant Biol.* **15**, 322–328.
- Thiam, A.R. and Beller, M. (2017) The why, when and how of lipid droplet diversity. *J. Cell Sci.* **130**, 315–324.
- Thiam, A.R. and Forêt, L. (2016) The physics of lipid droplet nucleation, growth and budding. *Biochim. Biophys. Acta* **1861**, 715–722.
- Thiam, A.R., Farese, R.V. Jr and Walther, T.C. (2013) The biophysics and cell biology of lipid droplets. *Nat. Rev. Mol. Cell Biol.* **14**, 775–786.
- Tsai, C.H., Zienkiewicz, K., Amstutz, C.L., Brink, B.G., Warakanont, J., Roston, R. and Benning, C. (2015) Dynamics of protein and polar lipid recruitment during lipid droplet assembly in *Chlamydomonas reinhardtii*. *Plant J.* **83**, 650–660.
- Vanhercke, T., Divi, U.K., El Tahchy, A. *et al.* (2017) Step changes in leaf oil accumulation via iterative metabolic engineering. *Metab. Eng.* **39**, 237–246.
- Viljoen, A., Dubois, V., Girard-Misguich, F., Blaise, M., Herrmann, J.L. and Kremer, L. (2017) The diverse family of MmpL transporters in mycobacteria: from regulation to antimicrobial developments. *Mol. Microbiol.* **104**, 889–904.
- Vizcaino, J.A., Csordas, A., del-Toro, N. *et al.* (2016) 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, 11033.
- Walter, P. and Johnson, A.E. (1994) Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol.* **10**, 87–119.
- Welte, M.A. and Gould, A.P. (2017) Lipid droplet functions beyond energy storage. *Biochim. Biophys. Acta* **1862**, 1260–1272. <https://doi.org/10.1016/j.bbalip.2017.07.006>
- Xing, S., Wallmeroth, N., Berendzen, K.W. and Grefen, C. (2016) Techniques for the analysis of protein–protein interactions *in vivo*. *Plant Physiol.* **171**, 727–758.
- Yang, H.J., Hsu, C.L., Yang, J.Y. and Yang, W.Y. (2012) Monodansylpentane as a blue-fluorescent lipid-droplet marker for multi-color live-cell imaging. *PLoS ONE*, **7**, e32693.
- Yu, S., Viswakarma, N., Batra, S.K., Rao, M.S. and Reddy, J.K. (2004) Identification of promethin and PGLP as two novel up-regulated genes in PPAR $\gamma$ -induced adipogenic mouse liver. *Biochimie*, **86**, 743–761.
- Zacharias, D.A., Violin, J.D., Newton, A.C. and Tsien, R.Y. (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science*, **296**, 913–916.
- Zhang, Z., Cheng, Z.J., Gan, L. *et al.* (2016) OsHSD1, a hydroxysteroid dehydrogenase, is involved in cuticle formation and lipid homeostasis in rice. *Plant Sci.* **249**, 35–45.



## 5. Discussion

### 5.1. The Proteome of Lipid Droplets in Plants and Beyond

In this thesis, I investigated the LD proteome from four different tissues: *Nicotiana tabacum* pollen tubes, and *Arabidopsis thaliana* siliques, seeds, and seedlings. Using label-free bottom-up proteomics, the proteome of LDs cannot only be accessed qualitatively, but also compared in a quantitative manner.

#### 5.1.1. The Composition of the Plant Lipid Droplet Proteome

LDs were first described in the late 1800, around the same time like the ER and Golgi apparatus (Garnier, 1897; Golgi, 1898; Murphy, 2012). For a long time, LDs were thought to be constricted to specific cell types, in which they serve as long-term storage for energy and carbon. Research in the last thirty years, however, has accumulated evidence of a dynamic organelle central to the lipid metabolism that is ubiquitous through taxa, organism, and tissues (Murphy, 2012). Despite this ubiquitous occurrence of LDs, their protein composition strongly differs across species and cell types.

The highly abundant plant LD proteins oleosins, caleosins and steroleosins were discovered already decades ago in seeds (Chen et al., 1999; Lin et al., 2002; Qu et al., 1986). Additionally, caleosins and oleosins have been found in pollen of different species, including olive (Zienkiewicz et al., 2010, 2011), lily (Jiang et al., 2007, 2008), and pine (Pasaribu et al., 2017). In Brassicaceae, a specific set of oleosin homologs, different from the ones found in seeds, have been identified in pollen (Kim et al., 2002; Ross and Murphy, 1996). Transcription analysis did detect expression of seed-like oleosins OLE7 and OLE8 in mature pollen grains of *Arabidopsis* (Honys and Twell, 2004), but these are not the major oleosins found in seeds (Manuscript II, Figure 9). In the Solanacea *Nicotiana tabacum*, we could show that homologs of the *Arabidopsis* seed oleosins are found also in the pollen tubes, but are distinct from the isoforms that were detected in tobacco seeds (Kretzschmar et al., 2018, Supplemental Figure 1). This distinction was less clear for caleosins, where some isoforms were shared between the two tissues investigated while others were unique for the tissues (Kretzschmar et al., 2018, Supplemental Figure 2). In *Arabidopsis* pollen tubes and mature pollen grains, CLO4 was detected by transcript analysis, and CLO8 on peptide level (Grobei et al., 2009; Honys and Twell, 2004). Tobacco CLO4 was also the isoform identified exclusively in pollen tubes and not in tobacco seeds, consistent with the *Arabidopsis* expression data. In *Arabidopsis* seeds, CLO1 and CLO2 are most abundant (Manuscript II, Figure 9; Næsted et al., 2000), and CLO3 was found in senescing leaves (Brocard et al., 2017; Takahashi et al., 2000).

**Table 2:** Contribution of LD-associated proteins in different tissues to the complete LD proteome. Relative iBAQ intensities were used to calculate the contribution of each LD-associated in percent of all presented LD-associated proteins.<sup>1</sup> Kretzschmar et al., 2018, <sup>2</sup> Manuscript II, RS – rehydrated seeds; StS – stratified seeds; 24 h, 36 h, 48 h, 60 h – seedlings after stratification

<i>N. tabacum</i> pollen tubes <sup>1</sup>		<i>A. thaliana</i> siliques <sup>2</sup>			<i>A. thaliana</i> seeds <sup>2</sup>			<i>A. thaliana</i> seedlings <sup>2</sup>				
proteins	% of all LD proteins	proteins	% of all LD proteins		proteins	% of all LD proteins		proteins	% of all LD proteins			
			<i>Phase I</i>	<i>Phase II</i>		<i>RS</i>	<i>StS</i>		<i>24 h</i>	<i>36 h</i>	<i>48 h</i>	<i>60 h</i>
NtOLE6a	8.97	OLE1	3.42	2.78	OLE1	9.00	14.59	OLE1	15.27	10.64	10.23	14.38
NtOLE6b	16.31	OLE2	12.62	17.45	OLE2	12.32	14.49	OLE2	15.14	11.65	11.71	12.36
NtCLO1a	40.32	OLE3	0.03	0.18	OLE3	0.11	0.13	OLE3	0.04	0.05	0.05	0.04
NtCLO1b/(d)	10.28	OLE4	20.64	24.16	OLE4	22.37	19.66	OLE4	14.23	22.53	22.10	19.14
NtCLO4g/h	0.04	OLE5	38.83	23.10	OLE5	14.73	19.90	OLE5	27.17	25.90	26.18	30.19
NtPux10a/b	0.57	OLE6	0.51	0.43	OLE6	0.17	0.23	OLE6	0.19	0.15	0.13	0.18
NtPux10c	3.82	OLE8	0.10	0.11	OLE7	0.01	0.01	OLE7	0	0.01	0	0
NtPux10d	0.25	CLO1	10.30	10.20	OLE8	0.46	0.51	OLE8	0.34	0.72	1.14	0.66
NtLDIPa	0.29	CLO2	1.50	1.95	CLO1	6.66	5.27	CLO1	5.08	3.96	3.38	1.99
NtLDIPb	0.47	CLO3	0.14	0.41	CLO2	1.64	1.58	CLO2	2.23	3.25	4.23	4.23
NtLDAP1a	7.15	CLO5	0	0.01	HSD1/1	30.96	22.48	CLO3	0	0	0.01	0.01
NtLDAP1b	1.78	HSD1/1	7.13	16.66	HSD5	1.05	0.69	HSD1/1	19.00	18.31	14.62	10.41
NtLDAP2a/b	0.43	HSD5	0.97	1.23	HSD6	0.03	0.01	HSD3	0	0	0.03	0.04
NtLDAP3a/b	1.29	HSD6	0.01	0.00	LDAP2	0	0.01	HSD5	0.59	0.35	0.22	0.14
NtCASa	1.04	LDAP1	0.91	0.12	LDIP	0.01	0.01	HSD6	0.01	0.04	0.06	0.02
NtCASb/(c)	4.14	LDAP2	0	0.25	PUX10	0.01	0	LDAP1	0	0.01	0.05	0.04
NtSMT1c/d	0.80	LDAP3	2.60	0.52	CAS/LAS	0.01	0	LDAP2	0.14	1.56	4.09	4.28
NtPTLD1a	0.70	SMT1	0.01	0.01	OBL1	0.05	0.05	LDAP3	0	0.05	0.19	0.31
NtPTLD1b	0.03	OBL1	0.04	0.06	SLDP1	0.02	0.01	LDIP	0.01	0	0.01	0.01
NtPTLD2a	0.44	SLDP1	0.01	0.01	SLDP2	0.31	0.29	PUX10	0.04	0.08	0.05	0.04
NtPTLD2b	0.19	SLDP2	0.22	0.31	LDPS1	0.09	0.10	SMT1	0	0.02	0.11	0.14
NtOBL1b/c	0.06	LDPS1	0	0.02				CAS/LAS	0.01	0.02	0.07	0.06
NtOBL1	0.67							OBL1	0.08	0.12	0.20	0.23
								OBL3	0	0	0.01	0.04
								α-DOX	0	0.01	0.06	0.07
								SLDP1	0	0.02	0.03	0.03
								SLDP2	0.32	0.44	0.56	0.61
								LIDL1	0	0	0.05	0.09
								LIDL2	0	0	0.01	0.01
								LDPS1	0.09	0.05	0.03	0.03
								LIME1/2	0	0.05	0.40	0.20

Despite low level of HSD6 expression being detected by Honys and Twell, we and others could not identify steroleosins in tobacco pollen tubes or *Arabidopsis* pollen grains on the protein level (Grobei et al., 2009), but only in the seeds of these species. This indicates that 1. Not a complete set of plant coat proteins is needed for proper LD formation and function, 2. In addition to coat protein function, the enzymatic activity of steroleosins is only needed in specific tissues (Kretzschmar et al., 2018, Supplemental Figure 3).

As shown in Table 2, the LD proteome of different tissues can be distinguished based on two factors: the diversity and the abundance of different LD proteins within the proteome. In *Arabidopsis thaliana* seedlings, we detected the most diverse proteome. Two of proteins that are described as part of this thesis, LIME and LIDL, were originally discovered in the proteome of *Arabidopsis* seedlings and were not expressed in the other tissues investigated. PUX10 was identified in pollen tubes of tobacco, and seeds and seedlings of *Arabidopsis* (Table 2). However, analysis of transgenic *Arabidopsis* expressing *PUX10* in fusion with GFP under the native *PUX10* promoter revealed that *PUX10* is also expressed in embryos (Kretzschmar et al., 2018, Figure 5). The SLDPs were detected in all seed-related tissues, including tobacco seeds (Kretzschmar et al., 2018, Suppl. Dataset 5 and 6). LDIP, originally discovered in seedlings (Pyc et al., 2017b), can also be found in pollen tubes and in seeds.

Even more interesting than the qualitative comparison between different tissues are the differences on a quantitative basis. The quantification of proteins by proteomics is biased on different levels: the efficiency of protein isolation and accessibility for tryptic digest, the varying ability of peptides to ionize and be detected in the mass spectrometer, and the bias created by processing of the data. Despite of this, it can give good estimates of changes in protein composition.

Pollen tubes, as mentioned above, do not contain steroleosins. More than half of LD proteome is instead contributed by caleosin. In seed-related tissues from *Arabidopsis*, the contribution of caleosins is smaller than 15 % at any investigated time point. When steroleosins are detected, like in all seed-related tissues, they contribute more to the complete LD proteome than caleosins, with the exception being Phase I siliques (Manuscript II, Figure 9). As caleosins are so abundant in pollen tubes, they are also more abundant than oleosins. The two oleosin isoforms detected in pollen tubes only contribute about 25 % to the complete LD proteome. In pollen tubes, the third most abundant protein family are LDAPs, which contribute with more than 10 % twice as much to the complete proteome than in any other tissue investigated.

Another interesting observation is that only in siliques and seedlings but not in pollen tubes or seeds, the most abundant single protein isoform is an oleosin. In seeds, HSD1 is the most abundant protein, and in pollen tubes, CLO1.

Given that these are the first comprehensive, quantitative studies of LD proteomes of different tissues, the functional basis of these differences is very speculative. The presence or absence of certain proteins can be interpreted as pinpointing towards specific functions of these proteins that are either required, or not needed. For example, as mentioned above, the absence of steroleosin from pollen tubes indicates that its putative function in brassinosteroid hormone signaling is not required for pollen tube growth. Pollen contain one of the highest concentration of brassinosteroids of all plant tissues, and brassinolide was first isolated from pollen (Grove et al., 1979; Takatsuto, 1994). The enzymatic activity and preferred substrates of HSD1 have been studied in a targeted manner (d'Andréa et al., 2007). While this clarified that HSD1 can catalyze the dehydrogenation of estradiol and cortisol, it did not react with cholesterol and other 3 $\beta$ -hydroxysteroids as substrate (d'Andréa et al., 2007). The *in vivo* substrate of HSD1, however, is not known. So far, it was also not possible to study knockout mutants of the *HSD1* gene, mainly because *HSD1* is the product of a very recent gene duplication, leaving two copies of the gene very close to each other in the genome (AT5G50600 and AT5G50700) that have not yet diverged on a sequence level. The same is true for other members of the gene family. In the future, it will be hopefully possible to further elucidate the function of this interesting protein family by creating knockout mutants by a CRISPR-CAS9 approach and by investigating the steroid composition of those mutants. This will also potentially reveal new interesting functions about LDs themselves.

It is possible that the function of HSDs or other coat proteins is partially taken over by LDAPs, explaining the high abundance of this protein family in pollen tubes compared to the other investigated tissues. It has been proposed that LDAPs might function as coat proteins in vegetative tissues (Pyc et al., 2017a). This idea is supported by the fact that the knockout of LDAP homologs in the rubber plant Russian dandelion (*Taraxacum brevicorniculatum*) resulted in a similar phenotype like the knockout of oleosins in Arabidopsis seeds. In both cases, rubber particles and LDs were bigger in the mutants (Hillebrand et al., 2012; Shimada et al., 2008). LDAPs were originally discovered in the oil-rich non-seed tissue of avocado mesocarp (Gidda et al., 2013; Horn et al., 2013), but expression analyses of different Arabidopsis tissues revealed that the mRNAs are also detected in dry and imbibed seeds (Gidda et al., 2016). We detected LDAP proteins in all investigated tissues but seeds, meaning in all cases that they share the LD surface with oleosins. This does not exclude their proposed function as coat proteins in oleosin-free tissues but indicates further functions beyond that. Analysis of Arabidopsis LDAP overexpression lines indicated an involvement of the proteins in the regulation of LD abundance during the day/night cycle, abiotic stress, and the regulation of seed oil mobilization during early stages of seedling establishment (Gidda et al., 2016). While these data indicate a systemic role of LDAPs in LD biology, the evidence is so far descriptive and does not further pinpoint toward a molecular concept. More studies, for example by the investigation of the Arabidopsis triple mutant, are necessary.

In conclusion, the proteome of seed and pollen LDs, but especially of seedling LDs is more diverse than originally observed. The diversity is observed on both a qualitative and a quantitative scale. The diversification of the proteome during seedling establishment could be indicative of a transition from seed LDs to vegetative LDs. While it seems like the proteomic differences between seed and vegetative LDs might not be as large as originally anticipated, these distinct populations of LDs probably still have very different core functions. The primary function of seed LDs is to provide an energy and carbon source for the heterotrophic growth of the plant seedling. For this, they have to be stable in relation to time and environmental conditions, especially during desiccation. Vegetative LDs have to be much more dynamic. They have to grow and shrink quickly in response to different stimuli, to provide or take up building blocks (fatty acids, sterols, other hydrophobic metabolites) for or from different metabolic pathways. Like this they could be connected to metabolism taking place also at different places inside the cell (plastids, ER, peroxisomes, mitochondria). Further quantitative proteomics studies of vegetative LDs under different conditions (senescence, pathogen attack, abiotic stresses) might help to elucidate differences and similarities to seed and seedling LDs and fill the gap left by the study in Manuscript II. Time points from older established seedlings might help to clarify when, for example, oleosins are finally degraded completely. This would, according to the current consensus, mark the point of transition from seed to vegetative LDs, and maybe even from heterotrophic to autotrophic growth of seedlings.

### 5.1.2. Conservation of Plant Lipid Droplet-associated Proteins throughout Evolution

As described in the introduction, the most common proteins found on LDs from different domains of life do not share much sequence similarity. However, they are to a certain extent functionally analogous. Both oleosins and PLINs are considered coat proteins. One of the main functions of those protein families embedded in the phospholipid monolayer of the LD is to decorate this monolayer and with this, passively inhibit the coalescence and fusion of the organelle. Thereby, they might also mark the membrane of the organelle in a recognizable manner, not only to us researchers, but also to cell interior.

Elucidating the function of newly identified LD-associated proteins is much easier if functions of related variants have been described. This was true for example in the study of PUX10, as shall be discussed later.

Therefore, the *in silico* analysis of plant LD-associated proteins concerning their evolutionary conservation could provide insights in potential functions of the proteins. Figure 5 shows the result of such an investigation, performed for one isoform of each protein family mentioned in this thesis. A collection of model organism were searched for homologs of this protein with the help of

the Basic Local Alignment Search Tool (BLAST-P). The best hit for each organism with an e-value smaller than  $10^{-10}$  was taken into consideration.

The conservation of the plant LD proteome, thus, varies substantially. Only two proteins were strongly evolutionary conserved: HSDs and Plant UBX-domain containing proteins (as UBX domain-containing proteins). For UBX domain-containing proteins, conserved functions in the cell's protein homeostasis have been shown (Schuberth and Buchberger, 2005). Five other proteins are not exclusively plant-specific: the recently described putative metabolic enzymes LIME and LIDL (Manuscript II), caleosins,  $\alpha$ -DOX1, and the lipase OBL1. While the closest homolog of OBL1 in *Neurospora crassa* is a lipase with an e-value of  $\sim 10^{-4}$ , closer homologs can be found in other fungi such as *Rhizopus microsporus*.

Especially interesting here is that homologs of  $\alpha$ -DOX1, caleosins and the lipase OBL1 are present in fungi. All three of them have been implemented in the production of oxylipins (Garbowicz et al., 2018; Shimada et al., 2014) and at least  $\alpha$ -DOX1 and the caleosin CLO3 act in a common pathway. Oxylipins are also formed by fungi (Brodhun and Feussner, 2011). This could be an indication that a whole LD-based module is conserved between plants and fungi.

PTLD and SLDP are the least conserved proteins. PTLD is plant specific, including the liverwort *Marchantia polymorpha*, but excluding both the moss *Physcomitrella patens*, and Brassicaceae. Full length SLDP is only found in Brassicaceae, in other seed plants only the central domain is conserved. Oleosins, LDAPs, LDIP, and LDPS are conserved in land plants. However full length LDPS is only found in Brassicaceae and in sunflower (of the species investigated), the other variants found in the remaining seed plant species lacked the N-terminal region. The observation that oleosins are less conserved than caleosins has led to the hypothesis that oleosins are in fact evolutionary descendants of caleosins (Partridge and Murphy, 2009).



evolution, the necessity to pack larger amounts of it in stable LDs was lost at certain points, and re-emerged later. These LDs were then coated with different proteins available in plants, algae, and animals. Apparently, the ability to synthesize and utilize stable LDs did give organisms an evolutionary advantage under certain conditions, given the omnipresence of LDs in today's life, and the importance they play for example in mammals and seed plants. Particularly visible is the evolutionary advantage coming from the generation of LDs in unicellular algae: Here, LDs are formed as response to nutrient starvation (Zienkiewicz et al., 2016). This process is hallmarked by growth arrest and the degradation of photosynthetic membranes resulting in massive carbon channeling into TAG. Upon resupply of the starved nutrient, cell division and photosynthesis can resume after LD degradation. Thus, algae capable of this process will have had a major survival advantage. A similar process has been observed for plant leaves, where LD formation is induced by different kind of abiotic stresses (Gidda et al., 2016; Kim et al., 2016a; Shimada and Hara-Nishimura, 2015). Thus, the LD production in algae and leaves, and the LD production in storage tissues like seeds and the adipocytes are fundamentally different responses: The first form in response to stress as a sink for carbon, the latter are synthesized particularly in times of abundance and serve as a source during conditions that are more limiting. Conversely, when conditions are limiting, algae will make LDs, and humans will degrade them.

To elucidate varying functions of LDs and LD-associated proteins, different research groups investigated LDs from different non-seed tissues, like avocado mesocarp (Horn et al., 2013) or tobacco pollen tubes (Kretzschmar et al., 2018). While this approach proved fruitful at the first glance, as previously unknown LD-associated protein families were detected, further studies indicated that most of the proteins newly identified were also present in the “classical” system of *Arabidopsis* seeds or seedlings.

In conclusion, the plant LD protein composition is very lowly conserved in other eukaryotes. This slows the elucidation of the protein functions as well as the function of LDs themselves.

### 5.1.3. Targeting of Lipid Droplet-associated Proteins

In principle, proteins interact with membranes in two possible ways: Via direct interaction with the membrane lipids, or via interaction with proteins that interact with the membrane lipids (Kory et al., 2016). Therefore, for all proteins that are somehow associated with a specific membrane in the cell, two defining questions should be answered: How does the protein recognize or find the specific membrane, and how does it interact with it in a stable manner. In many cases, the subcellular localization of proteins is mediated via specific amino acid sequences incorporated in the sequence of the protein.

Recently, LD proteins were divided into two different classes based on their LD targeting mechanism (Kory et al., 2016). According to this, class I proteins are targeted to LDs after (co-



translational) insertion into the ER membrane, a mechanism that was shown for oleosins (Abell et al., 1997). Class II proteins target the LD membrane from the cytosol, which was proposed for example for PLINs. The two classes are distinguished via their membrane domain topology: class I proteins contain a hydrophobic helix or hairpin that can also insert into a bilayer membrane, class II proteins more often have an amphipathic helix or hydrophobic sequences that might not be able to insert into a bilayer membrane.

However, whether each (plant) LD-associated protein can be clearly sorted in one of the two proposed groups is questionable. While oleosins are considered class I proteins, their N- and C-terminal domains are proposed to fold into extensive amphipathic helices (Tzen et al., 1992). Examples like CAS and LDAP, that both do not have a clear hydrophobic sequence (Kretzschmar et al., 2018, Suppl. Figure 25), indicate that LD-associated proteins outside of these two groups exist. For LDAP, no particular sequence was identified for correct LD targeting; rather the complete protein sequence was needed (Gidda et al., 2016). This indicates that a certain protein fold might in fact mediate LD localization. The localization of CAS, on the other hand, might be mediated through interaction with other LD-associated proteins or specific lipids.

In other LD-associated proteins, hydrophobic sequences can be identified. In most cases so far investigated, they contribute to LD localization, and are often even required for it. For example, a hydrophobic sequence in castor bean (*Ricinus communis*) OIL BODY LIPASE1 and 2 were trypsin-resistant, presumably because they were embedded in the membrane (Eastmond, 2004). For PUX10, we have proposed that a combination of hydrophobic domain and amphipathic helix is both necessary and sufficient for LD targeting, and no co-translational insertion signal could be identified (Kretzschmar et al., 2018). This would make it a class II protein. The observation, that more than just the hydrophobic sequence is necessary for correct targeting of PUX10 is consistent with an observation made with the mammalian homolog UBXD8. The hydrophobic sequence of UBXD8 alone cannot achieve correct membrane localization, but a construct containing the hydrophobic region including flanking regions (36 amino acids to both sides) achieves correct localization (Schrul and Kopito, 2016). However, we could observe PUX10 not only on the LD surface but also on the ER when the protein was overexpressed in tobacco pollen tubes. This is in conflict with what Kory et al., describe for class II proteins (Kory et al., 2016), but is again confirming what is known from the mammalian homolog UBXD8 (Olzmann et al., 2013). For UBXD8, the cofactors needed for its recruitment from the cytosol to the ER membrane were recently identified (Schrul and Kopito, 2016). Interestingly, these cofactors are distinct from the proteins involved in the canonical ER protein targeting pathways via the signal recognition particle and SEC61 translocon (Alberts et al., 2008), or the TRC/GET pathway (Favaloro et al., 2008; Stefanovic and Hegde, 2007). The cofactors bind UBXD8 after translation in the cytosol. The binding efficiency is partially dependent on the presence of the region N-terminal of the

hydrophobic sequence, again confirming our truncation results for PUX10, where we could not achieve proper membrane localization by the hydrophobic sequence alone (Kretzschmar et al., 2018, Figure 2D). The cofactors identified for correct UBXD8 localization are interestingly also chaperones for the insertion of peroxisomal membrane proteins (Schrul and Kopito, 2016). These proteins are conserved in plants, and an investigation in their involvement in the localization of PUX10 specifically, and other LD-associated proteins in general, can potentially answer many questions about the localization of LD proteins.

Particularly for class I proteins like oleosins, the question of how the proteins find the LD goes alongside another major puzzle of LD biology: How are sites of LD synthesis defined? For plant DGATs, the final enzyme in the synthesis of TAG from DAG and acyl-CoA, a heterogeneous ER localization pattern has been observed (Shockey et al., 2006). But is the presence of DGATs defining LD synthesis sites and lead to a recruitment of other factors of LD biogenesis? If yes, what are the reasons for the observed heterogeneity of DGAT localization? Proteins marking membrane microdomains (lipid rafts) have been identified on LD membranes in animals (Browman et al., 2007). Oleosins have been shown to have an increased affinity to monolayers (Abell et al., 2004). They might therefore passively migrate to TAG synthesis sites. Additionally to DGAT and oleosins, other ER-resident proteins influencing the size of LDs, like SEIPINs or FITs, find their way to the sites of LD synthesis. In conclusion, many open questions remain concerning the definition of LD synthesis sites and the regulation of LD synthesis components.

One caveat in the discovery of LD-associated proteins is that no consensus sequence for LD localization is known. The search for a common LD localization signal in plants was so far hampered by the small number of known LD proteins. With more LD-associated proteins identified, more data is available for bioinformatics analysis on this topic. If such a signal was identified, it could considerably speed up LD protein identification; making extensive proteomics studies no longer the only method available because candidate proteins could also be identified directly from their protein sequences.

## 5.2. Degradation of LDs and LD proteins

While the degradation of the TAG in seeds is relatively well understood, the same cannot be said for other LD components, may it be proteins or phospholipids, as discussed in the introduction. Here, I will shortly introduce the main cellular pathways for protein or organellar degradation and discuss how they contribute to the turnover of LDs as well as our contributions on that topic.

### 5.2.1. The Degradation of Lipid Droplet Proteins by the Ubiquitin Proteasome System

The Ubiquitin Proteasome System (UPS) is a ubiquitous, conserved and specific protein degradation system. While many more components are necessary, the two most important parts are ubiquitin (Ub) and the proteasome.

Ub is a 76 amino acid-long polypeptide. In the genome, it is usually present as polyubiquitin, meaning that each gene consists of 4 or more copies of Ub, or it is fused with two or three (in Arabidopsis) essential ribosomal proteins (Callis et al., 1995; Kimura and Tanaka, 2010). Ub is conjugated onto an  $\epsilon$ -amino group of a Lysine in the substrate protein in a three-step process that requires ATP and a specific set of enzymes (Alberts et al., 2008). Ubiquitination (or ubiquitylation) acts as a signal to the cellular environment, and it changes the fate of the protein it is conjugated to. The signal is variable and depends on how many single Ubs are added to a protein (mono- or multimonoubiquitination), on which amino acids they are conjugated, or whether and what kind of Ub chain is formed (Alberts et al., 2008). Polyubiquitin chains are created by conjugation of an activated Ub on a Lysine of a Ub already bound to a substrate protein. Homogenous chains of Lysine (K)-48 linked Ubs are well known to target the marked substrate protein for proteasomal degradation. However, less frequently also other homogenous Ub chains, formed with K11 or K29, have been shown to induce proteasomal degradation (Komander and Rape, 2012). Chains of K63, on the other hand, have been associated with autophagy (Zaffagnini and Martens, 2016).

The proteasome is an ATP-dependent protease that degrades proteins that have been marked by polyubiquitination. It is a 2.5 MDa protein complex consisting of multiple subunits that function in substrate recognition, Ub recycling, and proteolytic activity.

The proteasome is restricted to the cytosol and the nucleus. Protein degradation in the aqueous compartments of the organelles or their membranes must therefore be mediated in one of two ways: UPS-independent via organelle-specific proteases and regulation, or after retrotranslocation of the proteins into the UPS-active cytoplasm. The ER famously employs the latter mechanism: The ER-associated protein degradation (ERAD) pathway is based on the selective retrotranslocation of proteins destined for degradation from the ER lumen or membrane back into the cytoplasm. There, the proteins are received by an adapted UPS machinery that

mediates their ubiquitination and channeling towards the proteasome (Smith et al., 2011). In recent years, ERAD components have been studied in detail in yeast and mammals, but still the knowledge and understanding of the process is incomplete. In plants, elucidation of ERAD functions and components is even less exhaustive.

ERAD consists of multiple steps: substrate recognition, retrotranslocation, ubiquitination, substrate extraction, processing, and delivery to the proteasome. Functional Arabidopsis homologs of many essential yeast ERAD proteins involved in the above-mentioned steps have been identified (Liu and Li, 2014). However, one of the still missing components is a UBX-domain containing protein mediating the interaction between ubiquitinated ERAD substrates exiting the ER and cytosolic effector of the UPS, like the AAA-type ATPase CDC48. In yeast, Cdc48p interacts with ER membrane-bound ERAD complexes via a UBX domain-containing scaffold protein Ubx2p (Neuber et al., 2005). Interestingly, Ubx2p has an influence on a yeast cell's lipid metabolism and knockouts of *Ubx2* contain smaller LDs and have a generally reduced TAG content (Wang and Lee, 2012). The mammalian homolog of Ubx2p, UBXD8, is not only involved in ERAD but also in the turnover of a cofactor that controls the activity of adipocyte TAG lipase (ATGL), inhibiting the lipase activity and leading to bigger LDs (Olzmann et al., 2013). In plants, the family of UBX domain-containing proteins (termed PUX for plant UBX domain-containing proteins) has 16 members but is largely uninvestigated (Liu and Li, 2014). Two members, PUX1 and PUX7, have been shown to interact with CDC48 via their conserved UBX domain (Gallois et al., 2013; Rancour et al., 2004). Only PUX7 has been studied with regard to its subcellular localization, which appears nuclear, eliminating it as a candidate of the plant ERAD UBX protein (Gallois et al., 2013). We and others could identify another homolog, PUX10, that localizes to LDs (Kretzschmar et al. 2018, Figure 2A; Deruyffelaere et al. 2018). Analysis of PUX10 mutants revealed a delay in the degradation of LD coat proteins, and an accumulation of ubiquitin on the LDs. K48 ubiquitinations were also detected on ubiquitin, indicating that it was used as a marker for proteasomal degradation (Deruyffelaere et al., 2018; Kretzschmar et al., 2018). The unique LD membrane topology raises questions about how the degradation of LD-associated proteins might be regulated. In particular, LD coat proteins oleosins, caleosins, and steroleosins are tightly bound to the organelle by hydrophobic domains reaching into the TAG core (Chapman et al., 2012). Originally, it was assumed that this protein topology would require proteolysis of the cytosolic domains of oleosin, but it was recently shown that intact oleosins are extracted from the LD before degradation by the proteasome (Deruyffelaere et al., 2015). Therefore, a specific set of proteins is required to extract them if they are marked for degradation. CDC48 has been shown to be necessary for the final extraction of ERAD substrates from the ER membrane (Ye et al., 2001). It is possible that the extraction of proteins from the LD requires less energy than the extraction of a protein from a bilayer, when a hydrophilic domain has to cross through the hydrophobic

membrane. Therefore, for the extraction of proteins from bilayers like the ER, a channel is required to form a hydrophilic pore. For the extraction of monolayer proteins on the other hand, the force generated by AAA ATPases like CDC48 might be sufficient because no hydrophilic domain has to cross a hydrophobic layer.

We could show that PUX10 interacts with CDC48 via its UBX domain and thus helps recruit this cytosolic protein to the LD membrane (Kretzschmar et al. 2018, Figure 4). Interestingly, the effect of PUX10/UBXD8 depletion in Arabidopsis and HeLa cells had reverse effects. In plants, we observed a delayed LD turnover, although this was marked by a delay in protein degradation but no significant effect on TAG hydrolysis (Deruyffelaere et al., 2018; Kretzschmar et al., 2018). In a human cell line, depletion of UBXD8 led to an increased LD degradation (Olzmann et al., 2013). These observations are probably due to the protein degradation substrates that are controlled by PUX10 and UBXD8: Oleosin in plants and ATGL in humans.

Recently, a potential connection of LD protein degradation and ERAD was suggested for class I LD proteins like oleosins (Kory et al., 2016). Through membrane connections between the LD and the ER, these proteins could channel into the ER bilayer where they can become ERAD substrates. While oleosins degradation, particularly during seedling establishment, is better understood than the degradation of any other plant LD-associated protein, no experimental evidence exists to neither proof nor disproof this hypothesis. As mentioned earlier, so far no ER localized UBX-domain containing protein has been identified in plants. The involvement of an UBX-domain containing proteins in ERAD is conserved both in yeast and mammals, and can therefore be expected to also be present in plants. The best candidate protein to date is PUX10. While PUX10 is highest expressed in cells and developmental stages with high LD abundance, it in fact has a basal ubiquitous expression. Additionally, while full-length PUX10 preferably localizes to LDs, it also decorates the ER membrane (Kretzschmar et al., 2018, Figure 2B). Still, whether PUX10 is also the protein involved in the Arabidopsis ERAD remains to be investigated.

The potential function of PUX10 on vegetative LD turnover is also worth exploring. As mentioned above, I expect vegetative LDs to be more dynamic than seed LDs and thus respond faster to changing environmental conditions. This does also require the ability to quickly adapt the proteome. This adaptation can be reflected in the expression of certain proteins that will then localize to LDs, where they for example execute enzymatic functions. But adaptation is also reflected in the targeted degradation of LD-resident proteins that are no longer needed. Therefore, it might be worth exploring if plants lacking PUX10 had difficulties adapting to situations for which the involvement of vegetative LDs has been shown. This includes drought and temperature stress (Gidda et al., 2016; Kim et al., 2016a). Highly dynamic LDs are also observed in guard cells (McLachlan et al., 2016). There, blue light stimuli induce the TAG degradation, and with that presumably also the degradation of LD-associated proteins, to generate the energy needed for

stomatal opening. According to expression data, PUX10 expression is comparably high in these cells (Winter et al., 2007). Reduced or delayed stomatal opening caused by a lack of PUX10 could have different effects. On the one hand, closed stomata reduce water loss and should therefore increase drought resistance. On the other hand, reduced CO<sub>2</sub> levels lead to higher photorespiration rates. In extreme cases, this could also reduce rates of photosynthesis. Under laboratory conditions, no growth defects of *pux10* plants were observed.

In conclusion, we could contribute to the understanding how the degradation of LD-associated proteins is regulated. Further research will clarify how conserved this mechanism is in different plant species and tissues, and what other factors contribute upstream or downstream to the regulation of LD protein degradation. For example, it is still unclear at what point and how oleosins are ubiquitinated, or what other ubiquitination patterns or phosphorylations that were detected by Deruyffelaere et al., mean (Deruyffelaere et al., 2015, 2018).

### 5.2.2. The Degradation of Lipid Droplets by Autophagy

With PUX10, we could identify a component of a LD-localized protein degradation system. However, analysis of the mutant revealed that PUX10 is not essential for the proper degradation of the organelle (Deruyffelaere et al., 2018; Kretzschmar et al., 2018). Therefore, the degradation must be mediated in another way, either by redundant protein factors, or by alternative degradation pathways. Another major pathway of protein and organelle degradation in the cell is autophagy.

Autophagy, deriving from the Greek meaning “self-eating”, is a pathway for the degradation of cellular components via the vacuole (or lysosome in animals). It is a conserved eukaryotic process that is best known for degrading old organelles like mitochondria or peroxisomes. Autophagy is a multistep process that starts with the de novo formation of a double membrane (nucleation) that will expand and engulf the target organelle or macromolecule. After closure of the double membrane, the now mature autophagosome will fuse with the vacuole, leading to the degradation of the inner membrane and all components engulfed previously.

Different versions of autophagy can be distinguished. The term “autophagy” normally refers to macroautophagy, which has been identified in animals, yeast, and plants. The selective autophagy of organelles is achieved by macroautophagy and referred to by organelle-specific terms: mitophagy for mitochondria, pexophagy for peroxisomes, and lipophagy for LDs, to name a few.

Lipophagy was first observed in human hepatocyte cell culture (Singh et al., 2009). Both the pharmacological inhibition of autophagy and the knockdown of an autophagy gene led to an increase of intracellular TAG and LDs. Since its initial discovery, lipophagy has been studied extensively in animals and yeast, and is accepted to be a crucial pathway for lipid homeostasis in

these systems (Ward et al., 2016; Zechner et al., 2017). Due to the compelling evidence for the importance of this pathway in both animals and fungi, it is surprising that there is so little indication for this process in plants. This is true at least for cytosolic LDs as discussed in this thesis. For plastoglobuli, LDs of different composition accumulating in the chloroplast, vacuolar degradation has been known for quite some time and is even considered a hallmark of plant senescence (Elander et al., 2018; Guiamét et al., 1999). So far, there has been only one published incident of cytosolic LD in the vacuole during seed germination (Poxleitner et al., 2006), and there has been no follow-up study confirming and further explaining this observation. As highlighted in the introduction, the severe disruption of seedling establishment after knockout of two TAG lipases indicates that lipophagy, even if the components for it were present, is not used as a major pathway of neutral lipid breakdown (Eastmond, 2006; Kelly et al., 2011). However, LDs were found inside the vacuole also in lily pollen (Jiang et al., 2007). Whether lipophagy in pollen in general, or in lily pollen in particular, is metabolically relevant remains to be investigated.

Autophagy is also required for the differentiation of human lipid storage cells, as this process requires extensive cellular remodeling mediated by autophagy (Baerga et al., 2009; Zhang et al., 2009). In plants, there is only very limited, indirect evidence for the involvement of autophagy during LD biogenesis or cellular differentiation. For example, rice pollen grains lacking one essential autophagy gene contain less TAG and starch, and showed germination defects, leading to male sterility (Kurusu et al., 2014). In *Chlamydomonas reinhardtii*, a unicellular alga that normally accumulates LDs under nitrogen-starvation conditions, this process can be blocked by the pharmacological inhibition of autophagy. This indicates that autophagy provides building blocks required for extensive LD biogenesis (Couso et al., 2018).

## 6. Concluding Remarks

Lipid droplets (LDs) are ubiquitous but unique organelles. They consist of a hydrophobic matrix surrounded by a phospholipid monolayer that is decorated with a specific set of proteins. Such structures are found in all kingdoms of life and have recently been acknowledged as *bona fide* organelles central to the lipid metabolism.

While LD are conserved, the proteins associated with them vary to a large degree between animals and plants. Until recently, only about one dozen proteins or protein families have been shown to localize to plant LDs, a small number compared to both animal LDs and other plant organelles.

The discovery of new LD proteins is not only hampered by the lack of conservation between the plant and animal LD proteome, but also because no consensus sequence for LD localization is known in either domain of life.

We employ a three-step process to identify and characterize new plant LD proteins and with that, help to expand the knowledge about plant LD biology. LD proteins are identified via a label-free bottom-up proteomics approach of total protein fractions and LD-enriched fractions isolated from a specific plant tissue or developmental stage. By calculating enrichment factors, we are able to preselect significantly enriched proteins in the LD-enriched fraction for further analysis. These proteins are then subjected to a cell biological approach to study their subcellular localization in a transient expression system. Candidate proteins that localize to LDs in this approach are further characterized for their role in LD biology.

In this thesis, I present the expansion of the plant LD proteome by eight proteins and protein families. However, a number of factors still exist that limit progress in the research of LD biology. First, the identification of new LD protein candidates would be much facilitated if a consensus signal sequence responsible for LD targeting were known. Second, the characterization of LD-associated proteins has been slow and rather descriptive in recent years. Often, mutants display a phenotype in LD size or abundance, an observation that does not immediately allow conclusions about the molecular mechanisms in which the proteins are involved in LD biology. Lack of homology to the better understood animal LD-associated proteins further hamper progress in this regard. Combining the results of different descriptive studies will in the future hopefully allow a deeper understanding about the functions of individual LD proteins and LDs themselves in the variety of tissues and conditions they are present.

With the in-depth study of one LD-associated protein, PUX10, we could contribute a new piece in the puzzle that is the turnover of proteins on the LDs. PUX10 is involved in, but not essential for, the degradation of LD coat proteins during seedling establishment. Its potential role on vegetative LDs and the plant ERAD pathway remains to be investigated.



## 7. References

- 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. *Plant Cell* 9, 1481–1493.
- Abell, B.M., Hahn, M., Holbrook, L.A., and Moloney, M.M. (2004). Membrane topology and sequence requirements for oil body targeting of oleosin. *Plant J.* 37, 461–470.
- Al-Obaidi, J.R., Halabi, M.F., AlKhalifah, N.S., Asanar, S., Al-Soqeer, A.A., and Attia, M.F. (2017). A review on plant importance, biotechnological aspects, and cultivation challenges of jojoba plant. *Biol. Res.* 50, 1–9.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2008). *Molecular Biology of the Cell* (Garland Science).
- Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.
- Athenstaedt, K., Zweytick, D., Jandrositz, A., Kohlwein, S.D., and Daum, G. (1999). Identification and characterization of major lipid particle proteins of the yeast *Saccharomyces cerevisiae*. *J. Bacteriol.* 181, 6441–6448.
- 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 Cell Physiol.* 51, 1975–1987.
- Baerga, R., Zhang, Y., Chen, P.H., Goldman, S., and Jin, S. (2009). Targeted deletion of autophagy-related 5 (*atg5*) impairs adipogenesis in a cellular model and in mice. *Autophagy* 5, 1118–1130.
- Bannenberg, G., Martínez, M., Hamberg, M., and Castresana, C. (2009). Diversity of the enzymatic activity in the lipoxygenase gene family of *Arabidopsis thaliana*. *Lipids* 44, 85–95.
- Barbosa, A.D., Savage, D.B., and Siniosoglou, S. (2015). Lipid droplet-organelle interactions: Emerging roles in lipid metabolism. *Curr. Opin. Cell Biol.* 35, 91–97.
- 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. *J. Lipid Res.* 48, 837–847.
- Baud, S., Santos Mendoza, M., To, A., Harscoët, E., Lepiniec, L., and Dubreucq, B. (2007). WRINKLED1 specifies the regulatory action of LEAFY COTYLEDON2 towards fatty acid metabolism during seed maturation in *Arabidopsis*. *Plant J.* 50, 825–838.

- Baud, S., Dubreucq, B., Miquel, M., Rochat, C., and Lepiniec, L. (2008). Storage reserve accumulation in *Arabidopsis*: Metabolic and developmental control of seed filling. *Arab. B.* 6, e0113.
- Bewley, J.D., Bradford, K.J., Hilhorst, H.W.M., and Nonogaki, H. (2013). *Seeds - Physiology of development, germination and dormancy* (Springer).
- Binns, D., Januszewski, T., Chen, Y., Hill, J., Markin, V.S., Zhao, Y., Gilpin, C., Chapman, K.D., Anderson, R.G.W., and Goodman, J.M. (2006). An intimate collaboration between peroxisomes and lipid bodies. *J. Cell Biol.* 173, 719–731.
- Blée, E., Boachon, B., Burcklen, M., Guédard, M. Le, Hanano, A., Heintz, D., Ehlting, J., Herrfurth, C., Feussner, I., and Bessoule, 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 Physiol.* 166, 109–124.
- Boavida, L.C., and McCormick, S. (2007). Temperature as a determinant factor for increased and reproducible in vitro pollen germination in *Arabidopsis thaliana*. *Plant J.* 52, 570–582.
- 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 ACYLTRANSFERASE1 in plant sterol homeostasis and leaf senescence. *Plant Physiol.* 152, 107–119.
- Brasaemle, D.L., Subramanian, V., Garcia, A., Marcinkiewicz, A., and Rothenberg, A. (2009). Perilipin A and the control of triacylglycerol metabolism. *Mol. Cell. Biochem.* 326, 15–21.
- 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. *Front. Plant Sci.* 8.
- Brodhun, F., and Feussner, I. (2011). Oxylipins in fungi. *FEBS J.* 278, 1047–1063.
- Browman, D.T., Hoegg, M.B., and Robbins, S.M. (2007). The SPFH domain-containing proteins: more than lipid raft markers. *Trends Cell Biol.* 17, 394–402.
- Bucher, M., Brander, K.A., Sbicego, S., Mandel, T., and Kuhlemeier, C. (1995). Aerobic fermentation in tobacco pollen. *Plant Mol. Biol.* 28, 739–750.
- 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. *Plant Cell* 27, 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., et al. (2017). Mouse fat storage-inducing transmembrane protein 2 (FIT2) promotes lipid droplet accumulation in plants. *Plant Biotechnol. J.* 15, 824–836.

- Callis, J., Carpenter, T., Sun, C.-W., and Vierstrat, R.D. (1995). Structure and evolution of genes encoding polyubiquitin and ubiquitin-like proteins in *Arabidopsis thaliana* ecotype Columbia. *Genetics* 139, 921–939.
- Campbell, N.A. (1997). *Biologie* (Spektrum Akademischer Verlag).
- Cartwright, B.R., Binns, D.D., Hilton, C.L., Han, S., Gao, Q., and Goodman, J.M. (2015). Seipin performs dissectible functions in promoting lipid droplet biogenesis and regulating droplet morphology. *Mol. Biol. Cell* 26, 726–739.
- Cernac, A., and Benning, C. (2004). WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in *Arabidopsis*. *Plant J.* 40, 575–585.
- Chalhoub, B., Denoeud, F., Liu, S., Parkin, I.A.P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., et al. (2014). Early allopolyploid evolution in the post-neolithic *Brassica napus* oilseed genome. *Science*. 345, 950–953.
- Chapman, K.D., and Ohlrogge, J.B. (2012). Compartmentation of triacylglycerol accumulation in plants. *J. Biol. Chem.* 287, 2288–2294.
- Chapman, K.D., Dyer, J.M., and Mullen, R.T. (2012). Biogenesis and functions of lipid droplets in plants. *J. Lipid Res.* 53, 215–226.
- Chapman, K.D., Dyer, J.M., and Mullen, R.T. (2013). Commentary: Why don't plant leaves get fat? *Plant Sci.* 207, 128–134.
- 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. *Sci. China Ser. C Life Sci.* 52, 854–863.
- Chen, J.C.F., Tsai, C.C.Y., and Tzen, J.T.C. (1999). Cloning and secondary structure analysis of Caleosin, a unique calcium-binding protein in oil bodies of plant seeds. *Plant Cell Physiol.* 40, 1079–1086.
- Clough, S., and Bent, A. (1998). Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16, 735–743.
- Clouse, S.D., and Sasse, J.M. (2002). Brassinosteroids: Essential regulators of plant growth and development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 49, 427–451.
- 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. *Nat. Commun.* 9, 1332.

- Couso, I., Pérez-Pérez, M.E., Martínez-Force, E., Kim, H.-S., He, Y., Umen, J.G., and Crespo, J.L. (2018). Autophagic flux is required for the synthesis of triacylglycerols and ribosomal protein turnover in *Chlamydomonas*. *J. Exp. Bot.* *69*, 1355–1367.
- Czabany, T., Wagner, A., Zweytick, D., Lohner, K., Leitner, E., Ingolic, E., and Daum, G. (2008). Structural and biochemical properties of lipid particles from the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.* *283*, 17065–17074.
- d'Andréa, S., Canonge, M., Beopoulos, A., Jolivet, P., Hartmann, M.A., 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*, 222–229.
- 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 Cell Physiol.* *56*, 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*. *Plant Cell* *30*, 2116–2136.
- Eastmond, P.J. (2004). Cloning and characterization of the acid lipase from Castor beans. *J. Biol. Chem.* *279*, 45540–45545.
- Eastmond, P.J. (2006). SUGAR-DEPENDENT1 encodes a patatin domain triacylglycerol lipase that initiates storage oil breakdown in germinating *Arabidopsis* seeds. *Plant Cell* *18*, 665–675.
- Elander, P.H., Minina, E.A., and Bozhkov, P. V. (2018). Autophagy in turnover of lipid stores: Trans-kingdom comparison. *J. Exp. Bot.* *69*, 1301–1311.
- Favaloro, V., Spasic, M., Schwappach, B., and Dobberstein, B. (2008). Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. *J. Cell Sci.* *121*, 1832–1840.
- Fei, W., Alfaro, G., Muthusamy, B.-P., Klaassen, Z., Graham, T.R., Yang, H., and Beh, C.T. (2008). Genome-wide analysis of sterol-lipid storage and trafficking in *Saccharomyces cerevisiae*. *Eukaryot. Cell* *7*, 401–414.
- Fei, W., Shui, G., Zhang, Y., Krahmer, N., Ferguson, C., Kapterian, T.S., Lin, R.C., Dawes, I.W., Brown, A.J., Li, P., et al. (2011). A role for phosphatidic acid in the formation of “supersized” lipid droplets. *PLoS Genet.* *7*.
- Feussner, I., and Kindl, H. (1992). A lipoxygenase is the main lipid body protein in cucumber and

soybean cotyledons during the stage of triglyceride mobilization. *FEBS Lett.* **298**, 223–225.

Focks, N., and Benning, C. (2002). wrinkled1 : A novel, low-seed-oil mutant of *Arabidopsis* with a deficiency in the seed-specific regulation of carbohydrate metabolism. *Plant Physiol.* **118**, 91–101.

Gallois, J.-L., Drouaud, J., Lécureuil, A., Guyon-Debast, A., Bonhomme, S., and Guerche, P. (2013). Functional characterization of the plant ubiquitin regulatory X (UBX) domain-containing protein AtPUX7 in *Arabidopsis thaliana*. *Gene* **526**, 299–308.

Gao, G., Chen, F.-J., Zhou, L., Su, L., Xu, D., Xu, L., and Li, P. (2017). Control of lipid droplet fusion and growth by CIDE family proteins. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1862**, 1197–1204.

Garbowicz, K., Liu, Z., Alseekh, S., Tieman, D., Taylor, M., Kuhalskaya, A., Ofner, I., Zamir, D., Klee, H.J., Fernie, A.R., et al. (2018). Quantitative trait loci analysis identifies a prominent gene involved in the production of fatty acid-derived flavor volatiles in Tomato. *Mol. Plant* **11**, 1147–1165.

Garnier, C. (1897). Les filaments basaux des cellules glandulaires. *Bibliogr. Anat.* **5**, 278–289.

Gidda, S.K., Watt, S.C., Collins-Silva, J., Kilaru, A., Arondel, V., Yurchenko, O., Horn, P.J., James, C.N., Shintani, D., Ohlrogge, J.B., et al. (2013). Lipid droplet-associated proteins (LDAPs) are involved in the compartmentalization of lipophilic compounds in plant cells. *Plant Signal. Behav.* **8**, e27141.

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 Physiol.* **170**, 2052–2071.

Golgi, C. (1898). Intorno alla struttura delle cellule nervose. *Boll. Della Soc. Medico-Chirurgica Di Pavia* **13**, 3–16.

Granneman, J.G., Moore, H.-P.H., Krishnamoorthy, R., and Rathod, M. (2009). Perilipin controls lipolysis by regulating the interactions of AB-hydrolase containing 5 (Abhd5) and Adipose Triglyceride Lipase (Atgl). *J. Biol. Chem.* **284**, 34538–34544.

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. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1811**, 1165–1175.

Grobei, M.A., Qeli, E., Brunner, E., Rehrauer, H., Zhang, R., Roschitzki, B., Basler, K., Ahrens, C.H., and Grossniklaus, U. (2009). Deterministic protein inference for shotgun proteomics data provides new insights into *Arabidopsis* pollen development and function. *Genome Res.* **19**, 1786–1800.

Grove, M.D., Spencer, G.F., Rohwedder, W.K., Mandava, N., Worley, J.F., Warthen, J.D., Steffens, G.L., Flippen-Anderson, J.L., and Cook, J.C. (1979). Brassinolide, a plant growth-promoting steroid isolated from *Brassica napus* pollen. *Nature* **281**, 216–217.

- Guiamét, J.J., Pichersky, E., and Noodén, L.D. (1999). Mass exodus from senescing soybean chloroplasts. *Plant Cell Physiol.* *40*, 986–992.
- 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. *J. Biol. Chem.* *281*, 33140–33151.
- He, X., Sun, Y., and Zhu, R.L. (2013). The oil bodies of liverworts: Unique and important organelles in land plants. *CRC. Crit. Rev. Plant Sci.* *32*, 293–302.
- Hillebrand, A., Post, J.J., Wurbs, D., Wahler, D., Lenders, M., Krzyzanek, V., Prüfer, D., and Gronover, C.S. (2012). Down-regulation of small rubber particle protein expression affects integrity of rubber particles and rubber content in *Taraxacum brevicorniculatum*. *PLoS One* *7*, e41874.
- Honys, D., and Twell, D. (2004). Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol.* *5*, R85.
- 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 Physiol.* *162*, 1926–1936.
- Hsiao, E.S.L., and Tzen, J.T.C. (2011). Ubiquitination of oleosin-H and caleosin in sesame oil bodies after seed germination. *Plant Physiol. Biochem.* *49*, 77–81.
- Hsieh, K., and Huang, A.H.C. (2004). Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. *Plant Physiol.* *136*, 3427–3434.
- Huang, C.-Y., Chung, C.-I., Lin, Y.-C., Hsing, Y.-I.C., and Huang, A.H.C. (2009). Oil bodies and oleosins in *Physcomitrella* possess characteristics representative of early trends in evolution. *Plant Physiol.* *150*, 1192–1203.
- Itabe, H., Yamaguchi, T., Nimura, S., and Sasabe, N. (2017). Perilipins: a diversity of intracellular lipid droplet proteins. *Lipids Health Dis.* *16*, 1–11.
- Jacquier, N., Choudhary, V., Mari, M., Toulmay, A., Reggiori, F., and Schneider, R. (2011). Lipid droplets are functionally connected to the endoplasmic reticulum in *Saccharomyces cerevisiae*. *J. Cell Sci.* *124*, 2424–2437.
- 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 Cell Physiol.* *48*, 812–821.
- Jiang, P.L., Jauh, G.Y., Wang, C.S., and Tzen, J.T.C. (2008). A unique caleosin in oil bodies of lily pollen. *Plant Cell Physiol.* *49*, 1390–1395.

- Kadereit, B., Kumar, P., Wang, W.-J., Miranda, D., Snapp, E.L., Severina, N., Torregroza, I., Evans, T., and Silver, D.L. (2008). Evolutionarily conserved gene family important for fat storage. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 94–99.
- 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*, 4586–4598.
- Kaushik, S., and Cuervo, A.M. (2015). Degradation of lipid droplet-associated proteins by chaperone-mediated autophagy facilitates lipolysis. *Nat. Cell Biol.* *17*, 759–770.
- 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 Physiol.* *157*, 866–875.
- Kennedy, E.P. (1961). Biosynthesis of complex lipids. *Fed. Proc.* *20*, 934–940.
- 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 Physiol.* *170*, 2494–2510.
- Kim, H.U., Hsieh, K., Ratnayake, C., and Huang, A.H.C. (2002). A novel group of oleosins is present inside the pollen of *Arabidopsis*. *J. Biol. Chem.* *277*, 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*. *Plant J.* *85*, 758–771.
- Kim, Y.Y., Jung, K.W., Yoo, K.S., Jeung, J.U., and Shin, J.S. (2011). A stress-responsive caleosin-like protein, AtCLO4, acts as a negative regulator of ABA responses in *Arabidopsis*. *Plant Cell Physiol.* *52*, 874–884.
- Kimmel, A.R., Brasaemle, D.L., McAndrews-Hill, M., Sztalryd, C., and Londos, C. (2010). Adoption of PERILIPIN as a unifying nomenclature for the mammalian PAT-family of intracellular lipid storage droplet proteins. *J. Lipid Res.* *51*, 468–471.
- Kimura, Y., and Tanaka, K. (2010). Regulatory mechanisms involved in the control of ubiquitin homeostasis. *J. Biochem.* *147*, 793–798.
- Komander, D., and Rape, M. (2012). The ubiquitin code. *Annu. Rev. Biochem.* *81*, 203–229.
- Kory, N., Farese, R. V., and Walther, T.C. (2016). Targeting fat: Mechanisms of protein localization to lipid droplets. *Trends Cell Biol.* *26*, 535–546.
- Kovsan, J., Ben-Romano, R., Souza, S.C., Greenberg, A.S., and Rudich, A. (2007). Regulation of adipocyte lipolysis by degradation of the Perilipin protein. *J. Biol. Chem.* *282*, 21704–21711.

- Krahmer, N., Guo, Y., Wilfling, F., Hilger, M., Lingrell, S., Heger, K., Newman, H.W., Schmidt-Supprian, M., Vance, D.E., Mann, M., et al. (2011). Phosphatidylcholine synthesis for lipid droplet expansion is mediated by localized activation of CTP:phosphocholine cytidyltransferase. *Cell Metab.* *14*, 504–515.
- Kretzschmar, F.K., Mengel, L.A., Müller, A.O., Schmitt, K., Bliersch, K.F., Valerius, O., Braus, G.H., and Ischebeck, T. (2018). PUX10 is a lipid droplet-localized scaffold protein that interacts with CELL DIVISION CYCLE48 and is involved in the degradation of lipid droplet proteins. *Plant Cell* *30*, 2137–2160.
- Kuerschner, L., Moessinger, C., and Thiele, C. (2008). Imaging of lipid biosynthesis: How a neutral lipid enters lipid droplets. *Traffic* *9*, 338–352.
- Kurusu, T., Koyano, T., Hanamata, S., Kubo, T., Noguchi, Y., Yagi, C., Nagata, N., Yamamoto, T., Ohnishi, T., Okazaki, Y., et al. (2014). OsATG7 is required for autophagy-dependent lipid metabolism in rice postmeiotic anther development. *Autophagy* *10*, 878–888.
- Lass, A., Zimmermann, R., Oberer, M., and Zechner, R. (2011). Lipolysis - A highly regulated multi-enzyme complex mediates the catabolism of cellular fat stores. *Prog. Lipid Res.* *50*, 14–27.
- Li, F., Asami, T., Wu, X., Tsang, E.W.T., and Cutler, A.J. (2007). A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. *Plant Physiol.* *145*, 87–97.
- Lin, L.-J., Tai, S.S.K., Peng, C.-C., and Tzen, J.T.C. (2002). Steroleosin, a sterol-binding dehydrogenase in seed oil bodies. *Plant Physiol.* *128*, 1200–1211.
- Liu, Y., and Li, J. (2014). Endoplasmic reticulum-mediated protein quality control in Arabidopsis. *Front. Plant Sci.* *5*, 162.
- Lombard, J., López-García, P., and Moreira, D. (2012). The early evolution of lipid membranes and the three domains of life. *Nat. Rev. Microbiol.* *10*, 507–515.
- M'barek, K. Ben, Ajjaji, D., Chorlay, A., Vanni, S., Forêt, L., and Thiam, A.R. (2017). ER membrane phospholipids and surface tension control cellular lipid droplet formation. *Dev. Cell* *41*, 591–604.
- Magré, J., Delépine, M., Khallouf, E., Gedde-Dahl, T., Van Maldergem, L., Sobel, E., Papp, J., Meier, M., Mégarbané, A., Bachy, A., et al. (2001). Identification of the gene altered in Berardinelli-Seip congenital lipodystrophy on chromosome 11q13. *Nat. Genet.* *28*, 365–370.
- Masuda, Y., Itabe, H., Odaki, M., Hama, K., Fujimoto, Y., Mori, M., Sasabe, N., Aoki, J., Arai, H., and Takano, T. (2006). ADRP/adipophilin is degraded through the proteasome-dependent pathway during regression of lipid-storing cells. *J. Lipid Res.* *47*, 87–98.
- McLachlan, D.H., Lan, J., Geilfus, C.M., Dodd, A.N., Larson, T., Baker, A., Hōrak, H., Kollist, H., He, Z.,



- Graham, I., et al. (2016). The breakdown of stored triacylglycerols is required during light-induced stomatal opening. *Curr. Biol.* 26, 707–712.
- Mellema, S., Eichenberger, W., Rawyler, A., Suter, M., Tadege, M., and Kuhlemeier, C. (2002). The ethanolic fermentation pathway supports respiration and lipid biosynthesis in tobacco pollen. *Plant J.* 30, 329–336.
- Miranda, D.A., Kim, J.-H., Nguyen, L.N., Cheng, W., Tan, B.C., Goh, V.J., Tan, J.S.Y., Yaligar, J., KN, B.P., Velan, S.S., et al. (2014). Fat storage-inducing transmembrane protein 2 is required for normal fat storage in adipose tissue. *J. Biol. Chem.* 289, 9560–9572.
- Moellering, E.R., and Benning, C. (2010). RNA interference silencing of a major lipid droplet protein affects lipid droplet size in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* 9, 97–106.
- Müller, A.O., and Ischebeck, T. (2018). Characterization of the enzymatic activity and physiological function of the lipid droplet-associated triacylglycerol lipase AtOBL1. *New Phytol.* 217, 1062–1076.
- 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. *Plant J.* 89, 1055–1064.
- Murphy, D.J. (2001). The biogenesis and functions of lipid bodies in animals, plants and microorganisms. *Prog. Lipid Res.* 40, 325–438.
- Murphy, D.J. (2006). The extracellular pollen coat in members of the Brassicaceae: Composition, biosynthesis, and functions in pollination. *Protoplasma* 228, 31–39.
- Murphy, D.J. (2012). The dynamic roles of intracellular lipid droplets: From archaea to mammals. *Protoplasma* 249, 541–585.
- 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 Mol. Biol.* 44, 463–476.
- Nawamawat, K., Sakdapipanich, J.T., Ho, C.C., Ma, Y., Song, J., and Vancso, J.G. (2011). Surface nanostructure of *Hevea brasiliensis* natural rubber latex particles. *Colloids Surfaces A Physicochem. Eng. Asp.* 390, 157–166.
- Neuber, O., Jarosch, E., Volkwein, C., Walter, J., and Sommer, T. (2005). Ubx2 links the Cdc48 complex to ER-associated protein degradation. *Nat. Cell Biol.* 7, 993–998.
- Noll, F., May, C., and Kindl, H. (2000). Phospholipid monolayer of plant lipid bodies attacked by phospholipase A2 shows 80 nm holes analyzed by atomic force microscopy. *Biophys. Chem.* 86, 29–35.

- Nonogaki, H., Bassel, G.W., and Bewley, J.D. (2010). Germination—Still a mystery. *Plant Sci.* **179**, 574–581.
- Oh, S.K., Kang, H., Shin, D.H., Yang, J., Chow, K.S., Yeang, H.Y., Wagner, B., Breiteneder, H., and Han, K.H. (1999). Isolation, characterization, and functional analysis of a novel cDNA clone encoding a small rubber particle protein from *Hevea brasiliensis*. *J. Biol. Chem.* **274**, 17132–17138.
- Olzmann, J. a, Richter, C.M., and Kopito, R.R. (2013). Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1345–1350.
- Onal, G., Kutlu, O., Gozuacik, D., and Dokmeci Emre, S. (2017). Lipid Droplets in Health and Disease. *Lipids Health Dis.* **16**, 128.
- Osborne, T.B. (1924). *The vegetable protein* (London: Longmans, Green).
- Pagnon, J., Matzaris, M., Stark, R., Meex, R.C.R., Macaulay, S.L., Brown, W., O'Brien, P.E., Tiganis, T., and Watt, M.J. (2012). Identification and functional characterization of PROTEIN KINASE A phosphorylation sites in the major lipolytic protein, Adipose Triglyceride Lipase. *Endocrinology* **153**, 4278–4289.
- Partridge, M., and Murphy, D.J. (2009). Roles of a membrane-bound caleosin and putative peroxygenase in biotic and abiotic stress responses in Arabidopsis. *Plant Physiol. Biochem.* **47**, 796–806.
- Pasaribu, B., Chen, C.S., Liao, Y.K., Jiang, P.L., and Tzen, J.T.C. (2017). Identification of caleosin and oleosin in oil bodies of pine pollen. *Plant Physiol. Biochem.* **111**, 20–29.
- Platt, K.A., Huang, A.H.C., and Thomson, W.W. (1998). Ultrastructural study of lipid accumulation in tapetal cells of *Brassica napus* L. cv. Westar during microsporogenesis. *Int. J. Plant Sci.* **159**, 724–737.
- Poxleitner, M., Rogers, S.W., Lacey Samuels, A., Browse, J., and Rogers, J.C. (2006). A role for caleosin in degradation of oil-body storage lipid during seed germination. *Plant J.* **47**, 917–933.
- Prost, I., Dhondt, S., Rothe, G., Vicente, J., José Rodriguez, M., Kift, N., Carbonne, F., Griffiths, G., Esquerré-Tugayé, M.-T., Rosahl, S., et al. (2005). Evaluation of the antimicrobial activities of plant oxylipins supports their involvement in defense against pathogens. *Plant Physiol.* **139**, 1902–1913.
- Pyc, M., Cai, Y., Greer, M.S., Yurchenko, O., Chapman, K.D., Dyer, J.M., and Mullen, R.T. (2017a). Turning over a new leaf in lipid droplet biology. *Trends Plant Sci.* **22**, 596–609.
- Pyc, M., Cai, Y., Gidda, S.K., Yurchenko, O., Park, S., Kretzschmar, F.K., Ischebeck, T., Valerius, O., Braus, G.H., Chapman, K.D., et al. (2017b). Arabidopsis LDAP-interacting protein (LDIP) influences

lipid droplet size and neutral lipid homeostasis in both leaves and seeds. *Plant J.* **12**, 3218–3221.

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.). *Biochem. J.* **235**, 57–65.

Rahman, F., Hassan, M., Hanano, A., Fitzpatrick, D.A., McCarthy, C.G.P., and Murphy, D.J. (2018). Evolutionary, structural and functional analysis of the caleosin/psychoxygenase gene family in the Fungi. *BMC Genomics* **19**, 1–24.

Rajjou, L., Duval, M., Gallardo, K., Catusse, J., Bally, J., Job, C., and Job, D. (2012). Seed germination and vigor. *Annu. Rev. Plant Biol.* **63**, 507–533.

Rancour, D.M., Park, S., Knight, S.D., and Bednarek, S.Y. (2004). Plant UBX domain-containing protein 1, PUX1, regulates the oligomeric structure and activity of Arabidopsis CDC48. *J. Biol. Chem.* **279**, 54264–54274.

Ross, J.H.E., 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. *Plant J.* **9**, 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. *Plant J.* **92**, 129–146.

Rowe, E.R., Mimmack, M.L., Barbosa, A.D., Haider, A., Isaac, I., Ouberaï, M.M., Thiam, A.R., Patel, S., Saudek, V., Siniosoglou, S., et al. (2016). Conserved amphipathic helices mediate lipid droplet targeting of Perilipins 1–3. *J. Biol. Chem.* **291**, 6664–6678.

Santos Mendoza, M., Dubreucq, B., Miquel, M., Caboche, M., and Lepiniec, L. (2005). LEAFY COTYLEDON 2 activation is sufficient to trigger the accumulation of oil and seed specific mRNAs in Arabidopsis leaves. *FEBS Lett.* **579**, 4666–4670.

Schrul, B., and Kopito, R.R. (2016). Peroxin-dependent targeting of a lipid droplet-destined membrane protein to ER-subdomains. *Nat. Cell Biol.* **18**, 740–751.

Schuberth, C., and Buchberger, A. (2005). Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER-associated protein degradation. *Nat. Cell Biol.* **7**, 999–1006.

Schuldiner, M., and Bohnert, M. (2017). A different kind of love – lipid droplet contact sites. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1862**, 1188–1196.

Sharma, M., Gupta, S.K., and Mondal, A.K. (2012). Production and trade of major world oil crops. In *Technological Innovations in Major World Oil Crops, Volume 1: Breeding*, S.K. Gupta, ed. (New

York, NY: Springer New York), pp. 1–322.

Shen, B., Allen, W.B., Zheng, P., Li, C., Glassman, K., Ranch, J., Nubel, D., and Tarczynski, M.C. (2010). Expression of ZmLEC1 and ZmWRI1 increases seed oil production in maize. *Plant Physiol.* *153*, 980–987.

Shimada, T.L., and Hara-Nishimura, I. (2015). Leaf oil bodies are subcellular factories producing antifungal oxylipins. *Curr. Opin. Plant Biol.* *25*, 145–150.

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*. *Plant J.* *55*, 798–809.

Shimada, T.L., Takano, Y., Shimada, T., Fujiwara, M., Fukao, Y., Mori, M., Okazaki, Y., Saito, K., Sasaki, R., Aoki, K., et al. (2014). Leaf oil body functions as a subcellular factory for the production of a phytoalexin in *Arabidopsis*. *Plant Physiol.* *164*, 105–118.

Shimada, T.L., Takano, Y., and Hara-Nishimura, I. (2015). Oil body-mediated defense against fungi: From tissues to ecology. *Plant Signal. Behav.* *10*, e989036.

Shockey, J.M., Gidda, S.K., Chaptial, D.C., Kuan, J.-C., Dhanoa, P.K., Bland, J.M., Rothstein, S.J., Mullen, R.T., and Dyer, J.M. (2006). Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell* *18*, 2294–2313.

Singh, R., Kaushik, S., Wang, Y., Xiang, Y., Novak, I., Komatsu, M., Tanaka, K., Cuervo, A.M., and Czaja, M.J. (2009). Autophagy regulates lipid metabolism. *Nature* *458*, 1131–1135.

Smith, M.H., Ploegh, H.L., and Weissman, J.S. (2011). Road to ruin: Targeting proteins for degradation in the endoplasmic reticulum. *Science* (80-. ). *334*, 1086–1090.

Soni, K.G., Mardones, G.A., Sougrat, R., Smirnova, E., Jackson, C.L., and Bonifacio, J.S. (2009). Coatamer-dependent protein delivery to lipid droplets. *J. Cell Sci.* *122*, 1834–1841.

Stefanovic, S., and Hegde, R.S. (2007). Identification of a targeting factor for posttranslational membrane protein insertion into the ER. *Cell* *128*, 1147–1159.

Stone, S.L., Braybrook, S.A., Paula, S.L., Kwong, L.W., Meuser, J., Pelletier, J., Hsieh, T.-F., Fischer, R.L., Goldberg, R.B., and Harada, J.J. (2008). *Arabidopsis* LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 3151–3156.

Szymanski, K.M., Binns, D., Bartz, R., Grishin, N. V., Li, W.-P., Agarwal, A.K., Garg, A., Anderson, R.G.W., and Goodman, J.M. (2007). The lipodystrophy protein seipin is found at endoplasmic reticulum lipid droplet junctions and is important for droplet morphology. *Proc. Natl. Acad. Sci. U.*

S. A. *104*, 20890–20895.

Takahashi, S., Katagiri, T., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2000). An Arabidopsis gene encoding a Ca<sup>2+</sup>-binding protein is induced by abscisic acid during dehydration. *Plant Cell Physiol.* *41*, 898–903.

Takatsuto, S. (1994). Brassinosteroids: Distribution in plants, bioassays and microanalysts by gas chromatography-mass spectrometry. *J. Chromatogr. A* *658*, 3–15.

Tarnopolsky, M.A., Rennie, C.D., Robertshaw, H.A., Fedak-Tarnopolsky, S.N., Devries, M.C., and Hamadeh, M.J. (2007). Influence of endurance exercise training and sex on intramyocellular lipid and mitochondrial ultrastructure, substrate use, and mitochondrial enzyme activity. *Am. J. Physiol. Integr. Comp. Physiol.* *292*, R1271–R1278.

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 Physiol.* *176*, 1531–1546.

Thazar-Poulot, N., Miquel, M., Fobis-Loisy, I., and Gaude, T. (2015). Peroxisome extensions deliver the Arabidopsis SDP1 lipase to oil bodies. *Proc. Natl. Acad. Sci. U. S. A.* *112*, 4158–4163.

Theodoulou, F.L., and Eastmond, P.J. (2012). Seed storage oil catabolism: A story of give and take. *Curr. Opin. Plant Biol.* *15*, 322–328.

Thiam, A.R., and Forêt, L. (2016). The physics of lipid droplet nucleation, growth and budding. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* *1861*, 715–722.

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. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 13244–13249.

Thiam, A.R., Farese, R. V., and Walther, T.C. (2013b). The biophysics and cell biology of lipid droplets. *Nat. Rev. Mol. Cell Biol.* *14*, 775–786.

Turchetto-Zolet, A.C., Maraschin, F.S., De Morais, G.L., Cagliari, A., Andrade, C.M.B., Margis-Pinheiro, M., and Margis, R. (2011). Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. *BMC Evol. Biol.* *11*, 263.

Tzen, J.T.C., and Huang, A.H.C. (1992). Surface structure and properties of plant seed oil bodies. *J. Cell Biol.* *117*, 327–335.

Tzen, J.T.C., Lie, G.C., and Huang, A.H.C. (1992). Characterization of the charged components and

their topology on the surface of plant seed oil bodies. *J. Biol. Chem.* **267**, 15626–15634.

Vindigni, J.D., Wien, F., Giuliani, A., Erpapazoglou, Z., Tache, R., Jagic, F., Chardot, T., Gohon, Y., and Froissard, M. (2013). Fold of an oleosin targeted to cellular oil bodies. *Biochim. Biophys. Acta - Biomembr.* **1828**, 1881–1888.

Walther, T.C., and Farese Jr., R. V. (2012). Lipid droplets and cellular lipid metabolism. *Annu. Rev. Biochem.* **81**, 687–714.

Wang, C.-W., and Lee, S.-C. (2012). The ubiquitin-like (UBX)-domain-containing protein Ubx2/Ubx8 regulates lipid droplet homeostasis. *J. Cell Sci.* **125**, 2930–2939.

Wang, C., and St. Leger, R.J. (2007). The *Metarhizium anisopliae* perilipin homolog MPL1 regulates lipid metabolism, appressorial turgor pressure, and virulence. *J. Biol. Chem.* **282**, 21110–21115.

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 Biol.* **16**, 1–12.

Ward, C., Martinez-Lopez, N., Otten, E.G., Carroll, B., Maetzel, D., Singh, R., Sarkar, S., and Korolchuk, V.I. (2016). Autophagy, lipophagy and lysosomal lipid storage disorders. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1861**, 269–284.

van Wijk, K.J., and Kessler, F. (2017). Plastoglobuli: Plastid microcompartments with integrated functions in metabolism, plastid developmental transitions, and environmental adaptation. *Annu. Rev. Plant Biol.* **68**, 253–289.

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**, e718.

Xu, G., Sztalryd, C., and Londos, C. (2006). Degradation of perilipin is mediated through ubiquitination-proteasome pathway. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1761**, 83–90.

Yamaguchi, T., Fujikawa, N., Nimura, S., Tokuoka, Y., Tsuda, S., Aiuchi, T., Kato, R., Obama, T., and Itabe, H. (2015). Characterization of lipid droplets in steroidogenic MLTC-1 Leydig cells: Protein profiles and the morphological change induced by hormone stimulation. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1851**, 1285–1295.

Yang, X., Lu, X., Lombès, M., Rha, G.B., Chi, Y.-I., Guerin, T.M., Smart, E.J., and Liu, J. (2010). The G0/G1 switch gene 2 regulates adipose lipolysis through association with Adipose Triglyceride Lipase. *Cell Metab.* **11**, 194–205.

Ye, Y., Meyer, H.H., and Rapoport, T.A. (2001). The AAA ATPase Cdc48/p97 and its partners

transport proteins from the ER into the cytosol. *Nature* **414**, 652–656.

Zaffagnini, G., and Martens, S. (2016). Mechanisms of selective autophagy. *J. Mol. Biol.* **428**, 1714–1724.

Zechner, R., Madeo, F., and Kratky, D. (2017). Cytosolic lipolysis and lipophagy: two sides of the same coin. *Nat. Rev. Mol. Cell Biol.* **18**, 671–684.

Zhang, Y., Goldman, S., Baerga, R., Zhao, Y., Komatsu, M., and Jin, S. (2009). Adipose-specific deletion of autophagy-related gene 7 (atg7) in mice reveals a role in adipogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19860–19865.

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. *J. Exp. Bot.* **64**, 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. *J. Exp. Bot.* **61**, 1537–1546.

Zienkiewicz, K., Zienkiewicz, A., Rodríguez-García, M.I., and Castro, A.J. (2011). Characterization of a caleosin expressed during olive (*Olea europaea* L.) pollen ontogeny. *BMC Plant Biol.* **11**, 122.

Zienkiewicz, K., Du, Z.Y., Ma, W., Vollheyde, K., and Benning, C. (2016). Stress-induced neutral lipid biosynthesis in microalgae — Molecular, cellular and physiological insights. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1861**, 1269–1281.

# Acknowledgements

Research is a collective effort; I am more aware of this now than I was before I started this project. So there are many people that deserve a Thank you on these pages, and I hope I will do them all justice.

First of all, I would like to thank my supervisor of many years, Dr. Till Ischebeck. Thank you for first convincing me of this project, motivating me during the hard times, and cheering with me during the high times. You always believed in me and in this project!

Thank you also to Prof. Dr. Ivo Feußner, who welcomed me here in his wonderful department. Heike Lott I would like to thank for her support and knowledge of bureaucracy.

I would like to thank Dr. Alexander Stein and Prof. Dr. Christiane Gatz for being in my Thesis Advisory Committee and their support and honesty throughout my PhD.

Furthermore, I am thankful to Prof. Dr. Gerhard Braus for taking part in my examination board but also for providing us with access to his Proteomics platform. Also to Prof. Dr. Jörg Stülke and PD Dr. Marcel Wiermer, I am grateful for being members of my examination board.

Scientifically, the collective efforts of people leading to this dissertation are acknowledged by authorship of the publications contained in it. But a mere authorship does not do justice to Dr. Oliver Valerius and Dr. Kerstin Schmitt, who measured endless proteomics samples, who were both trusting and critical, and helped me emerge myself in the fascinating field of Proteomics.

Similarly, the impact of Prof. Dr. Robert Mullen, Prof. Dr. Kent Chapman, and Dr. John Dyer, and their students, most importantly Dr. Michal Pyc and Nathan Donor, is hard to put into words. Representative of the whole plant lipid community, which is warm, supportive and welcoming, I would like to thank them for sharing their ideas and passions, and letting me be part of their fascinating projects that both challenged and excited me. Also Prof. Dr. Sabine d'Andrea is a perfect example of this community, and I am grateful for the way we cooperated and collaborated.

Furthermore, I would like to thank the people that graciously granted us access to their confocal microscopes, and worked to maintain them: Prof. Dr. Steven Johnson and Dr. Florian Wegwitz, and Prof. Dr. Jörg Großhans and Johannes Sattmann.

My everyday lab life would have been much harder without the support of Dr. Anna Müller, Dr. Ellen Hornung, Sabine Freitag, and Susanne Mester. Anna, especially during my early PhD days, you taught me all the things I needed to know, and I am glad that we were not only good colleagues but are also friends. Ellen und Sabine, ihr seid beide fantastisch in eurem Job und könnt selbst die komischsten Fragen beantworten (wovon ich einige hatte), aber ihr seid vor allem auch das Herz und die Seele dieser Abteilung, und eure Unterstützung geht weit über das Wissenschaftliche hinaus. Danke! Susanne, ohne dich hätte ich hier wohl wenige Experimente erfolgreich abgeschlossen, vielen Dank für deine Pflege des Tabaks und der Arabidopsen.

The whole AG Ischebeck deserves a big Thank you! Athanas, Elisa, Patricia, Maurice, Alex, Antony, Max, Philipp, you made working in the lab always fun, you were always happy to help out here and there, and to discuss science (among other things). A special thanks again to those people who



have worked on this project and helped me out: my labrotation students Maurice and Antony, Patricia, Laura, Kathi, Lara, and Kirsten.

The Department of Plant Biochemistry is a great and supportive environment to work in. I would like to thank the Post Docs that help me and answered many questions: Dr. Kirstin Feußner, Dr. Amélie Kelly, Dr. Agnieszka Zienkiewicz, Dr. Krzysztof Zienkiewicz, and Dr. Martin Fulda. Thank you to the PhD student community for making my time in this department so worthwhile: Jasmin, Athanas, Elisa, Milena, Yi-Tse, Dimi, Hanno, Kathi and Sven. Especially Elisa I would like to thank for two crazy years with many shared beds and poster prizes and chats.

Outside the Lab, I am grateful for the friendship of Cathie, Julia, Madhobi, and Marija. Thank you for night-outs and night-ins, discussion about literally anything, a lot of room for complaints. Your friendship has helped me grow so much as a person, and I am proud of the person you helped me become. You always believed in me, and in turn I believe in you so much, and I wish you all the best for your future (although we all agree on our general thoughts about this future).

Der Studienstiftung des Deutschen Volkes möchte ich danken für die Finanzierung meines Promotionsstipendiums, und für die fantastischen Events, bei denen ich teilnehmen durfte. Die Erfahrungen, die ich dort gesammelt habe, werden mich hoffentlich noch lange begleiten.

Der Graduate School for Molecular Biology, und ganz besonders Dr. Steffen Burkhardt und Kerstin Grüniger, gebührt ein besonderer Dank. Danke für die finanzielle Unterstützung, danke für die Organisation des Programms als ganzen und der Unterstützung unzähliger Events von Culture Nights bis Horizons. Auf euch ist immer Verlass. Ihr haltet den Laden am Laufen, und eure Leidenschaft für diese Arbeit ist inspirierend und beispielhaft. Ich werde mich immer gerne an meine Zeit in Göttingen und als Molbio zurückerinnern.

Zuletzt gebührt mein Dank den Menschen, die schon mein Leben lang hinter mir stehen und das auch noch für den Rest davon tun werden: Meiner Familie. Mama, Papa, Kathi und der Rest der Familie. Nichts was ich hier schreiben kann wird dem Dank gerecht werden, den ihr verdient habt! Ohne eure Unterstützung und euren Glauben daran, dass ich alles schaffen kann, wäre ich nicht hier.

Hanno, auch du bist jetzt ein Teil meiner Familie. Dafür, und dass ich ein Teil deiner wunderbaren Familie sein darf, bin ich unendlich dankbar. Deine Herzenswärme und dein Humor bringen mich auch durch die schwersten Zeiten. Danke!