### Title

Tandem oleosin genes in a cluster acquired in Brassicaceae created tapetosomes and conferred additive benefit of pollen vigor




### Journal

Proceedings of the National Academy of Sciences of the United States of America, 110(35)

### ISSN

0027-8424



### Authors

Huang, Chien Yu

Chen, Pei-Ying

Huang, Ming-Der

et al.



### Publication Date

2013-08-27

### DOI

10.1073/pnas.1305299110

Peer reviewed

# Tandem oleosin genes in a cluster acquired in Brassicaceae created tapetosomes and conferred additive benefit of pollen vigor

Chien Yu Huang[a], Pei-Ying Chen[b], Ming-Der Huang[b], Chih-Hua Tsou[b], Wann-Neng Jane[b], and Anthony H. C. Huang[a,b,1]

[a]Center for Plant Cell Biology, Department of Botany and Plant Sciences, University of California, Riverside, CA 92521; and [b]Institute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan

Edited by June B. Nasrallah, Cornell University, Ithaca, NY, and approved July 12, 2013 (received for review March 19, 2013)

**During evolution, genomes expanded via whole-genome, segmental, tandem, and individual-gene duplications, and the emerged redundant paralogs would be eliminated or retained owing to selective neutrality or adaptive benefit and further functional divergence. Here we 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 five to nine 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. Transfer-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. 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. We 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.**

evolutionary appearance of organelle | tandem gene cluster

In evolution, genomes expanded via whole-genome, segmental, tandem, and individual-gene duplications (1–3). 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 (4) and homeobox (5) in animals, as well as lectins (6) and other stress-resistant proteins (7, 8) 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 (9–11). 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) (11), or glycine-rich-protein genes because one or two of the paralogs encode oleosins with short glycine-rich repeats at the nonconserved C termini (10, 12). These earlier terms were used before the function of the oleosins in anthers was delineated (13, 14). All of the nine T-oleosin paralogs locate on chromosome 5; eight are 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 (10, 11). This rapid evolution applies to the sequences encoding the two 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 two exons, and the location of the exons/intron differs from those of the seed- and seed/pollen-expressed oleosin paralogs (9). The T-oleosin gene cluster has been found in several genera of Brassicaceae (10, 11); whether it is present outside of Brassicaceae has not been determined.

Oleosins are structural proteins on lipid droplets in seeds, pollen, and tapeta (15–17). The protein has a hallmark of ~72 uninterrupted nonpolar residues that form a hairpin penetrating into the lipid droplets. This hairpin, together with the adjacent

Author contributions: C.Y.H., C.-H.T., and A.H.C.H. designed research; C.Y.H., P.-Y.C., M.-D.H., C.-H.T., W.-N.J., and A.H.C.H. performed research; C.Y.H., P.-Y.C., M.-D.H., C.-H.T., W.-N.J., and A.H.C.H. analyzed data; and C.Y.H. and A.H.C.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. KC777372, KC777373, KC777374, KF270089, and KF270090).

[1]To whom correspondence should be addressed. E-mail: anthony.huang@ucr.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1305299110/-/DCSupplemental.

amphipathic N- and C-terminal motifs, stabilizes the hydrophobic lipid droplet in the cytoplasm. In the tapetum of Brassicaceae, T-oleosins are components of the abundant organelles called the tapetosomes (14, 18, 19). Each tapetosome contains numerous oleosin-coated alkane lipid droplets associated ionically with many flavonoid-containing, endoplasmic reticulum-derived vesicles (14). 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 surface (20). Similarly, mutational loss of pollen-coat flavonoids in *Arabidopsis* results in pollen being more susceptible to UV radiation (14).

The tandem T-oleosin gene cluster produces abundant oleosins in the tapetum. We investigated the clustered paralogs in conferring functional neutrality, adaptive benefit, and/or quantitative benefit. We probed the existence of the T-oleosin gene cluster outside Brassicaceae and found that the closest family, Cleomaceae (21–23), did not possess the T-oleosin gene cluster, tapetosomes, and dehydration-tolerant pollen. We 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

**All Paralogs of the *A. thaliana* Tandem T-Oleosin Gene Cluster Expressed.** The arrangement of the nine T-oleosin paralogs in chromosome 5 is shown in Fig. 1 (http://www.arabidopsis.org/ in 2013), which includes refinements from those reported in 2002 (9) and 2004 (10, 11). Each of the nine T-oleosin paralogs produced transcripts, as revealed by RT-PCR (9). The transcripts were present only in the tapetum at a midstage of anther development before the tapetum underwent programmed cell death (9). This developmental specificity was supported by microarray results (Fig. S1, from http://www.weigelworld.org/).

The transcript levels of the nine 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. S1), which is an approximation of quantity. In http://mpss.udel.edu/at/, the massively parallel signature sequencing (MPSS) transcript-per-million values in inflorescence were 17,006, 13,644, 1,738, 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 sequencing by synthesis (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 our RT-PCR analysis (Fig. S2). Overall, the reported order of levels of transcripts was *T3* ≥ *T5* ~ *T4* ~ *T6* > other paralogs.

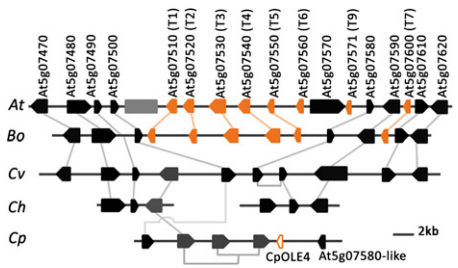


**Fig. 1.** Presence of the T-oleosin gene cluster in Brassicaceae but absence in Cleomaceae and Caricaceae. Primary transcripts are indicated with filled arrows (orange for oleosin genes and black for nonoleosin genes) in genomic DNA sequences from *A. thaliana* (*At*) and *Brassica oleracea* (*Bo*) of Brassicaceae, *C. violacea* (*Cv*), and *C. hassleriana* (*Ch*) of Cleomaceae, and *C. 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.

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 (9) and on the pollen surface of wild-type Ws2 (12, 20). Similar findings of T3-ortholog oleosin being more abundant than other individual T-oleosins in tapetosomes and pollen coat were obtained in *Brassica napus* (13, 18) and *Brassica rapa* (19). The higher amount of T3 oleosin determined visually by SDS/PAGE could be due to its high molecular mass (53 kDa) compared with T5 (10 kDa), T4 (22 kDa), and T6 (15 kDa) oleosins, as well as an 8-kDa 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 on 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.** Our observations did not suggest that individual paralogs of the T-oleosin gene cluster contributed to functional divergence. Thus, we tested whether the individual paralogs conferred quantitative benefits rather than functional neutrality. We focused on the three 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 transfer DNA (T-DNA) mutant, *ΔT5*, had a knockout of *T5* transcript (*Methods* and Fig. S2). In *ΔT5*, the tapetosomes observed with transmission electron microscopy (TEM) were of similar sizes but had numerous electron-transparent structures compared with those in the wild type (Fig. 2*A*). We used siRNA to knock down *T1*, *T3*, *T6*, and/or *T4* in *ΔT5* (Fig. S2). 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*A*). Another transformant, *si-5*/*ΔT5*, possessed obscure structures that resembled fragmented tapetosomes (Fig. S2). A different T-DNA mutant, *ΔT6*, had knockout of *T6* oleosin transcript (*Methods*) and defective tapetosomes (Fig. S2) similar to those of *ΔT5* (Fig. 2*A*).

With a wild-type Ws2 background, a T-DNA mutant, *ΔT3*, had knockout of *T3* transcript (20). In *ΔT3*, the tapetosomes were smaller and did not contain numerous internal electron-dense droplets, compared with those in the wild type (Fig. 2*A*). The tapetum cells and the pollen surface were viewed with confocal laser scanning microscopy (CLSM) (Fig. 2*B*). The tapetosomes in the wild type contained both lipid droplets (stained with Nile Red) and flavonoids [stained with diphenylboric acid 2-aminoethyl ester (DPBA)], as previously reported (14). In *ΔT3*, the tapetosomes were small (containing both lipid droplets and flavonoids), and some lipid droplets 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*B*), 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. We tested the mutant pollen for dehydration tolerance resulting from the loss of coat materials. We placed the pollen grains at 25% relative humidity and 24 °C for 1–3 h and allowed them to germinate on agar medium (Fig. 2*C*). Pollen of either of the two 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*C*). Among the various mutants, there was a correlation in the levels of loss of T-oleosin transcripts (Fig. S2), tapetosome structures, and pollen hydration tolerance (Table 1).

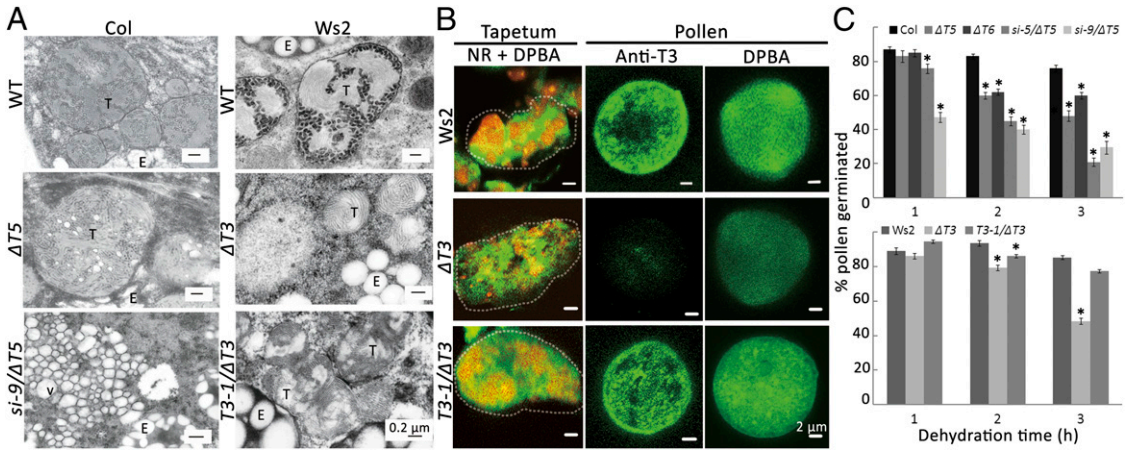


**Fig. 2.** Effect of mutational loss of T-oleosins on morphological features of tapetosomes and pollen dehydration tolerance in *Arabidopsis*. (*A*) TEM of tapetosomes in WT (Col and Ws2), T-DNA mutants (*ΔT5* and *ΔT3*), T-oleosin silenced mutant (*si-9/ΔT5*), and *T3* complementation mutant (*T3-1/ΔT3*). T, E, and V, tapetosome, elaioplast (part of one elaioplast packed with electron-transparent spherical plastoglobules), and aggregated vesicles, respectively. (Scale bars, 0.2 μm.) (*B*) CLSM of tapetum cells and pollen of Ws2, *ΔT3*, and *T3-1/ΔT3*. Samples were detected for T3 oleosin (with antibodies, pseudocolor in green), lipids [with Nile Red (NR), in red; many scattered lipid droplets (*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 of Ws2, *ΔT3*, and *T3-1/ΔT3*. White dotted line marks the circumference of the cell. (Scale bars, 2 μm.) (*C*) Test of dehydration tolerance of pollen from WT and T-oleosin mutants. Pollen was placed at 25% relative humidity and 24 °C for 1–3 h and allowed to germinate. Significant difference was calculated from the germination % of each mutant compared with WT (*t* test, $*P \leq 0.0001$). TEM images of *Si-9/ΔT5* and *Si-5/ΔT5* are shown in *A* and Fig. S2, respectively.

**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 *A* and *B*), pollen coat T3 oleosin, and flavonoids (Fig. 2*B*), as well as pollen tolerance to dehydration (Fig. 2*C* and Table 1). Therefore, the tapetum and pollen phenotypes of *ΔT3* resulted from loss of the *T3* gene.

**Diverse Plant Families, Including Cleomaceae, the Family Closest to Brassicaceae, Did Not Possess the T-Oleosin Gene Cluster, Tapetum Oleosins, and Tapetosomes.** We surveyed plant families with four approaches to obtain complementary information: detection of (*i*) tapetosomes with TEM, (*ii*) the T-oleosin gene cluster via public genomic information and our own sequencing effort, (*iii*) oleosin transcripts in anthers at the mid stage of development, and (*iv*) transcript-deduced oleosin sequences similar to those of Brassicaceae T-oleosins.

We used TEM to survey diverse plant orders (Fig. S3) 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. S3*A*). Species in families other than Brassicaceae had no tapetosomes (Fig. S3*B*). Instead, lipid droplets 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 we have examined. We 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 (24), contained solitary lipid droplets and vesicles dispersed in the cytoplasm (Fig. 3*A*). 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. S3*B*).

We 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*, we selected several 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 our expectations. We then subjected *Cleome* genomic DNA to Illumina sequencing and obtained two segments that had the highest similarity with the Brassicaceae T-oleosin gene cluster region (Fig. 1). Each of these two segments contained two 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 two segments. Assembly of the Illumina reads and performance of PCR with genomic DNA failed to link the two segments. Apparently, the two fragments were not closely linked. Other less-similar *Cleome* segments are shown in Fig. S4; the presence of these segments could reflect the early triplication of the *Cleome* genome. Additionally, 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/), we obtained the segment that most resembled the Brassicaceae T-oleosin gene cluster region (Fig. 1); 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 three different families in Brassicales. Apparently *Cleome* and *Carica* do not have the T-oleosin gene cluster.

When we 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 we obtained a continuous DNA segment (Fig. 1) with high resemblance of the *Arabidopsis* T-oleosin cluster region but without an oleosin gene. This finding reinforces our conclusion that Cleomaceae lacks the T-oleosin gene cluster.

We explored whether T-oleosin transcripts corresponding to the *Arabidopsis* T-oleosin transcripts of a defined temporal

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

| Species/line | Presence of tapetosome (Yes/No) | Phenotype of tapetosome | Pollen germination (%)* |
|---|---|---|---|
| *Arabidopsis* | | | |
| WT (Col) | Yes | >2 μm | 76 ± 2.0[†] |
| ΔT6 | Yes | >2 μm; with clear holes | 60 ± 2.5[†‡] |
| ΔT5 | Yes | >2 μm; with clear holes | 48 ± 2.2[†‡] |
| Si9/ΔT5 | No | Large aggregates of vesicles | 29 ± 3.0[†‡] |
| Si5/ΔT5 | Yes | <0.2 μm fragmented tapetosomes | 21 ± 2.2[†‡] |
| WT (Ws2) | Yes | 1–2 μm | 85 ± 1.4[†] |
| ΔT3 | Yes | <0.5 μm | 48 ± 1.9[†‡] |
| T3-1/ΔT3 | Yes | 0.8–1 μm | 78 ± 1.6[†‡] |
| *Cleome* | | | |
| WT | No | Lipid droplets separated from vesicles | 15 ± 2.0[§] |
| T3-3/Ch | Yes/No | Primitive tapetosomes | 70 ± 1.7[‡§] |

*Germination after dehydration for 3 h for *Arabidopsis* and 4 h for *Cleome*. Details are in Figs. 2 and 3.
[†]For *Arabidopsis*, for each set of experiments, mutants are compared with the WT, Col, or Ws2.
[‡]Significant difference from the germination % of WT (*t* test, $P < 0.0001$).
[§]For *Cleome*, the T3 transformed mutant is compared with the WT.

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 (9) (Fig. S1). This differential developmental pattern occurred because the tapetum matures and undergoes programmed cell death before the pollen matures (25). We 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. S1). Rice, with available anther transcriptomes (26), was used for further comparison. Rice also had pollen oleosin transcripts appearing at a late stage of anther

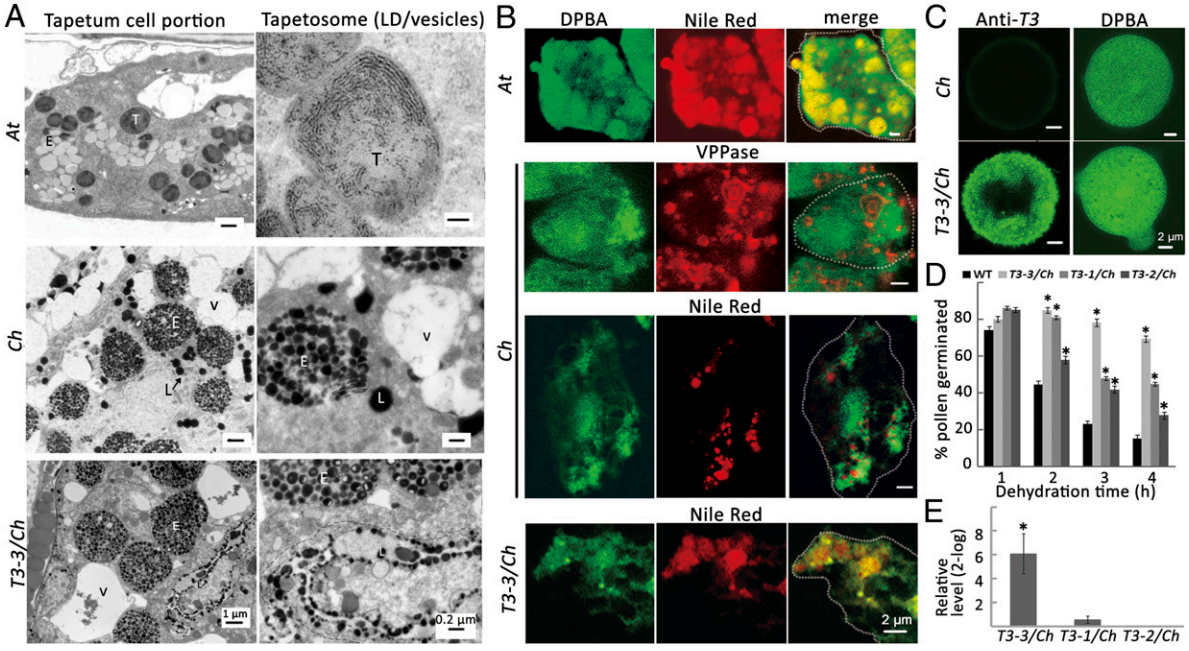


**Fig. 3.** Effect of transformation of *C. hassleriana* with *Arabidopsis* T3-oleosin gene. (*A*) TEM of tapetum cells in *Arabidopsis* and *Cleome* WT and *T3-3/Ch* (*Cleome* transformed with *Arabidopsis T3* gene). T, E, V, and L, tapetosome, elaioplast (packed with electron-dense plastoglobules), vacuole, and lipid droplet, respectively. Lipid droplets 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. (Scale bars, 1 μm on *Left*, 0.2 μm on *Right*.) (*B*) Immuno CLSM of tapetum cells of *Arabidopsis* and *Cleome* WT and *T3-3/Ch*. Samples were detected for flavonoids (with DPBA, pseudocolor in green), lipid droplets [with Nile Red, in red; many scattered lipid droplets (*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. (Scale bars, 2 μm.) (*C*) Immuno CLSM of pollen of *Cleome* WT and *T3-3/Ch*. Samples were detected for T3 oleosin (with antibodies, pseudocolor 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*. (Scale bars, 2 μm. (*D*) Test of dehydration tolerance of pollen of *Cleome* WT and three transformants, *T3-3/Ch*, *T3-1/Ch*, and *T3-2/Ch*, with decreasing levels of *T3* gene expression (*E*). Pollen was placed at 25% relative humidity and 24 °C for 1–4 h and allowed to germinate. Significant difference was calculated from the germination % of mutant compared with WT (*t* test, $*P < 0.0001$). (*E*) Quantitative real-time RT-PCR of *T3* transcript in the three 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.

development but no oleosin transcript at an earlier stage. Rice and maize do not have tapetum tapetosomes (26, 27).

Finally, we constructed a phylogenetic tree based on oleosin amino acid sequences and examined whether *Cleome* and *Carica*, as well as rice, possessed oleosins similar to those of Brassicaceae T-oleosins (Fig. S5). In the tree, *Arabidopsis* T-oleosins form 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.

***C. hassleriana* Transformed with an *Arabidopsis* T-Oleosin Gene Produced Primitive Tapetosomes and Dehydration-Tolerant Pollen.** We tested whether *Cleome* transformed with the *Arabidopsis T3* oleosin gene would generate the downstream gene activities and have adaptive benefit. In nontransformed *Cleome*, lipid droplets and flavonoid-containing vesicles dispersed in the cytoplasm of tapetum cells (Fig. 3*A*). CLSM revealed that the lipid droplets (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 colocate in the cytoplasm (Fig. 3*B*). This arrangement differed from that in Brassicaceae, in which the lipid droplets and flavonoid-containing vesicles colocated in tapetosomes (Fig. 3 *A* and *B*). In the transformed *Cleome* line, *T3-3/Ch*, lipid droplets were not dispersed in the cytoplasm but all attached to the surface of membranes/vesicles to form packages several micrometers wide (Fig. 3*A*). CLSM revealed that the lipid droplets and flavonoid-containing vesicles colocated in packages. We considered these packages primitive tapetosomes. Nontransformed *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*C*). Nontransformed *Cleome* pollen was susceptible to dehydration treatment, with germination decreased from 70% to 15% (Fig. 3*D* and Table 1). *T3-1/Ch* pollen became tolerant to the dehydration treatment (Fig. 3*D*). 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 *D* and *E*). The finding reiterates the quantitative aspect of T-oleosin in exerting adaptive advantage.

## 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. We reveal that paralogs of a tandem oleosin gene cluster in Brassicaceae exert no apparent functional divergence but confer quantitative benefit. In our studies, we overcame the traditional technical difficulties of lacking knockout mutants of tightly linked paralogs in a tandem cluster. We used T-DNA knockout mutants for each of the three most active oleosin paralogs plus siRNA to knock down additional paralogs. The loss of any one of the three 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% to 4% of the total inflorescence transcripts (MPSS data; *Results*), and the oleosin proteins could represent a similar percentage of the total proteins (*Results*). This high percentage of oleosin transcripts and proteins was obtained from findings with whole inflorescences, which included the tapetum, microspores, and numerous other cell types (25). The percentage could be as high as 10% in the tapetum cells, 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 present study, we 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 (20).

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 lipid droplets 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 lipid droplets (with surface oleosins) and vesicles associate ionically (14). Association of lipid droplets 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 (21).

We propose the evolution of oleosins in the tapetum of Brassicaceae as follows. After the partition of Brassicaceae and Cleomaceae >20 Mya (23), 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 Mya (21–23), because *A. thaliana* has only one tandem T-oleosin gene cluster. The new oleosin in the ancestor Brassicaceae species aggregated lipid droplets 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. We 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 we 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 (28) remains to the elucidated.

## Methods

**Plant Materials.** Seeds of *A. thaliana* wild type (Col-0 and Ws-2) and T-DNA insertion mutants *ΔT3* (CS11664) (20), *ΔT5* (CS104700), and *ΔT6* (SALK_134093) were obtained from the *Arabidopsis* Biological Research Center (Ohio State University). *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. Plants were grown at 22 °C in growth rooms with a 8-h/16-h day/night cycle. Flowering *C. hassleriana* from a local flower market were maintained in greenhouses at 26/18 °C. The name *C. hassleriana* (named earlier as *Cleome spinosa*) is used in this report because of its general recognition, even though some researchers have renamed it *Tarenaya hassleriana* (29). *C. 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 *A. thaliana* and *C. hassleriana*.** The constructs, pC1300/T3 (for *T3* complementation test, S1) and pHG/TOLE (for T-oleosin RNA silencing, S1) (Fig. S6), were transformed into *Arabidopsis ΔT3* and *ΔT5*, respectively, via floral dipping (30). 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. *Cleome* seeds were allowed to germinate on MS medium at 30/23 °C and a 16-h/8-h day/night cycle for 14 d. The hypocotyl was cut into 8- to 10-mm sections and immersed with Agrobacterium LBA4404 harboring the *pC2300/T3ole* construct (Fig. S6) 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 (24). After 4 wk, regenerated plantlets were transferred to MS medium with 0.1 mg·L$^{-1}$ 1-naphthaleneacetic acid (NAA) and carbenicillin. Transformed plants were identified by content of the *T3* gene fragment with PCR.

**Pollen Germination Test.** Pollen of *Arabidopsis* and *Cleome* were collected from open flowers and placed at 25% relative humidity and 24 °C. *Arabidopsis*

pollen grains were placed on 9-cm Petri dishes containing an 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] (31) at 28 °C for 1–3 h. *Cleome* pollen grains were placed on a glass slide with a liquid germination medium [10% (wt/vol) sucrose, 1.3 mM $H_3BO_3$, 2.9 mM $KNO_3$ and 9.9 mM $CaCl_2$, pH 5.8] at 28 °C (32) 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 the Lecia Automatic Freeze-Substitution System and embedded in LR Gold resin (Structural Probe). For chemical fixation, anthers were fixed with 2.5% glutaraldehyde, 4% paraformaldehyde, and 0.1 M K-phosphate (pH 7.0) at 4 °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 °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 (70–90 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 a 1:50 dilution of the IgG fraction, 1% (wt/vol) milk powder, and 1× PBS. Unless otherwise stated, each wash was performed with PBST [1× PBS and 0.05% (wt/wt) Tween-20] for 10 min. Tissues were fixed in 4% paraformaldehyde, 1× PBS, and 0.15 M sucrose at 4 °C for 16 h. For flavonoid localization with 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 (14). Samples were washed with 10% glycerol twice and stained with DPBA [saturated DPBA (0.5%, wt/vol), 0.01% Triton X-100, and 10% glycerol] at 20 °C for 2 h. Samples were washed twice with 10% glycerol. For double labeling with DPBA for flavonoids and Nile Red for lipids, the samples after the above 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 above DPBA staining were blocked with 3% milk in 1× PBS at 20 °C for 1 h. They were treated with rabbit IgG against V-PPase (Agrisera) in 1% milk and 1× PBS at 20 °C for 2 h. After two washes, samples were treated with a mixture of DPBA and secondary antibodies (0.5% DPBA, 0.01% Triton X-100, 10% glycerol, 1× 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 500–530 nm, 565–615 nm, and 650–750 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). 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 S1), and the results were normalized with Actin transcript as reference.

**Sequencing of *Cleome* DNA.** *Cleome* leaves were ground in liquid nitrogen, and the resulting powder was extracted for DNA with preheated extraction buffer [100 mM Tris·HCl (pH 7), 1.4 M NaCl, 20 mM EDTA, 2% cetyltrimethylammonium bromide, 1% polyvinylpyrrolidone, 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 Yourgene Bioscience. Sequences of the assembled fragments of *C. hassleriana* (KC777373, KC777374) and *C. papaya* (KC777372) have been deposited in the National Center for Biotechnology Information.

**ACKNOWLEDGMENTS.** We thank Drs. Wen-Dar Lin and Caroline Hsing for advice on database studies; Ms. Lin-Yun Kuang for advice on *Cleome* transformation; Drs. Daphne Preuss, Aretha Fiebig, Anna Dobritsa, and Jake Mayfield for *Arabidopsis grp*17 (named *ΔT3* mutant in this report, which was eventually obtained from the Arabidopsis Biological Resource Center); Dr. Eric Schranz for a BAC library of *C. hassleriana*; and Dr. Shin-Han Shiu for advice on tandem paralog evolution. This work was supported by a US Department of Agriculture National Research Initiative grant, the Institute of Plant and Microbial Biology, Academia Sinica, and a National Science Council grant.

1. Birchler JA (2012) Insights from paleogenomic and population studies into the consequences of dosage sensitive gene expression in plants. *Curr Opin Plant Biol* 15(5): 544–548.
2. Ding Y, Zhou Q, Wang W (2012) Origins of new genes and evolution of their novel functions. *Annu Rev Ecol Evol Syst* 43:345–363.
3. 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.
4. Piertney SB, Oliver MK (2006) The evolutionary ecology of the major histocompatibility complex. *Heredity (Edinb)* 96(1):7–21.
5. Duboule D (2007) The rise and fall of Hox gene clusters. *Development* 134(14): 2549–2560.
6. Jiang SY, Ma Z, Ramachandran S (2010) Evolutionary history and stress regulation of the lectin superfamily in higher plants. *BMC Evol Biol* 10:79.
7. Leister D (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. *Trends Genet* 20(3):116–122.
8. 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(2):993–1003.
9. 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(25):22677–22684.
10. 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(9):3286–3291.
11. 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(4):659–669.
12. Mayfield JA, Fiebig A, Johnstone SE, Preuss D (2001) Gene families from the *Arabidopsis thaliana* pollen coat proteome. *Science* 292(5526):2482–2485.
13. 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(6):889–899.
14. 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(2):582–596.
15. Huang AHC (1992) Oil bodies and oleosins in seeds. *Annu Rev Plant Physiol Plant Mol Biol* 43:177–200.
16. Hsieh K, Huang AHC (2004) Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. *Plant Physiol* 136(3):3427–3434.
17. Frandsen GI, Mundy J, Tzen JT (2001) Oil bodies and their associated proteins, oleosin and caleosin. *Physiol Plant* 112(3):301–307.
18. Wu SSH, et al. (1997) Isolation and characterization of neutral-lipid-containing organelles and globuli-filled plastids from *Brassica napus* tapetum. *Proc Natl Acad Sci USA* 94(23):12711–12716.
19. Ting JTL, Wu SSH, Ratnayake C, Huang AHC (1998) Constituents of the tapetosomes and elaioplasts in *Brassica campestris* tapetum and their degradation and retention during microsporogenesis. *Plant J* 16(5):541–551.
20. Mayfield JA, Preuss D (2000) Rapid initiation of *Arabidopsis* pollination requires the oleosin-domain protein GRP17. *Nat Cell Biol* 2(2):128–130.
21. Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2002) Brassicales. *Plant Systematics* (Sinauer Associates, Sunderland, MA), 2nd Ed, pp 402–405.
22. Schranz ME, Mitchell-Olds T (2006) Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell* 18(5):1152–1165.
23. Barker MS, Vogel H, Schranz ME (2009) Paleopolyploidy in the Brassicales: Analyses of the *Cleome* transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales. *Genome Biol Evol* 1:391–399.
24. Newell CA, et al. (2010) *Agrobacterium tumefaciens*-mediated transformation of *Cleome gynandra* L., a $C_{(4)}$ dicotyledon that is closely related to *Arabidopsis thaliana*. *J Exp Bot* 61(5):1311–1319.
25. Goldberg RB, Beals TP, Sanders PM (1993) Anther development: Basic principles and practical applications. *Plant Cell* 5(10):1217–1229.
26. 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(2):694–707.
27. 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(4):1548–1561.
28. Tantikanjana T, Nasrallah ME, Nasrallah JB (2010) Complex networks of self-incompatibility signaling in the Brassicaceae. *Curr Opin Plant Biol* 13(5):520–526.
29. Iltis 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.
30. Clough SJ, Bent AF (1998) Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16(6):735–743.
31. 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 tip-localized calcium influx. *Plant Cell* 11(9):1731–1742.
32. Wang H, Jiang L (2011) Transient expression and analysis of fluorescent reporter proteins in plant pollen tubes. *Nat Protoc* 6(4):419–426.