Title
Evolution and Functions of Oleosins and Oleosin-Coated Oil Bodies in Plants

Author
Huang, Chien-Yu

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Evolution and Functions of Oleosins and Oleosin-Coated Oil Bodies in Plants

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Plant Biology

by

Chien-Yu Huang

August 2013

Dissertation Committee:
Dr. Anthony Huang, Chairperson
Dr. Eugene A. Nothnagel
Dr. Zhenbiao Yang

The Dissertation of Chien-Yu Huang is approved:

Committee Chairperson

University of California, Riverside

Acknowledgements

I sincerely thank my advisor, Dr. Anthony Huang for his guidance and support with which I make first achievement in my research career. Dr. Huang as a generous mentor not only gives me opinions on research but also advises me on philosophy of life. I also want to thank my dissertation committee members, Dr. Eugene A. Nothnagel and Dr. Zhenbiao Yang, for their kind advice and guidance. Additionally, I thank my qualifying committee members Dr. Patricia S. Springer, and Dr. Chia-en Chang, for their time and opinions on my academic activities.

I would like to thank Dr. Chih-Hua Tsou for advice on and input into my dissertation research project and Dr. Wann-Neng Jan for assistance with electron microscopy. I also want to thank Ms. Pei-Ying Chen, Dr. Ming-Der Huang, Ms. Chun-I Chuang, Mr. Yao-Cheng Lin and Dr. Yue-Ie Caroline Hsing for providing me with and teaching me many experimental techniques. I also thank my other lab members for their friendship, encouragement and intellectual discussions. Without their inputs, I would not have completed my dissertation.

Most importantly, I would like to thank the strong support from my family, especially my husband.

Dedications

To my family

ABSTRACT OF THE DISSERTATION

Evolution and Functions of Oleosins and Oleosin-Coated Oil Bodies in Plants

by

Chien-Yu Huang

Doctor of Philosophy, Graduate Program in Plant Biology
University of California, Riverside, August 2013
Dr. Anthony Huang, Chairperson

Eukaryotes and prokaryotes contain neutral lipids in subcellular droplets as food reserves and/or for other purposes. Among all of the lipid droplets in diverse organisms, those in plant seeds are the most prominent. Lipid droplets in higher plants are present in seeds, pollens, fruits and flowers. Lipids from plants are important for human and farm-animal nutrition as well as industrial uses. Plant storage lipid droplets are stabilized with abundant structural proteins called oleosins, which contain a long conserved central hydrophobic hairpin structure penetrating into the matrix of the droplet. Plant lipid droplets with oleosins covering the surface are also termed oil bodies. The expression of oleosin genes is regulated individually in a tissue-specific manner in coordination with the biogenesis of oil bodies. This dissertation aims at delineating the evolution and functions of oleosins and oil bodies in primitive and advanced plant

species. I have found that the primitive plant moss, *Physcomitrella patens,* possesses 3 oleosin genes. I have delineated the early evolution of these oleosin genes and the storage functions of the oleosin- enclosed oil bodies. In addition to seeds, tapeta of Brassicaceae also contain numerous oleosin-coated oil bodies, which are associated with flavonoid-containing vesicles to form abundant tapetosomes. Species in Brassicaceae possess a cluster of 4-7 oleosin genes specifically expressed in tapeta. I have established quantitative adaptive benefit of the oleosin gene cluster members conferring dehydration-tolerant pollen. I have found that Cleomaceae, the family closest to Brassicacea, does not possess the oleosin gene cluster or tapetosomes. Transformation of Cleomecacea with an Arabidopsis oleosin gene generates primitive tapetosomes and dehydration-tolerant pollen.

TABLE OF CONTENTS

List of Figures

List of Tables

GENERAL INTRODUCTIONS

Subcellular neutral lipids are usually densely-packed high-energy reserves, which include triacylglycerols (TAGs) and steryl esters (SEs). Eukaryotes and prokaryotes contain neutral lipids in subcellular droplets as food reserves and/or for other purposes. These lipid droplets (LDs) are present in seeds, pollens, fruits and flowers of higher plants; the vegetative and reproductive organs of lower plants, algae, fungi, and nematodes; mammalian organs/tissues such as mammalian glands and adipose tissues; and bacteria (Hsieh and Huang, 2004; Martin and Parton, 2006; Goodman, 2008; Rajakumari et al., 2008).

Among all of the LDs in diverse organisms, those in plant seeds are the most prominent, and their studies pioneered those in non-plant organisms. Seed LDs contain a matrix of hydrophobic TAGs covered and stabilized by a layer of phospholipids (PLs) and abundant structural proteins called oleosins. Each oleosin molecule contains a conserved central hydrophobic motif that could form a long hairpin structure penetrating into the matrix of a LD. Plant LDs with oleosins covering the surface are also termed oil bodies (OBs). The oleosin gene was first cloned in 1987 (Vance and Huang, 1987). Plant lipids (oils or TAGs) are important for human and farm-animal nutrition as well as industrial uses (soap, detergents, lubricant, etc.).

Classical breeding and genetically engineering have enhanced the quality and quantity of plant storage lipids. Seeds of most plant species have OBs storing TAGs as energy reserves for future germination and post germination growth before the seedling can acquire energy via photosynthesis (Napier et al, 1996; Frandsen et al, 2001; Murphy, 2001; Hsieh and Huang, 2004).

Other non-plant organisms, such as mammals and yeast, also possess LDs in cells. Nevertheless, the components of these LDs are different from those of OBs in plant seeds. Mammals and yeast do not possess oleosins but have other structural and metabolic proteins on the surface of the LDs (Martin and Parton, 2006; Goodman, 2008; Rajakumari et al., 2008). These surface proteins differ from oleosins (Ting et al., 1997) and do not contain a long hydrophobic stretch. The latter aspect indicates that these surface proteins do not penetrate into the lipid matrix. Moreover, seed OBs possess only TAGs as the matrix lipids, whereas the LDs in mammals and yeast contain both TAGs and SEs. The evolutionary relationship between plant OBs and non-plant LDs is not clear.

The structure and biogenesis of seed OBs

The spherical subcellular OB is composed of TAGs in the matrix enclosed by a layer of PLs and oleosins. The size of seed OBs is approximately 0.5-2 μm in diameter. Plant seed OBs contain 0.6–3% proteins (Tzen *et al.*, 1993). Among OB-associated proteins, oleosins dominate (Huang, 1996). Oleosins essentially cover the whole surface area of an OB and prevents the OBs from fusing in the cytosol. Therefore, the small OBs have a large surface area per unit TAG for rapid lipase binding and catalysis during germination. The central part of the oleosin molecule is a characteristic long hydrophobic stretch, which is highly conserved among several hundred known oleosins. The hydrophobic stretch of an oleosin molecule forms a hairpin penetrating into the matrix TAGs for stable anchorage. This property makes OBs more stable in the aqueous cytosol.

Seed OBs, including their component TAGs, PLs and oleosins, are synthesized on ER. The last enzyme and the only one unique to the synthesis of TAGs, diacylglycerol acyltransferase (DAGAT), and enzymes for the synthesis of precursor diacylglycerols and PLs, are associated with rough endoplasmic reticulum (RER). TAGs newly synthesized in endoplasmic reticulum (ER) are sequestered in the hydrophobic region (i.e., the acyl region of the PL bilayer). Simultaneously, the newly assembled

oleosin-mRNA is associated with RER via its long hydrophobic hairpin stretch, and the newly synthesized oleosin resides on the ER surface. Continuation of TAG and oleosin accumulation at a region of ER forms a budding OB, which is enclosed by a single layer of PLs and oleosins. The OB is released into the cytosol and becomes a solitary entity. Seed OBs are stable because their surface is shielded by a layer of PLs and oleosins (Loer and Herman, 1993; Abell et al., 2002; Beaudoin and Napier, 2002).

In plant seeds, the thick seed coat is an obstacle for microscopy observation of oleosins and TAGs on the ER and the OBs in the internal storage cells and embryos. Results of immunodetection of oleosins by transmission electron microscopy (TEM) (Herman, 1987), biochemical isolation and enzymatic analysis of ER subfractions (Lacey et al., 1999) and fluorescence detection of TAG-synthesizing enzymes (Shockey et al., 2006) suggest that these major OB components are synthesized in specific ER subdomains. This suggestion raises the possibility that each ER subdomain produces an OB with only one kind of oleosin. Alternatively, oleosin and TAG synthesis could occur in non-specific, extended regions of ER (Hsieh and Huang, 2004).

An early release of the LD from ER generates a smaller LD, and vice versa. The size and shape of an LD are determined in part or completely by the relative amount

or rate of synthesis of TAGs and oleosins. High-oil maize kernels (having a high TAG-to-oleosin ratio) generated by breeding have larger, spherical LDs, whereas low-oil kernels have smaller LDs with irregularly shaped surface (Ting et al. 1996). Subsequent studies with knockout and knockdown oleosin mutants of Arabidopsis confirmed the importance of the TAG-to-oleosin ratio in determining the size of LDs (Siloto et al. 2006). A special mechanism may exist for the physical release of a budding LD from ER. Oleosins accumulated on the bud surface may interact among themselves to produce a physical force of constriction at the neck of the bud, thereby releasing the LD. Or, the physical release may require specific proteins (e.g., dynamins) or actions by ER or cytosol.

All the known >500 oleosins of diverse plant species contain a conserved central hydrophobic motif composed of ~72 non-polar residues that could form a long hairpin structure. In the middle of the hairpin, the turn is composed of 12 most-conserved residues (PX_5SPX_3P) and termed a proline knot. The 3 proline and 1 serine residues in the proline knot are completely conserved without a single exception among all known oleosins. No other lipid-related protein in any organism has such a long, or even half as long, hydrophobic stretch of 72 residues. The

amphipathic N and C termini are less conserved and highly variable in length (Hsieh and Huang, 2004).

Knockout and knockdown oleosin mutants in Arabidopsis further demonstrated the biological function of oleosins in seeds, and the mutants had delayed seed germination and decreases in freezing tolerance (Siloto et al. 2006; Shimada et al., 2008).

OBs and oleosins in flowers

Oleosin-coated LDs are also present in pollen (Kim et al., 2002) and the tapeta of *Brassica* and Arabidopsis (Hsieh and Huang, 2005, 2007). In pollen, OBs store oils, which probably provide acyl moieties for membrane growth during germination and tube elongation.

Pollen development in the anther is a highly coordinated and specialized process that involves drastic differentiation. The tapetum is a single layer of nutritive cells enclosing the anther locule, in which microspores develop and mature to become pollen. Among all anther sporophytic cells, the tapetum cells are most active in metabolism and control the development of pollen grains (Polowick and Sawhney, 1993). At the early stage of anther development, the tapetum cells contain abundant RER and secretory vesicles for active secretion. At the mid stage of anther

development, the cells store ingredients to be deposited onto the maturing pollen to form a pollen coat. In Brassicaceae, the tapetum cells at this mid stage contain two abundant storage organelles, the elaioplasts and the tapetosomes (Platt et al., 1998). Elaioplasts, of 3 to 4 μm in diameter, are specialized plastids for the storage of lipids. Elaioplasts are filled with small spherical LDs of steryl esters enclosed by structural proteins. Elaioplasts of a similar morphology can be found in nontapetum cells, such as those in fruits and petals.

The tapetosomes are unique to the tapetum cells. Oleosins present in tapetum cells of anthers in Arabidopsis and Brassica were discovered from unintended gene cloning results (de Oliveira et al., 1993; Roberts et al., 1994). In tapetum cells, oleosin-coated LDs are associated with flavonoid-containing vesicles to form tapetosomes of 2-3 μm in diameter. Tapetosomes temporarily store lipids and flavonoids, which are to be deposited onto the maturing pollen as pollen coat for water-proofing and UV protection, respectively. Among the 17 oleosin genes in Arabidopsis, 9, termed T-oleosin genes, have expression only in the tapetum (Kim et al., 2002). All the 9 T-oleosin genes are on chromosome 5, and 8 are in tandem. This T-oleosin gene cluster is also present in other genera of Brassicaceae, including *Boechera drummondii, Brassica olerecea, Olimarabidopsis pumila*, *Capsella rubella*,

and *Sisymbrium irio* (Fiebig et al., 2004; Schein et al., 2004). The tapetosome is not ubiquitous in plant tapetum cells. For example, maize and rice have no tapetosomes in the tapeta, and their pollen coat has no oleosin. Nevertheless, pollen of all plant species has surface lipids and flavonoids derived from tapeta. Also, tapetum cells in diverse plant species contain small LDs. The evolution of tapetosomes is intriguing.

LDs without oleosins in vegetative tissues of plants

Oleosin genes are highly expressed in seeds and some floral tissues like anthers and pollens during lipid accumulation but are not so in most vegetative and fruit tissues of plants. The fleshy mesocarp of some fruits including the oil-rich oil palm (*Elaeis guineensis*), olive (*Olea europaea*), avocados (*Persea americana*) contains much larger subcellular lipid particles of TAGs. These lipid particles do not have surface oleosins and are more diverse in size, of about 10-50 μm in diameter (Murphy 2001; Hsieh and Huang, 2004) and could occupy the bulk of the cell volume. Mesocarp lipids are for attraction to animals for seed dispersion and thus are not required to be in small entities as the LDs in seeds, which are for rapid mobilization. Most likely, TAGs are synthesized in ER as those in seeds, but without co-synthesis of oleosins. As a consequence, the budding LD enclosed only by PLs becomes larger

(and/or fuses with adjacent budding LDs) before it is released from ER to the cytosol (Murphy 2001; Hsieh and Huang, 2004).

Some other vegetative tissues in plants also accumulate lipids, such as roots/tubers (e.g., cotton, nutsedge), and stems (e.g., Mongolian oil wood) (Carlsson et al, 2011). The regulatory mechanism of LDs accumulation in these tissues is unknown.

Evolution of LDs and oleosins

Prokaryotes, in general, do not store TAGs as food reserves. An exception occurs in *Actinomyces*, which produce TAGs under certain nutritional and other environmental conditions (Hänisch et al. 2006). TAGs likely evolved as efficient food reserves in a primitive eukaryote by acquisition of one enzyme, DAGAT, which evolved from one of the existing acyltransferases (Turchetto-Zolet et al., 2011). The enzyme diverted diacylglycerols from the ubiquitous PL metabolic pathway to TAGs. Initially, the hydrophobic TAGs were produced by DAGAT and were present between the two PL layers of ER membrane. Seeds of some plant species on occasions still have some TAGs present in the hydrophobic region between the PL bilayers in ER. Excess TAGs accumulated in ER membrane would interfere in the normal functioning

of ER. This problem was overcome by removal of the TAGs from ER via budding to solitary LDs. The LDs, each containing a TAG matrix enclosed by a layer of PLs originated from ER, would be unstable in the aqueous cytosol. In yeasts, the droplets are more stable because they have a coat of amphipathic proteins, especially the TAG synthesizing and hydrolyzing enzymes, and the semi-stability would allow the droplets to undergo dynamic metabolic fluxes (Szymanski et al, 2007). In mammals, the droplets were modified to different forms with proteins and membranes, such that they are also semi-stable and amenable to rapid metabolic fluxes (Goodman JM, 2008, 2009). In plants, the droplets were stabilized by the evolutionary appearance of oleosins, whose long hydrophobic hairpin stabilizes the droplets such that they are amenable to prolonged storage in desiccated seeds.

In each plant species, the expression of oleosin genes is regulated individually in a tissue-specific manner. For example, the 17 oleosin genes in *Arabidopsis thaliana* are specifically expressed in seed, pollen and tapetum cells (Kim et al, 2002).

Objectives of the dissertation research

This dissertation aims at delineating the evolution and functions of OBs and oleosins in primitive and advanced plant species. To explore the evolutionary trends

of oleosin-coated OBs in plants, I searched genes encoding oleosins in all organisms whose genomes have been sequenced. I found 3 oleosin genes in the moss *Physcomitrella patens* and none in algae or other primitive organisms. Comparing to advanced plant species, *P. patens* possesses fewer oleosin genes, which is an advantage for gene functional study. In my preliminary observation, *Physcomitrella patens* has abundant OBs in the gametophyte tissue, which is composed of a single layer of cells. Theses characteristics make *P. patens* an ideal system for functional and biogenesis study of OBs and the oleosins. In my dissertation research, I took advantage of this simple system to study oleosin gene evolution, oleosin targeting, and LD function. About 1/3 of this dissertation deals with studies with *P. patens*.

The tapetum of Brassicacea contains abundant oleosin-coated LDs. Theses LDs are associated with flavonoid-containing vesicles to form tapetosomes. Evolution of the T-oleosin gene cluster and tapetosomes is unclear. I investigated the evolution of the T-oleosin gene cluster and tapetosomes within and outside of Brassicaceae and found the closest family, Cleomaceae (Barker et al, 2009; Schranz and Mitchell-Olds, 2006), did not possess the T-oleosin gene cluster. About 2/3 of the dissertation deals with this investigation.

REFERENCES

Abell BM, Hahn M, Holbrook LA, Moloney MM (2004) membrane topology and sequence requirements for oil body targeting of oleosin. Plant J **37:** 461-470

Abell BM, High S, Moloney MM (2002) Membrane protein topology of oleosin is constrained by its long hydrophobic domain. J Biol Chem **277:** 8602–8610

Barker MS, Vogel H, Schranz ME (2009) Paleopolyploidy in Brassicales: analyses of the *Cleome* transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales. Genome Biol Evol **1:** 391-399

Beaudoin F, Napier JA (2002) Targeting and membrane-insertion of a sunflower oleosin in vitro and in *Saccharomyces cervisiae*: The central hydrophobic domain contains more than one signal sequence, and directs oleosin insertion into the endoplasmic reticulum membrane using a signal anchor sequence mechanism. Planta **215:** 293–303

Carlsson AS, Yilmaz JL, Green AG, Stymne S, and Hofvander P (2011). Replacing fossil oil with fresh oil – with what and for what? Eur J Lipid Sci Technol **113:** 812-831

de Oliveira DE, Franco LO, Simoens C, Seurinck J, Coppieters J, Botterman J, van Montagu M (1993) Inflorescence-specific genes from *Arabidopsis thaliana* encoding glycine-rich proteins. Plant J **3:**495-507

Fiebig A, Kimport R, Preuss D (2004) Comparisons of pollen coat genes across Brassicaceae species reveal rapid evolution by repeat expansion and diversification. Proc Natl Acad Sci USA **101:** 3286–3291

Goodman JM (2008) The gregarious lipid droplets. J Biol Chem **42:** 28005-28009.

Goodman JM (2009) Demonstrated and inferred metabolism associated with cytosolic lipid droplets. J Lipid Res **50:** 2148-2156

Hänisch, J., Wältermann, M., Robenek, H. and Steinbüchel, A (2006) Eukaryotic lipid body proteins in oleogenous Actinomycetes and their targeting to intracellular triacylglycerol Inclusions: impact on models of lipid body biogenesis. Appl Environ Microbiol **72**: 6743-6750

Herman EM (1987) Immunogold-localization and synthesis of an oil-body membrane protein in developing soybean seeds. Planta **172:** 336-345

Hsieh K, Huang AHC (2004) Endoplasmic reticulum, oleosins, and oils in seeds and tapetum. Plant Physiol **136:** 3427-3434

Hsieh K. Huang AHC (2005) Lipid-rich tapetosomes in *Brassica* tapetum are composed of oleosin-coated oil droplets and vesicles, both assembled in and then detached from the endoplasmic reticulum. Plant J **43:** 889-899

Hsieh K, Huang AHC (2007) Tapetosomes in *Brassica* tapetum accumulate endoplasmic reticulum-derived flavonoids and alkanes for delivery to the pollen surface. Plant Cell **19:** 582-596

Huang AHC (1996) Oleosin and oil bodies in seeds and other organs. Plant Physiol **110:**1055–1061

Jauh GY, Phillips TE, Rogers JC (1999) Tonoplast intrinsic protein isoforms as markers for vacuolar functions. Plant Cell **11:** 1867–1882

Kim DH, Eu YJ, Yoo CM, Kim YW, Pih KT, Jin JB, Kim SJ, Stenmark H, and Hwang I (2001) Trafficking of phosphatidylinositol 3-phosphate from the trans-Golgi network to the lumen of the central vacuole in plant cells. Plant Cell **13:** 287–301

Kim HU, Hsieh K, Ratnayake C, Huang AHC (2002) Expression of *Arabidopsis* oleosin genes and characterization of their encoded oleosins. J Biol Chem **277:** 22677-22684

Lacey DJ, Beaudoin F, Dempsey CE, Shewry PR, Napier JA (1999) The accumulation of triacylglycerols within the endoplasmic reticulum of developing seeds of *Helianthus annuus*. Plant J **17:** 397-405

Lee MS, Mullen RT, Trelease RN (1997) Oilseed isocitrate lyases lacking their essential type 1 peroxisomal targeting signal are piggybacked to glyoxysomes. Plant Cell **9:**185-197

Lersten NR, Czlapinski AR, Curtis JD, Freckmann R, Horner HT (2006) Oil bodies in leaf mesophyll cells of angiosperms: overview and a selected survey. Am J Bot **93:** 1731-1739.

Loer DS, Herman EM (1993) Cotranslational integration of soybean (*Glycine max*) oil body membrane protein oleosin into microsomal membranes. Plant Physiol **101:** 993–998

Lorenz H, Hailey DW, Lippincott-Schwartz J (2006) Fluorescence protease protection of GFP chimeras to reveal protein topology and subcellular localization. Nature Methods **3:** 205-210

Martin S, Parton RG (2006) Lipid droplets: a unified view of a dynamic organelle. Nature Review **7:** 373-378

Murphy DJ (2001) The biogenesis and functions of lipid bodies in animals, plants and microorganisms. Plant J **13:** 1-16

Napier JA, Stobart AK, Shewry PR (1996) The structure and biogenesis of plant oil bodies: the role of the ER membrane and the oleosin class of proteins. Plant Mol Biol **31:** 945-956

Platt KA, Huang AHC, Thomson WW (1998) Ultrastructural study of lipid accumulation in tapetal cells of *Brassica napus* L. cv. Westar during microsporogenesis. Int J Plant Sci **159:** 724–737

Polowick PL, Sawhney VK (1993). Differentiation of the tapetum during microsporogenesis in tomato (*Lycopersicon esculentum* Mill.), with special reference to the tapetal cell wall. Ann of Bot **72**:595–605

Rajakumari S, Grillitsch K, Daum G (2008) Synthesis and turnover of non-polar lipids in yeast. Prog in Lip Res **47:** 157-171.

Roberts LS, Gerster J, Allard S, Cass L, Simmonds J (1994) Molecular characterization of two *Brassica napus* genes related to oleosins which are highly expressed in the tapetum. Plant J **6:** 927–933

Schein M, Yang Z, Mitchell-Olds T, Schmid KJ (2004) Rapid evolution of a pollen-specific oleosin-like gene family from *Arabidopsis* thaliana and closely related species. Mol Biol Evol **21:**659–669

Schranz ME, Mitchell-Olds T (2006) Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. Plant Cell **18:** 1152-1165

Shimada TL., 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

Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT and Dyer JM (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.

Szymanski KM. , Binns, Bartz R , Grishin NV, Li WP, Agarwal AK, Garg A , Anderson RG, and Goodman JM (2007) The lipodystrophy protein seipin is found at endoplasmic reticulum lipid droplet junctions and is important for droplet morphology. Proc Natl Acad Sci USA **104 :** 20890 – 20895 .

Ting JTL, Balsamo RA, Ratnayake C, Huang AHC (1997) Oleosin of plant seed oil bodies is correctly targeted to the lipid bodies in transformed yeast. J Biol Chem **272:** 3699-3706

Turchetto-Zolet AC, **Maraschin FS**, **de Morais GL**, **Cagliari A**, **Andrade CMB**, **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 JTC, Cao YZ, Laurent P, Ratnayake C, Huang AHC (1993) Lipids, proteins, and structure of seed oil bodies from diverse species. Plant Physiol **101:**267–276

Vance VB, Huang AHC (1987) The major protein from lipid bodies of maize. Characterization and structure based on cDNA cloning. J Biol Chem **262:** 11275-11279

CHAPTER ONE

Oil bodies and oleosins in *Physcomitrella* possess characteristics representative of early trends in evolution

ABSTRACT

Searches of sequenced genomes of diverse organisms revealed that the moss *Physcomitrella patens* is the most primitive organism possessing oleosin genes. Microscopy examination of *Physcomitrella* revealed that oil bodies (OBs) were abundant in the photosynthetic vegetative gametophyte and the reproductive spore. Chromatography illustrated the neutral lipids in OBs isolated from the gametophyte to be largely steryl esters and triacylglycerols, and SDS-PAGE showed the major proteins to be oleosins. RT-PCR revealed the expression of all 3 oleosin genes to be tissue specific. This tissue specificity was greatly altered via alternative splicing, a control mechanism of oleosin gene expression unknown in higher plants. During the production of sex organs at the tips of gametophyte branches, the number of OBs in the top gametophyte tissue decreased concomitant with increases in the number of peroxisomes and level of transcripts encoding the glyoxylate cycle enzymes; thus, the OBs are food reserves for gluconeogenesis. In spores during germination, peroxisomes adjacent to OBs, along with transcripts encoding the glyoxylate cycle enzymes, appeared; thus, the spore OBs are food reserves for gluconeogenesis and equivalent to seed OBs. The one-cell-layer gametophyte could be observed easily with confocal microscopy for the subcellular OBs and other structures. Transient

expression of various gene constructs transformed into gametophyte cells revealed that all OBs were linked to the endoplasmic reticulum (ER), that oleosins were synthesized in extended regions of the ER and that two different oleosins were co-located in all OBs.

INTRODUCTION

Eukaryotes and prokaryotes contain neutral lipids in subcellular droplets as food reserves and/or for other purposes (Hsieh and Huang, 2004; Martin and Parton, 2006; Goodman, 2008; Rajakumari et al., 2008). These LDs are present in seeds, pollens, fruits and flowers of higher plants; the vegetative and reproductive organs of lower plants, algae, fungi, and nematodes; mammalian organs/tissues such as mammalian glands and adipose tissues; and bacteria. Among all these LDs, oil bodies (OBs) in seeds are the most prominent and have been extensively studied.

Seeds of diverse plant species store oils (triacylglycerols [TAGs]) as food reserves for germination and postgermination growth (Napier et al, 1996; Frandsen et al, 2001; Murphy, 2001; Hsieh and Huang, 2004). The TAGs are present in small subcellular, spherical OBs of approximately 0.5-2 μm in diameter. Each OB has a matrix of TAGs surrounded by a layer of phospholipids (PLs) and the structural protein oleosins. The massive oleosins completely cover the surface of the OBs and prevent them from coalescence; so, a large surface area per unit TAG is available for lipase binding and catalysis during germination. Each oleosin molecule has a characteristic long central hydrophobic stretch, which forms a hairpin penetrating into the matrix TAGs for stable anchorage.

Other than being present in the seeds of plants, oleosin-coated OBs are also present in pollen (probably for storage of acyl moieties for tube elongation [Kim et al., 2002]) and the tapeta of *Brassica* and Arabidopsis (Hsieh and Huang, 2005, 2007). Inside each tapetum cell, many oleosin-coated OBs associate with numerous flavonoid-containing vesicles to form large subcellular particles termed tapetosomes, each 2-3 μm in diameter. Tapetosomes temporarily store lipids and flavonoids, which are deposited onto the maturing pollen as a pollen coat for water-proofing and UV protection, respectively. In fruits of some species, such as olive, avocado and oil palm, the fleshy mesocarp possesses much larger (10-50 μm diameter) subcellular lipid particles of TAGs, which are devoid of surface oleosins and apparently are for attracting animals for seed dispersal (Murphy 2001; Hsieh and Huang, 2004). OBs are also present, although generally in low abundance, in leaves of diverse plant species, and their structures and functions are unknown (Lersten et al., 2006).

Oleosins of all plant species contain a conserved central hydrophobic hairpin of ~72 residues flanked by less conserved amphipathic N and C termini of highly variable length (Hsieh and Huang, 2004). Within the hairpin, the turn consists of 12 most-conserved residues (PX_5SPX_3P), of which the 3 proline and 1 serine residues (termed the proline knot) are completely conserved without a single exception

among hundreds of examined oleosins of various species. Paralogs of oleosin genes are present within each species and individuals and are expressed in a tissue-specific manner. For example, Arabidopsis has 17 oleosin genes, which are selectively expressed in seed, pollen and the tapetum (Kim et al, 2002).

Oleosins and TAGs are generally believed to be synthesized on the ER inside a seed cell. Whereas the nascent oleosins are attached to the ER surface via the long hydrophobic hairpin stretch, TAGs are sequestered between the 2 PL layers of the ER membrane. These oleosins and TAGs migrate to and are eventually concentrated in confined ER regions, which are detached to form mature OBs (Napier et al, 1996; Murphy, 2001; Abell et al., 2004; Hsieh and Huang, 2004). What is uncertain is the location of the ER on which oleosins and TAGs are synthesized. These major OB components could be synthesized in specific ER subdomains, as interpreted from results of immunodetection of oleosins with transmission electron microscopy (TEM) (Herman, 1987), biochemical isolation and enzymatic analysis of ER subfractions (Lacey et al., 1999) and fluorescence detection of TAG-synthesizing enzymes with confocal laser scanning microscopy (CLSM) (Shockey et al., 2006). Existence of OB-synthesizing ER subdomains would raise the possibility that each subdomain produces an OB with only 1 of the several oleosin isoforms in the cell. Alternatively,

oleosin and TAG synthesis could occur in non-specific, extended regions of the ER (Hsieh and Huang, 2004).

LDs in cells of non-plant organisms, such as mammals and yeast, also possess surface proteins with structural and/or metabolic functions (Martin and Parton, 2006; Goodman, 2008; Rajakumari et al., 2008). These surface proteins are not related to oleosins (Ting et al., 1997) and do not possess a long hydrophobic stretch. Their polypeptides wrap around rather than penetrate into the LDs. Whereas seed OBs possess only TAGs as the matrix lipids, the LDs in mammals and yeast contain both TAGs and steryl esters (SEs). The evolutionary relationship of plant OBs and non-plant LDs is unknown.

To explore the evolutionary trends of OBs and oleosins in primitive plants, algae and fungi, I searched for genes encoding oleosins in these organisms having completely sequenced genomes. Only the moss *Physcomitrella* (3 paralogs) and the primitive fern (fern ally) *Selaginella* (8 paralogs) possess genes encoding oleosins. Primitive plants, including bryophytes (mosses) and ferns, contain neutral lipids and oil bodies (Swanson et al., 1976; Pihakaski et al., 1987; Dembitsky1993). I chose the more primitive *Physcomitrella* for more intensive study. The 3 oleosin genes are expressed in a tissue-specific manner, which is further regulated via alternative

splicing, a process unknown with oleosin genes in higher plants. OBs in the dehydrated spore resemble those in seeds in being food reserves for germination. OBs in the non-dehydrated, photosynthetic gametophyte, although harboring surface oleosins, possess both TAGs and SEs esters and thus resemble more the LDs in mammals and yeasts. The gametophyte OBs are for gluconeogenesis when food reserves are needed, such as during sex organ production. The 1-cell-layer gametophyte can be used for transient expression of oleosin genes for convenient microscopy exploration. The approach shows that all OBs are linked to extended regions of the ER, on which different oleosins are synthesized concurrently. Here I report my findings.

RESULTS

Abundant OBs are present in the photosynthetic gametophyte and dehydrated spore

I used light microscopy and TEM to observe OBs in *Physcomitrella* cells throughout the life cycle (Fig. 1-1). Cells of the protonema, which were young tissues grown from spore after germination for 10 d, contained no or few OBs. The cells had conspicuous plastids with large starch grains. Cells of the predominate mature gametophyte, which was the conspicuous photosynthetic branches, contained numerous OBs. These OBs could be observed after Sudan Black staining and were as numerous as the larger chloroplasts (~150 per 100 μm x 100 μm). The spherical OBs had heterogeneous sizes, of <1 μm to several micrometers in diameter. Cells of the antheridium, the male reproductive structure, had 1-2 OBs per cell; they were strongly electron dense after osmium fixation, presumably possessing highly unsaturated lipids. Early cells of the archegonium, the female sex structure, contained no or few OBs. Some internal cells of the archegonium differentiated into spore mother cells and began to accumulate OBs, even before meiosis. After meiosis, the spore continued to accumulate OBs, and the mature, dehydrated spore were

packed with OBs of various sizes, from 0.2 to 3 μm in diameter. Upon germination, the spore became less hydrated, and peroxisomes (glyoxysomes) appeared.

OBs isolated from the photosynthetic gametophyte contain oleosins, SEs and TAGs

I could not collect enough spores, which were of minute sizes (20 μm in diameter), and then crack their hard shell gently for isolation of the internal OBs. Thus, I isolated the OBs from the photosynthetic gametophyte after gentle homogenization of the cells and floatation centrifugation. TEM of the floated OB fraction revealed OBs of heterogeneous sizes, ranging from 0.5 to 5μm (Fig. 1-2A). SDS-PAGE showed that the OB fraction was highly enriched with protein(s) of ~17 kD (Fig. 1-2B). The protein was extracted from the gel and subjected to trypsin digestion. The resulting fragments were analyzed with Matrix-assisted Laser Desorption/Ionization-Time of Flight (MALDI-TOF) mass spectrometry and identified with use of the predicted protein database derived from the *Physcomitrella* genome (http://moss.nibb.ac.jp/). They corresponded to the N-terminal fragments of OLE1 and OLE2 (whole proteins predicted to be 13-21 kD; to be described). Thus, the ~17-kD proteins highly enriched in the OB fraction were OLE1 and OLE2.

The neutral lipids of the OB fraction were analyzed with thin layer chromatography (TLC) and visualized after sulfuric acid spraying (Fig. 1-2C). The major lipids were SEs and TAGs, and diacylglycerols (DAGs) were in lesser amounts. The SEs were identified with HPLC-mass spectrometry (HPLC-MS) to be lanosterol esters (42.7% of all SEs), stigmasterol esters (30.32%), β-sitosterol esters (19.97%) and campesterol esters (8.89%). The TAGs contained the common acyl moieties of oleic (O), linoleic (L), linolenic (Ln) and palmitic (P) acids and were O/L/Ln (28.68% of all TAGs), L/L/O and O/O/L (not separated, 20.73%), P/L/L (20.73%), L/L/L (16.90%), L/L/Ln (7.76%) and O/O/O (5.21%).

Physcomitrella is the most primitive organism possessing oleosins

I used (A) the conserved proline knot sequence (PX_5SPX_3P) and (B) the complete hairpin sequence of oleosins from different plant species as queries to search for oleosins in the JGI Genomic database. Attention was paid to organisms whose genomes have been completely sequenced. The searched primitive species included lycophytes (*Selaginella moellendorffii),* bryophytes (*Physcomitrella patens*), algae and related organisms (*Aureococcus anophagefferens, Chlamydomonas reinhardtii, Chlorella sp.,* Emiliania huxleyi, *Micromonas pusilla, Ostreococcus lucimarinus,*

Phaeodactylum tricornutum and *Volvox carteri*), fungi and related organisms (*Aspergillus niger, Batrachochytrium dendrobatidis, Cochliobolus heterostrophus, Laccaria bicolor, Mycosphaerella fijiensis, Nectria haematococca, Phanerochaete chrysosporium, Phycomyces blakesleeanus*, *Saccharomyces cerevisiae*, *Pichia stipitis, Postia placenta and Sporobolomyces roseus) and* oomycetes (Phytophthora ramorum). Among these species, only the primitive lycophyte *Selaginella moellendorffii* (8 genes) and the moss *Physcomitrella patens* (3 genes) had oleosin genes. Clearly, no oleosin gene was present in algae, fungi and oomycetes. Thus, the moss *Physcomitrella* was the most primitive organism found to possess oleosin genes.

An unrooted phylogenetic tree of oleosins in species with completely sequenced genomes was constructed on the basis of their conserved hairpin sequences plus the moderately conserved sequences immediately flanking the hairpin (Fig. 1-3). The tree includes 17 oleosins from Arabidopsis, 6 from rice, 8 from *Populus*, 8 from *Selaginella* and 3 from *Physcomitrella*. The oleosin genes in the higher plants Arabidopsis, rice and *Populus* have had more variations. *Physcomitrella* has the least variations and the fewest oleosin genes. A pileup of the amino acid sequences of these oleosins revealing the conserved and non-conserved residues is in Fig. 1-4.

The 3 oleosin genes are expressed in a tissue-specific manner, which can be altered via alternative splicing

RT-PCR with use of gene-specific primers was performed to examine the levels of transcripts encoding oleosins and related proteins in various tissues throughout the life cycle of *Physcomitrella* (Fig. 1-5). For each of the 3 oleosin transcripts, the RT-PCR primers detected the sequence encoding the oleosin hairpin region and thus would detect both oleosin isoforms generated via alternative splicing (see next paragraph). *OLE*1 and *OLE*2 transcripts were present in all tissues, and their levels in zygotes and spore increased during spore maturation. The *OLE*3 transcript was present only in the spore samples. Transcripts of genes encoding malate synthase and isocitrate lyase (1 gene each per haploid genome) were also present in all tissues but at higher levels in mature and germinated spore. *Physcomitrella* has 4 genes encoding putative DAG acyltransferase with use of acyl-CoA as the acyl donor (*DAGAT*1a, b, c and d), 2 genes encoding putative DAG acyltransferase with use of PLs as the acyl donor (*DAGAT*2a and b) and 1 gene encoding a putative steryl acyltransferase (*SEAT*). These genes were annotated on the basis of their sequence similarities with the annotated genes in yeast (Rajakumari et al., 2008). The transcripts of these acyltransferases were present at different levels in diverse tissues.

Only the expression of *DAGAT*1d and *SEAT* had a clear pattern, similar to that of *OLE*1 and *OLE*2, of increasing in level from zygotes to maturing spore. Thus, the data strongly suggest that *DAGAT*1d and *SEAT* encode the acyltransferases for synthesis of the storage TAGs and SEs, respectively, in maturing spore.

*OLE*1 and *OLE*2 could each produce 2 different transcripts via alternative splicing, which would result in 2 oleosin isoforms of different sizes (Fig. 1-6). The alternative splicing sites occurred downstream of the sequence encoding the hydrophobic hairpin stretch, and thus the resulting 2 oleosin isoforms still possessed the hairpin stretch and the structural characteristics of an oleosin. Alternative splicing of *OLE*1 gave OLE1a of 15.3 kD (predicted) and OLE1b of 16.1 kD, and that of *OLE*2 gave OLE2a of 12.5 kD and OLE2b of 21.7 kD. Importantly, transcripts encoding OLE1a and OLE2a were restricted to spore, whereas those encoding OLE1b and OLE2b were ubiquitous. Thus, the tissue-specific presence of the machineries of alternative splicing (Barbazuk et al., 2008) allows for substantial changes in expression and thus, presumably, differentiation of the oleosin genes and oleosin functions. Preliminary testing of *OLE*3 for alternative splicing generated negative results and the predicted OLE3 had 13.8 kD.

OBs in spore are equivalent metabolically to those in seeds

During the life cycle of *Physcomitrella*, haploid spores were produced via meiosis. Each spore became dehydrated and packed with OBs (Fig. 1-1C) and could stay dormant or germinate in favorable conditions. Many of these physiological aspects are similar to those of seeds. In spores that had just germinated, peroxisomes appeared adjacent to the OBs (Fig. 1-1C). These peroxisomes were most likely the glyoxysomes, in reference to those in germinated seeds (Pracharoenwattana and Smith, 2008). Attempts to use antibodies against cotton malate synthase to detect the enzyme in *Physcomitrella* spore peroxisomes (glyoxysomes) via immunofluorescence microscopy were unsuccessful, presumably because the antibody-antigen reaction was not strong enough. Nevertheless, the levels of transcripts encoding malate synthase and isocitrate lyase, 2 marker enzymes of the glyoxysomes, in spore increased substantially during germination (Fig. 1-5). Thus, the spore OBs are present as food reserves for future gluconeogenesis via the glyoxysomes and other metabolic machineries.

OBs in the photosynthetic gametophyte serve as food reserves and are mobilized via the glyoxysomes, as exemplified in the production of sex organs

The vegetative gametophyte was induced to produce sex organs by switching the culture temperature from 25^0C to 15^0C. Within a 7-d period, brown antheridia (to produce sperms) and greenish archegonia (not easily visible) were produced (Fig. 1-7). OBs and peroxisomes in the uppermost leafy tissue were observed during this period of induction with BODIPY dye (for OBs) and antibodies against cotton catalase (for peroxisomes). During the 7-d period, the OB number rapidly deceased, concomitant with increased number of peroxisomes (glyoxysomes) (Fig. 1-7).

The above changes in OBs and peroxisomes occurred only in the uppermost leafy tissues but not in the middle leafy tissues of a standup branch (Fig. 1-8A). During the 7-d period, the OB number in the uppermost leafy tissue decreased by 80%, concomitant with a marked increase in peroxisome number. No such changes of the 2 organelles occurred in the mid leafy tissues of a standup branch. In both the uppermost and mid leafy tissues, the number of chloroplasts remained unchanged. Thus, mobilization of lipid reserves to initiate production of sex organs in a standup branch required only OBs in the uppermost leafy tissue. Presumably, a longer sustained sexual reproduction process would require mobilization of the lipid reserves in the lower leafy tissues of a branch. The plant was cultivated in a

sugar-rich medium, and thus there was a lesser need for mobilizing all lipid reserves in a branch for sexual reproduction.

During the 7-d period, transcripts encoding malate synthase and isocitrate lyase, markers of glyoxysomes, in the uppermost leafy tissue rapidly increased in level, concomitant with a decrease in levels of transcripts encoding OLE1 and OLE2 (Fig. 1-8B). These changes in transcript levels did not occur in the mid leafy tissue.

All of the OBs in a gametophyte cell are linked to extended regions of the ER, on which different oleosins are synthesized concurrently

The leafy tissue of the gametophyte consists of only 1 cell layer. I tried to establish the leafy gametophyte as a transient expression system for cells that contain abundant OBs, that can be transformed easily with bombardment and that can be observed clearly with CLSM. Such a plant system has not been previously established (Miao and Jiang, 2007).

When cells were transformed with *GFP* or *RFP* driven by a 35S promoter, GFP or RFP was observed in the cytosol and was not associated with specific subcellular structures (Fig. 1-9A). When *GFP* was attached to the 3'-terminus of a complete *OLE*1a open reading frame, OLE1a-GFP initially appeared in a network and the

associated droplets (Fig. 1-9B). Gradually, from 7, 10 to 12 h, proportionally less OLE1a-GFP was present in the network and more in the associated droplets. The network and the associated droplets were the ER and OBs, respectively, because after co-transformation with *OLE1a-GFP* and *BiP-RFP* (chaperone binding protein [BiP], an ER marker, from Arabidopsis [Kim et al., 2001]), OLE1a-GFP overlapped with BiP-RFP in the network and was highly enriched in the droplets (Fig. 1-9C). When *OLE*1a-*GFP* was used, OLE1a-GFP and the lipid dye Nile Red superimposed in all the droplets (Fig. 1-9D). When *OLE*1a-*RFP* and *OLE*2a-*GFP* were co-transformed, their encoded proteins appeared in all the droplets (Fig. 1-9E). The overall findings indicate that different oleosins are synthesized in extended regions of the ER and move to the associated OBs.

DISCUSSION

The OBs in both *Physcomitrella* and seeds apparently are similar in having a matrix of oils enclosed by a layer of oleosins and presumably also PLs. However, *Physcomitrella* OBs have the following early evolutionary trends. (A) The sizes of OBs in both the photosynthetic gametophyte and mature spore vary substantially, from 0.2 to 5 μm in diameter. Seed OBs have a narrow size range within a species (Tzen et al., 1993). *Physcomitrella* might not have evolved a mechanism to control the sizes of OBs. OBs in the photosynthetic gametophyte may resemble the LDs in yeast and mammal cells in that the droplet sizes are related to the metabolic conditions of the cells. (B) Among all examined plant species, *Physcomitrella* has the fewest number of oleosin genes with minimal diversification. Alternative splicing that alters the tissue-specific expression of the oleosin genes in *Physcometrella* may represent a mechanism for gene diversification. (C) Whereas seed OBs contain mostly TAGs, *Physcomitrella* OBs possess largely SEs and TAGs and some DAGs. This lipid composition of *Physcomitrella* OBs is similar to that of the intracellular and/or extracellular LDs in yeast and mammals (with SEs and TAGs) and the extracellular LDs in insects (largely DAGs; Ryan, 1994). (D) All the OBs within a *Physcomitrella* gametophyte cell apparently are physically linked to the ER, a trend suspected to

occur in yeast, mammals and other non-plant organisms (Martin and Parton, 2006; Goodman, 2008; Rajakumari et al., 2008). Actually, this characteristic may be normal for LDs in all non-dehydrated vegetative cells, in which metabolic dynamics are expected. No such information is available for the OBs in not-yet-dehydrated, maturing seed cells, and the OBs in mature seed cells are apparently not linked to the ER, which disappears during dehydration. (E) Unlike mature seeds, the *Physcomitrella* gametophyte possesses oleosin-coated OBs in highly hydrated vegetative cells; this situation is similar to the LDs in yeast and mammalian cells. Oleosins on seed OBs may protect the OBs from dehydration (Napier et al, 1996; Murphy, 2001; Hsieh and Huang, 2004). This idea could be applied to the oleosins on OBs in the vegetative non-dehydrated *Physcomitrella* gametophyte, because many moss tissues can undergo extreme and prolonged dehydration and still resuscitate upon water uptake. The evolutionary acquisition of oleosins, which are absent in algae, would represent one of the desiccation and stress tolerance features adapted by *Physcomitrella* (Rensing et al., 2008). Whether during early evolution, OBs coated with oleosins appeared first in vegetative cells (photosynthetic gametophyte and then leaves) or in desiccated sexual organs (spore and then seed) is unknown.

LDs are present in green leaves of diverse species, although they are less obvious and abundant (Lernsten et al., 2006). Whether these LDs possess oleosins and other characteristics and play a similar physiological role of food storage as do *Physcomitrella* gametophyte OBs remains to be elucidated. No or few (authentic or simply background) oleosin transcripts are present in high-quality Massively Parallel Sinature Sequencing (MPSS) or sequencing by synthesis (SBS) leaf transcriptomes of Arabidopsis (Meyers et al., 2004) and rice (Nobuta et al., 2007). Regardless, oleosin encoded by an Arabidopsis gene transformed into tobacco leaf cells targets to LDs via the ER (Wahlroos et al., 2003). In some species, the leaf LDs may contain hydrophobic secondary metabolites (e.g., rubber droplets in guayule) instead of TAGs. In some other species, the leaf LDs may be remnants of OBs in primitive plants and may be induced to proliferate under special situations, such as starvation and senescence. Under the latter situations, glyoxysomes and other machineries appear and convert degraded lipids into sugar for internal use or for export to non-senescing tissues (Pracharoenwattana and Smith, 2008). Overall, leaf LDs in diverse species, unlike those in seeds, are heterogeneous in structures and lipid contents and have diverse functions under different developmental, physiological and environmental conditions.

The haploid *Physcomitrella* spore is genetically and physiologically similar to the haploid pollen in higher plants. Both the spore and pollen also contain storage OBs coated with oleosins. However, the function of *Physcomitrella* spore OBs is for gluconeogenesis via the glyoxysomes. Pollen OBs are not metabolized via the glyoxysomes and likely act as reserves of acyl moieties for synthesis of new plasma membrane during pollen tube elongation.

Physcomitrella can be easily transformed and examined with CLSM for transient expression of genes that are involved in storage TAG and SE metabolism and contain abundant OBs. Such a plant system was not previously available (Miao and Jiang, 2007), and transient expression of genes in *Physcomitrella* has been performed only with juvenile protonema (e.g., Marella et al., 2006), which contain few or no OBs (Fig. 1-1A). Earlier, maturing embryos of flax and microspore cultures of *Brassica* were used for transient expression of modified oleosin genes, and the transformed plant materials were examined with in vitro biochemical analyses (Abell et al., 2004). The OB-containing internal cells in an embryo may not be transformed easily with bombardment or observed clearly with CLSM, and the microspore culture is highly artificial, and the microspore (pollen) OBs are not for gluconeogenesis. Another major advantage of the *Physcomitrella* transient expression system is that

the growth condition of the plant can be altered easily for study of storage lipid metabolism and packaging.

The sizes of OBs in the *Physcomitrella* gametophyte and spore are highly variable, whereas those of OBs in seeds of individual species are more confined. Presumably, *Physcomitrella* has not evolved a mechanism to control the coordinate synthesis of TAGs and oleosins, and thus the sizes of OBs, within the same cell. In seeds, OB sizes are directly related to the ratio of TAGs to oleosins, as in kernels of maize lines that were bred for high or low oils (Ting et al., 1996) and seeds of Arabidopsis mutants whose oleosin genes were knocked out or down (Siloto et al., 2006; Shimada et al., 2008). In maize lines bred for low oils (resulting in a low ratio of oils to oleosins), the OBs are not only smaller but also have irregularly shaped surface that could accommodate more surface oleosins per unit of matrix TAGs. In high-oil maize lines, the OBs are spherical and substantially larger. Similarly, in mutant Arabidopsis seeds with lesser amounts of oleosins, the OBs are considerably bigger. Occasional dumbbell-shaped OBs are present and looked upon as fusing detached OBs that do not have sufficient surface oleosins. An alternative explanation is that the apparent fusion occurred among budding OBs that were still attached to the ER.

In maturing seeds, whether oleosins and TAGs are synthesized in extended regions or restricted subdomains of the ER is uncertain. Earlier, OB synthesis in restricted ER subdomains has been suggested on the basis that immuno-TEM reveals more oleosins in the ER near ER-OB structures (Herman, 1987) and that a low-density, isolated subfraction (ER-OB structures) could synthesize more TAGs in vitro (Lacey et al., 1999). However, these results could also be interpreted as OBs being synthesized in extended ER regions, such that there is a concentration gradient in the ER, with more oleosins near the budding OBs; the abundant oleosins in these ER-OB structures, upon isolation, would facilitate in vitro TAG synthesis. More recently, fluorescence microscopy revealed specific DAGAT for TAG synthesis in highly defined ER subdomains in transformed BY2 cells (Shockey et al., 2006); uncertainty exists because BY2 cells do not contain OBs. On the contrary, in cells of both *Physcomitrella* gametophyte (current report) and *Brassica* tapetum (Hsieh and Huang, 2005), oleosins are synthesized in extended regions of the ER. Nevertheless, neither of these systems is directly related to maturing seeds, which is specialized to produce massive OBs. It is also possible that oleosins are synthesized in extended ER regions and diffuse to restricted TAG-synthesizing ER subdomains from which nascent OBs detach to become solitary entities.

MATERIALS AND METHODS

Plant Materials

Spore of *Physcomitrella patens* subsp. *patens* was kindly provided by Dr. Eugene Nothnagel of the University of California, Riverside. Gametophytes were grown axenically on a solid Knop's medium containing 125 mg.L^{-1} KNO_3, 125 mg.L^{-1} KH_2PO_4, 125 mg.L^{-1} $MgSO_47H_2O$, 500 mg.L^{-1} $Ca(NO_3)_24H_2O$ and 10 g.L^{-1} glucose supplemented with 1 ml.L^{-1} 1000X Hunter's "metal 49" micronutrients (76 mg.L^{-1} 5-sulfosalicylic acid dihydrate, 7 g.L^{-1} $Fe(NH_4)_2(SO_4)_26H_2O$, 3.04 g.L^{-1} $MnSO_4H_2O$, 2.2g.L^{-1} $ZnSO_47H_2O$, 0.025 mg.L^{-1} $(NH_4)_6Mo_7O_{24}4H_2O$, 616 mg.L^{-1} $CuSO_45H_2O$, 238 mg.L^{-1} $CoSO_47H_2O$, 57.2 mg.L^{-1} H_3BO_3, 18 mg.L^{-1} Na_3VO_4; Basile, 1978) and 1.2% (w/v) agar, pH 4.6. Plants were cultured at 25±1^{0}C under a 16-h light (60~100 µE $m^{-2}S^{-1}$)/8-h dark cycle. Sexual development was carried out with cold stimulation. Cultures of 45 d were half-submerged in water and cultured at 15^{0}C. After the cold treatment for 60 d, mature sporophytes were harvested from the apex. From these sporophytes, spore was collected.

Tissues for RT-PCR analysis were protonema (tissue grown from spore after germination for 10 d); mature gametophytes (60-d-old culture) subdivided into top, middle, and bottom leafy tissues; antheridia; antheridia and archegonia; zygote;

young, maturing and old sporophytes (S1, S2 and S3 obtained after 32, 45 and 56 d of cold stimulation, respectively), mature spore and germinating spore (on a solid Knops' medium covered with a layer of cellophane at 25^0C and with continuous light for 2 d).

Staining of OBs in situ

OBs in situ were stained with Sudan Black B, Nile Red (Greenspan et al., 1985) or 4,4-difluoro-1,3,5,7,8-pentamethyl-4-bora-3a,4a-diaza-*s*-indacene (BODIPY® 493/503, D-3922 from Invitrogen Corp., Carlsbad, CA). For Sudan Black staining, fresh tissues were placed in 70 % (v/v) propylene glycol for 5 min, transferred to a saturated Sudan Black B solution (in 70 % propylene glycol) for 10 min, washed with 50 % propylene glycol twice and observed with light microscopy. For Nile-Red or BODIPY staining, fresh tissue or fixed tissue (after immunofluorescence treatment, to be described) were placed in a solution consisting of Nile Red stock (100 µg/ml acetone) or BODIPY® 493/503 stock (10 mg/ml DMSO) diluted 100x with 1x phosphate buffered saline (PBS, [10 mM K phosphate, pH 7.4, 138 mM NaCl and 2.7 mM KCl]) for 10 min, washed with PBS twice, and observed with a Ziess LSM 510 META NLO confocal microscope. Nile Red and BODIPY® 493/503 were excited with the 543 and

488 nm lines, and its emission was detected with filter BP 565-615 and BP 500-530, respectively.

Isolation of OBs from the Gametophyte

All procedures were performed at 4^0C. Fresh, 60-d-old gametophytes were soaked in a grinding medium (0.6 M sucrose, 0.1 M HEPES-NaOH, 4 mM dithiothretol, pH 7.5) for 20 min and chopped with a razor blade and then ground with a mortar and pestle. The ground sample was filtered through a layer of Nitex cloth (50 µm x 50 µm) to yield a total extract. The total extract was placed at the bottom of a centrifuge tube, and a lighter solution (0.4 M sucrose, 0.1 M HEPES-NaOH, pH 7.5) was loaded above the extract. The tube was centrifuged at 18,000 rpm for 45 min in a Beckman SW28 rotor. Floated OBs at the top of the gradient were collected with use of a spatula.

Analysis of Lipids

Lipids in the total gametophyte extract and isolated OB fraction were extracted with 1.2x volume of lipid extraction buffer (chloroform/heptanes/methanol, 4/3/2,

v/v/v). The extract was evaporated to dryness with a stream of nitrogen gas and re-dissolved in ether or acetone for TLC or HPLC, respectively.

Lipid samples were applied to TLC plates (silica gel 60A; Whatman), which were developed in hexane:diethyl ether:acetic aicd (80:20:2, v/v/v). Lipids on the plates were visualized after sulfuric acid spray. HPLC-APCI/MS was performed with Agilent 1100 series liquid chromatography coupled with ThermoFinnigan LCQ Advantage ion trap mass spectrometer (San Jose, CA) with an APCI interface. HPLC was carried out with a 5-μl sample (20 mg/ml) and a RP-18 column (Phenomenex Luna 3μ C18, 150 mm×2.0 mm) at 30 °C. Elution was performed with isocratic acetone-acetonitrile (1:1, v/v) at a flow-rate of 0.2 ml/min. The peaks were analyzed with DAD detection at 205 nm and then MS detection. Ionization was performed in the positive ion mode for all analyses.

Electron Microscopy

Tissues (cut into ~2 mm x 2 mm pieces) and the isolated OB fraction (in 0.4 M sucrose and 0.01M K-phosphate buffer [pH 7.4]) were fixed with 2.5% glutaraldehyde, 4% paraformaldehyde and 0.1 M K-phosphate (pH 7.0) at 4^0C for 24 h. The samples were washed with 0.1 M K-phosphate buffer for 10 min 2x and then treated with 1%

OsO_4, 0.1 M K-phosphate (pH 7.0) at room temperature for 4 h. The fixed samples were rinsed with 0.1 M K-phosphate buffer and dehydrated through an acetone series and embedded in Spurr medium. Ultrathin sections (70 to 90 nm) were obtained with a Lecia Reichert Ultracut S or Lecia EM UC6 ultramicrotome (Wetzlar, Germany). Sections were stained with uranyl acetate and lead citrate and examined with a Philips CM 100 transmission electron microscope at 80 KV.

Immunofluorescence CLSM

All antibody treatments were performed with 1:50 dilution of the IgG fraction (isolated and resuspended into the same original anti-catalase serum volume), 1% (w/v) milk powder and 1x PBS. Each wash was performed with PBST (1x PBS and 0.05% [w/w] Tween-20) for 10 min. Tissues were fixed in 4% paraformaldehyde, 1x PBS and 0.15 M sucrose at 4^0C for 16 h. After 2 washes, the tissues were treated with 1% cellulase R10 (Yakult in Honsa, Japan) in 1 x PBS for 20 min at 25^0C. After 2 washes, the tissues were incubated with 1x PBS with 0.1 % Tween-20 for 20 min at 25^0C. After 2 washes, the tissues were treated with a blocking solution (3% milk, 1x PBS) at 25^0C for 1 h and then rabbit antibodies against cotton seed catalase at 4^0C for 16 were. After 3 washes, the tissues were treated with cyanine 3-conjugated donkey

antibodies against rabbit IgG (Jackson Immuno Research Laboratories, west Grove, PA) for 1 h at 25^{0}C. After 3 washes, the tissues were stained with BODIPY® 493/503 for OB staining. The tissues were placed on a slide and observed with a LSM 510 META confocal microscope (Carl Ziess in Jena, Germany). BODIPY® 493/503, cyanines 3 and chloroplast were excited with the Argon 488-, HeNe 543- and Argon 488-nm lines, respectively, and the emissions were detected with emission filters of BP 500–530, BP 565–615 and BP 650–710 nm, respectively.

SDS-PAGE and Identification of Oleosins with MALDI-TOF

Proteins in the total cellular extract and the isolated OB fraction were separated with 12 % (w/v) SDS-PAGE (Wu et al., 1997). The gel was stained with Coomassie Blue. The gel containing the visible ~17-kD proteins of the OB fraction was cut. The proteins were extracted and subjected to trypsin digestion and mass spectrometry (MALDI-TOF) analysis with Voyager DE-STR (PerSeptive Biosystems, Framingham, MA).

Searches for Oleosin Genes of *Physcomitrella* and Other Organisms

Sequences of the conserved hairpin domain of oleosins and several complete

oleosins of Arabidopsis (Kim et al., 2002) were used as query sequences for the BLAST program (tblastn) against genome and transcriptome databases of *Physcomitrella patens* (http://www.cosmoss.org/). Three oleosin genes, *PpOLE*1, *PpOLE*2 and *PpOLE*3, were found on scaffold 84, scaffold 21 and scaffold 180, respectively. Similar searches yielded oleosin genes of *Oryza sativa* (from http://rice.plantbiology.msu.edu/), *Arabidopsis thaliana*, *Selaginella mutica* and *Populus trichocarpa* (from JGI Eukaryotic Genomics database, http://www.jgi.doe.gov/). An oleosin of *Pinus ponderosa*, *Pinus-OLE*, was obtained from an earlier study (Lee et al., 1994). A phylogenetic tree of the above oleosins was constructed on the basis of protein sequence similarities (of the conserved hairpin sequence plus the moderately conserved sequences immediate flanking the hairpin) and constrictions with the Clustal method and a distance method (Neighbor-Joining) and PHYLIP with 1,000 bootstrap replicates.

Reverse Transcription-PCR analyses

RNA was extracted from tissues with use of an RNeasy *Mini* Kit (Qiagen, Valencia, CA). Total RNA (2 µg) was first treated with DNaseI (Invitrogen) for 35 min at 37^0C. The proteins were removed with PCI (*Phenol*/chloroform/isoamyl alcohol

[25:24:1], pH 4.5), and the phenol with chloroform-isoamyl alcohol (24/1, v/v). The RNA was precipitated with 0.1x volume of 3 M NaOAc (pH 5.2) and 2.5x volume of absolute ethanol at -20^{0}C for 16 h. The RNA (1µg) was used to make cDNA with the SuperScriptIII RT-PCR system (Invitrogen). The RNA was pre-incubated with oligo(dT)$_{12,18}$ and dNTPs at 65^{0}C for 5 min and then placed on ice. cDNA Synthesis Mix was added to the RNA, and the mixture was incubated at 50^{0}C for 1 h. The reaction was terminated by heating at 75^{0}C for 5 min. PCR was carried out with use of 0.2 µl of the cDNA as template and DyNazyme DNA polymerase with dNTP and primers. From the sequence information of the genes and their transcripts, primers were selected and synthesized. Primer pairs for amplifying full-length cDNA of *PpOLE*1a, *PpOLE*2a and *PpOLE*3 are shown in Table 1-1. Amplified DNA fragments were subcloned into pGEM-T Easy (Promega, Madison, WI) and subjected to DNA sequencing with use of M13 forward and reverse primers. Primer pairs for amplification of specific gene fragments in the study of gene expression patterns are shown in Supplemental Table 1-1. Amplified DNA fragments of ~200bp were analyzed on a 1.8% agarose gel.

Transient Expression assays

DNA sequences encoding the complete coding region of *PpOLE*1a and *PpOLE*2a were amplified by PCR with use of primers shown in Table 1-1. The resulting coding fragments were digested with BamHI and cloned into the expression site of a GFP expression vector (Chiu et al. 1996) or an RFP expression vector (Lee et al., 2001) to be driven by a CaMV 35S promoter. A BIP-RFP expression vector of a similar construct (Kim et al., 2001) was obtained from Dr. David Ho, Institute of Plant and Microbial Biology, Taipei. Transformation of the gametophyte was carried out with particle bombardment. Sixty-day-old gametophyte tissues were placed on solid Knop's medium. Plasmid DNA (5 μg) was coated onto the surface of 1.25 mg 1.6-nm gold particles, which would be used for 6 different shootings. The gold particles were bombarded with 900 psi under 28-in Hg vacuum onto the gametophyte from a distance of 6 cm in PDS-1000 (BIO-RAD, Hercules, CA). After bombardment, the tissues were left on the culture medium and observed with CLSM at time intervals. GFP and RFP were excited with the Argon 488- and HeNe 543-nm lines, and their emissions were detected by emission filters of BP 500–530 and BP 565–615, respectively.

REFERENCES

Abell BM, Hahn M, Holbrook LA, Moloney MM (2004) Membrane topology and sequence requirements for oil body targeting of oleosin. Plant J **37:** 461-470.

Barbazuk WB, Fu Y, McGinnis KM (2008) Opportunities and challenges: genome-wide analyses of alternative splicing in plants. Genome Res **18:** 1381-1392.

Chiu W, Niwa Y, Zeng W, Hirano T, Kobayashi H, Sheen J (1996) Engineered GFP as a vital reporter for plants. Curr Biol **6:** 325-330.

Dembitsky VM (1993) Lipids in bryophytes. Progress Lipid Research **32:** 281-356.

Frandsen GI, Mundy J, Tzen JT (2001) Oil bodies and their associated proteins, oleosin and caleosin. Physiol Plant **112:** 301-307.

Goodman JM (2008) The gregarious lipid droplets. J Biol Chem **42:** 28005-28009.

Greenspan P, Mayer EP, Fowler SD (1985) Nile Red: A selective fluorescent stain for intrcellular lipid droplets. J Cell Biol **100:** 965-973.

Herman EM (1987) Immunogold-localization and synthesis of an oil-body membrane protein in developing soybean seeds. Planta **172:** 336-345.

Hsieh K, Huang AHC (2004) Endoplasmic reticulum, oleosins, and oils in seeds and tapetum. Plant Physiol **136:** 3427-3434.

Hsieh K. Huang AHC (2005) Lipid-rich tapetosomes in *Brassica* tapetum are composed of oleosin-coated oil droplets and vesicles, both assembled in and then detached from the endoplasmic reticulum. Plant J **43:** 889-899.

Hsieh K, Huang AHC (2007) Tapetosomes in *Brassica* tapetum accumulate endoplasmic reticulum-derived flavonoids and alkanes for delivery to the pollen surface. Plant Cell **19:** 582-596

Kim DH, Eu YJ, Yoo CM, Kim YW, Pih KT, Jin JB, Kim SJ, Stenmark H, and Hwang I. (2001). Trafficking of phosphatidylinositol 3-phosphate from the trans-Golgi network to the lumen of the central vacuole in plant cells. Plant Cell **13**: 287–301.

Kim HU, Hsieh K, Ratnayake C, Huang AHC (2002) Expression of *Arabidopsis* oleosin genes and characterization of their encoded oleosins. J Biol Chem **277:** 22677-22684.

Lacey DJ, Beaudoin F, Dempsey CE, Shewry PR, Napier JA (1999) The accumulation of triacylglycerols within the endoplasmic reticulum of developing seeds of *Helianthus annuus*. Plant J **17:** 397-405.

Lee K, Bih FY, Learn GH, Ting JTL, Sellers C, Huang AHC (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.

Lee YJ, Kim DH, Kim YW and Hwang I (2001) Identification of a signal that distinguishes between the chloroplast outer envelope membrane and the endomembrane system *in vivo*. Plant Cell **13**: 2175–2190.

Lersten NR, Czlapinski AR, Curtis JD, Freckmann R, Horner HT (2006) Oil bodies in leaf mesophyll cells of angiosperms: overview and a selected survey. Am J Bot **93:** 1731-1739.

Marella HH, Sakata Y, Quatrano RS (2006) Characterization and functional analysis of ABSCISIC ACID INSENSITIVE3-like genes from *Physcomitrella patens*. Plant J **46:** 1032-1044.

Martin S, Parton RG (2006) Lipid droplets: a unified view of a dynamic organelle. Nature Review **7:** 373-378.

Meyers BC, Lee DK, Vu TH, Tej SS, Edberg SB, Matvienko M, Tindell LD (2004) *Arabidopsis* MPSS. An online resource for quantitative expression analysis. Plant Physiol **135:** 801-813.

Miao YS, Jiang L (2007) Transient expression of fluorescent fusion proteins in protoplasts of suspension cultured cells. Nature Protocols **2:** 2348-2353.

Murphy DJ (2001) The biogenesis and functions of lipid bodies in animals, plants and microorganisms. Plant J **13:** 1-16.

Napier JA, Stobart AK, Shewry PR (1996) The structure and biogenesis of plant oil bodies: the role of the ER membrane and the oleosin class of proteins. Plant Mol Biol **31:** 945-956.

Nobuta K, Venu RC, Lu C, Belo A, Vemaraju K, Kulkarni K, Wang W, Pillay M, Green PJ, Wang GL, Meyers BC (2007) An expression atlas of rice mRNAs and small RNAs. Nat Biotechnol **25:** 473-477.

Pihakaski K, Pihakaski S, Karunen P (1987) Seasonal changes in leaf lipids of *Diphensia lapponica* with special reference to storage lipid bodies. Nordic J Bot **7:** 281-292.

Pracharoenwattana I, Smith SM (2008) When is a peroxisome not a peroxisome? Trends in Plant Sci **13:** 522-525.

Rajakumari S, Grillitsch K, Daum G (2008) Synthesis and turnover of non-polar lipids in yeast. Prog in Lip Res **47:** 157-171.

Rensing SA, and >50 coauthors (2008) The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. Science **319:** 64-69.

Ryan RO (1994) The structures of insect lipoproteins. Curr. Opinion Structural Biol **4:** 499-506.

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

Shockey JM, Gidda SK, Chapital DC, Kuan JC, Dhanoa PK, Bland JM, Rothstein SJ, Mullen RT, Dyer JM (2006) Tung tree DGAT1 and DGAT2 have nonredundant

functions in triacylglycerols biosynthesis and are localized to different subdomains of the endoplasmic reticulum. Plant Cell **18:** 2294-2313.

Siloto RMP, Findlay K, Lopez-Villalobos A, Yeung EC, Nykiforuk CL, Moloney MM (2006) The accumulation of oleosins determines the size of seed oilbodies in *Arabidopsis*. Plant Cell **18:** 1961-1974.

Swanson ES, Anderson WH, Gellerman JL, Schlenk H (1976) Ultrastructure and lipid composition of mosses. The Bryologist **79:** 339-349.

Ting JTL, Balsamo RA, Ratnayake C, Huang AHC (1997) Oleosin of plant seed oil bodies is correctly targeted to the lipid bodies in transformed yeast. J Biol Chem **272:** 3699-3706.

Tzen JTC, Cao YZ, Laurent P, Ratnayake C, Huang AHC (1993) Lipids, proteins, and structure of seed oil bodies from diverse species. Plant Physiol **101:** 267-276.

Wahlroos T, Soukka J, Denesyuk A, Wahlroos R, Korpela T, Kilby NJ (2003) Oleosin expression and trafficking during oil body biogenesis in tobacco leaf cells. Genesis **35:** 125-132.

Wu SSH, Platt KA, Ratnayake C, Wang TW, Ting JTL, Huang AHC (1997) Isolation and characterization of novel neutral-lipid-containing organelles and globuli-filled plastids from *Brassica napus* tapetum. Proc Natl Acad Sci USA **94:** 12711-12716.

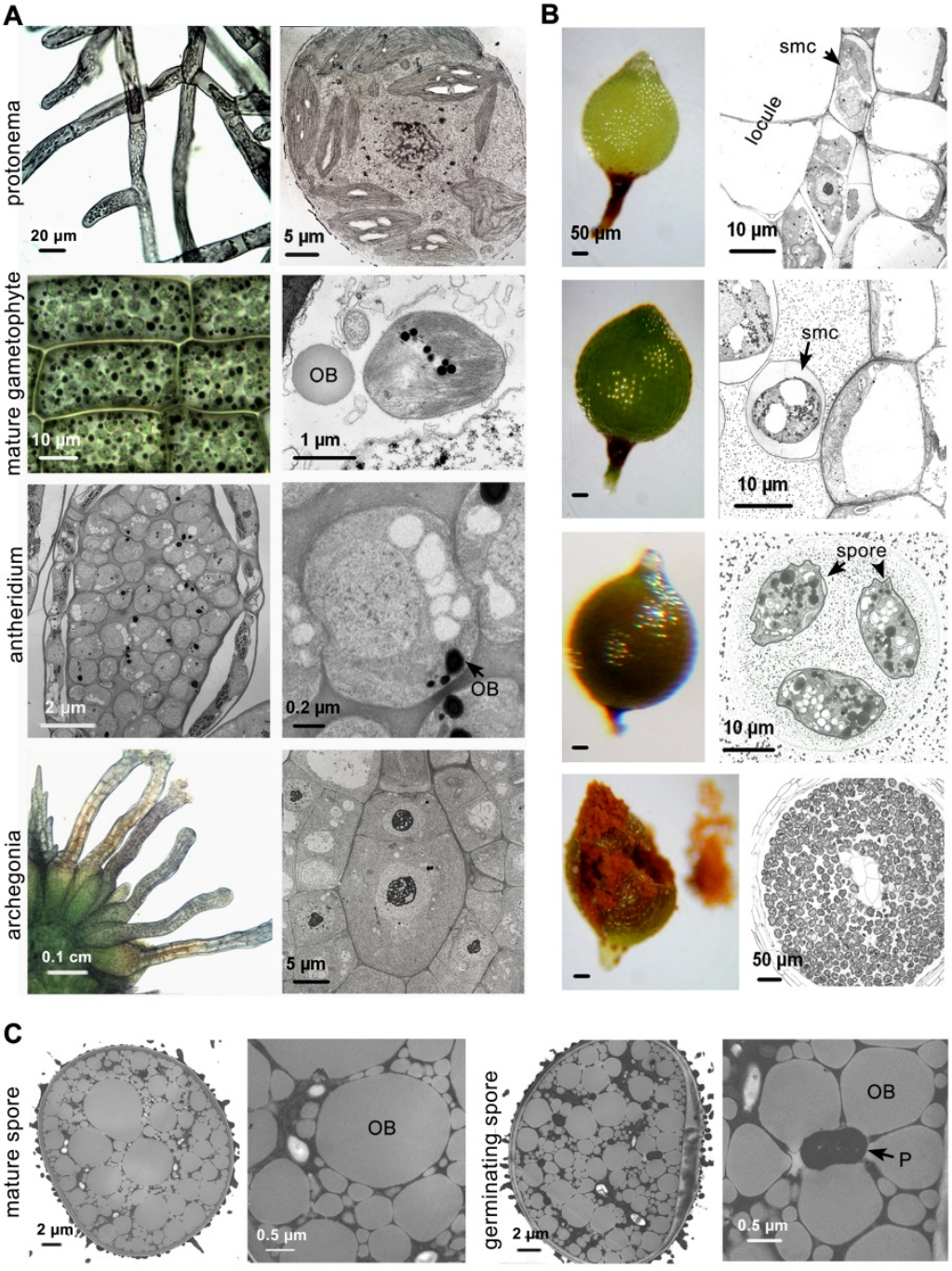

Fig. 1-1. Microscopy images of different tissues of *Physcomitrella* during its life cycle, showing the presence or absence of subcellular OBs.

Samples were photographed with light microscopy (in color) or transmission electron microscopy (black and white).

A. Haploid cells (from top to bottom) of protonema (immature tissue grown from spore after germination for 10 d), mature gametophytes (dark OBs [stained with

Sudan Black B] among green chloroplasts), antheridium (dark OBs) and archegonia (no OBs).

B. Sporophytes during maturation (from top to bottom). Left column shows the whole sporophyte, and right column reveals diploid cells enclosing the locule in which spore-mother cells (smc) became spore.

C. Mature spore packed with OBs but no peroxisomes (left 2 images) and germinating spore with numerous peroxisomes (p) (right 2 images).

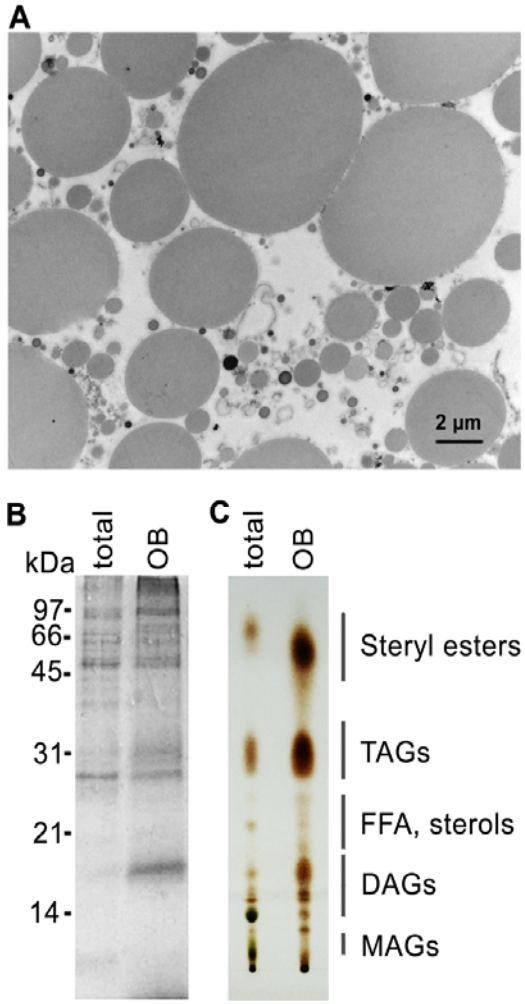

Fig. 1-2. Analyses of an OB fraction isolated from mature, green gametophyte branches.

A. TEM of the OB fraction, showing spherical OBs of heterogeneous sizes.

B. SDS-PAGE gel of proteins in total extract and OB fraction. Molecular marker positions are indicated.

C. TLC plate of neutral lipids in total extract and OB fraction. Approximate locations of lipid groups, steryl esters (cholesteryl palmitate as marker), TAGs (trioloein), FFA (free fatty acids, oleic acid), sterols (cholesterol), DAGs (1,3- and 1,2-diolein) and MAGs (1-monoolein) are indicated.

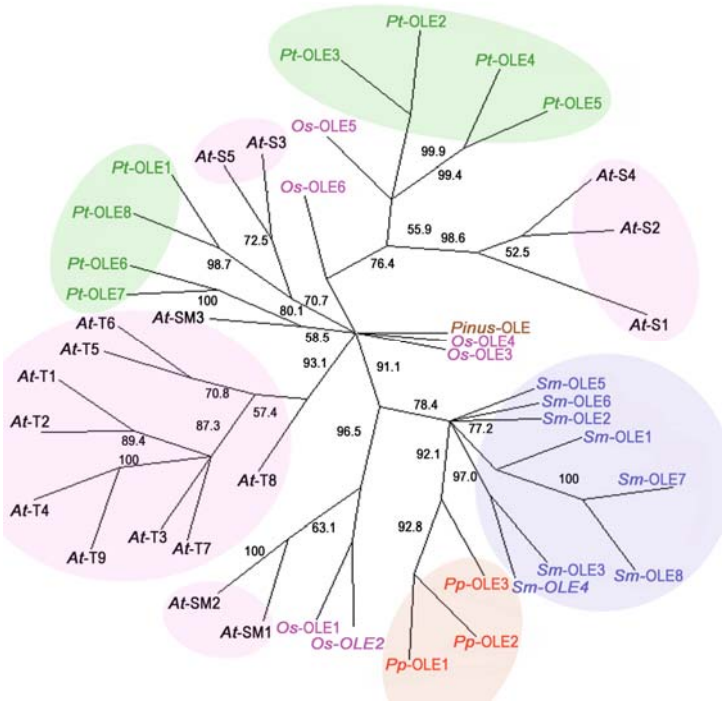

Fig. 1-3. An un-rooted phylogenetic tree of oleosins from *Physcomitrella* and several other representative species constructed on the basis of their predicted amino acid sequences.

Arabidopsis thaliana (*At*, a non-woody dicot), *Populus trichocarpa* (*Pt*, a woody dicot), *Oryza sativa* (*Os*, a monocot), *Selaginella moellendorffii* (*Sm*, a primitive fern) and *Physcomitrella patens* (*Pp*) genomes have been completely sequenced, and all their oleosins are included. One pine (*Pinus ponderosa*) oleosin is used to represent gymnosperm proteins. Nomenclature of the Arabidopsis oleosins follows that reported (Kim et al., 2002); S, T and SM denote oleosins present specifically in seed, tapetum, and seed-and-microspore, respectively. The phylogenetic tree was constructed from aligned sequences of oleosins (the conserved hairpin sequence plus its immediately flanking semi-conserved sequences) by a distance method (Neighbor-Joining) with use of PHYLIP and 1,000 bootstrap replicates. Bootstrap values higher than 50 are indicated.

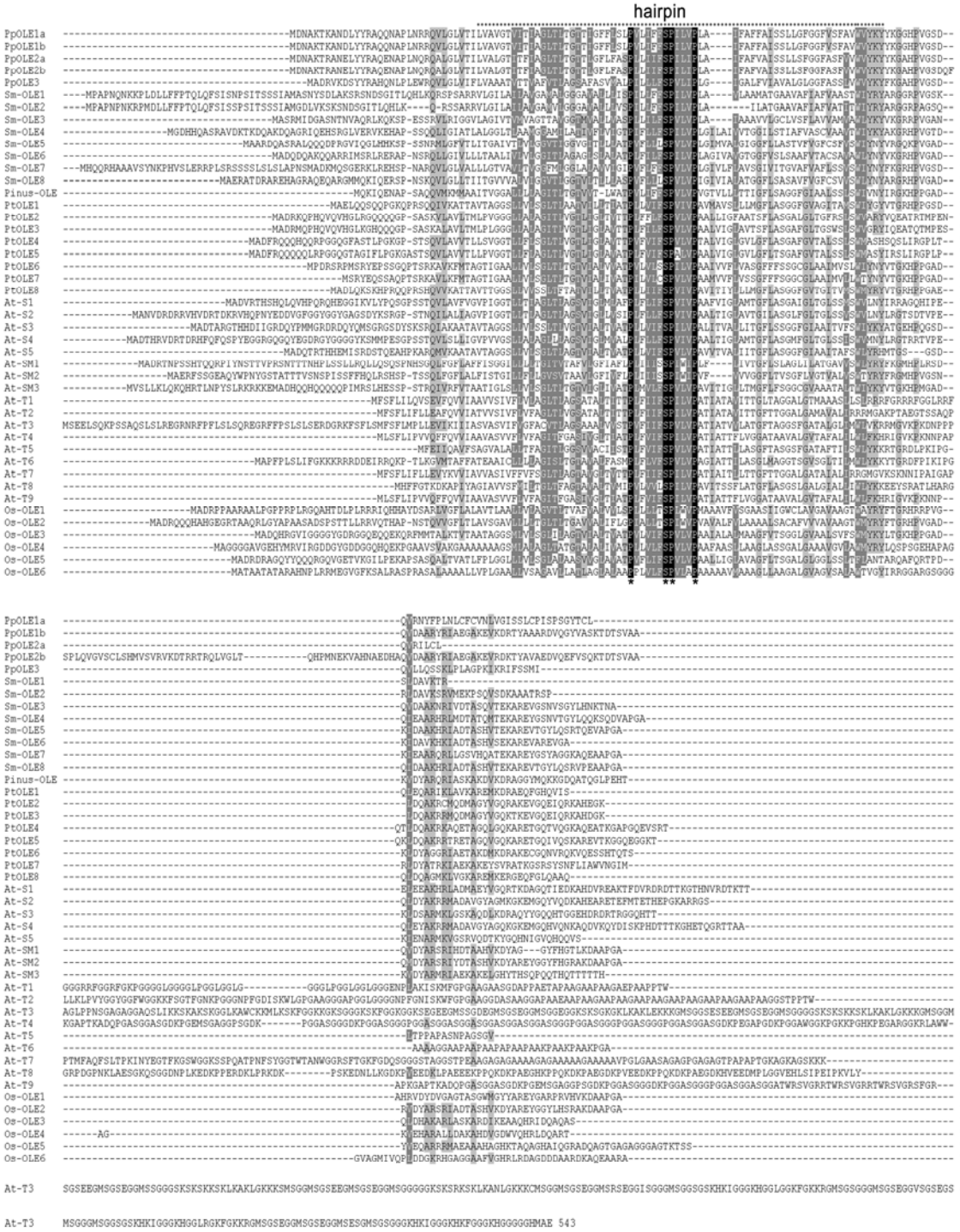

Fig. 1-4. A pile-up of 45 oleosins from *Physcomitrella patens* (Pp), *Arabidopsis thaliana* (At), *Populus trichocarpa* (Pt), *Oryza sativa* (Os), *Selaginella moellendorffii* (Sm) and *Pinus ponderosa* (Pinus).

The hairpin sequences are dotted. The identical residues in proline knot were indicated with asterisks.

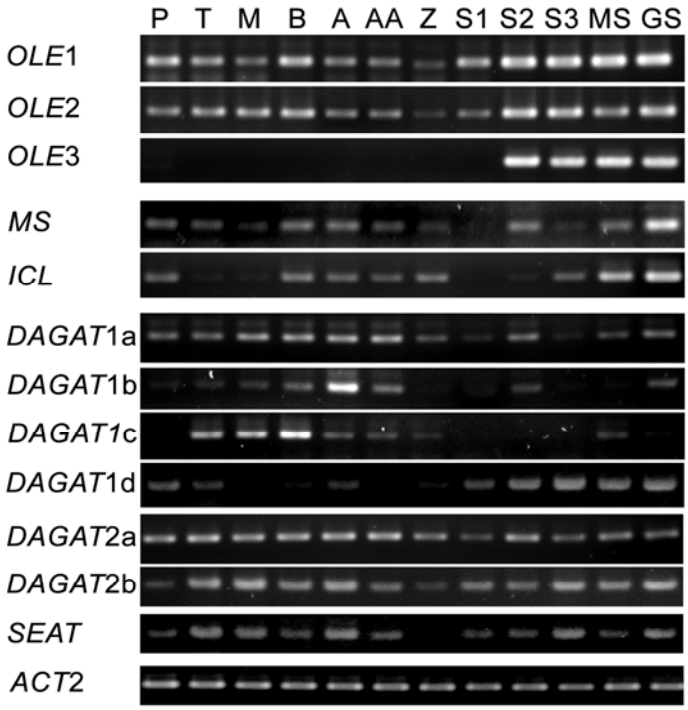

Fig. 1-5. RT-PCR of transcripts encoding oleosins and related proteins in various tissues

Tissues shown from left to right are protonema (P); top, middle and bottom leafy tissues of gametophyte branches (T, M, and B, respectively); antheridia (A); antheridia and archegonia (AA); zygotes (Z), sporophytes of early, middle and late developmental stages (S1, S2 and S3, respectively); and mature spore (MS) and germinated spore (GS). Transcripts are those encoding oleosins (*OLE*1, 2 and 3), malate synthase (*MS*), isocitrate lyase (*ICL*), DAG acyltransferases (*DAGAT*1 [4 paralogs] and *DAGAT*2 [2 paralogs] utilizing acyl CoA and PLs as the acyl donors, respectively) and steryl acyltransferase (*SEAT*). Approximately equal amounts of the transcript encoding actin (*ACT*2) were present in the samples.

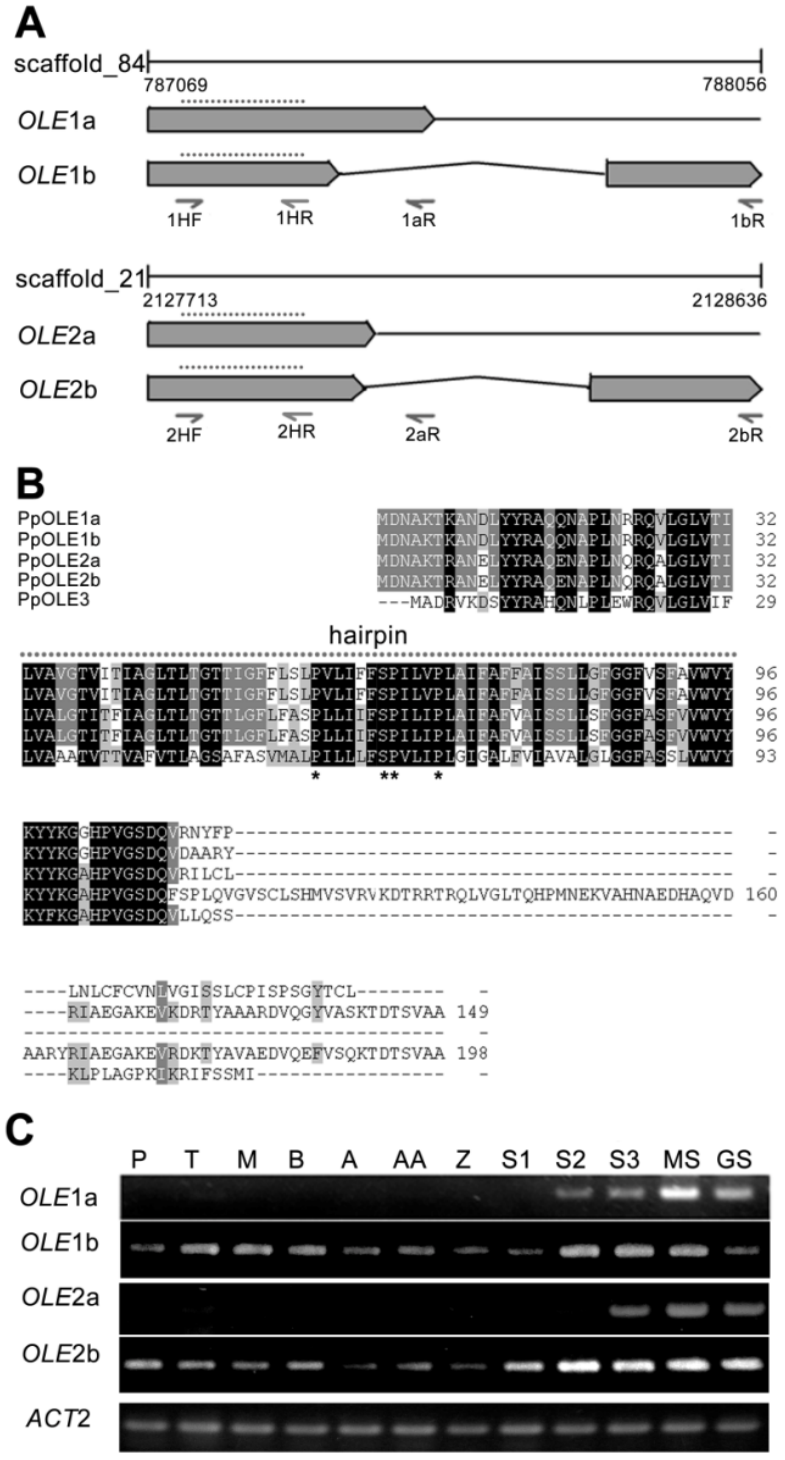

Fig. 1-6. Gene structures of *OLE*1 and *OLE*2 and alternative splicing of each gene resulting in 2 tissue-specific transcripts.

A. Arrangement of *OLE*1 and *OLE*2 in 2 scaffolds. Occurrence of 2 open reading frames (shaded boxes) in each of the 2 genes via alternative splicing is indicated. The

dotted lines represent the sequences encoding the hairpin region. Primers for RT-PCR are shown.

B. A pile-up of OLE1a, OLE1b, OLE2a, OLE2b and OLE3. The hairpin sequences in the second row are dotted, and the 3P and 1S in the central hairpin turn, PX_5SPX_3P, are highlighted with asterisks.

C. RT-PCR of transcripts of *OLE*1a and *OLE*1b, *OLE*2a and *OLE*2b, as well as *ACT*2 (a loading control) in various tissues (see Fig. 4 legend for labels). Primers were 1HF and 1aR for *OLE*1a and 1HF and 1bR for *OLE*1b (see **A**); and 2HF and 2aR for *OLE*2a and 2HF and 2bR for *OLE*2b.

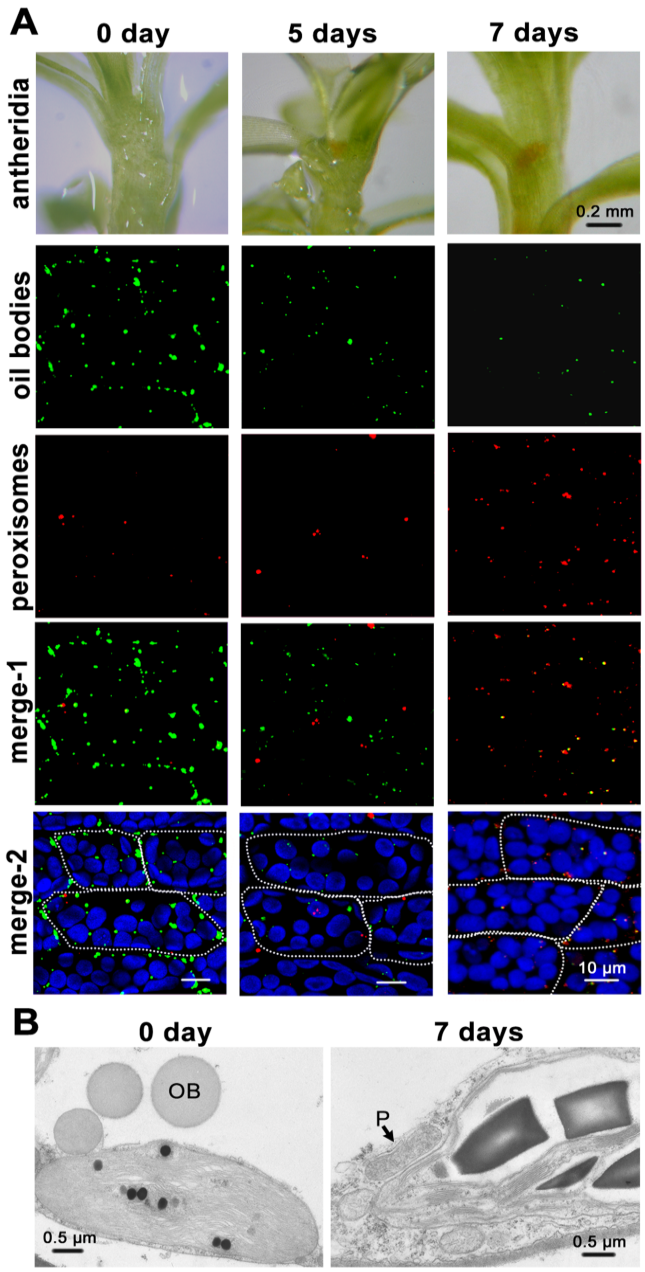

Fig. 1-7. Images of the gametophyte after induction of sporophyte development

The gametophyte was examined after induction of sporophyte development on cold treatment for 0, 5 and 7 d. **A.** The uppermost row shows images of the tip of gametophyte, which was producing antheridia (brown color) and archegonia (not visible). The subsequent rows are fluorescence CLSM images of several cells in an uppermost leafy tissue of a branch. The cells were examined for OBs with use of the lipid probe BODIPY 493/503 (green), peroxisomes with rabbit anti-catalase antibodies and then anti-rabbit IgG antibodies conjugated to Cy3 (red) and chloroplasts with autofluorescence (blue). Each column shows identical cells after the indicated days of cold treatment. Photos were taken to reveal OBs or peroxisomes alone, or in combination (merge-1 for OBs and peroxisomes, and merge-2 for chloroplasts also and with dotted lines to outline the cell circumference).

B. TEM pictures of portions of cells in an uppermost leafy tissue of a gametophyte branch showing the presence of OBs (OB) at d 0 and a peroxisome (p) at d 7.

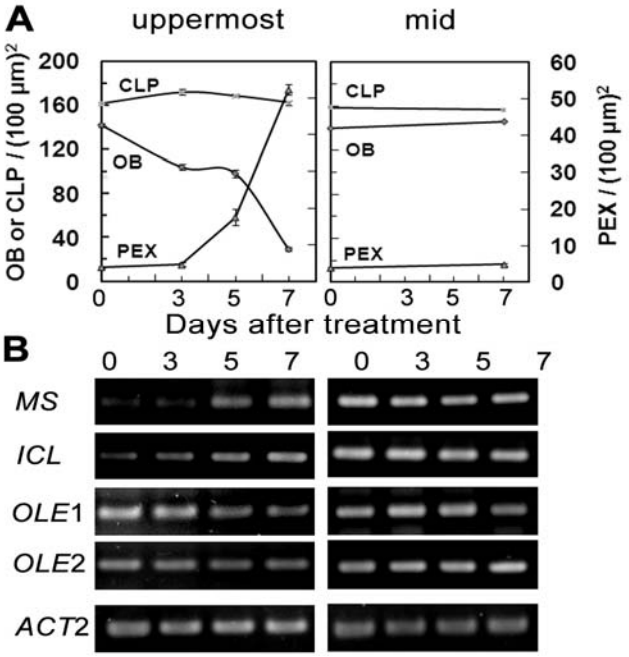

Fig. 1-8. Changes in the number of organelles and levels of gene transcripts in the uppermost and mid leafy tissues of gametophyte branches upon induction of sporophyte development for 0, 3, 5 and 7 d

A. Number of OBs, peroxisomes (PEX) and chloroplasts (CLP) per cell area.

B. RT-PCR results of transcripts encoding malate synthase (*MS*), isocitrate lyase (*ICL*), OLE1 (*OLE*1), OLE2 (*OLE*2) and actin (*ACT*2, as a loading control).

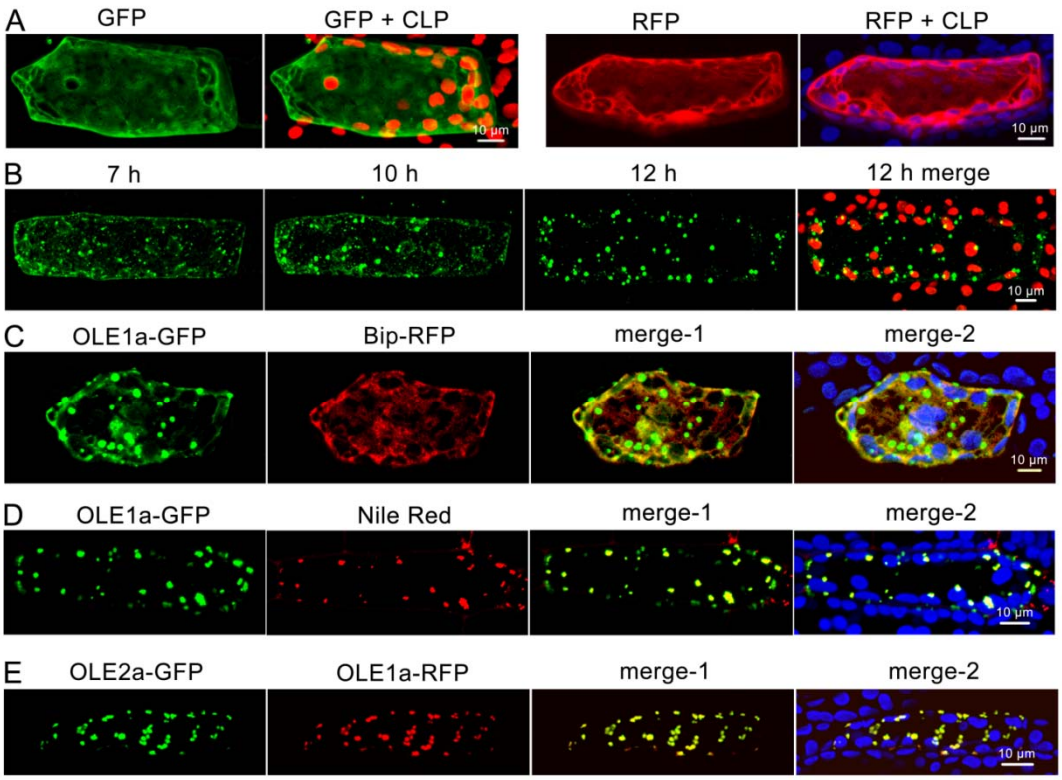

Fig. 1-9. Transient expression of various *GFP* and *RFP* constructs in individual gametophyte cells.

Chloroplast autofluorescence is shown in red or blue in merged pictures. The speed of transient expression varied among experiments, and CLSM images were obtained at 6-8 (early), 8-10 (mid) and 10-12 (late time point) were after bombardment.

A. Expression of control *GFP* or *RFP* not attached to *OLE1a* at a late time point.

B. Expression of *OLE*1a-*GFP* at time intervals. GFP (green) was present largely in a cellular network at an early time point but was associated more with subcellular droplets at a late time point.

C. Co-expression of *OLE*1a-*GFP* and *BiP-RFP*. GFP (green) and RFP (red) at a mid time point are shown.

D. Expression of *OLE*1a-GFP. GFP (green) and OBs (red, stained with Nile Red) at a late time point are shown.

E. Co-expression of *OLE*1a-*GFP* and *OLE*2a-*RFP*. GFP (green) and RFP (red) at a late time point are shown.

Table 1-1. primer sequences

Primer name	Sequences (5'→3')
For full-length cDNA of *PpOLE*1a, *PpOLE*2a, and *PpOLE*3	
ppOLE1aF	ACGTTCGGCAACATGGATAA
ppOLE1aR	TCTCGTTACTCGGAATAAATTA
ppOLE2aF	GCTCTGCAACATGGACAAT
ppOLE2aR	TTCATAAGCAAAGGATTCG
ppOLE3F	GAGATGGCCGACAGGGTC
ppOLE3R	GGCTGATTATATCATGGAAGA
For specific gene fragments (see figures 4 and 5)	
PpOLE1HF	ATTGGTTACCATCCTCGTCG
PpOLE 1HR	GTGACCACCCTTGTAGTATTT
PpOLE 1aR	CTCACTACAGACAAGTATACCCCG
PpOLE 1bR	CTAAGCCGCGACGCTG
PpOLE 2HF	CGCTGGGATTGGTGAC
PpOLE 2HR	GCGCCCTTGTAGTACTTGTA
PpOLE 2aR	CTAAACTGAATCCAAACACGAGA
PpOLE 2bR	CTGCAACACTGGTGTCCGT
PpOLE 3HF	ATCTTTTTGGTTGCGGC
PpOLE 3HR	AGCCAACGGGATGAGCG
MSF	TCCTCTTACGGTGACGAATC
MSR	AGAAGATACACCTGCAGCATG
ICLF	GTGCAGGACATCCAGCG
ICLR	CGTACACCTCCCCACTG
DGAT1aF	CTGCTACAAATGGTGGAAACCTACAGGA
DGAT1aR	GATCAATACACAAATAATGGGC
DGAT1bF	GCTATGGACGAGGGCG
DGAT1bR	AATAGCTGCGTTAAACTTTCTAC
DGAT1cF	GCCTTCAATGAAAACCAGC
DGAT1cR	CTTGTGGTTTGGATCTGGC
DGAT1dF	CGGGTCGCCTCTAGTTCC
DGAT1dR	CGGGTCGCCTCTAGTTCC
DGAT2aF	ACCGTGGAAAGGAAAGACA
DGAT2aR	ATACGTTCTGACCATTCGC
DGAT2bF	CATGACCAGAGATAGCAGAGTG
DGAT2bR	CAGATGAAGCTTGCAGTCG

SEATF	CAAGTGGATGGTGCGGC
SEATR	TGGGTGATCTCAGTTCAGGTAC
ACT2F	GCGAAGAGCGAGTATGACGAG
ACT2R	AGCCACGAATCTAACTTGTGATG
For transient expression DNA constructs of *OLE1* and *OLE2* with *BamH*I end sites	
OLE1BamHIF	ATCGGGATCCATGGATAATGCCAAAACC (for GFP and RFP)
OLE1BamHIR	ATCGGGATCCCAGACAAGTATACCCCGAAG (for RFP)
OLE1BamHIrfpR	ATCGGGATCCAGACAAGTATACCCCGAAG (for RFP)
OLE2BamHIF	ATCGGGATCCATGGCCGACAGGGTCA (for GFP)
OLE2BamHIR	ATCGGGATCCTATCATGGAAGAGAAAATTCTT (for GFP)

CHAPTER TWO

Tandem oleosin genes in a cluster in Brassicaceae confer additive benefit of pollen vigor

ABSTRACT

During evolution, genomes expanded via whole-genome, segmental, tandem and individual-gene duplications, and the emerged redundant paralogs would be eliminated, or retained due to selective neutrality or adaptive benefit and further functional divergence. Here, I show that tandem paralogs can contribute adaptive quantitative benefit and thus have been retained in a lineage-specific manner. In Brassicaceae, a tandem oleosin gene cluster of 5-9 paralogs encodes ample tapetum-specific oleosins located in abundant organelles called tapetosomes in flower anthers. Tapetosomes coordinate the storage of lipids and flavonoids and their transport to the adjacent maturing pollen as the coat to serve various functions. T-DNA and siRNA mutants of *Arabidopsis thaliana* with knockout and knockdown of different tandem oleosin paralogs had quantitative and correlated loss of organized structures of the tapetosomes, pollen-coat materials and pollen tolerance to dehydration. Complementation with the knockout paralog restored the losses.

INTRODUCTION

In evolution, genomes expanded via, whole-genome, segmental, tandem and individual-gene duplications (Birchler, 2012; Cannon et al., 2004; Ding et al., 2012). The mechanisms of these duplications have been explored. The newly emerged paralogs would be eliminated over time, unless they conferred adaptive neutrality or benefit to the organism. Further substitutions in the retained paralogs have resulted in refinement and divergence of functions. The refinement could give rise to gene expression in restricted cells or tissues and under particular physiological conditions and environmental cues. The divergence could include new structural properties, binding specificity or enzymic reactions of the encoded proteins. The retained paralogs in modern genomes appear as numerous gene families, each consisting of a few to hundreds of members. Members of each family may be present as individual paralogs dispersed throughout the whole genome, in linkage with neighboring unrelated genes of the original duplicated genome or segment, and/or as tandem repeats, depending on the original mechanisms of duplication and the subsequent retention, deletion and rearrangement.

Tandem paralogs are ubiquitous in sequenced genomes. Paralogs that have been studied intensively include genes encoding major histocompatibility complex

(Piertney and Oliver, 2005) and homeobox (Duboule, 2007) in animals, as well as lectins (Jiang et al., 2010) and other stress-resistant proteins (Leister, 2004; Hanada et al., 2008) in plants. The functional study of tandem paralogs has focused on their presence in diverse organisms and their differential expression in specific tissues and under various physiological conditions and environmental cues. The findings have led to the delineation of the mechanism of gene duplication, deletion and retention due to adaptive neutrality or benefit, and additional functional divergence. Studies of paralogs in tandem, unlike those dispersed throughout the genome, have inherent technical difficulties because of the lack of knockout mutants of more than one of the tightly linked paralogs. As a consequence, studies involving loss-of-function and even gain-of-function approaches are difficult. In addition, a seldom-addressed but important conceptual issue is whether tandem paralogs offer quantitative adaptive benefit in boosting gene expression, such that they have been retained and expanded in number.

A tandem oleosin gene cluster in *Arabidopsis thaliana* has been studied in terms of evolution and function. *A. thaliana* has 17 oleosin genes, 5 expressing in seed, 3 in both seed and pollen, and 9 in the tapetum of flower anthers (Kim et al., 2002; Fiebig et al., 2004; Schein et al., 2004). The oleosin paralogs specifically

expressed in the tapetum are termed T-oleosin genes. They were termed genes encoding oleosin-like proteins because the proteins possessed oleosin sequence characteristics (next paragraph) (Schein et al., 2004), or glycine-rich-protein genes because 1 or 2 of the paralogs encode oleosins with short glycine-rich repeats at the non-conserved C termini (Fiebig et al., 2004; Mayfield et al., 2001). These earlier terms were used before the function of the oleosins in anthers was delineated (Hsieh and Huang, 2005, 2007). All of the 9 T-oleosin paralogs locate on chromosome 5, with 8 tightly linked in a 30-kb locus, and the remaining one is 23 mb downstream of the cluster.

Comparisons of the T-oleosin paralogs in species of Brassicaceae show that the paralogs evolved more rapidly than did their neighbor genes, as is expected for reproductive genes (Fiebig et al., 2004; Schein et al., 2004). This rapid evolution applies to the sequences encoding the 2 less-conserved N- and C-terminal motifs of oleosin but not to the sequence encoding the essential central hairpin. Each of the T-oleosin paralogs contains 2 exons, and the location of the exons/intron differs from those of the seed- and seed/pollen-expressed oleosin paralogs (Kim et al., 2002). The T-oleosin gene cluster has been found in several genera of Brassicaceae (Fiebig et al.,

2004; Schein et al., 2004); whether it is present outside of Brassicaceae has not been determined.

Oleosins are structural proteins on LDs in seeds, pollen and tapeta (Huang, 1992; Hsieh and Huang, 2004; Frandsen et al., 2001). The protein has a hallmark of ~72 uninterrupted nonpolar residues that form a hairpin penetrating into the LDs. This hairpin, together with the adjacent amphipathic N- and C-terminal motifs, stabilizes the hydrophobic LD in the cytoplasm. In the tapetum of Brassicaceae, T-oleosins are components of the abundant organelles called the tapetosomes (Hsieh and Huang, 2007, Wu et al., 1997; Ting et al., 1998). Each tapetosome contains numerous oleosin-coated alkane LDs associated ionically with many flavonoid-containing, endoplasmic reticulum-derived vesicles (Hsieh and Huang, 2004). In the tapetum cells, the tapetosomes temporarily store alkanes and flavonoids, which will be discharged to the adjacent maturing pollen grains, forming the pollen coat. The coat waterproofs the pollen grain and protects it against UV radiation. Mutational loss of a major T-oleosin in Arabidopsis leads to pollen having a delay in hydration on the stigma (Mayfield and Preuss, 2000). Similarly, mutational loss of pollen-coat flavonoids in Arabidopsis results in pollen being more susceptible to UV radiation (Hsieh and Huang, 2007).

The tandem T-oleosin gene cluster produces abundant oleosins in the tapetum. I investigated the clustered paralogs in conferring functional neutrality, adaptive benefit and/or quantitative benefit.

RESULTS

All paralogs of the *A. thaliana* tandem T-oleosin gene cluster expressed

The arrangement of the 9 T-oleosin paralogs in chromosome 5 is shown in Fig.2-1 (http://www.*arabidopsis*.org/ in 2013), which includes refinements from those reported in 2002 (Kim et al., 2002) and 2004 (Fiebig et al., 2004; Schein et al., 2004). Each of the 9 T-oleosin paralogs produced transcripts, as revealed by RT-PCR (Kim et al., 2002). The transcripts were present only in the tapetum at a mid-stage of anther development before the tapetum underwent programmed cell death (Kim et al., 2002). This developmental specificity was supported by microarray results (Fig.2-2, from http://www.weigelworld.org/).

The transcript levels of the 9 T-oleosin paralogs varied. According to http://www.weigelworld.org/ data for wild-type Columbia, the order of microarray transcript intensity was *T3* ≈ *T5* > *T4* > *T6* > other paralogs (Fig. 2-2), which is an approximation of quantity. In http://mpss.udel.edu/at/, the MPSS transcript-per-million (TPM) values in inflorescence were 17,006, 13,644, 1738, 300, 29, and 0 for *T5*, *T3*, *T4*, *T7*, *T8* and all other paralogs, respectively (*T6* and *T2* had 0 values because they lacked the restriction site for this MPSS analysis). The SBS transcripts-per-35-million values were 86,846, 14,073, 10,523, 7,078, 5,203, 1,180,

601, 257 and 99 for *T3*, *T5*, *T4*, *T1*, *T6*, *T2*, *T7*, *T8* and *T9*, respectively. *T9* transcript was not detected in my RT-PCR analysis (Fig. 2-3). Overall, the reported order of levels of transcripts was *T3* ≥ *T5* ≈ *T4* ≈ *T6* > other paralogs.

The relative level of oleosin proteins, as visually determined by SDS-PAGE and Coomassie blue staining, was T3 oleosin ≥ other individual oleosins in tapetosomes from wild-type Columbia (Kim et al., 2002) and on the pollen surface of wild-type Ws2 (Mayfield et al., 2001; Mayfield and Preuss, 2000). Similar findings of T3-ortholog oleosin being more abundant than other individual T-oleosins in tapetosomes and pollen coat were obtained in *Brassica napus* (Hsieh and Huang, 2005, Wu et al., 1997) and *Brassica rapa* (Ting et al., 1998). The higher amount of T3 oleosin determined visually by SDS-PAGE could be due to its high molecular mass (53 kDa) as compared with T5 (10 kDa), T4 (22 kDa) and T6 (15 kDa) oleosins, as well as an 8-kD portion of these T-oleosins being the conserved nonpolar hairpin motif that was less reactive to Coomassie blue. The overall findings suggest that the relative numbers of T-oleosin molecules are in the order of T3 ≥ T5 ≈ T4 ≈ T6 > others.

Arabidopsis mutants of various T-oleosin paralogs had similar phenotypes of reduced levels of transcripts, tapetosome structures, pollen coat and pollen dehydration tolerance

My observations did not suggest that individual paralogs of the T-oleosin gene cluster contributed to functional divergence. Thus, I tested whether the individual paralogs conferred quantitative benefits rather than functional neutrality. I focused on the 3 most active T-oleosin paralogs, *T5*, *T6* and *T3*, whose transcripts together accounted for ~90% of all transcripts of T-oleosin paralogs.

With a wild-type Columbia background, a T-DNA mutant, *ΔT5*, had a knockout of *T5* transcript (see Methods and Fig. 2-3). In *ΔT5*, the tapetosomes observed with transmission electron microscopy (TEM) were of similar sizes but had numerous electron-transparent structures as compared with those in the wild type (Fig. 2-4A). I used siRNA to knock down *T1*, *T3*, *T6* and/or *T4* in *ΔT5* (Fig. 2-3). Among the 50 transformants examined, one (*si-9/ΔT5*) had no apparent tapetosomes in the tapetum cell, which instead contained several aggregates of electron transparent vesicles (Fig. 2-4A). Another transformant, *si-5/ΔT5*, possessed obscure structures that resembled fragmented tapetosomes (Fig. 2-4A). A different T-DNA mutant, *ΔT6*, had knockout of

T6 oleosin transcript (see Methods) and defective tapetosomes similar to those of *ΔT5* (Fig. 2-4A).

With a wild-type Ws2 background, a T-DNA mutant, *ΔT3*, had knockout of *T3* transcript (Mayfield and Preuss, 2000). In *ΔT3*, the tapetosomes were smaller and did not contain numerous internal electron-dense droplets, as compared with those in the wild type (Fig. 2-4A). The tapetum cells and the pollen surface were viewed with confocal laser scanning microscopy (CLSM) (Fig. 2-4B). The tapetosomes in the wild type contained both LDs (stained with Nile Red) and flavonoids (stained with DPBA), as previously reported (Hsieh and Huang, 2007). In *ΔT3*, the tapetosomes were small (containing both LDs and flavonoids), and some LDs were outside of the tapetosomes. The coat of wild-type pollen contained T3 oleosin and flavonoids, whereas that of *ΔT3* pollen had no T3 oleosin and reduced flavonoids. Although *ΔT3* and the wild type had similar flavonoid contents in the tapetum cells (Fig. 2-4B), the *ΔT3* pollen coat possessed reduced flavonoid content, and so lesser tapetum flavonoids were transferred to the pollen surface.

The knockout oleosin mutant *ΔT3* and the lesser-studied *ΔT5* and *ΔT6* mutants had defective organization of tapetosomes in the tapetum and reduced levels of oleosins and flavonoids in the pollen coat. I tested the mutant pollen for dehydration

tolerance resulting from the loss of coat materials. I placed the pollen grains at 25% relative humidity and 24 ^{0}C for 1-3 h and allowed them to germinate on agar medium (Fig. 2-5). Pollen of either of the 2 wild types was similar in percentage germination before and after the dehydration treatment, but pollen of the various T-oleosin mutants had significantly reduced percentage germination (Fig. 2-5). Among the various mutants, there was a correlation in the levels of loss of T-oleosin transcripts (Fig. 2-3), tapetosome structures and pollen hydration tolerance (Table 2-1).

Losses of tapetosome structures, pollen coat and pollen dehydration tolerance in *ΔT3* were restored after transformation with the *T3* gene

In this test for functional complementation, the transformants examined largely recovered from the losses of tapetosome structures (Fig. 2-4), pollen coat T3 oleosin and flavonoids (Fig. 2-4B) as well as pollen tolerance to dehydration (Fig. 2-5; Table 2-1). Therefore, the tapetum and pollen phenotypes of *ΔT3* resulted from loss of the *T3* gene.

DISCUSSION

In evolution, newly emerged paralogs via diverse mechanisms of gene duplication would be eliminated, unless they conferred adaptive neutrality or benefit to the organism. I reveal that paralogs of a tandem oleosin gene cluster in Brassicaceae exert no apparent functional divergence but confer quantitative benefit.

In my studies, I overcame the traditional technical difficulties of lacking knockout mutants of tightly linked paralogs in a tandem cluster. I used T-DNA knockout mutants for each of the 3 most active oleosin paralogs plus siRNA to knock down additional paralogs. The loss of any one of the 3 paralogs resulted in a similar defective phenotype, and the loss of more than one paralog generated more severe defective phenotypes. In Arabidopsis, the tapetum oleosin transcripts together represent 3-4% of the total inflorescence transcripts (MPSS data; see Results), and the oleosin proteins could represent a similar percentage of the total proteins (see Results). This high percentage of oleosin transcripts and proteins were obtained from findings with whole inflorescences, which included the tapetum, microspores and numerous other cell types (Goldberg et al., 1993). The percentage could be as high as 10% in the tapetum cells, the site to which the T-oleosin transcripts were restricted. The abundant oleosins produced from several tandem paralogs additively maintain

the predominant tapetosomes, which confer adaptive benefit to the organism. In the current study, I tested the phenotype of pollen dehydration tolerance. Other adaptive benefits conferred by the tandem oleosin paralogs are possible. Earlier, it was shown that pollen without T3 oleosin from Arabidopsis *ΔT3* exhibits a delay in hydration on the stigma (Mayfield and Preuss., 2000).

MATERIALS AND METHODS

Plant Materials

Seeds of *A. thaliana* wild type (Col-0 and Ws-2) and T-DNA insertion mutants of T-oleosin genes *ΔT3* (CS11664) (20), *ΔT5* (CS104700) and *ΔT6* (SALK_134093) were obtained from the Arabidopsis Biological Research Center (Ohio State U.). *ΔT5* has the T-DNA inserted in the intron adjacent to the 3'-end of the second exon. *ΔT6* has the T-DNA inserted into the second exon. Both *ΔT5* and *ΔT6* do not produce transcripts. Plant*s* were grown at 22 °C in growth rooms with 8-h/16-h day/night cycle.

Transformation of *A. thaliana*

The constructs, pC1300/T3 (for *T3* complementation test) and pHG/TOLE (for T-oleosin RNA silencing) (Fig. 2-6), were transformed into Arabidopsis *ΔT3* and *ΔT5*, respectively, via floral dipping (Clough and Bent, 1998). The transformants were selected on 1/2 MS agar medium with hygromycin (30 mg. l^{-1}) and identified by content of the *hyg* fragment with PCR.

Pollen Germination Test

Pollen of Arabidopsis were collected from open flowers and placed at 25% relative humidity and 24 ^{0}C. Arabidopsis pollen grains were placed on 9-cm Petri dishes with agar medium (0.5% agar, 18% sucrose, 0.01% boric acid, 1 mM $MgSO_4$, 1 mM $CaCl_2$ and 1 mM $Ca(NO_3)_2$, pH 7) (Li et al., 1999) at 28 °C for 1-3 h. Pollen grains were observed for germination by microscopy. At least 500 pollen grains of each line were counted.

Electron Microscopy

Anthers were fixed via a high-pressure freezing or chemical fixation procedure. For the former, anthers were fixed in a high pressure freezer (Leica EM PACT2). Fixed materials were subjected to freeze substitution in ethanol (containing 0.2% glutaraldehyde and 0.1% uranyl acetate) in Lecia Automatic Freeze-Substitution System and embedded in LR Gold resin (Structural Probe, west Chester, PA). For chemical fixation, anthers were fixed with 2.5% glutaraldehyde, 4% paraformaldehyde and 0.1 M K-phosphate (pH 7.0) at 4^{0} C for 24 h. The materials were washed with 0.1 M K-phosphate buffer (pH 7.0) for 10 min twice and then treated with 1% OsO_4 in 0.1 M K-phosphate (pH 7.0) at 24 ^{0}C for 4 h. The fixed

materials were rinsed with 0.1 M K-phosphate buffer (pH 7.0), dehydrated through an acetone series and embedded in Spurr resin. Ultrathin sections (70 to 90 nm) were stained with uranyl acetate and lead citrate and examined with a Philips CM 100 TEM at 80 KV.

Immunofluorescence and Chemical Staining with CLSM

All antibody treatments were performed with 1:50 dilution of the IgG fraction, 1% (w/v) milk powder and 1x PBS. Unless otherwise stated, each wash was performed with PBST (1x PBS and 0.05% [w/w] Tween-20) for 10 min. Tissues were fixed in 4% paraformaldehyde, 1x PBS and 0.15 M sucrose at 4° C for 16 h. For flavonoid localization with diphenylboric acid 2-aminoethyl ester (DPBA), fixed anther sections and mature pollen were washed twice and treated with 1 M HCl at 70 °C (14). Samples were washed with 10% glycerol twice and stained with DPBA (saturated DPBA [<0.5%, w/v], 0.01% Triton X-100 and 10% glycerol) at 20 °C for 2 were then washed twice with 10% glycerol. For double labeling with DPBA for flavonoids and Nile Red for lipids, the samples after the DPBA staining were treated with Nile Red. For double labeling with DPBA for flavonoids and antibodies against the vacuole marker protein V-PPase, the samples after the DPBA staining were

blocked with 3% milk in 1x PBS at 20 °C for 1 h. They were treated with rabbit IgG against V-PPase (Agrisera, Sweden) in 1% milk and 1x PBS at 20 °C for 2 h. After 2 washes, samples were treated with a mixture of DPBA and secondary antibodies (0.5% DPBA, 0.01% Triton X-100, 10% glycerol, 1x PBS, and 1:100 cyanine 5–conjugated goat IgG against rabbit IgG) at 20 °C for 2 h. Samples were washed twice. The labeled samples were observed with a Leica SP2 confocal microscope. DPBA-flavonoid, Nile Red, and Cyanine 5 were excited with 488 nm, 543 nm and 633 nm; and the emissions were detected at 500 to 530 nm, 565 to 615 nm and 650 to 750 nm, respectively.

RT-PCR of T-oleosin transcripts

RNA was extracted from Arabidopsis inflorescences with use of the Illustra RNAspin *Mini* Kit (GE Health Care, Piscataway, NJ). The RNA (1 µg) was used for reverse-transcription of cDNA with the SuperScriptIII RT-PCR system (Invitrogen). PCR had 0.2 µl of the cDNA as template and DyNazyme DNA polymerase with dNTP and primer sets (Table 2-2).

Construction of plasmids p1300/T3 and pHG/TOLE.

T3 oleosin gene (including promoter and terminator) was amplified with the primers T3proFKpnI and T3terRSalI and inserted into pCAMBIA1300 (CAMBIA) via the KpnI and SalI sites to generate p1300/T3. The plasmids were transformed into, Arabidopsis *ΔT3* to generate *T3/ΔT3* in a complementary test. The details showed in Fig.2-6A. The sequence of the attB/siTOLE fragment was designed to generate siRNA against the conserved regions of *T1* (high similarity to *T4*), *T3* and *T6*. Primers with attB1 and attB2 sequences and primers a-f (representing T1FattB1, T1RT6, T6F, T6RT3, T3F, and T3RattB2, respectively) were used for amplification and ligation by PCR to generate the attB/siTOLE fragment. This fragment was recombined with pHELLSGATE (Wesley et al., 2001) via BP clonase recombination reactions according to the manufacturer's instructions (Invitrogen). Then, the *CaMV35S* promoter was substituted with the *T5* promoter (amplified by primers T5proFSacI and T5proRXhoI) via SacI and XhoI sites. Finally, the *T5* promoter with the hairpin region cassette was ligated to pCAMBIA1302 via the SacI and PmlI sites to generate a hygromycin resistance cassette (pHG/TOLE). The details showed in Fig.2-6B.

REFERENCES

Birchler JA (2012) Insights from paleogenomic and population studies into the consequences of dosage sensitive gene expression in plants. Curr Opin Plant Biol **15:** 544-548

Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biol **4:**10

Clough SJ and Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis* thaliana. Plant J **16**: 735

Ding Y, Zhou Q, Wang W (2012) Origins of new genes and evolution of their novel functions. Annu Rev of Ecol Evol and Syst **43:** 345-363

Duboule D (2007) The rise and fall of Hox gene clusters. Development **134:** 2549-2560

Fiebig A, Kimport R, Preuss D (2004) Comparisons of pollen coat genes across Brassicaceae species reveal rapid evolution by repeat expansion and diversification. Proc Natl Acad Sci USA **101:** 3286-3291

Goldberg RB, Beals TP, and Sanders PM (1993) Anther development: basic principles and practical applications. Plant Cell **5:** 1217-1229

Frandsen GI, Mundy J, Tzen JT (2001) Oil bodies and their associated proteins, oleosin and caleosin. Physiol Plant **112:** 301-307

Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu SH (2008) Importance of lineage-specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant Physiol **148:** 993-1003

Hsieh K, Huang AHC (2004) Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. Plant Physiol **136:** 3427-3434

Hsieh K, Huang AHC (2005) Lipid-rich tapetosomes in *Brassica* tapetum are composed of oleosin-coated oil droplets and vesicles, both assembled in and then detached from the endoplasmic reticulum. Plant J **43:** 889-899

Hsieh K, Huang AHC (2007) Tapetosomes in Brassica tapetum accumulate endoplasmic reticulum–derived flavonoids and alkanes for delivery to the pollen surface. Plant Cell **19:**582-596

Huang AHC (1992) Oil bodies and oleosins in seeds. Annu Rev Plant Physiol Plant Mol Biol **43:** 177-200

Jiang SY, Ma Z, Ramachandran S (2010) Evolutionary history and stress regulation of the lectin superfamily in higher plants. BMC Evol Biol **10:** 79

Kim HU, Hsieh K, Ratnayake C, Huang AHC (2002). A novel group of oleosins is present inside the pollen of *Arabidopsis*. J Biol Chem **277:** 22677-22684

Leister D (2004) Tandem and segmental gene duplication and recombination in the evolution of plant diseases resistance genes. Trends Genet **20:** 116-122

Li H, Lin YK, Heath RM, Zhu MX, Yang ZB (1999) Control of pollen tube tip growth by a Rop GTPase-dependent pathway that leads to the tip-localized calcium influx. Plant Cell 11: 1731–1742

Mayfield JA, Preuss D (2000) Rapid initiation of *Arabidopsis* pollination requires the oleosin-domain protein GRP17. Nat Cell Biol **2:** 128-130.

Mayfield JA, Fiebig A, Johnstone SE, Preuss D (2001) Gene families from the *Arabidopsis* thaliana pollen coat proteome. Science **292:** 2482-2485

Piertney SB, Oliver MK (2005) The evolutionary ecology of the major histocompatibility complex. Heredity **96:** 7-21

Schein M, Yang Z, Mitchell-Olds T, Schmid KJ (2004) Rapid evolution of a pollen-specific oleosin-like gene family from *Arabidopsis thaliana* and closely related species. Mol Biol Evol **21:** 659-669

Ting JTL, Wu SSH, Ratnayake C, Huang AHC (1998) Constituents of the tapetosomes and elaioplasts in *Brassica campestris* and their degradation and retention during microsporogenesis. Plant J **16:** 541-551

Wesley SV, Helliwell CA, Smith NA, Wang MB, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA, Robinson SP, Gleave AP, Green AG and Waterhouse, PM (2001) Constructs for efficient, effective and high throughput gene silencing in plants. Plant J **27:** 581–590

Wu SSH, Platt KA, Ratnayake C, Wang TW, Ting JTL, Huang AHC (1997) Isolation and characterization of novel neutral-lipid-containing organelles and globuli-filled plastids from *Brassica napus* tapetum. Proc Natl Acad Sci USA **94:** 12711-12716

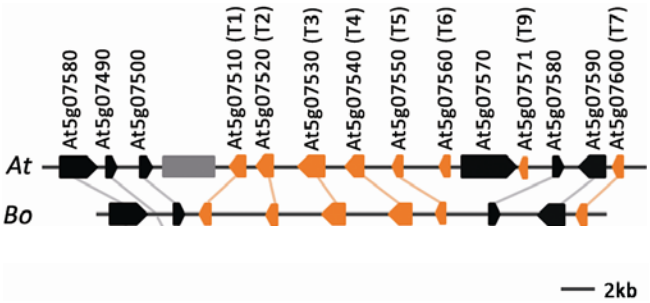

Fig. 2-1. T-oleosin gene clusters in Brassicaceae

Primary transcripts are indicated with filled arrows (orange for oleosin genes and black for non-oleosin genes) in genomic DNA sequences from *Arabidopsis thaliana* (At) and *Brassica oleracea* (Bo) of Brassicaceae. *At* has the T8 oleosin (At5g61610) gene 23 mb downstream of T7. Gray filled box indicates transposon.

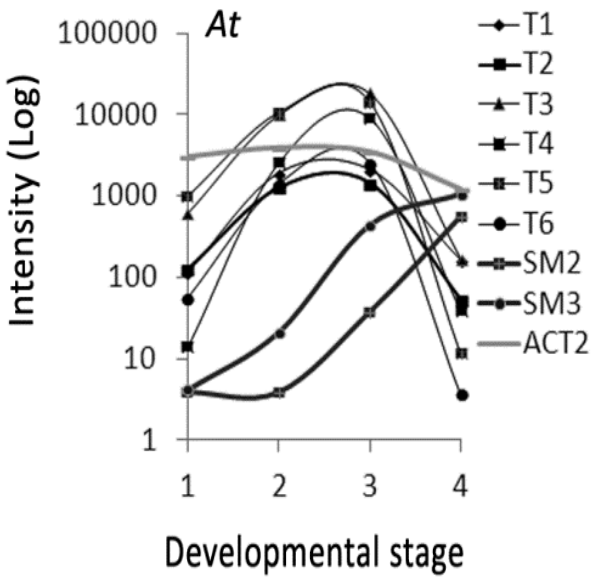

Fig. 2-2. Levels of oleosin transcripts in inflorescences of Arabidopsis

Stages 1-3 represent inflorescences or anthers containing microspores of progressive morphologic features (tetrad, early vacuolated and late vacuolated stages, respectively), and stage 4 denotes mature pollen. The first 3 stages were of whole inflorescences of designated flower stages 9, 10/11, and 11 (AtGenExpress Visualization Tool AVT, http://jsp.weigelworld.org/expviz/expviz.jsp). Values of T1-6 (tapetum-specific oleosin) and SM2-3 (seed- and microspore-specific oleosin) transcripts obtained from AtGenExpress are expressed in log microarray intensity. T7-9 transcripts were not represented or detected in the microarrays.

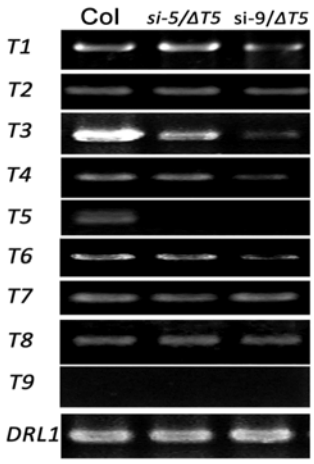

Fig. 2-3. RT-PCR analysis of mRNA levels of the 9 T-oleosin genes in Col WT and T-oleosin silencing mutants

T-oleosins were silenced by siRNA. The si9 and si5 were designed to knockdown T-1, T-3, T-4 and T-6. A tapetum-specific transcript (*DRL*1 of At4g34520) was used as a reference. RNA was extracted from inflorescences.

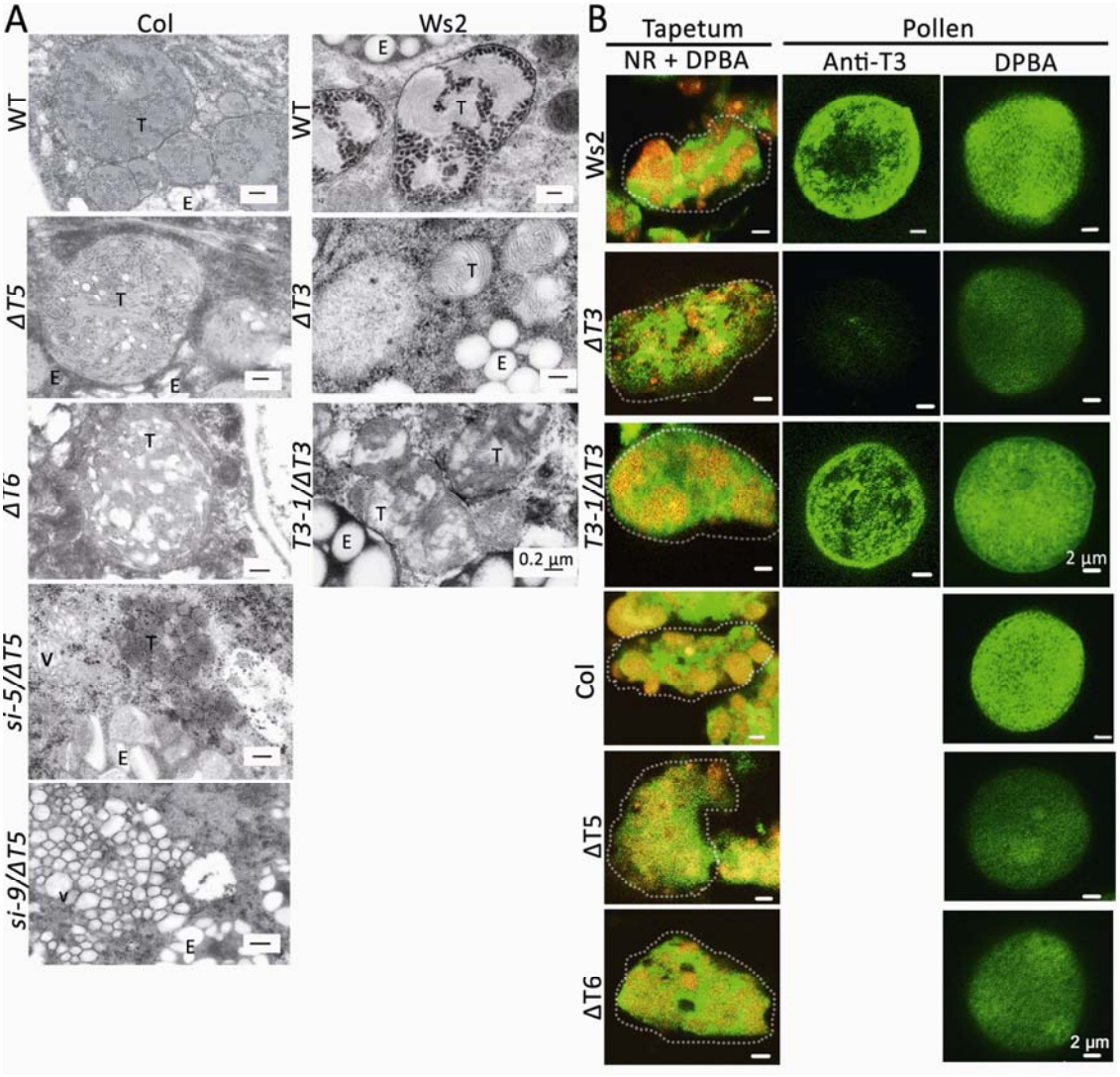

Fig. 2-4. Effect of mutational loss of T-oleosins on morphological features of tapetosomes and pollen dehydration tolerance in Arabidopsis.

A. TEM of tapetosomes in WT (Col and Ws2), T-DNA mutants (*ΔT5, ΔT6* and *ΔT3*), T-oleosin silenced mutant (*si5/ΔT5* and *si9/ΔT5*) and *T3* complementation mutant (*T3/ΔT3*). Bars represent 0.2 µm. T, E and V indicate tapetosome, elaioplast and aggregated vesicles, respectively.

B. CLSM of tapetum cells and pollen of WT (Col and Ws2), T-DNA mutants (*ΔT5, ΔT6* and *ΔT3*). Samples were detected for T3 oleosin (with antibodies, pseudo-color in green), lipids (with Nile Red [NR], in red ; many scattered LDs [panel A] were too small to be revealed by CLSM) and flavonoids (with DPBA, in green). The same settings (laser power and detection gain) were used for direct comparison with the WT. White dotted line reveals the circumference of the cell. Bars represent 2 µm.

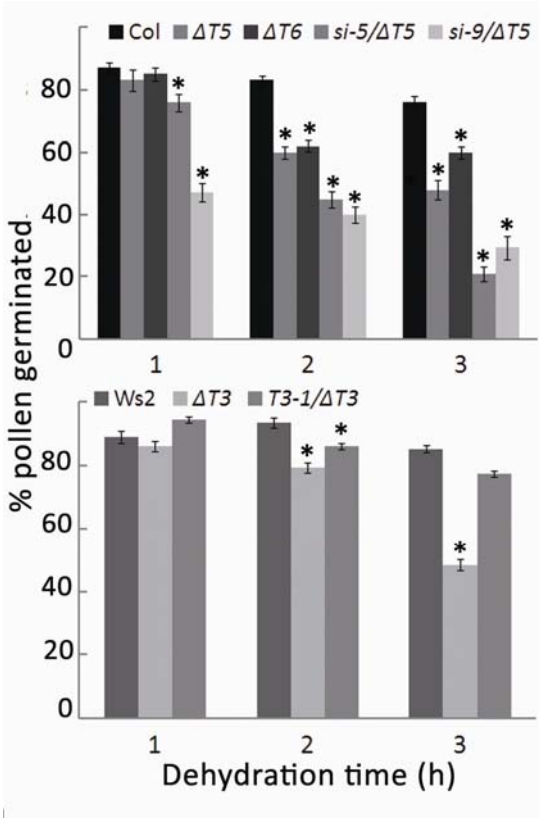

Fig. 2-5. Effect of mutational loss of T-oleosins on pollen dehydration tolerance in Arabidopsis.

Test of dehydration tolerance of pollen from WT and T-oleosin mutants. Pollen was placed at 25% relative humidity at 24 ^{0}C for 1-3 h and allowed to germinate. Significant difference was caculated from the germination % of each mutant compared to WT (t-test, *P$\leqq$0.0001).

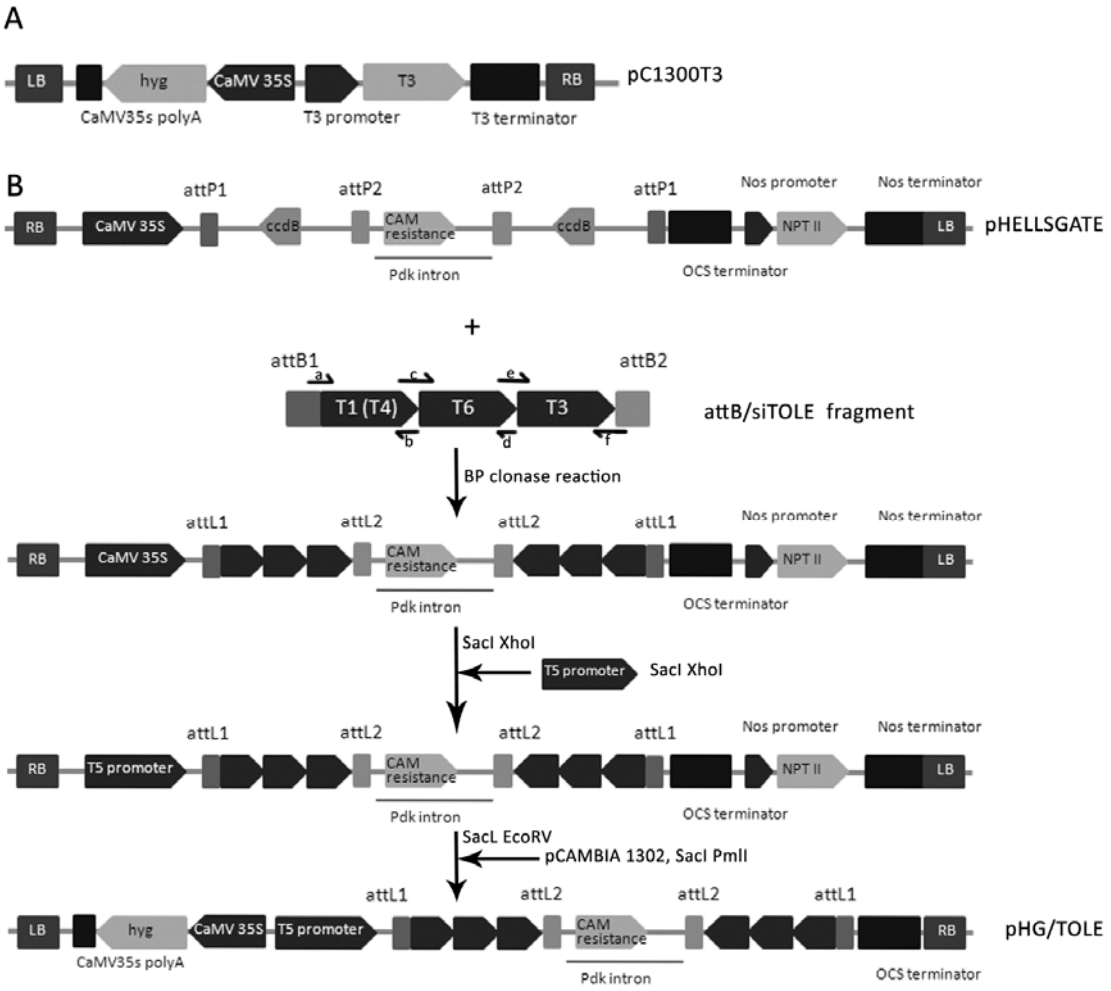

Fig. 2-6. Construction of plasmids p1300/T3 and pHG/TOLE.

A. *T3* oleosin gene (including promoter and terminator) was subcloned into pCAMBIA1300 (CAMBIA) to generate p1300/T3. The plasmids were transformed into, Arabidopsis *ΔT3* to generate *T3/ΔT3* in a complementary test.

B. The sequence of the attB/siTOLE fragment was designed to generate siRNA against the conserved regions of *T1* (high similarity to *T4*), *T3* and *T6*.

Table 2-1. Additive phenotypes of tapetosomes and pollen germination in T-oleosin mutants in Arabidopsis

	Presence of tapetosome (Yes/No)	Phenotype of tapetosome	Pollen germination (%)**
WT (Col)	Y	>2 μm	76±2.0[a]
ΔT6	Y	>2 μm; with clear holes	60±2.5[a]*
ΔT5	Y	>2 μm; with clear holes	48±2.2[a]*
Si9/ΔT5	N	large aggregates of vesicles	29±3.0[a]*
Si5/ΔT5	Y	<0.2 μm fragmented tapetosomes	21±2.2[a]*
WT (Ws2)	Y	1~2 μm	85±1.4[b]
ΔT3	Y	<0.5 μm	48±1.9[b]*
T3-1/ΔT3	Y	0.8~1 μm	78±1.6[b]*

**Germination after dehydration for 3 h Details are in Figs. 2-4.

* Significant difference from the germination % of WT (t-test, $P < 0.0001$). [a,b]For Arabidopsis, for each set of experiments, mutants are compared with the WT, Col or Ws2.

Table 2-2. Sequences of primers used in chapter two

primer	sequence
T1rtF	5'-TTTGGGGGACTTAGGAGATTT-3'
T1rtR	5'-AATTGTAGTCTTGTGTTTCTTTAT-3'
T2rtF	5'-ATGGGAGCTAAGCCGACTGC-3'
T2rtR	5'-TTACCATGTGGGTGGTGTACT-3'
T3rtF	5’-GTATGTCCGGAGGTGGTATG-3’
T3rtR	5'-TTACTCCGCCATGTGGCCAC-3'
T4rtF	5'-AGGATTGGAGTAAAACCGAAG-3'
T4rtR	5'-CTCCATATGATCATCCTTACCA-3'
T5rtF	5'-CTACCAAAAATTCCAGGATTGA-3’
T5rtR	5'-CAATCTCAATGATATTACTTCATT-3’
T6rtF	5’-ACTTTCCAATTAAAATTCCAGGT-3’
T6rtR	5’-GCACGGCATACATCAATAGAA-3’
T7rtF	5'-GTCGAAGAACAATATACCTGCA-3'
T7rtR	5'-CTTTTTCTTTGAGCCGGCCTT-3'
T8rtF	5’-GAAAGTGGCAAGCAATCAGGA-3’
T8rtR	5’-GTACCTTTGGTATCTCAGGGA-3’
T9rtF	5’GCCTCCATGC CGCTCTTCCT-3’
T9rtR	5’-GGAAGACGACCTGTTGCAC-3’
T1FattB1	5'-AA AAAGCAGGCTGTCGGAGGTTTTTCAGG-3'
T1RT6	5'-GCCGTCATAACACCCTTGCGGCCATCGTTCCGAG-3'
T6F	5'-CAAGGGTGTTATGACGGC-3'
T6RT3	5'-TAACCTCCAACAGTGGCAGACCCGACACACCAG-3'
T3F	5'-TGCCACTGTTGGAGGTTA-3'
T3RattB2	5'-A GAAAGCTGGGTGTGCTGTCGCTCCAAAA-3'
T5proFSacI	5’-ATGCGAGCTCGATTTGACGCGAGTGCAT-3’
T5proRXhoI	5’-ATCG CTCGAGCTTTGGGATGATCAAGTTAAAC-3’
T3ProFKpnI	5’-GGGGTACCGGATGTTTGTATTAGTTTCTCAC-3’
T3terRSalI	5’-ATCGTCGACGGTCCAAGCTATTTCTCAAC-3’

CHAPTER THREE

The evolutionary change of oil bodies to tapetosomes occurred after the divergence between Brassicaceae and Cleomaceae

ABSTRACT

Cleomaceae is the family closest to Brassicaceae. *Cleome* species did not contain the tandem oleosin gene cluster, tapetum oleosin transcripts, tapetosomes or pollen tolerant to dehydration. *Cleome hassleriana* transformed with an Arabidopsis oleosin gene for tapetum expression possessed primitive tapetosomes and pollen tolerant to dehydration. I propose that during early evolution of Brassicaceae, a duplicate oleosin gene mutated from expression in seed to the tapetum. The tapetum oleosin generated primitive tapetosomes that organized stored lipids and flavonoids for their effective transfer to the pollen surface for greater pollen vitality. The resulting adaptive benefit led to retention of tandem-duplicated oleosin genes for production of more oleosin and modern tapetosomes.

INTRODUCTION

Pollen development is a highly coordinated and specialized process that involved both structural and morphological changes and cellular differentiation in the tissues of the anther. The tapetum is a single layer of nutritive cells within the anther locule which controls the development of pollen grains (Polowick and Sawhney, 1993). The tapetum cells store ingredients to be deposited onto the maturing pollen to form pollen coat. In Brassicaceae, the tapetum cells at this mid stage contain two abundant storage organelles, the elaioplasts and the tapetosomes (Platt et al., 1998). Elaioplasts are a type of plastids which is specialized for the storage of lipids in plants.Unlike elaioplasts of similar morphology can be found in nontapetum cells, such as fruit and petal, the tapetosomes are unique to the tapetum cells.

Oleosins presenting in tapetum cells of anthers in Arabidopsis and Brassica were discovered from gene cloning results (de Oliveira et al., 1993; Roberts et al., 1994). In tapetum cells, oleosin-coated OBs are associated with flavonoid-containing vesicles to form tapetosomes of 2-3 μm in diameter. Tapetosomes temporarily store lipids and flavonoids, which are deposited onto the maturing pollen as pollen coat for water-proofing and UV protection, respectively. Arabidopsis possess a T-oleosin gene

cluster which contains 9 tapetum specific expressed T-oleosin (Kim et al., 2002). All the 9 T-oleosin genes are tandem repeats on chromosome 5. This T-oleosin gene cluster is also present in some other plants in Brassicaceae, including *Boechera drummondii, Brassica olererecea, Olimarabidopsis pumila*, *Capsella rubella*, and *Sisymbrium irio* (Fiebig et al., 2004; Schein et al., 2004). The tapetosome is not ubiquitous in higher-plant tapetum cells. For example, maize and rice has no tapetosomes, and its pollen coat has no oleosin. Nevertheless, pollen of all plant species has surface flavonoids and lipid derived from tapeta. Also, tapetum cells in diverse plant species contain small LDs. The evolutionary trend of tapetosomes is intriguing.

In this study, the existence of the T-oleosin gene cluster and tapetosome outside Brassicaceae were investigated. The result revealed the closest family, Cleomaceae (Barker et al., 2009; Judd et al., 2002; Schranz et al., 2006), did not possess the T-oleosin gene cluster, tapetosomes and dehydration-tolerant pollen. I thus examined whether transforming a *Cleome* species with an Arabidopsis T-oleosin gene could generate tapetosomes, which in turn would produce a pollen coat conferring pollen dehydration tolerance.

RESULTS

Diverse plant families, including Cleomaceae, the family closest to Brassicaceae, did not possess the T-oleosin gene cluster, tapetum oleosins and tapetosomes

I surveyed plant families with 4 approaches to obtain complementary information: detection of (A) tapetosomes with TEM, (B) the T-oleosin gene cluster via public genomic information and my own sequencing effort, (C) oleosin transcripts in anthers at the mid stage of development, and (D) transcript-deduced oleosin sequences similar to those of Brassicaceae T-oleosins.

I used TEM to survey diverse plant orders (Fig. 3-1 and 3-2) for the presence of tapetosomes in tapetum cells of anthers of progressive developmental stages. These plant orders ranged from Poales (of monocots) to Brassicales, which includes the Brassicaceae family. Species of the genera *Brassica*, *Capsella*, *Raphanus* and *Cardamine* of Brassicaceae all possessed abundant tapetosomes with varying internal structures (Fig. 3-1). Species in families other than Brassicaceae had no tapetosomes (Fig. 3-2). Instead, LDs and vesicles, presumably containing temporarily stored pollen-coat lipids and flavonoids, respectively, were dispersed in the cytoplasm. Many of the species possessed easily-recognizable elaioplasts housing massive plastoglobules similar to those in Brassicaceae species. The absence of tapetosomes

in species other than those in Brassicaceae should apply to what I have examined. I focused on the families in Brassicales, which include Cleomaceae, the family closest to Brassicaceae in phylogeny, and Caricaceae, for which the genome sequence of *Carica papaya* is available.

The tapetum cells of *Cleome hassleriana*, a readily available ornamental plant of Cleomaceae used in C3/C4 photosynthesis studies (Newell et al., 2010), contained solitary LDs and vesicles dispersed in the cytoplasm (Fig. 3-2 and 3-7A). They also possessed conspicuous plastoglobules-packed elaioplasts similar to those in Brassicaceae. The tapetum cells of *C. papaya* contained no identifiable tapetosomes and elaioplasts (Fig. 3-2).

I attempted to obtain *Cleome* and *Carica* DNA segments syntenic to the Arabidopsis and *Brassica* DNA segments of the T-oleosin gene cluster region. With *Cleome*, I selected several Bacterial Artificial Chromosomes (BACs) from a BAC library with positive PCR identification of possessing genes adjacent to the *Arabidopsis* T-oleosin gene cluster. Sequencing of these *Cleome* BACs revealed that they did not meet my expectations. I then subjected *Cleome* genomic DNA to Illumina sequencing and obtained 2 segments that had the highest similarity with the Brassicaceae T-oleosin gene cluster region (Fig. 3-3). Each of these 2 segments contained 2

orthologs of the same 5'→3' direction of the Arabidopsis genes bordering the tandem T-oleosin gene cluster. No nearby oleosin genes were present within these 2 segments. Assembly of the Illumina reads and performance of PCR with genomic DNA failed to link two other less-similar 2 segments. *Cleome* segments are also shown in Fig. 3-3. Apparently, the 2 fragments were not closely linked. The presence of these segments could reflect the early triplication of the *Cleome* genome. Also, none of the oleosin genes in the *Cleome* genomic sequence has neighboring genes closely related to the neighboring genes of the tandem oleosin gene cluster in Arabidopsis. From the *Carica* genome sequence (http://genome.jgi.psf.org/), I obtained the segment that most resembled the Brassicaceae T-oleosin gene cluster region (Fig. 3-3); the resemblance was low. For both *Cleome* and *Carica*, no apparent tapetum-related T-oleosin gene was found, and the oleosin gene (*Cp*OLE4) in the *Carica* fragment encodes an oleosin most similar to the Arabidopsis seed S3 oleosin in sequence. Thus, the regions syntenic to the Brassicaceae T-oleosin gene cluster region differ greatly among the species from 3 different families in Brassicales. Apparently *Cleome* and *Carica* do not have the T-oleosin gene cluster.

When I began this project in early 2010, no *Cleome* genome sequence was available. An unassembled *Cleome violacea* genome sequence became available in

September 2012 (http://genome.jgi.psf.org/). From it I obtained a continuous DNA segment (Fig. 3-3) with high resemblance of the Arabidopsis T-oleosin cluster region but without an oleosin gene. This finding reiterates that Cleomaceae lacks the T-oleosin gene cluster.

I explored whether T-oleosin transcripts corresponding to the Arabidopsis T-oleosin transcripts of a defined temporal development pattern were present in anthers of *Cleome* and *Carica*. In *Arabidopsis* developing anthers, which contain tapetum cells and maturing pollen, the levels of tapetum T-oleosin transcripts peaked before the appearance of pollen oleosin transcripts (Kim et al., 2002; Fig. 3-4). This differential developmental pattern occurred because the tapetum matures and undergoes programmed cell death before the pollen matures (Goldberg et al., 1993). I generated transcriptomes of *Cleome* and *Carica* anthers of progressive developmental stages and found only putative pollen oleosin transcripts at a late stage of anther development (which represented mature pollen) and no oleosin transcripts at earlier stages when pollen was immature (Fig. 3-4). Rice, with available anther transcriptomes (Huang et al., 2009), was used for further comparison. Rice also had pollen oleosin transcripts appearing at a late stage of anther development

but no oleosin transcript at an earlier stage. Rice and maize do not have tapetum tapetosomes (Huang et al., 2009; Li et al., 2012).

Finally, I constructed a phylogenetic tree based on oleosin amino acid sequences (Fig. 3-5) and examined whether *Cleome* and *Carica*, as well as rice, possessed oleosins similar to those of Brassicaceae T-oleosins (Fig. 3-6). In the tree, Arabidopsis T-oleosins forms a clade distinct from those of its seed and pollen. None of the oleosins deduced from transcripts of *Cleome*, *Carica* and rice falls into the clade of Arabidopsis T-oleosins. Apparently, these other species do not have T-oleosins.

Cleome hassleriana transformed with an Arabidopsis T-oleosin gene produced primitive tapetosomes and dehydration-tolerant pollen.

I tested whether *Cleome* transformed with the Arabidopsis *T3* oleosin gene would generate the downstream gene activities and have adaptive benefit. In non-transformed *Cleome*, LDs and flavonoid-containing vesicles dispersed in the cytoplasm of tapetum cells (Fig. 3-7A). CLSM revealed that the LDs, (stained with Nile Red), vacuoles (detected with V-PPase antibodies) and vesicles (stained with DPBA for flavonoids; individual vesicles in the cytoplasm being too small to be observed by

CLSM) did not co-locate in the cytoplasm (Fig. 3-7B). This arrangement differed from that in Brassicaceae, in which the LDs and flavonoid-containing vesicles co-located in tapetosomes (Fig. 3-7A and B). In the transformed *Cleome* line, *T3-3/Ch*, LDs were not dispersed in the cytoplasm but all attached to the surface of membranes/vesicles to form packages of several micrometer width (Fig. 3-7A). CLSM revealed that the LDs and flavonoid-containing vesicles co-located in packages. I considered these packages primitive tapetosomes. Non-transformed *Cleome* pollen surface did not have Arabidopsis T3 oleosin but possessed some flavonoids, whereas *T3-3/Ch* pollen surface had T3-oleosin and more flavonoids (Fig. 3-7C). Non-transformed *Cleome* pollen was susceptible to dehydration treatment, with germination decreased from 70% to 15% (Fig. 3-7D; Table 3-1). *T3-1/Ch* pollen became tolerant to the dehydration treatment (Fig. 3-7D). The degree of tolerance was proportional to the level of *T3* gene expression in various transformants (*T3-3/Ch, T3-1/Ch* and *T3-2/Ch* with decreasing levels of *T3* transcript; Fig. 3-7D, E). The finding reiterates the quantitative aspect of T-oleosin in exerting adaptive advantage.

DISCUSSION

Cleomaceae, the family closest to Brassicaceae, did not possess the T-oleosin gene cluster, tapetosomes and dehydration-tolerant pollen. *Cleome* transformed with an Arabidopsis T-oleosin gene had a tapetum-pollen phenotype similar to that of Brassicaceae. This phenotype included aggregated cytoplasmic LDs and flavonoid-containing vesicles as primitive tapetosomes, more flavonoids on pollen and dehydration-tolerantant pollen. Generation of primitive tapetosomes with the mere addition of T-oleosin in transformed *Cleome* is consistent with the structure of a modern tapetosome, in which LDs (with surface oleosins) and vesicles associate ionically (Hsieh and Huang, 2007). Association of LDs and vesicles in the cell could represent the initial event in the Brassicaceae species when it first acquired tapetum T-oleosin. This and other acquisitions may reflect that the flowers underwent substantial structural changes after the divergence between Brassicaceae and Cleomaceae (Jydd et al., 2002).

I propose the evolution of oleosins in the tapetum of Brassicaceae as follows. After the partition of Brassicaceae and Cleomaceae >20M years ago (Barker et al., 2009), a redundant oleosin gene for seed/pollen expression mutated at its promoter or another regulatory mechanism for tapetum expression. This gain of tapetum

oleosin probably occurred after the *At*-α polyploidy event ~23 M years ago (Barker et al., 2009; Jydd et al., 2002; Schranz et al., 2006), because *A. thaliana* has only one tandem T-oleosin gene cluster. The new oleosin in the ancestor Brassicaceae species aggregated LDs and flavonoid-containing vesicles to form primitive tapetosomes, which conferred adaptive benefit in generating vigorous pollen. The subsequently generated tandem paralogs produced more oleosins, thus resulting in the production of modern tapetosomes and conference of a greater adaptive benefit. As a consequence, the tandem T-oleosin gene cluster was retained. I consider unlikely that the tapetum oleosin from a single gene or tandem genes evolved before the partitioning of Brassicaceae and Cleomaceae and that Cleomaceae lost the gene or tandem genes, because I could observe the possible adaptive benefit of *Cleome* having acquired a single T-oleosin gene. Whether the expanded T-oleosin paralogs have conferred an additional benefit of promoting pollen self-incompatibility (Tantikanjana et al., 2010) remains to the elucidated.

During evolution, the new oleosin in the ancestor Brassicaceae species aggregated LDs and flavonoid-containing vesicles to form primitive tapetosomes, which conferred adaptive benefit in generating vigorous pollen. The subsequently generated tandem paralogs produced more oleosins, thus resulting in the production

of modern tapetosomes and conference of a greater adaptive benefit. As a consequence, the tandem T-oleosin gene cluster was retained.

MATERIALS AND METHODS

Plant Materials

Flowering *Cleome hassleriana* from a local flower market were maintained in greenhouses at 26/18 °C. The name *Cleome hassleriana* (named earlier as *Cleome spinosa*) is used in this report because of its general recognition, even though some researchers have renamed it to *Tarenaya hassleriana* (weltis and Cochrane, 2007). *Carica papaya* trees were grown in a farm at Taipei. Developing anthers and mature pollen of the above plants, as well those of other species from flower markets in Taipei, were used.

Transformation of *Cleome hassleriana*

Cleome seeds were allowed to germinate on MS medium at 30/23 °C and 16-h/8-h day/night cycle for 14 d. The hypocotyl was cut into 8-10 mm sections and immersed with Agrobacterium LBA4404 harboring the *pC2300/T3ole* construct (Fig. 3-8) in a medium with 20 $g.l^{-1}$ sucrose, 5.9 $g.l^{-1}$ Na_3 citrate and 0.2 mM acetosyringone, pH 5.5, for 30 min. The infected tissues were transferred to MS medium with 0.2 mM acetosyringone for 2 d. The explants were transferred to a regeneration medium with carbenicillin and kanamycin for selection of transformed

tissues (Newell et al., 2010). After 4 weeks, regenerated plantlets were transferred to MS medium with 0.1 $mg.l^{-1}$ NAA and carbenicillin. Transformed plants were identified by content of the *T3* gene fragment with PCR.

Pollen Germination Test

Pollen of *Cleome* were collected from open flowers and placed at 25% relative humidity and 24 ^{0}C. Pollen grains were placed on a glass slide with liquid germination medium (10% sucrose, 1.3 mM H_3BO_3, 2.9 mM KNO_3 and 9.9 mM $CaCl_2$, pH 5.8) at 28 °C (Wang and Jiang, 2011) for 1-4 h. Pollen grains were observed for germination by microscopy. At least 500 pollen grains of each line were counted.

Electron Microscopy

Anthers were fixed via a high-pressure freezing or chemical fixation procedure. For the former, anthers were fixed in a high pressure freezer (Leica EM PACT2). Fixed materials were subjected to freeze substitution in ethanol (containing 0.2% glutaraldehyde and 0.1% uranyl acetate) in Lecia Automatic Freeze-Substitution System and embedded in LR Gold resin (Structural Probe, west Chester, PA). For chemical fixation, anthers were fixed with 2.5% glutaraldehyde, 4%

paraformaldehyde and 0.1 M K-phosphate (pH 7.0) at 4^0 C for 24 h. The materials were washed with 0.1 M K-phosphate buffer (pH 7.0) for 10 min twice and then treated with 1% OsO_4 in 0.1 M K-phosphate (pH 7.0) at 24 ^{0} C for 4 h. The fixed materials were rinsed with 0.1 M K-phosphate buffer (pH 7.0), dehydrated through an acetone series and embedded in Spurr resin. Ultrathin sections (70 to 90 nm) were stained with uranyl acetate and lead citrate and examined with a Philips CM 100 TEM at 80 KV.

Immunofluorescence and Chemical Staining with CLSM

All antibody treatments were performed with 1:50 dilution of the IgG fraction, 1% (w/v) milk powder and 1x PBS. Unless otherwise stated, each wash was performed with PBST (1x PBS and 0.05% [w/w] Tween-20) for 10 min. Tissues were fixed in 4% paraformaldehyde, 1x PBS and 0.15 M sucrose at 4° C for 16 h. For flavonoid localization with diphenylboric acid 2-aminoethyl ester (DPBA), fixed anther sections and mature pollen were washed twice and treated with 1 M HCl at 70 °C (for Arabidopsis) or 0.8 M HCl at 56 °C (for *Cleome*) for 30 min (Hsieh and Huang, 2007). Samples were washed with 10% glycerol twice and stained with DPBA (saturated DPBA [<0.5%, w/v], 0.01% Triton X-100 and 10% glycerol) at 20 °C for 2 were then

washed twice with 10% glycerol. For double labeling with DPBA for flavonoids and Nile Red for lipids, the samples after the DPBA staining were treated with Nile Red. For double labeling with DPBA for flavonoids and antibodies against the vacuole marker protein V-PPase, the samples after the DPBA staining were blocked with 3% milk in 1x PBS at 20 °C for 1 h. They were treated with rabbit IgG against V-PPase (Agrisera, Sweden) in 1% milk and 1x PBS at 20 °C for 2 h. After 2 washes, samples were treated with a mixture of DPBA and secondary antibodies (0.5% DPBA, 0.01% Triton X-100, 10% glycerol, 1x PBS, and 1:100 cyanine 5–conjugated goat IgG against rabbit IgG) at 20 °C for 2 h. Samples were washed twice. The labeled samples were observed with a Leica SP2 confocal microscope. DPBA-flavonoid, Nile Red, and Cyanine 5 were excited with 488 nm, 543 nm and 633 nm; and the emissions were detected at 500 to 530 nm, 565 to 615 nm and 650 to 750 nm, respectively.

RT-PCR of T-oleosin transcripts

RNA was extracted from Arabidopsis inflorescences with use of the Illustra RNAspin *Mini* Kit (GE Health Care, Piscataway, NJ). The RNA (1 μg) was used for reverse-transcription of cDNA with the SuperScriptIII RT-PCR system (Invitrogen). Quantitative real-time RT-PCR of *T3* transcript was performed with iQ SYBR Green

Supermix (Bio-Rad) and gene-specific primers (Table 3-2), andthe results were normalized with actin transcript as reference.

Sequencing of *Cleome* DNA

Cleome leaves were ground in liquid nitrogen, and extracted for DNA with pre-heated extraction buffer (100 mM Tris-HCl pH 7, 1.4 M NaCl, 20 mM EDTA, 2% CTAB, 1% PVP and 1% 2-mercaptoethanol) at 65 °C, followed by extraction with an equal volume of chloroform and then treated with RNase A (Sigma). Whole genome sequencing was performed with Illumina Genome Analyzer IIx by Ymygene Bioscience, Taipei. Sequences of the assembled fragments of *C. hassleriana* (KC777373, KC777374) and *C. papaya* (KC777372) have been deposited in NCBI.

REFERENCES

Barker MS, Vogel H, Schranz ME (2009) Paleopolyploidy in Brassicales: analyses of the *Cleome* transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales. Genome Biol Evol **1:** 391-399

de Oliveira DE, Franco LO, Simoens C, Seurinck J, Coppieters J, Botterman J, van Montagu M (1993) Inflorescence-specific genes from *Arabidopsis* thaliana encoding glycine-rich proteins. Plant J **3:**495-507

Fiebig A, Kimport R, Preuss D (2004) Comparisons of pollen coat genes across Brassicaceae species reveal rapid evolution by repeat expansion and diversification. Proc Natl Acad Sci USA **101:** 3286-3291

Goldberg, R.B., Beals, T.P., and Sanders, P.M (1993) Anther development: basic principles and practical applications. Plant Cell **5:** 1217-1229

Hsieh K, Huang AHC (2007) Tapetosomes in Brassica tapetum accumulate endoplasmic reticulum–derived flavonoids and alkanes for delivery to the pollen surface. Plant Cell **19:**582-596

Huang MD, Wei FJ, Wu CC, Hsing YIC, Huang AHC (2009) Analyses of advanced rice anther transcriptomes reveal global tapetum secretory functions and potential proteins for lipid exine formation. Plant Physiol **149:** 694-707

Weltis HH, Cochrane TS (2007) Studies in the Cleomaceae V: A new genus and ten new combinations for the flora of North America. NOVON **17:** 447-451.

Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2002) Brassicales. In Plant Systematics. 2nd edition. (Sinauer Asso., Inc. publishers), pp 402-405

Kim HU, Hsieh K, Ratnayake C, Huang AHC (2002). A novel group of oleosins is present inside the pollen of *Arabidopsis*. J Biol Chem **277:** 22677-22684

Li YB, Suen DF, Huang CY, Kung SY, Huang AHC (2012) The maize tapetum employs diverse mechanisms to synthesize and store proteins and flavonoids and transfer them to the pollen surface. Plant Physiol **158:** 1548-1561

Newell CA, Brown NJ, Liu Z, Pflug A, Gowik U, Isthoff P, Hibberd JM (2010) *Agrobacterium tumefaciens*-mediated transformation of *Cleome gynandra* L., a C_4 dicotyledon that is closely related to *Arabidopsis thaliana*. J Exp Bot **61:**1311–1319

Platt KA, Huang AHC, Thomson WW (1998) Ultrastructural study of lipid accumulation in tapetal cells of *Brassica napus* L. cv. westar during microsporogenesis. Int J Plant Sci **159:** 724–737

Polowick PL, Sawhney VK (1993). Differentiation of the tapetum during microsporogenesis in tomato (*Lycopersicon esculentum* Mill.), with special reference to the tapetal cell wall. Ann of Bot **72**:595–605

Schein M, Yang Z, Mitchell-Olds T, Schmid KJ (2004) Rapid evolution of a pollen-specific oleosin-like gene family from *Arabidopsis thaliana* and closely related species. Mol Biol Evol **21:** 659-669

Schranz ME, Mitchell-Olds T (2006) Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. Plant Cell **18:** 1152-1165

Tantikanjana T, Nasrallah ME, Nasrallah JB (2010) Complex networks of self-incompatibility signaling in the Brassicaceae. Current Opinion Plant Biol **13:** 520-526

Roberts LS, Gerster J, Allard S, Cass L, Simmonds J (1994) Molecular characterization of two Brassica napus genes related to oleosins which are highly expressed in the tapetum. Plant J **6:** 927–933

Wang H, Jiang L (2011) Transient expression and analysis of fluorescent reporter proteins in plant pollen tubes. Nat Protoc **6:** 419-426.

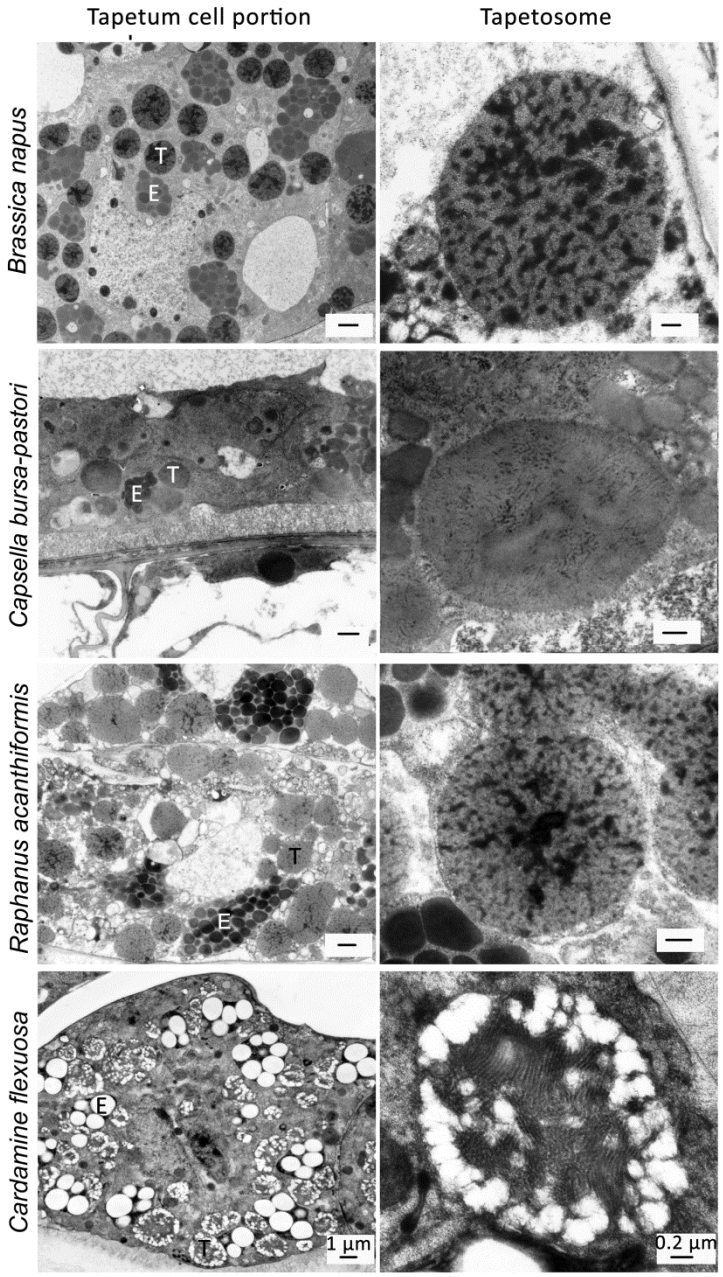

Fig. 3-1. TEM of tapetum cells of various species of Brassicaceae.

Tapeta were at a developmental stage immediate before their programmed cell death. Anthers were fixed with a chemical fixation procedure except those of *Cardamine flexuosa*, which was fixed with a freezing fixation procedure. For each species, low (left panel, scale bar representing 1 µm) and high magnification (right panel, scale bar representing 0.2 µm) images are shown. Elaioplasts (E) and tapetosomes (T) are labeled.

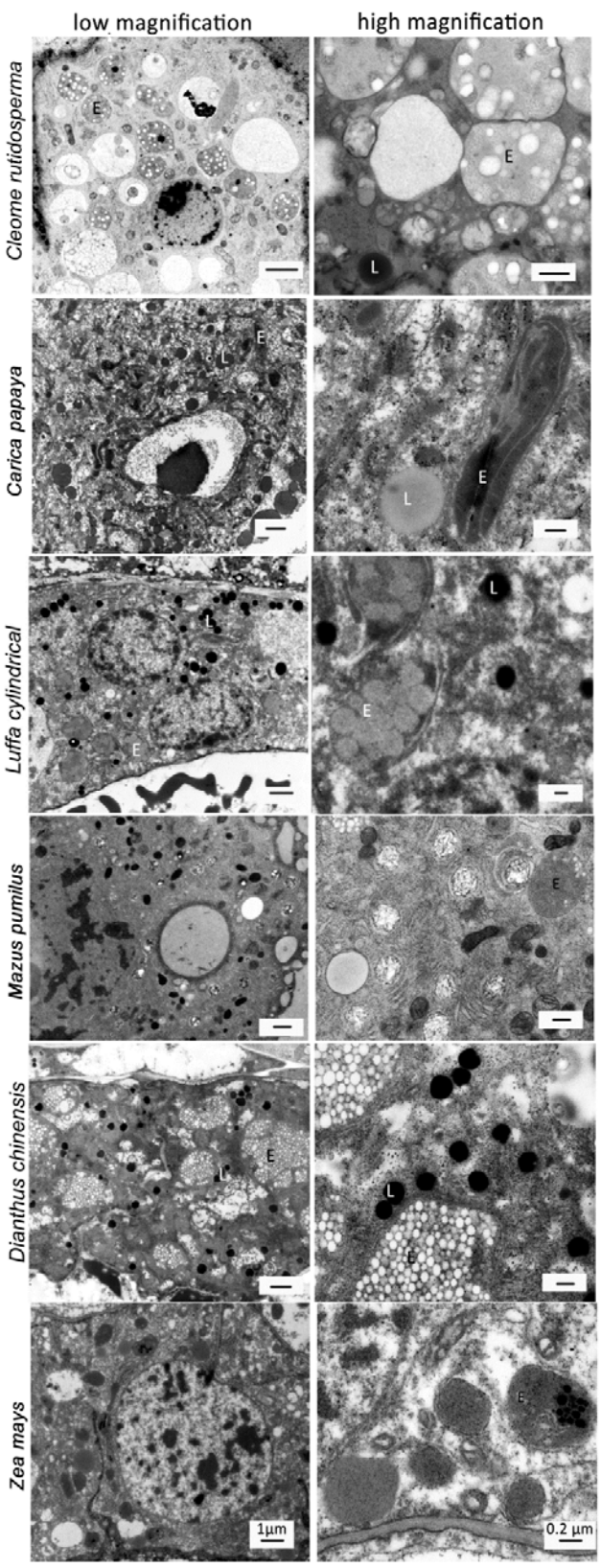

Fig. 3-2. TEM of tapetum cells of various species outside of Brassicaceae.

Cleome and *Carica* are in the Order Brassicales but not in the Family Brassicaceae. Tapeta were at a developmental stage immediate before their programmed cell death. Anthers of *Diarthus chinensis* were subjected to a chemical fixation procedure, and those of the other species were treated with a freezing fixation procedure. For each species, low (left panel, scale bar representing 1 µm) and high magnification (scale bar representing 0.2 µm) images are shown. Solitary lipid LDs (L) and elaioplasts (E) are indicated. No tapetosome-like subcellular structures were observed.

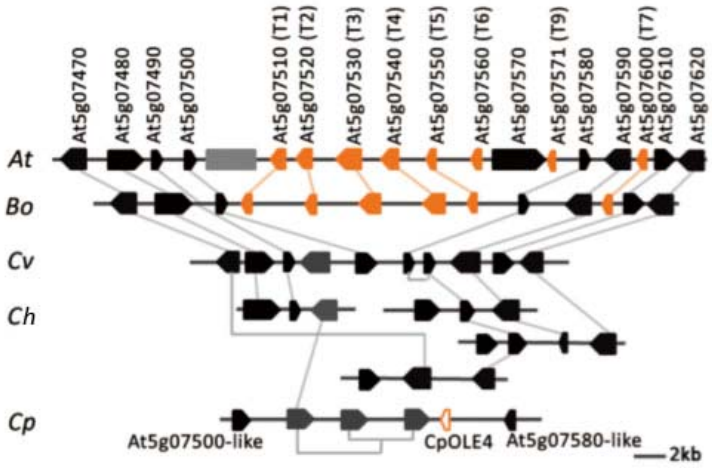

Fig. 3-3. Presence of the T-oleosin gene cluster in Brassicaceae but absence in Cleomaceae and Caricaceae.

Primary transcripts are indicated with filled arrows (orange for oleosin genes and black for non-oleosin genes) in genomic DNA sequences from *Arabidopsis thaliana* (*At*) and *Brassica oleracea* (*Bo*) of Brassicaceae, *Cleome violacea* (*Cv*) and *Cleome hassleriana* (*Ch*) of Cleomaceae, and *Carica papaya* (*Cp*) of Caricaceae. Orthologs among the genomes are related by gray lines. *CpOLE4* encodes a seed oleosin. *At* has the T8 oleosin gene (At5g61610) 23 mb downstream of T7. Gray filled box indicates transposon.

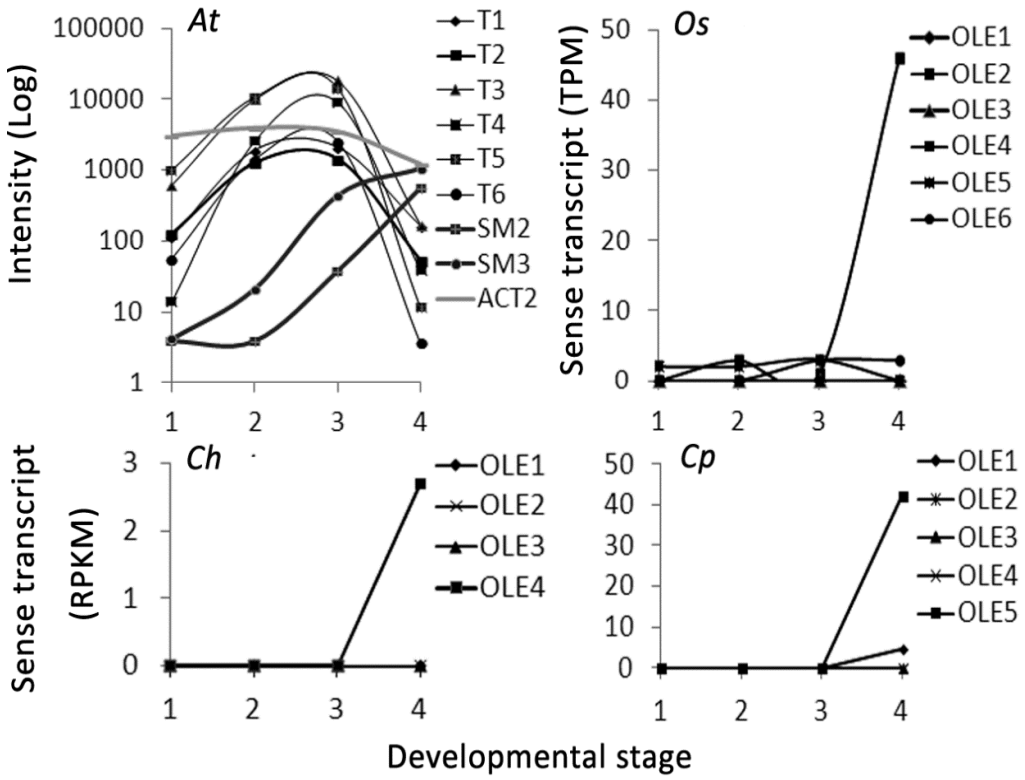

Fig. 3-4. Levels of oleosin transcripts in inflorescences or anthers of different developmental stages from various species.

Stages 1-3 represent inflorescences or anthers containing microspores of progressive morphologic features (tetrad, early vacuolated and late vacuolated stages, respectively), and stage 4 denotes mature pollen.

For *Arabidopsis thaliana (At)*, the first 3 stages were of whole inflorescences of designated flower stages 9, 10/11, and 11 (AtGenExpress Visualization Tool AVT, http://jsp.weigelworld.org/expviz/expviz.jsp). For *Oryza sativa (Os)*, the data were obtained from anthers (http://rice.sinica.edu.tw). For *Cleome hassleriana (Ch)* and *Carica papaya (Cp)*, I obtained anthers with the above-described microspore morphologic features. The *Cleome* anther lengths were 4~5 mm, 5~6 mm and 6~8 mm for stages 1, 2 and 3, respectively, and the *Carica* anther lengths were 2~3 mm, 3~4 mm and 4~5 mm. From these anthers, RNA was extracted with use of the Illustra RNAspin *Mini* Kit (GE Health Care, Piscataway, NJ, USA). RNA-seq was performed with Illumina Genome Analyzer IIx by Ymygene Bioscience (Taipei). Sequences of oleosins of *Ch* and *Cp* can be found in Fig. 3-5.

In Arabidopsis, values of *T1*-6 (tapetum-specific oleosin) and *SM2-3* (seed- and microspore-specific oleosin) transcripts obtained from AtGenExpress are expressed in log microarray intensity. *T7-9* transcripts were not represented or detected in the

microarrays. *Oryza* anthers had 6 oleosin transcripts with levels expressed in SBS transcripts per million (TPM). *Cleome* and *Carica* anthers had 4 and 5 oleosin transcripts, respectively, with levels expressed in SBS Read Per Kilobase of transcript per Million mapped reads (RPKM).

Fig. 3-5. Aligned sequences of oleosins of *Arabidopsis thaliana* (At), *Cleome hassleriana* (*Ch*), *Carica papaya (Cp)* and *Oryza sativa (Os)*.
Sequences were used for construction of the phylogenetic tree shown in Fig. 3-6. Stars indicate the 4 completely conserved proline residues in the proline-knot motif (the hairpin turn). Dark and light shading labels the highly and lightly conserved residues, respectively. The longer T-oleosins have short repeats at the C termini, which have undergone rapid evolutionary changes (Schein et al., 2004) because of the lack of structural constraints on functionality. For example, the longest T3 is glycine-rich (26%) as noted (Fiebig et al., 2004) but also serine-rich (16%) and lysine-rich (11%).

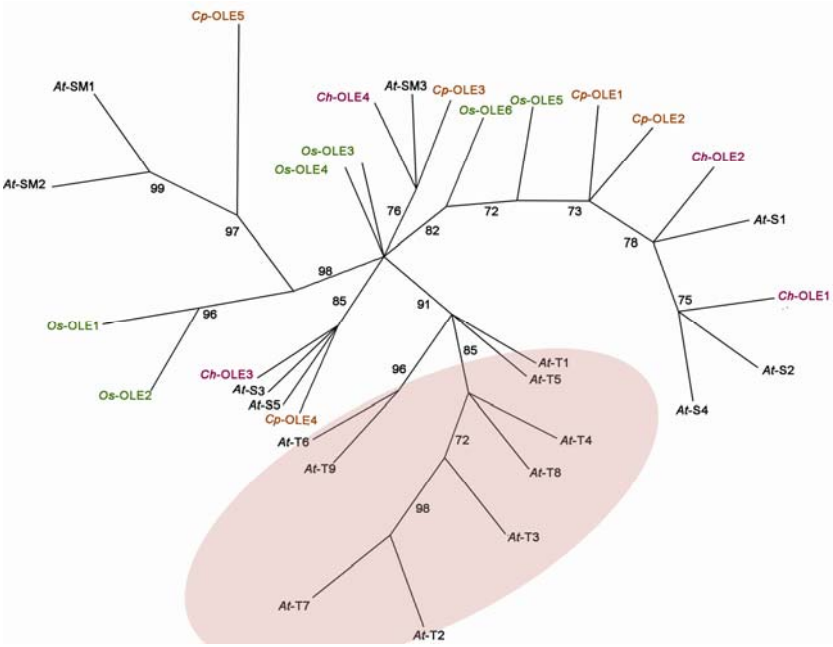

Fig. 3-6. An unrooted phylogenetic tree of oleosins from *Arabidopsis thaliana (At), Cleome hassleriana (Ch), Carica papaya (Cp)* and *Oryza sativa (Os).*

The tree was constructed according to the oleosin sequences shown in Fig. 3-5 by a distance method (neighbor-joining) with use of PHYLIP and 1000 bootstrap replicates. Bootstrap values > 70 are indicated.

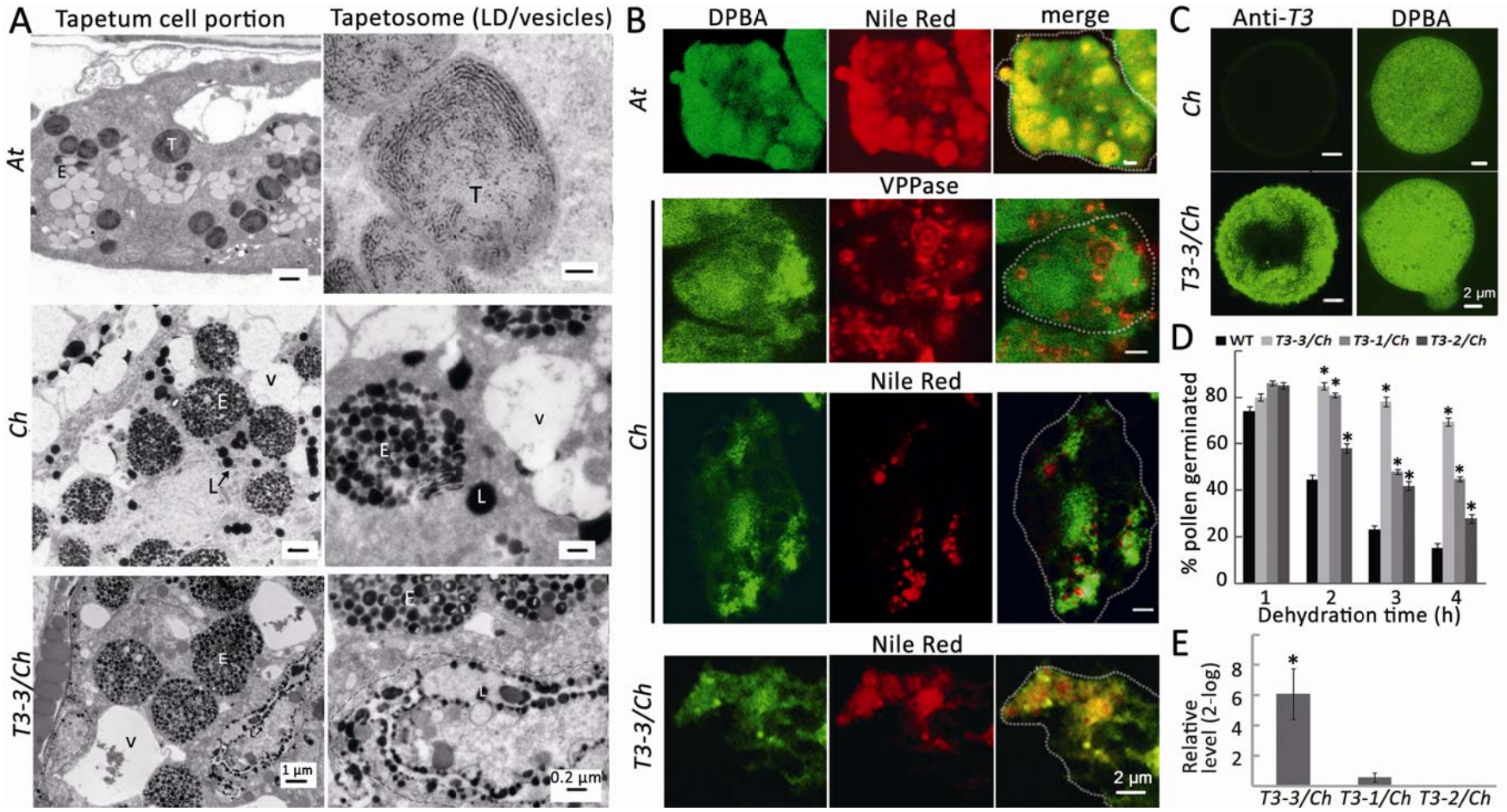
A
Tapetum cell portion
Tapetosome (LD/vesicles)
At
Ch
T3-3/Ch
T
E
V
L
1 µm
0.2 µm
B
DPBA
Nile Red
merge
At
VPPase
Ch
Nile Red
T3-3/Ch
2 µm
C
Anti-T3
DPBA
Ch
T3-3/Ch
2 µm
D
WT
T3-3/Ch
T3-1/Ch
T3-2/Ch
% pollen germinated
0
20
40
60
80
1
2
3
4
Dehydration time (h)
E
Relative level (2-log)
2
4
6
8
T3-3/Ch
T3-1/Ch
T3-2/Ch

Fig. 3-7. Effect of transformation of *Cleome hassleriana* with Arabidopsis T3-oleosin gene.

A. TEM of tapetum cells in Arabidopsis and *Cleome* WT and *T3-3/Ch* (*Cleome* transformed with Arabidopsis T3 oleosin gene). Bars on left and right panels represent 1 μm and 0.2 μm, respectively. T, E, V, and L indicate tapetosome, elaioplast (packed with electron-dense plastoglobules), vacuole and LD, respectively. LDs in *Cleome* WT appear as electron-dense spheres (arrows) scattered in the cytoplasm, and those in *T3-3/Ch* are confined along membranous structures as primitive tapetosomes.

B. Immuno CLSM of tapetum cells of Arabidopsis and *Cleome* WT and *T3-3/Ch*. Samples were detected for flavonoids (with DPBA, pseudo-color in green), LDs (with Nile Red, in red; many scattered LDs [panel A] were too small to be revealed by CLSM) and vacuoles (with antibodies against VPPase, in red). White dotted line marks the circumference of the cell in the merge image. Bars represent 2 μm.

C. Immuno CLSM of pollen of *Cleome* WT and *T3-3/Ch*. Samples were detected for T3 oleosin (with antibodies, pseudo-color in green) and flavonoids (with DPBA, in green). The same settings (laser power and detection gain) were used for direct comparison of *Cleome* WT and *T3-3/Ch*. Bar represents 2 μm.

D. Test of dehydration tolerance of pollen of *Cleome* WT and 3 transformants, *T3-3/Ch*, *T3-1/Ch* and *T3-2/Ch* with decreasing levels of T3 gene expression (Panel E). Pollen was placed at 25% relative humidity and 24 ^{0}C for 1-4 h and allowed to germinate. Significant difference was calculated from the germination % of mutant compared to WT (t-test, *$P < 0.0001$).

E. Quantitative real-time RT-PCR of *T3* transcript in the 3 transformants. *T3* levels in *T3-3/Ch* (t-test, *P<0.001) and *T3-1/Ch* relative to the lowest level in *T3-2/Ch* are shown.

Fig. 3-8. Construction of plasmids p2300/T3

T3 oleosin gene (including promoter and terminator) was amplified with the primers T3proFKpnI and T3terRSalI and inserted into pCAMBIA2300 (CAMBIA) via the KpnI and SalI sites to generate p2300/T3. The plasmids were transformed into *Cleome hassleriana* to generate *T3*/*Ch* for probing *T3* effect on *Cleome* tapetum cell structures.

Table 3-1. Phenotypes of tapetosomes and pollen germination in T-oleosin transformed *Cleome.*

	Presence of tapetosome (Yes/No)	Phenotype of tapetosome	Pollen germination (%)**
Cleome			
WT	N	lipid droplets separated from vesicles	15±2.0
T3-3/Ch	Y/N	primitive tapetosomes	70±1.7*

**Germination after dehydration for 4 h for *Cleome*. Details are in Fig. 3-7.

* Significant difference from the germination % of WT (t-test, $P < 0.0001$). The T3 transformed mutant (T3-3/Ch) is compared with the WT.

Table 3-2. Sequences of primers used in capter three.

primer	sequence
T3rtF	5’-GTATGTCCGGAGGTGGTATG-3’
T3rtR	5'-TTACTCCGCCATGTGGCCAC-3'
T3ProFKpnI	5’-GGGGTACCGGATGTTTGTATTAGTTTCTCAC-3’
T3terRSalI	5’-ATCGTCGACGGTCCAAGCTATTTCTCAAC-3’

GENERAL CONCLUSIONS

Plants lipids (vegetable oils or TAGs) are important sources of calories for human and farm-animal nutrition. Vegetable oils are used extensively in general cooking. The high energy density of TAGs has also made these molecules an attractive source of biodiesel by transesterification of their component fatty acids. Futhermore, humans use vegetable oils in a variety of bio-based industrial products, including lubricants, soap, detergents, and drying oils (Lu et al, 2011). Therefore, classical breeding and genetic engineering have been used to enhance the quality and quantity of plant lipids (Weselake et al., 2009).

In plant LDs, oleosin acts as a stabilizer. The oleosin-coated OBs provide stable sinks for lipid storage in the cytosol. The unique structures of oleosin play a pivotal role in lipid storage and metabolism. In higher plants, seeds are the major organs that synthesize OBs. Due to the natural property of seeds of having a thick coat and the internal storage cells with densely packed cytoplasm, microscopy studies of the biogenesis of OBs in seeds are difficult. Although LDs are present in green leaves of diverse plant species, they are not prominent in the cytoplasm (Lernsten et al., 2006). No or few oleosin transcripts are present in high-quality MPSS or SBS leaf transcriptomes of Arabidopsis (Meyers et al., 2004) and rice (Nobuta et al., 2007).

The haploid *Physcomitrella* spore and gametophyte is genetically and physiologically similar to the haploid pollen in higher plants. Both the spore and pollen also contain storage OBs coated with oleosins. In this dissertation, I have shown that the gametophyte of *Physcomitrella* stores TAG and SE in abundant OBs. Moreover, the gametophyte consisting of a single layer of cells can be easily transformed and examined with CLSM for transient expression of genes. Such a plant system for OB studies was not previously available (Miao and Jiang, 2007), and transient expression of genes in *Physcomitrella* was performed only with juvenile protonema (e.g., Marella et al., 2006), which contain few or no OBs. Earlier, maturing embryos of flax and microspore cultures of *Brassica* were used for stable transformation of modified oleosin genes, and the transformed plant materials were examined with in vitro biochemical analyses (Abell et al., 2004). The OB-containing internal cells in an embryo may not be transformed easily with bombardment or observed clearly with CLSM, the microspore culture is highly artificial, and the microspore (pollen) OBs are not for gluconeogenesis as those in seeds and *Physcomitrella* are. Overall, *in vivo* studies of the topology of oleosin protein structure will be feasible with the *Physcomitrella* gametophyte.

Evolutionarily advanced plants usually have more copies of oleosin genes, e.g,

Arabidopsis with 17 and rice with 5. Although the expression of oleosin genes is in a tissue specific manner, most of the organs (seed, tapeta) with oleosin-coated OBs have specific oleosins genes expressed (Kim et al, 2002; Nobuta et al, 2007). *Physcomitrella* possesses 3 oleosin genes, and only 2 of them have expression in the gametophyte. A *Physcomtrilla* mutant with total knockout of the 2 oleosins can be created via homologous recombination for functional study. Another advantage of the *Physcomitrella* transient expression system is that the growth condition of the plant can be altered easily for study of storage lipid metabolism and packaging.

In addition to seed, the tapetum in the anthers of Brassicaceae is another tissue containing abundant OBs in higher plants. Oleosin-coated alkane LDs associated ionically with many flavonoid-containing, ER-derived vesicles form the abundant tapetosomes in tapetum cells of Brassicacea (Hsieh and Huang, 2007; Ting et al, 1998; Wu et al, 1997). Oleosin genes that are specifically expressed in tapetum cells are termed T-oleosin genes, and are clustered as 4-9 tandem repeats. Various genera and species in Brassicaceae contain the T-oleosin gene cluster. In evolution, genomes expanded via whole-genome, segmental, tandem and individual-gene duplications (Cannon, 2004; Ding, 2012). The duplicated genes would be eliminated over time, unless they conferred adaptive neutrality or benefit to the organism. The tandem

T-oleosin gene cluster produces abundant oleosins in the tapetum. In this dissertation, I have shown the clustered paralogs of T-oleosin genes conferring quantitative adaptive benefit, such that they have been retained through evolution.

In the tapetum cells, the tapetosomes are the warehouses of temporarily stored alkanes and flavonoids, which will be discharged to the surface of adjacent maturing pollen grains, forming the pollen coat. The coat waterproofs the pollen grain and protects it against UV radiation. Not all pollen coats of diverse plants, such as maize and rice, have oleosins. In this dissertation, I have delineated the evolution and functions of tapetosomes and T-oleosin genes. Cleomaceae, the family closest to Brassicaceae, does not possess tapetosomes, the tandem oleosin gene cluster and the downstream gene activities. Compared to pollen of Brassicacea, pollen of Cleomaceae is susceptible to dehydration damage. *Cleome* transformed with an Arabidopsis T-oleosin gene has a tapetum-pollen phenotype similar to that in Brassicaceae. This phenotype includes the aggregation of cytoplasmic LDs and flavonoid-containing vesicles to form primitive tapetosomes, presence of more flavonoids on pollen, and pollen being more resistant to dehydration damage. Aggregation of LDs and vesicles in the cell could represent the initial event in the Brassicaceae species when it first acquired tapetum T-oleosin. The evolutionary

appearence of tapetosomes may be related to the changes of floral structure and of pollination mechanism after the divergence between Brassicaceae and Cleomaceae.

Oleosin-coated OBs and oleosin genes are unique to and important in plants. In my dissertation, I delineate the evolution of OBs and oleosin genes from the primitive moss to higher plants. The structure of the oleosin protein of oleosin protein is still unclear. I have established the gametophyte of *Physcomitrella* to be an excellent system to study OBs. In the future, researchers can take advantage of this system to study the oleosin protein structures and functions. Additionally, my study of the T-oleosin tandem repeat cluster provides the first documentation of the quantitative benefit effect of tandem repeat genes. Tapetosomes enable pollen grains to be more tolerant to environment at stresses, such that the T-oleosin cluster has be retained. However, It is still uncertain that if Brassicasea is the only family that acquired T-oleosin cluster and tapetosomes. This uncertainty should be explored. During evolution, the anthers also have other mechanisms evolved to produce stress-resistance pollen.

REFERENCES

Abell BM, Hahn M, Holbrook LA, Moloney MM (2004) membrane topology and sequence requirements for oil body targeting of oleosin. Plant J **37:** 461-470

Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biol **4:**10

Ding Y, Zhou Q, Wang W (2012) Origins of new genes and evolution of their novel functions. Annu Rev of Ecol Evol and Syst **43:** 345-363

Hsieh K, Huang AHC (2007) Tapetosomes in brassica tapetum accumulate endoplasmic reticulum–derived flavonoids and alkanes for delivery to the pollen surface. Plant Cell **19:**582-596

Kim HU, Hsieh K, Ratnayake C, Huang AHC (2002). A novel group of oleosins is present inside the pollen of *Arabidopsis*. J Biol Chem **277:** 22677-22684

Lersten NR, Czlapinski AR, Curtis JD, Freckmann R, Horner HT (2006) Oil bodies in leaf mesophyll cells of angiosperms: overview and a selected survey. Am J Bot **93:** 1731-1739

Lu C, Napier, JA, Clemente, TE and Cahoon, EB (2011) New frontiers in oilseed biotechnology: meeting the global demand for vegetable oils for food, feed, biofuel and industrial applications. Curr Opin Biotechnol. **22:** 252–259

Meyers BC, Lee DK, Vu TH, Tej SS, Edberg SB, Matvienko M, Tindell LD (2004) *Arabidopsis* MPSS. An online resource for quantitative expression analysis. Plant Physiol **135:** 801-813

Marella HH, Sakata Y, Quatrano RS (2006) Characterization and functional analysis of ABSCISIC ACID INSENSITIVE3-like genes from *Physcomitrella patens*. Plant J **46:** 1032-1044

Miao YS, Jiang L (2007) Transient expression of fluorescent fusion proteins in protoplasts of suspension cultured cells. Nature Protocols **2:** 2348-2353

Nobuta K, Venu RC, Lu C, Belo A, Vemaraju K, Kulkarni K, Wang W, Pillay M, Green PJ, Wang GL, Meyers BC (2007) An expression atlas of rice mRNAs and small RNAs. Nat Biotechnol **25:** 473-477

Ting JTL, Wu SSH, Ratnayake C, Huang AHC (1998). Constituents of the tapetosomes and elaioplasts in *Brassica campestris* and their degradation and retention during microsporogenesis. Plant J **16:** 541-551

Weselake RJ, Taylor DC, Rahman MH, Shah S, Laroche A (2009) Increasing the flow of carbon into seed oil. Biotechnol Adv **27:** 866–878

Wu SSH, Platt KA, Ratnayake C, Wang TW, Ting JTL, Huang AHC (1997) Isolation and characterization of novel neutral-lipid-containing organelles and globuli-filled plastids from *Brassica napus* tapetum. Proc Natl Acad Sci USA **94:** 12711-12716