A seed-like proteome in oil-rich tubers

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

[1]*Department of Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences and Göttingen Center for Molecular Biosciences (GZMB), University of Göttingen, 37077, Göttingen, Germany,*
[2]*Department of Applied Bioinformatics, Göttingen Center for Molecular Biosciences (GZMB) and Campus Institute Data Science (CIDAS), Institute for Microbiology and Genetics, University of Göttingen, 37077, Göttingen, Germany,*
[3]*Department for Molecular Microbiology and Genetics, Genetics and Göttingen Center for Molecular Biosciences (GZMB) and Service Unit LCMS Protein Analytics, Institute for Microbiology, University of Göttingen, 37077, Göttingen, Germany,*
[4]*Department of Plant Biochemistry, Service Unit for Metabolomics and Lipidomics, Göttingen Center for Molecular Biosciences (GZMB), University of Göttingen, 37077, Göttingen, Germany,*
[5]*Department of Plant Breeding, SLU Alnarp, Swedish University of Agricultural Sciences, Box 190, SE-234 22, Lomma, Sweden, and*
[6]*Green Biotechnology, Institute of Plant Biology and Biotechnology (IBBP), University of Münster, 48143, Münster, Germany*

SUMMARY

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

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

INTRODUCTION

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

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

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

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

RESULTS

Tubers of yellow nutsedge are enriched in hallmark proteins of seeds

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

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

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

We additionally compared the expression levels of the genes coding for the tuber-specific proteins by investigating the recently published transcriptomes of yellow nutsedge and the respective homologs in purple nutsedge (Ji et al., 2021; Figure 2b) In this study, tubers were harvested 20, 50, and 90 days after tuber formation. All of the seven 'seed' protein-coding genes had at least 10 times higher transcript levels in yellow nutsedge than in purple
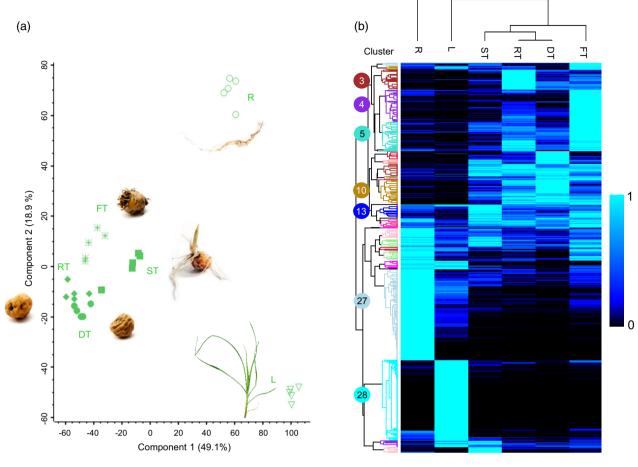
Figure 1. Developmental tuber stages have a distinct proteome.
Comparative proteomics of total protein extracts of tubers (fresh [FT], dry [DT], rehydrated for 48 h [RT], and sprouted [ST]), leaves (L), and roots (R) (Data S1). (a) Principal component analysis plot. (b) The values were normalized setting the average of the stage with the highest abundance to 1 and these values were hierarchically clustered. Clusters are labeled with numbers in circles. Proteins sorted by cluster can be found in Data S3. $n = 5$ biological replicates per stage.

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

Proteomes of tubers of yellow nutsedge and Arabidopsis seeds correlate

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

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

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

Second, hierarchical clustering of the stages showed again a similarity of Arabidopsis seed and yellow but not purple nutsedge tuber stages. Furthermore, separating the proteins into 30 clusters (Figure 3b and Data S8) yielded two clusters (3 and 4), comprising 78 proteins with a marked elevation in seeds rehydrated for 30 min and seeds stratified for 74 h at 4°C, and in dry and rehydrated tubers.
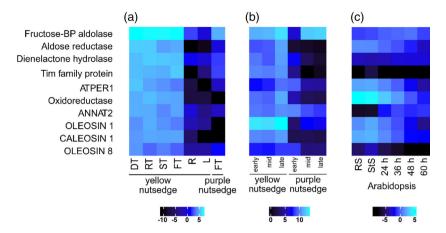
Figure 2. Highly abundant and tuber-specific proteins display higher expression in yellow than in purple nutsedge tubers and have seed-enriched homologs in Arabidopsis.

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

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

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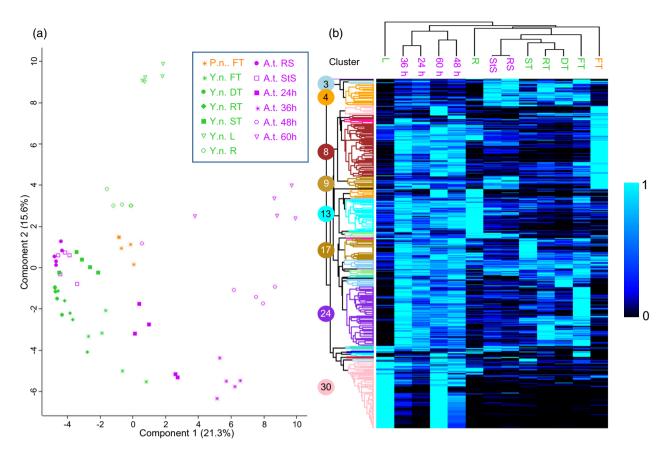
Homologs to proteins with important functions in seeds and seedlings are enriched in tubers

GO term analyses based on the total abundance of proteins (refer to Data S9 for yellow nutsedge, Data S10 for purple nutsedge, and Data S11 for a comparison of both species to Arabidopsis) bolstered the similarities of yellow nutsedge tubers and Arabidopsis seeds (Figure 4). For instance, the terms 'embryo development', 'lipid storage', 'maintenance of seed dormancy', and 'seed germination' were especially high in seeds and tubers but reduced in older stages. The term 'maintenance of seed dormancy' only comprises one protein, a homolog to the Arabidopsis protein 1-CYSTEINE PEROXIREDOXIN 1 (PER1), which enhances primary seed dormancy by suppressing abscisic acid (ABA) catabolism and gibberellic acid biosynthesis (Chen, Ruan, et al., 2020). Another protein, homolog to GEM-RELATED 5 (GER5, AT5G13200), displays a similar pattern (Data S3). *GER5* is an ABA-responsive gene that also regulates germination in Arabidopsis seeds (Baron et al., 2014). Based on these results, we tested if ABA levels were increased in nutsedge tubers. Indeed, ABA levels (51 pmol g^{-1} fresh weight) in dry tubers decreased to 7% (Figure 5; Data S12) of the dry tuber levels during sprouting. Also in roots and leaves of yellow nutsedge the levels were much lower. In the tubers of purple nutsedge, no ABA could be detected. The tubers were not enriched in any other hormones we could detect in comparison to other tissues, with jasmonoyl-isoleucine (JA-Ile) and salicylic acid (SA) levels being especially high in roots and leaves, respectively.

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

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

families are enriched in tubers and decrease during sprouting (Data S9, GO term 45735), Arabidopsis seeds contain a far greater amount of storage proteins amounting to more than half of the total protein based on the proteomic data; this amount is far lower in tubers (<0.5%). Tubers on the other hand contain many proteins associated with protein

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

Figure 4. GO term comparison.

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

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

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

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

Lipid droplet proteins known from seeds abound in tubers of yellow nutsedge

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

Homologs to known LD proteins made up approximately 60–70% of the proteins in LD-enriched fractions (Data S16), with only approximately 5% being found in total extracts (purple nutsedge tubers for comparison only contained 0.2% LD proteins in the total extracts). With one exception, all these homologs were significantly enriched in the LD fraction (Figure 6a), indicating that these homologs are also LD-associated. The most abundant among these proteins were the oleosins (Figure 6b), especially the isoforms with higher sequence similarity to the
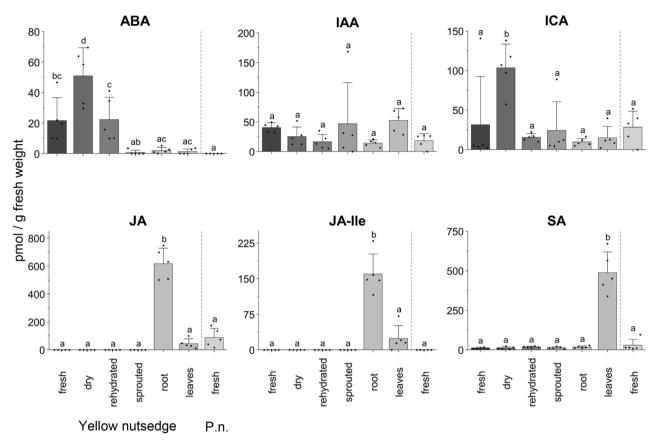
Figure 5. Hormone analysis.
Hormone levels were determined by UPLC-nanoESI-MS/MS in different tissues of yellow nutsedge and fresh tubers of purple nutsedge (P.n.). $n = 5$ for each stage. Error bars indicate standard deviation. ABA, abscisic acid; IAA, indole-3-acetic acid; ICA, indole carboxylic acid; JA, jasmonic acid; JA-Ile, jasmonoyl-isoleucine; SA, salicylic acid.

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

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

detected in tubers (Figure S5b). Interestingly, seed lipid droplet protein (SLDP), another LD-associated protein (Kretzschmar et al., 2020) that is seed- and seedling-specific in Arabidopsis (Klepikova et al., 2016) and is part of a membrane contact site between LDs and the plasma membrane (Krawczyk et al., 2022; Scholz et al., 2022), was also detected in tubers.
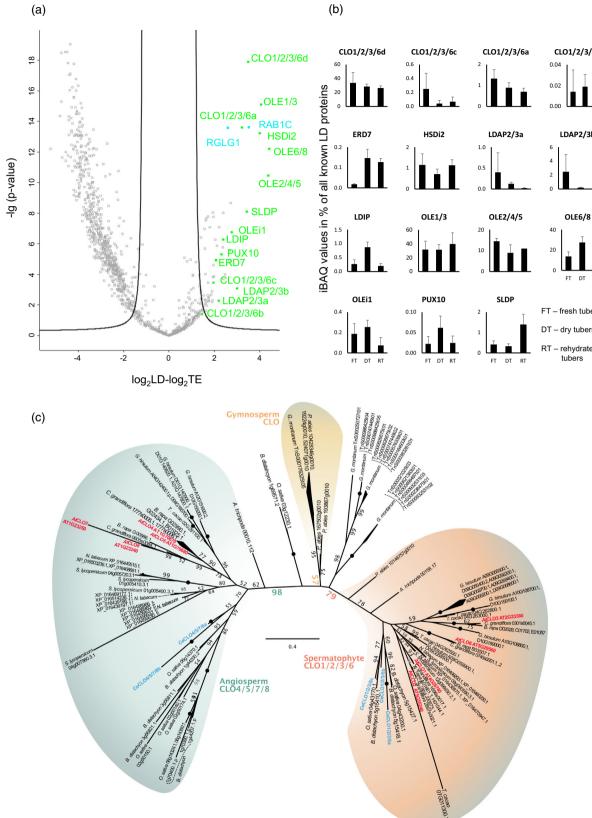
Caleosins (CLOs), in comparison, have been found in seeds, pollen, and leaves (Chen et al., 1999; Jiang et al., 2008; Shimada & Hara-Nishimura, 2015). In Arabidopsis, CLO1 and CLO2 are the dominant caleosins of

Figure 6. LD proteomes of yellow nutsedge tubers resemble those of seeds and seedlings.
(a) A volcano plot was constructed to visualize proteins, which are significantly and consistently LD-enriched in yellow nutsedge. All 15 LD-enriched samples were compared to all total cellular fractions of fresh, dry, and rehydrated tubers (Data S1). The log₂-transformed values and P-values were calculated. Known LD proteins are indicated in green, while interesting candidates are shown in blue (see Data S16 for a list of all candidates). Black lines indicate a false discovery rate of 0.01. (b) The riBAQ values of LD-associated proteins in the LD-enriched fraction were calculated as a percentage of the riBAQ of all known LD-associated proteins (see Data S9 for data). (c) Unrooted maximum likelihood phylogeny of caleosin homologs detected in the predicted proteomes of spermatophyte genomes and the yellow nutsedge transcriptomes. Yellow nutsedge and Arabidopsis caleosins are depicted in blue and red, respectively. The angiosperm and gymnosperm depicted protein clades contain only proteins from species of the respective plant clades while the spermatophyte depicted protein clade contains proteins from both plant clades. Abbreviations: CLO, caleosin; ERD, early responsive to dehydration; HSD, steroleosin; LDAP, LD-associated protein; LDIP, LDAP-interacting protein; OLE, oleosin; PUX, plant UBX domain-containing protein; RAB, Rab GTPase; RGLG, RING domain ligase, SLDP, seed LD protein.

(a)

(b)

(c)

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

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

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

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

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

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

There are striking similarities between Arabidopsis seeds and tubers of yellow but not purple nutsedge. It is therefore conceivable that oil accumulation and desiccation tolerance of both organs are based on a similar proteome and the underlying gene expression pattern that could be regulated by a similar network of hormones and transcription factors, as for example the transcription factors ABSCISIC ACID INSENSITIVE3 (ABI3), WRINKLED1 (WRI1), and LEAFY COTYLEDON1 (LEC1), which show increased expression levels in the tubers of yellow nutsedge in comparison to purple nutsedge (Figure S7; Ji et al., 2021).

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

DISCUSSION

Yellow nutsedge tubers contain a seed-like proteome

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

These comparable functions are mirrored by striking similarities to seed (Data S5) but not tuber proteomes (Figure S3) of other dicotyledons on several levels. The tubers of yellow nutsedge contain LEA proteins that are known to protect the cells during desiccation (Hanin et al., 2011) and homologs to other seed-specific proteins in Arabidopsis like PER1 and a plastidial aldehyde reductase, whose function in seeds has so far not been elucidated (Data S5). The tubers harbor LD-associated proteins (Figure 6) like oleosins highly prevalent in seeds (Huang, 2018) that have to the best of our knowledge so far not been found in the proteomes of vegetative tissues from flowering plants (Brocard et al., 2017; Horn et al., 2013). We also found a homolog to SLDP1 and steroleosins that have so far only been found in seeds and seedlings at the proteome level (Ischebeck et al., 2020). As steroleosins are implicated in brassinosteroid metabolism, this warrants future work on their role in development, dormancy, and/or sprouting of yellow nutsedge. Furthermore, similar to seeds (Chen, Li, et al., 2020), tuber dormancy and desiccation tolerance might be regulated by the phytohormone ABA. ABA levels
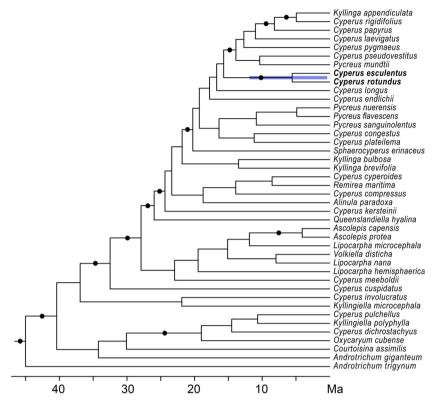
Figure 7. Yellow and purple nutsedge diverged approximately 5.6 million years ago according to a Bayesian molecular clock analysis.
Excerption of the time-calibrated phylogeny of the Cyperaceae inferred with a Bayesian relaxed molecular clock on three plastid markers (*rbcL*, *trnL* intron, and *trnL–trnF* spacer; 2460 aligned nucleotides) and 267 species. The divergence between yellow and purple nutsedge was inferred to be 5.6 million years ago (95% highest posterior density interval, 0.3–12 million years ago). Dots represent branches with >90% ultrafast bootstrap support. The full tree is depicted in Figure S8 and Data S18.

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

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

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

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

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

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

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

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

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

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

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

EXPERIMENTAL PROCEDURES

Plant material

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

Isolation of total and LD-enriched protein fractions

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

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

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

Peptide sample preparation

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

LC-MS/MS analysis

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

Generation of a protein library of yellow nutsedge for peptide identification

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

Generation of a protein library of purple nutsedge for peptide identification

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

The RNA-seq reads from all 15 samples were filtered for quality, base contents, and adapters and subsequently trimmed for polyG/polyA using fastp (v0.20.1). The resulting 1023 million reads were assembled into a *de novo* transcriptome assembly for purple nutsedge using TRINITY (v2.11.0). All filtered reads were mapped back to the transcriptome assembly and the back-mapping rate was calculated using Bowtie2 (v2.4.2) and RSEM (v1.2.28). The completeness of the transcriptome was assessed using BUSCO (v5.0.0) with the embryophyta_odb10 (2019/11/27) dataset in transcriptome mode.

The transcriptome assembly was annotated using the Trinotate pipeline by the following steps.

First, the longest ORFs in transcriptome assembly were identified using Transdecoder (v5.5.0). The resulting candidate coding sequences were further refined based on sequence and domain homology to known protein sequences in the Swiss-Prot (access date: 22-02-21) and PFAM (v34.0) databases, using NCBI BLAST+ (v2.11.0) and HMMER (v3.3.2), respectively.

In addition, RNA-seq data of purple nutsedge tubers (Ji et al., 2021) were downloaded from NCBI (SRA) and coding sequences were predicted using TransDecoder v.5.5.0 with default settings.

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

Calculation of riBAQ values

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

Analysis of yam and potato data

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

BLAST

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

PCA

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

Hierarchical clustering

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

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

Analysis of gene expression

RNA-seq data of purple and yellow nutsedge tubers (Ji et al., 2021) were downloaded from NCBI (SRA). Transcriptomes were assembled *de novo* using Trinity v2.11.0. Gene expression was estimated with RSEM after read mapping with BOWTIE2, followed by cross-sample normalization (trimmed mean of M [TMM]) and retaining the most abundant isoform per gene, following established protocols (Haas et al., 2013). Homologs for the 10 most abundant proteins identified by proteomics were searched on the two transcriptome assemblies (containing only the most highly expressed isoforms) using tBLASTn and an e-value threshold of 1e−6. Homology was confirmed upon multiple sequence alignment (mafft; Katoh & Standley, 2013) and when necessary confirmed by maximum likelihood phylogenetic inference (IQ-TREE; Nguyen et al., 2015). For each species, TMM expression values were added for protein fragments, isoforms, and close homologs, averaged across time points and replicates, and summarized on a heatmap after log$_2$ transformation.

Pearson correlation

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

GO term enrichment

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

Quantitative GO term analysis

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

Volcano plots

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

Analysis of LD proteins

All proteins that were homologous to a known LD protein from Arabidopsis were considered as LD proteins. The abundance of each LD protein was divided by the total abundance of all LD proteins in the sample.

Phytohormone measurements

Phytohormones were extracted with methyl-tert-butyl ether, reversed phase-separated using an ACQUITY UPLC system (Waters Corp., Milford, MA, USA), and analyzed by nanoelectrospray ionization (nanoESI) (TriVersa Nanomate; Advion BioSciences, Ithaca, NY, USA) coupled with an AB Sciex 4000 QTRAP tandem mass spectrometer (AB Sciex) employed in scheduled multiple reaction monitoring modes (Herrfurth & Feussner, 2020) with the following modifications. For quantification, 10 ng D$_4$-SA, 10 ng D$_6$-ABA, 10 ng D$_5$-jasmonic acid (JA; all three from C/D/N Isotopes Inc., Pointe-Claire, Canada), and 20 ng D$_5$-IAA (Eurisotop, Freising, Germany) were added at the beginning of the extraction procedure. For ABA, indole-3-acetic acid (IAA), indole carboxylic acid (ICA), and SA analysis, the following mass transitions were included: 137/93 (declustering potential [DP] −25 V, entrance potential [EP] −6 V, collision energy [CE] −20 V) for SA, 141/97 (DP −25 V, EP −6 V, CE −22 V) for D$_4$-SA, 160/116 (DP −40 V, EP −6.5 V, CE −22 V) for ICA, 174/130 (DP −35 V, EP −9 V, CE −14 V) for IAA, 179/135 (DP −35 V, EP −9 V, CE −14 V) for D$_5$-IAA, 263/153 (DP −35 V, EP −4 V, CE −14 V) for ABA, and 269/159 (DP −30 V, EP −5 V, CE −16 V) for D$_6$-ABA. JA-Ile was quantified with D$_5$-JA as internal standard.

Phylogenetic analyses

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

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

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

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

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

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

Estimation of divergence times

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

AUTHOR CONTRIBUTIONS

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

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

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

SUPPORTING INFORMATION

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

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

Figure S2. PCA plot of yellow and purple nutsedge.

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

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

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

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

Figure S7. Expression levels of transcription factors.

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

Data S1 Normalized iBAQ values of yellow nutsedge.

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

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

Data S4 List of tuber-enriched proteins.

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

Data S6 Normalized iBAQ values of purple nutsedge.

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

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

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

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

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

Data S12 Hormone analysis.

Data S13 Analysis of dried purple nutsedge tubers.

Data S14 Analysis of potato data.

Data S15 Analysis of yam data.

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

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

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

REFERENCES

Alejo-Jacuinde, G., Gonzalez-Morales, S.I., Oropeza-Aburto, A., Simpson, J. & Herrera-Estrella, L. (2020) Comparative transcriptome analysis suggests convergent evolution of desiccation tolerance in *Selaginella* species. *BMC Plant Biology*, **20**, 468.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Amborella Genome, P. (2013) The Amborella genome and the evolution of flowering plants. *Science*, **342**, 1241089.

Argout, X., Salse, J., Aury, J.M., Guiltinan, M.J., Droc, G., Gouzy, J. *et al.* (2011) The genome of *Theobroma cacao*. *Nature Genetics*, **43**, 101–108.

Aubert, Y., Vile, D., Pervent, M., Aldon, D., Ranty, B., Simonneau, T. *et al.* (2010) RD20, a stress-inducible caleosin, participates in stomatal control, transpiration and drought tolerance in *Arabidopsis thaliana*. *Plant and Cell Physiology*, **51**, 1975–1987.

Banks, J.A., Nishiyama, T., Hasebe, M., Bowman, J.L., Gribskov, M., DePamphilis, C. *et al.* (2011) The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science*, **332**, 960–963.

Barajas-Lopez, J.D., Tiwari, A., Zarza, X., Shaw, M.W., Pascual, J.S., Punkkinen, M. *et al.* (2021) EARLY RESPONSE TO DEHYDRATION 7 remodels cell membrane lipid composition during cold stress in Arabidopsis. *Plant & Cell Physiology*, **62**, 80–91.

Baron, K.N., Schroeder, D.F. & Stasolla, C. (2014) GEm-related 5 (GER5), an ABA and stress-responsive GRAM domain protein regulating seed development and inflorescence architecture. *Plant Science*, **223**, 153–166.

Bartels, D. (2005) Desiccation tolerance studied in the resurrection plant *Craterostigma plantagineum*. *Integrative and Comparative Biology*, **45**, 696–701.

Battaglia, M., Olvera-Carrillo, Y., Garciarrubio, A., Campos, F. & Covarrubias, A.A. (2008) The enigmatic LEA proteins and other hydrophilins. *Plant Physiology*, **148**, 6–24.

Baud, S., Dichow, N.R., Kelemen, Z., d'Andrea, S., To, A., Berger, N. *et al.* (2009) Regulation of HSD1 in seeds of *Arabidopsis thaliana*. *Plant & Cell Physiology*, **50**, 1463–1478.

Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K. *et al.* (2017) Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell*, **171**, 287–304.

Brocard, L., Immel, F., Coulon, D., Esnay, N., Tuphile, K., Pascal, S. *et al.* (2017) Proteomic analysis of lipid droplets from arabidopsis aging leaves brings new insight into their biogenesis and functions. *Frontiers in Plant Science*, **8**, 894.

Chapman, K.D., Dyer, J.M. & Mullen, R.T. (2012) Biogenesis and functions of lipid droplets in plants. *Journal of Lipid Research*, **53**, 215–226.

Chen, H., Ruan, J., Chu, P., Fu, W., Liang, Z., Li, Y. *et al.* (2020) AtPER1 enhances primary seed dormancy and reduces seed germination by suppressing the ABA catabolism and GA biosynthesis in Arabidopsis seeds. *The Plant Journal*, **101**, 310–323.

Chen, J.C.F., Tsai, C.C.Y. & Tzen, J.T.C. (1999) Cloning and secondary structure analysis of caleosin, a unique calcium-binding protein in oil bodies of plant seeds. *Plant and Cell Physiology*, **40**, 1079–1086.

Chen, K., Li, G.J., Bressan, R.A., Song, C.P., Zhu, J.K. & Zhao, Y. (2020) Abscisic acid dynamics, signaling, and functions in plants. *Journal of Integrative Plant Biology*, **62**, 25–54.

Cheng, M.C., Hsieh, E.J., Chen, J.H., Chen, H.Y. & Lin, T.P. (2012) Arabidopsis RGLG2, functioning as a RING E3 ligase, interacts with AtERF53 and negatively regulates the plant drought stress response. *Plant Physiology*, **158**, 363–375.

Cheng, S., Xian, W., Fu, Y., Marin, B., Keller, J., Wu, T. *et al.* (2019) Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell*, **179**, 1057–1067.

Cormier, F., Lawac, F., Maledon, E., Gravillon, M.C., Nudol, E., Mournet, P. *et al.* (2019) A reference high-density genetic map of greater yam (*Dioscorea alata* L.). *Theoretical and Applied Genetics*, **132**, 1733–1744.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N. & Mann, M. (2014) Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular Proteomics*, **13**, 2513–2526.

Cox, J. & Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized ppb-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, **26**, 1367–1372.

Das Gupta, M. & Tsiantis, M. (2018) Gene networks and the evolution of plant morphology. *Current Opinion in Plant Biology*, **45**, 82–87.

de Vries, J., Curtis, B.A., Gould, S.B. & Archibald, J.M. (2018) Embryophyte stress signaling evolved in the algal progenitors of land plants. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, E3471–E3480.

de Vries, J., de Vries, S., Curtis, B.A., Zhou, H., Penny, S., Feussner, K. *et al.* (2020) Heat stress response in the closest algal relatives of land plants reveals conserved stress signalling circuits. *Plant Journal*, **103**, 1025–1048.

de Vries, J. & Ischebeck, T. (2020) Ties between stress and lipid droplets pre-date seeds. *Trends in Plant Science*, **25**, 1203–1214.

Defelice, M.S. (2002) Yellow nutsedge *Cyperus esculentus* L.—snack food of the gods. *Weed Technology*, **16**, 901–907.

Deruyffelaere, C., Bouchez, I., Morin, H., Guillot, A., Miquel, M., Froissard, M. *et al.* (2015) Ubiquitin-mediated proteasomal degradation of oleosins is involved in oil body mobilization during post-germinative seedling growth in Arabidopsis. *Plant and Cell Physiology*, **56**, 1374–1387.

Deruyffelaere, C., Purkrtova, Z., Bouchez, I., Collet, B., Cacas, J.L., Chardot, T. *et al.* (2018) PUX10 associates with CDC48A and regulates the dislocation of ubiquitinated oleosins from seed lipid droplets. *Plant Cell*, **30**, 2116–2136.

Doner, N.M., Seay, D., Mehling, M., Sun, S., Gidda, S.K., Schmitt, K. *et al.* (2021) *Arabidopsis thaliana* EARLY RESPONSIVE TO DEHYDRATION 7 localizes to lipid droplets via its senescence domain. *Frontiers in Plant Science*, **12**, 658961.

Escudero, M. & Hipp, A. (2013) Shifts in diversification rates and clade ages explain species richness in higher-level sedge taxa (Cyperaceae). *American Journal of Botany*, **100**, 2403–2411.

Fernandez-Pozo, N., Haas, F.B., Meyberg, R., Ullrich, K.K., Hiss, M., Perroud, P.F. *et al.* (2019) PEATmoss (Physcomitrella expression atlas tool): a unified gene expression atlas for the model plant *Physcomitrella patens*. *The Plant Journal*, **102**, 165–177.

Fernández-Santos, R., Izquierdo, Y., López, A., Muñiz, L., Martínez, M., Cascón, T. *et al.* (2020) Protein profiles of lipid droplets during the hypersensitive defense response of Arabidopsis against pseudomonas infection. *Plant & Cell Physiology*, **61**, 1144–1157.

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

Gao, B., Li, X., Zhang, D., Liang, Y., Yang, H., Chen, M. *et al.* (2017) Desiccation tolerance in bryophytes: the dehydration and rehydration transcriptomes in the desiccation-tolerant bryophyte *Bryum argenteum*. *Scientific Reports*, **7**, 7571.

Gidda, S.K., Park, S., Pyc, M., Yurchenko, O., Cai, Y., Wu, P. *et al.* (2016) Lipid droplet-associated proteins (LDAPs) are required for the dynamic regulation of neutral lipid compartmentation in plant cells. *Plant Physiology*, **170**, 2052–2071.

Gish, W. & States, D.J. (1993) Identification of protein coding regions by database similarity search. *Nature Genetics*, **3**, 266–272.

Guo, L.-M., Li, J., He, J., Liu, H. & Zhang, H.-M. (2020) A class I cytosolic HSP20 of rice enhances heat and salt tolerance in different organisms. *Scientific Reports*, **10**, 1383.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J. *et al.* (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, **8**, 1494–1512.

Han, C., Yin, X., He, D. & Yang, P. (2013) Analysis of proteome profile in germinating soybean seed, and its comparison with rice showing the styles of reserves mobilization in different crops. *PLoS One*, **8**, e56947.

Hanin, M., Brini, F., Ebel, C., Toda, Y., Takeda, S. & Masmoudi, K. (2011) Plant dehydrins and stress tolerance: versatile proteins for complex mechanisms. *Plant Signaling & Behavior*, **6**, 1503–1509.

Harrison, C.J. & Morris, J.L. (2018) The origin and early evolution of vascular plant shoots and leaves. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **373**, 20160496.

Herrfurth, C. & Feussner, I. (2020) Quantitative Jasmonate profiling using a high-throughput UPLC-NanoESI-MS/MS method. *Methods in Molecular Biology*, **2085**, 169–187.

Hofvander, P., Ischebeck, T., Turesson, H., Kushwaha, S.K., Feussner, I., Carlsson, A.S. *et al.* (2016) Potato tuber expression of Arabidopsis WRINKLED1 increase triacylglycerol and membrane lipids while affecting central carbohydrate metabolism. *Plant Biotechnology Journal*, **14**, 1883–1898.

Horn, P.J., James, C.N., Gidda, S.K., Kilaru, A., Dyer, J.M., Mullen, R.T. *et al.* (2013) Identification of a new class of lipid droplet-associated proteins in plants. *Plant Physiology*, **162**, 1926–1936.

Hsiao, E.S.L. & Tzen, J.T.C. (2011) Ubiquitination of oleosin-H and caleosin in sesame oil bodies after seed germination. *Plant Physiology and Biochemistry*, **49**, 77–81.

Hu, T.T., Pattyn, P., Bakker, E.G., Cao, J., Cheng, J.F., Clark, R.M. *et al.* (2011) The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. *Nature Genetics*, **43**, 476–481.

Huang, A.H.C. (2018) Plant lipid droplets and their associated proteins: potential for rapid advances. *Plant Physiology*, **176**, 1894–1918.

Huang, C.-Y., Chung, C.-I., Lin, Y.-C., Hsing, Y.-I.C. & Huang, A.H. (2009) Oil bodies and oleosins in Physcomitrella possess characteristics representative of early trends in oil bodies. *Plant Physiology*, **150**, 1192–1203.

Iqbal, J., Hussain, S., Ali, A. & Javaid, A. (2012) Biology and management of purple nutsedge (*Cyperus rotundus* L.). *The Journal of Animal & Plant Sciences*, **22**, 384–389.

Ischebeck, T. (2016) Lipids in pollen — they are different. *Biochimica et Biophysica Acta*, **1861**, 1315–1328.

Ischebeck, T., Krawczyk, H.E., Mullen, R.T., Dyer, J.M. & Chapman, K.D. (2020) Lipid droplets in plants and algae: distribution, formation, turnover and function. *Seminars in Cell & Developmental Biology*, **108**, 82–93.

Ischebeck, T., Valledor, L., Lyon, D., Gingl, S., Nagler, M., Meijon, M. *et al.* (2014) Comprehensive cell-specific protein analysis in early and late pollen development from diploid microsporocytes to pollen tube growth. *Molecular & Cellular Proteomics*, **13**, 295–310.

Jayson, C.B.K., Arlt, H., Fischer, A.W., Lai, Z.W., Farese, R.V., Jr. & Walther, T.C. (2018) Rab18 is not necessary for lipid droplet biogenesis or turnover in human mammary carcinoma cells. *Molecular Biology of the Cell*, **29**, 2045–2054.

Ji, H., Liu, D. & Yang, Z. (2021) High oil accumulation in tuber of yellow nutsedge compared to purple nutsedge is associated with more abundant expression of genes involved in fatty acid synthesis and triacylglycerol storage. *Biotechnology for Biofuels*, **14**, 54.

Jiang, P.-L., Jauh, G.-Y., Wang, C.-S. & Tzen, J.T.C. (2008) A unique caleosin in oil bodies of lily pollen. *Plant & Cell Physiology*, **49**, 1390–1395.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A. & Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, **14**, 587–589.

Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.

Klepikova, A.V., Kasianov, A.S., Gerasimov, E.S., Logacheva, M.D. & Penin, A.A. (2016) A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *The Plant Journal*, **88**, 1058–1070.

Krawczyk, H.E., Sun, S., Doner, N.M., Yan, Q., Lim, M.S.S., Scholz, P. *et al.* (2022) SEED LIPID DROPLET PROTEIN1, SEED LIPID DROPLET PROTEIN2 and LIPID DROPLET PLASMA MEMBRANE ADAPTOR mediate lipid droplet-plasma membrane tethering. *The Plant Cell*, **34**, 2424–2448.

Kretzschmar, F.K., Doner, N., Krawczyk, H.E., Scholz, P., Schmitt, K., Valerius, O. *et al.* (2020) Identification of low-abundance lipid droplet proteins in seeds and seedlings. *Plant Physiology*, **182**, 1236–1245.

Kretzschmar, F.K., Mengel, L.F., Müller, A., Schmitt, K., Blersch, K.F., Valerius, O. *et al.* (2018) PUX10 is a lipid droplet-localized scaffold protein that interacts with CDC48 and is involved in the degradation of lipid droplet proteins. *The Plant Cell*, **30**, 2137–2160.

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R. *et al.* (2012) The Arabidopsis information resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, **40**, D1202–D1210.

Lang, D., Ullrich, K.K., Murat, F., Fuchs, J., Jenkins, J., Haas, F.B. *et al.* (2018) The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *The Plant Journal*, **93**, 515–533.

Le, T.N. & McQueen-Mason, S.J. (2006) Desiccation-tolerant plants in dry environments. *Reviews in Environmental Science and Bio/Technology*, **5**, 269–279.

Lebecka, R., Kistowski, M., Dębski, J., Szajko, K., Murawska, Z. & Marczewski, W. (2019) Quantitative proteomic analysis of differentially expressed proteins in tubers of potato plants differing in resistance to *Dickeya solani*. *Plant and Soil*, **441**, 317–329.

Li, F., Asami, T., Wu, X., Tsang, E.W. & Cutler, A.J. (2007) A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. *Plant Physiology*, **145**, 87–97.

Li, F., Fan, G., Lu, C., Xiao, G., Zou, C., Kohel, R.J. *et al.* (2015) Genome sequence of cultivated upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. *Nature Biotechnology*, **33**, 524–530.

Li, F.W., Brouwer, P., Carretero-Paulet, L., Cheng, S., de Vries, J., Delaux, P.M. *et al.* (2018) Fern genomes elucidate land plant evolution and cyanobacterial symbioses. *Nature Plants*, **4**, 460–472.

Li, F.W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z. *et al.* (2020) Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. *Nat Plants*, **6**, 259–272.

Linssen, J.P., Cozijnsen, J.L. & Pilnik, W. (1989) Chufa (*Cyperus esculentus*): a new source of dietary fibre. *Journal of the Science of Food and Agriculture*, **49**, 291–296.

Liu, S., Liu, Y., Yang, X., Tong, C., Edwards, D., Parkin, I.A. *et al.* (2014) The *Brassica oleracea* genome reveals the asymmetrical evolution of polyploid genomes. *Nature Communications*, **5**, 3930.

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y.O. & Borodovsky, M. (2005) Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research*, **33**, 6494–6506.

Lu, K.J., van 't Wout Hofland, N., Mor, E., Mutte, S., Abrahams, P., Kato, H. *et al.* (2020) Evolution of vascular plants through redeployment of ancient developmental regulators. *Proceedings of the National Academy of Sciences of the United States of America*, **117**, 733–740.

Maumus, F. & Quesneville, H. (2014) Ancestral repeats have shaped epigenome and genome composition for millions of years in *Arabidopsis thaliana*. *Nature Communications*, **5**, 4104.

Mi, H., Muruganujan, A., Huang, X., Ebert, D., Mills, C., Guo, X. *et al.* (2019) Protocol update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nature Protocols*, **14**, 703–721.

Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J.H. *et al.* (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). *Nature*, **452**, 991–996.

Neale, A.D., Blomstedt, C.K., Bronson, P., Le, T.N., Guthridge, K., Evans, J. *et al.* (2000) The isolation of genes from the resurrection grass *Sporobolus stapfianus* which are induced during severe drought stress. *Plant, Cell & Environment*, **23**, 265–277.

Nguyen, L.-T., Schmidt, H.A., von Haeseler, A. & Minh, B.Q. (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, **32**, 268–274.

Nishiyama, T., Sakayama, H., de Vries, J., Buschmann, H., Saint-Marcoux, D., Ullrich, K.K. *et al.* (2018) The Chara genome: secondary complexity and implications for plant terrestrialization. *Cell*, **174**, 448–464.e424.

Nystedt, B., Street, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.C., Scofield, D.G. *et al.* (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, **497**, 579–584.

Oliver, M.J., Farrant, J.M., Hilhorst, H.W.M., Mundree, S., Williams, B. & Bewley, J.D. (2020) Desiccation tolerance: avoiding cellular damage during drying and rehydration. *Annual Review of Plant Biology*, **71**, 435–460.

Oliver, M.J., Tuba, Z. & Mishler, B.D. (2000) The evolution of vegetative desiccation tolerance in land plants. *Plant Ecology*, **151**, 85–100.

Ouyang, S., Zhu, W., Hamilton, J., Lin, H., Campbell, M., Childs, K. *et al.* (2007) The TIGR Rice Genome annotation resource: improvements and new features. *Nucleic Acids Research*, **35**, D883–D887.

Pardo, J., Man Wai, C., Chay, H., Madden, C.F., Hilhorst, H.W.M., Farrant, J.M. *et al.* (2020) Intertwined signatures of desiccation and drought tolerance in grasses. *Proceedings of the National Academy of Sciences*, **117**, 10079–10088.

Peres, A.L.G., Soares, J.S., Tavares, R.G., Righetto, G., Zullo, M.A., Mandava, N.B. *et al.* (2019) Brassinosteroids, the sixth class of phytohormones: a molecular view from the discovery to hormonal interactions in plant development and stress adaptation. *International Journal of Molecular Sciences*, **20**, 331.

Pham, G.M., Hamilton, J.P., Wood, J.C., Burke, J.T., Zhao, H., Vaillancourt, B. *et al.* (2020) Construction of a chromosome-scale long-read reference genome assembly for potato. *Gigascience*, **9**, 1–11.

Pyc, M., Cai, Y., Gidda, S.K., Yurchenko, O., Park, S., Kretzschmar, F.K. *et al.* (2017) Arabidopsis lipid drop-associated protein (LDAP) - interacting protein (LDIP) influences lipid droplet size and neutral lipid homeostasis in both leaves and seeds. *The Plant Journal*, **92**, 1182–1201.

Pyc, M., Gidda, S.K., Seay, D., Esnay, N., Kretzschmar, F.K., Cai, Y. *et al.* (2021) LDIP cooperates with SEIPIN and LDAP to facilitate lipid droplet biogenesis in Arabidopsis. *Plant Cell*, **33**, 3076–3103.

Rappsilber, J., Mann, M. & Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols*, **2**, 1896–1906.

Rippin, M., Becker, B. & Holzinger, A. (2017) Enhanced desiccation tolerance in mature cultures of the streptophytic green alga *Zygnema circumcarinatum* revealed by transcriptomics. *Plant and Cell Physiology*, **58**, 2067–2084.

Scholz, P., Chapman, K.D., Mullen, R.T. & Ischebeck, T. (2022) Finding new friends and revisiting old ones - how plant lipid droplets connect with other subcellular structures. *The New Phytologist*. https://doi.org/10.1111/nph.18390

Sharma, S. & Deswal, R. (2021) Dioscorea Alata tuber proteome analysis uncovers differentially regulated growth-associated pathways of tuber development. *Plant & Cell Physiology*, **62**, 191–204.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols*, **1**, 2856–2860.

Shimada, T.L. & Hara-Nishimura, I. (2015) Leaf oil bodies are subcellular factories producing antifungal oxylipins. *Current Opinion in Plant Biology*, **25**, 145–150.

Shimada, T.L., Hayashi, M. & Hara-Nishimura, I. (2018) Membrane dynamics and multiple functions of oil bodies in seeds and leaves. *Plant Physiology*, **176**, 199–207.

Shimada, T.L., Shimada, T., Takahashi, H., Fukao, Y. & Hara-Nishimura, I. (2008) A novel role for oleosins in freezing tolerance of oilseeds in *Arabidopsis thaliana*. *The Plant Journal*, **55**, 798–809.

Sierro, N., Battey, J.N., Ouadi, S., Bakaher, N., Bovet, L., Willig, A. *et al.* (2014) The tobacco genome sequence and its comparison with those of tomato and potato. *Nature Communications*, **5**, 3833.

Slotte, T., Hazzouri, K.M., Agren, J.A., Koenig, D., Maumus, F., Guo, Y.L. *et al.* (2013) The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. *Nature Genetics*, **45**, 831–835.

Stoller, E.W., Nema, D.P. & Bhan, V.M. (1972) Yellow nutsedge tuber germination and seedling development. *Weed Science*, **20**, 93–97.

Stoller, E.W. & Sweet, R.D. (1987) Biology and life cycle of purple and yellow nutsedges (*Cyperus rotundus* and *C. esculentus*). *Weed Technology*, **1**, 66–73.

The International Brachypodium Initiative. (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*, **463**, 763–768.

The International Wheat Genome Sequencing Consortium. (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, 361, eaar7191.

The Tomato Genome Consortium. (2012) The tomato genome sequence provides insights into fleshy fruit evolution. *Nature*, **485**, 635–641.

Turesson, H., Marttila, S., Gustavsson, K.-E., Hofvander, P., Olsson, M.E., Bülow, L. *et al.* (2010) Characterization of oil and starch accumulation in tubers of *Cyperus esculentus* var. sativus (Cyperaceae): a novel model system to study oil reserves in nonseed tissues. *American Journal of Botany*, **97**, 1884–1893.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T. *et al.* (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods*, **13**, 731–740.

VanBuren, R., Wai Ching, M., Zhang, Q., Song, X., Edger, P.P., Bryant, D. *et al.* (2017) Seed desiccation mechanisms co-opted for vegetative desiccation in the resurrection grass *Oropetium thomaeum*. *Plant, Cell & Environment*, **40**, 2292–2306.

Vanhercke, T., Divi, U.K., El Tahchy, A., Liu, Q., Mitchell, M., Taylor, M.C. *et al.* (2017) Step changes in leaf oil accumulation via iterative metabolic engineering. *Metabolic Engineering*, **39**, 237–246.

Vanhercke, T., Dyer, J.M., Mulle, R.T., Kilaru, A., Rahman, M.M., Petrie, J.R. *et al.* (2019) Metabolic engineering for enhanced oil in biomass. *Progress in Lipid Research*, **74**, 103–129.

Vizcaíno, J.A., Deutsch, E.W., Wang, R., Csordas, A., Reisinger, F., Rios, D. *et al.* (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nature Biotechnology*, **32**, 223–226.

Wan, T., Liu, Z.M., Li, L.F., Leitch, A.R., Leitch, I.J., Lohaus, R. *et al.* (2018) A genome for gnetophytes and early evolution of seed plants. *Nat Plants*, **4**, 82–89.

Wang, S., Li, L., Li, H., Sahu, S.K., Wang, H., Xu, Y. *et al.* (2020) Genomes of early-diverging streptophyte algae shed light on plant terrestrialization. *Nat Plants*, **6**, 95–106.

Wang, X., Wang, H., Wang, J., Sun, R., Wu, J., Liu, S. *et al.* (2011) The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, **43**, 1035–1039.

Wehmeyer, N., Hernandez, L.D., Finkelstein, R.R. & Vierling, E. (1996) Synthesis of small heat-shock proteins is part of the developmental program of late seed maturation. *Plant Physiology*, **112**, 747–757.

Wills, G.D. (1987) Description of purple and yellow nutsedge (*Cyperus rotundus* and *C. esculentus*). *Weed Technology*, **1**, 2–9.

Yang, Z. (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586–1591.

Yang, Z., Ji, H. & Liu, D. (2016) Oil biosynthesis in underground oil-rich storage vegetative tissue: comparison of *Cyperus esculentus* tuber with oil seeds and fruits. *Plant & Cell Physiology*, **57**, 2519–2540.