# Evolution and development of orchid flowers and fruits

Anita Dirks-Mulder

# Evolution and development of orchid flowers and fruits

Anita Dirks-Mulder

2020

Anita Dirks-Mulder, 2020. Evolution and development of orchid flowers and fruits

**Ph.D. Thesis** at Leiden University, the Netherlands, 2020

**Cover design:** Anita Dirks-Mulder.
**Cover photograph:** 3D image of an *Erycina pusilla* fruit, 4 weeks after pollination
**Layout:** Anita Dirks-Mulder
**Printed by:** Proefschriftmaken.nl



The research described in this thesis was financially supported as indicated in each chapter.

# Evolution and development of orchid flowers and fruits

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 5 februari 2020
klokke 15.00 uur

door

Anita Dirks-Mulder
geboren te ‘s Gravenhage
in 1965

**Promotor**
Prof. dr. Erik F. Smets
Naturalis Biodiversity Center, Leiden University & KU Leuven, Belgium

**Second promotor**
Prof. dr. Barbara Gravendeel
Radboud University, University of Applied Sciences Leiden & Naturalis Biodiversity Center

**Promotion committee**
Prof. dr. Gilles P. van Wezel (chairman)
Leiden University

Prof. dr. Remko Offringa (secretary)
Leiden University

Prof. dr. Gerco C. Angenent
Wageningen University & Research

Prof. dr. Annette Becker
Justus-Liebig-Universität Gießen, Germany

Prof. dr. Vera van Noort
Leiden University & KU Leuven, Belgium

Prof. dr. Philipp Schlüter
Universität Hohenheim, Germany

Ik ben bijna waar ik zijn moet
Bijna waar ik zijn moet
Bijna op mijn plaats
Die ruimte is van mij
En mocht ik het niet halen
Dan was ik toch dichtbij
Ik ben bijna waar ik zijn moet

(BLØF, Bijna waar ik zijn moet)

**Contents**

# Chapter 1

## General introduction

## Evolution of flowering plants

Flowering plants or angiosperms dominate the terrestrial flora. With more than 350,000 species they make up about 90% of all living plant species. This was not always the case, as flowering plant first evolved 150 million years ago (mya) from non-flowering plants (Soltis *et al*., 2008;Doyle, 2012;Brown and Smith, 2018;Magallon *et al*., 2019). The non-flowering and thus more ancient group of living seed plants, commonly known as gymnosperms contain just over 1000 species. The gymnosperms possess male and female cones, while the angiosperms evolved several key innovations such as the flowers containing stamens (male reproductive organs) and carpels (female reproductive organs with ovules) surrounded by sepals and petals.

Understanding the mode and mechanisms of angiosperm evolution is a central challenge of the field of plant evolutionary developmental biology (evo-devo). Explanation of the apparently ''sudden'' origin ~150 mya and the early diversification of the angiosperms, as revealed by fossils, proved to be difficult because fossils with reproductive structures intermediate between ancestral gymnosperms and angiosperms are lacking. The sudden appearance of angiosperms in the fossil record puzzled many scientists including Charles Darwin. In 1879 he wrote a letter to Joseph Hooker, the director of the Royal Botanic Gardens in Kew, UK, stating that ''The rapid development, as far as we can judge, of all the higher plants within recent geological time is an abominable mystery" (Darwin *et al*., 1903). Molecular dating methods suggest that flowering plants are much more ancient (Magallon, 2010;Magallon *et al*., 2019), estimating the origin of crown group angiosperms between 140 and 130 mya (Crane *et al*., 1995;Smith *et al*., 2010;Doyle, 2012;2014) but the current discrepancy between the oldest fossils and age estimates of molecular clock analyses is probably caused by the fact that the fossil record is far from complete.

Sauquet *et al*. (2017) made a reconstruction of the earliest common ancestor of flowering plants by modeling the distribution of floral organs present in modern angiosperms on their phylogeny. These authors concluded that the ancestral flower, like most modern flowers, was bisexual, had multiple whorls of petal-like organs arranged in concentric circles similar to a modern *Magnolia* flower. This ancestral flower went through a series of simplifications, in which organs were reduced or merged until it settled on an optimal and stable morphology. Once this basal morphology had evolved, it further diversified in selected clades into for instance bilateral symmetry. The latter is believed to optimize efficiency of interaction with pollinators (Sauquet *et al*., 2017). It is not known which animals might have eaten or pollinated the presumed ancestral flowers. Studies from fossilized dinosaur dung from 100.5-113.0 mya show that these animals ate angiosperms (Vajda *et al*., 2016).

Angiosperm flowers exhibit an enormous diversity, but usually contain four floral organs: sepals and petals - forming the perianth - stamens and carpels. The evolution of reward systems for animal-pollinated flowers allowed for species diversifications in many different clades by a high variety of pollinating animals and other vectors. Within the angiosperms, orchids are one of the most species-rich

families. They are very diverse in terms of floral shape, size and color. Reproductive isolation of many orchid species occurred through highly specialized interactions with pollinators. Unique floral organs, such as modified sepals and petals, a callus on a modified median petal (the lip) and a gynostemium with wing-like structures (stelidia), allow efficient pollen transfer via specific body parts of pollinators. Orchid flowers are composed of five whorls of three parts each, including two perianth whorls (sepals and petals), two staminal whorls (gynostemium and stelidia) and one carpel whorl (**Figure 1**).

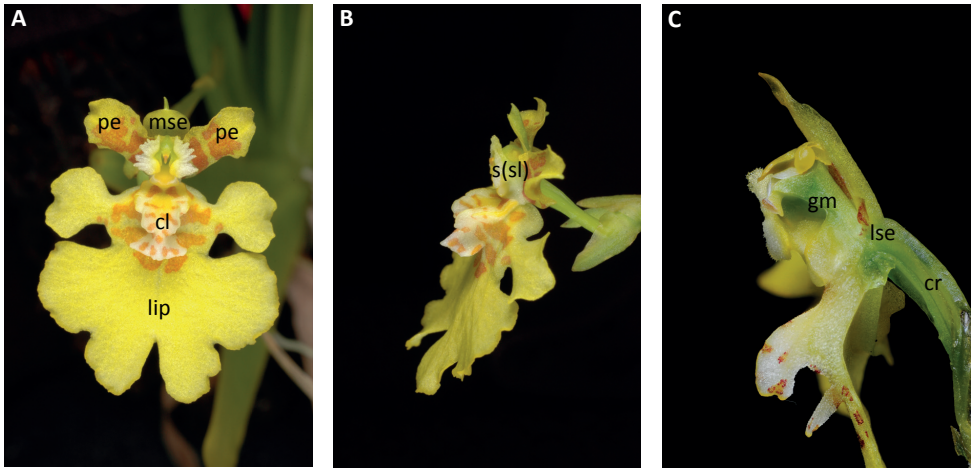



**Figure 1. Flower of the orchid species *Erycina pusilla*. (A)** Frontal view. **(B)** Lateral view. **(C)** Transversal section through the flower (one petal and part of the lip were removed). cl = callus; mse = median sepal; lse = lateral sepal; pe = petal; gm=gynostemium; cr = carpel; s(sl) = stelidium (photos by Joel McNeal (A,B) and Jean Claessens (C)).

"Why are orchids so diverse?" is a question that scientists have been wondering about for many centuries. Charles Darwin for instance wrote an entire book about the various contrivances by which orchids are fertilized by insects (Darwin, 1862). To explain the abominable mystery of the origin and evolution of orchids, similar to the angiosperms in general, fossil surveys and molecular clock analyses were carried out (**Figure 2**). Ramirez *et al*. (2007) dated a fossil orchid pollinarium, carried by a worker bee preserved in Dominican amber from 15-20 mya, and concluded that the most recent ancestor of extant orchids lived 76-84 mya (**Figure 2 C-D)** . Another fossilized orchid pollinarium, this time carried by a fungal gnat preserved in Baltic amber from 45-55 mya (Poinar and Rasmussen, 2017), further confirmed this estimate of the origin of the orchids. Radiation of orchids began around 73 mya (Magallon *et al*., 2019). Multiple hypotheses exist about the main drivers behind the high diversity of modern orchid species. These include: (i) the evolution of highly specific pollination interactions, in which pollinia are deposited on very specific body parts of a few species of pollinators only, (ii) symbiotic associations with species-specific groups of mycorrhizal fungi (important for germination and seedling development), (iii) colonization of different epiphytic habitats and (iiii)

the development of multiple types of photosynthesis including Crassulacean Acid Metabolism (Gravendeel *et al*., 2004;Silvera *et al*., 2009;Givnish *et al*., 2015). All these factors likely contributed to the high diversity of orchids observed today.

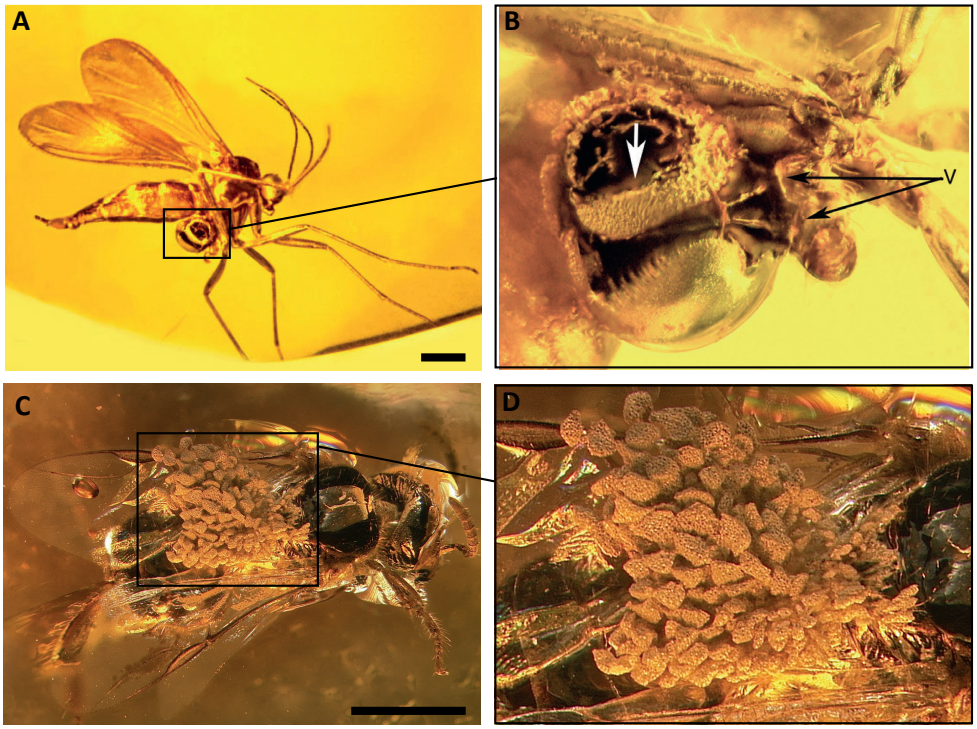



**Figure 2**. **Undisputed fossils of orchid flowers. (A-B)** *Succinanthera baltica* (Poinar and Rasmussen, 2017). **(C-D)** *Meliorchis carribea* (Ramirez *et al*., 2007).White arrow = pollinium. V = viscidia. Scale bars: 1 mm.

## Floral organ identity genes

Moyroud *et al*. (2017) eventually solved part of the puzzle of how angiosperm flowers evolved by studying the gymnosperm *Welwitschia mirablis*, a gymnosperm species with separate male and female reproductive structures organized in cones. These authors studied the genetic circuits that control the development of *Welwitschia* reproductive units and compared these circuits to those active in the cones. They discovered that the same developmental genes play not only a central role in the development of flowers but also in the cones of this gymnosperm species. A similar genetic cascade was found in angiosperms and two other gymnosperm genera: *Pinus* and *Picea*, indicating that this cascade was inherited from their last common ancestor. The results of this study show that flowers did not appear all of a sudden but that the developmental genes involved were probably already present, being inherited and reused during plant evolution. I will now summarize what we currently know about the most

important developmental genes driving angiosperm and orchid flower evolution.

The majority of the genes involved in floral organ identity belong to the family of MADS-box transcription factors. Their different interactions, resulting in the different floral organs, was first explained by the "ABC" model in *Arabidopsis thaliana* and *Antirrhinum majus* (Coen and Meyerowitz, 1991). According to this model, the development of sepals, petals, stamens and carpels are specified by class A, B and C MADS-box genes. Mutations in each class exhibit homeotic transformations of organ identity in two adjacent floral whorls: A class mutants have sepals transformed into carpels and petals into stamens; B class mutants have petals transformed into sepals and stamens into carpels; and C class mutants have stamens transformed into petals and carpels transformed into sepals. In Arabidopsis, genes corresponding to the three classes have been well characterized, *APETALA1 (AP1)* and *APETALA2 (AP2)* represent the A class, *APETALA3 (AP3)* and *PISTILLATA (PI)* the B class and *AGAMOUS (AG)* the C class. Later the ABC model was extended by D class MADS-box genes *SHATTERPROOF (SHP)* and *SEEDSTICK (STK)*, involved in ovule development (Angenent and Colombo, 1996), and E class MADS-box genes *SEPALLATA (SEP)* (Theissen, 2001), needed for petal, stamen and carpel development. MADS is actually an acronym derived from the initials of four loci: MCMI of the yeast *Saccharomyces cerevisiae*, AG of Arabidopsis, DEF of Antirrhinum and SRF of humans (*Homo sapiens*). MADS-box genes can be found in all eukaryotes and while the human genome contains only a few of these genes, most angiosperm genomes contain more than a hundred. Martinez-Castilla and Alvarez-Buylla (2003) recovered 104 MADS-box genes from the Arabidopsis genome, which can be divided in M-type and MIKC-type based on their protein domain structure (**Figure 3**). The M-type proteins only have the MADS-domain in common and are involved in seed and female gametophyte development (Masiero *et al*., 2011), while the MIKC-type genes share four conserved domains (Alvarez-Buylla *et al*., 2000;Henschel *et al*., 2002;Nam *et al*., 2004). The MADS (M) domain contains around 60 amino acids and is involved in DNA binding and protein dimerization. The intervening (I) and keratin-like (K) domains are critical for dimerization and tetramerization with other MADS-domain proteins. The C-terminal (C) region contains short, highly conserved clade specific motifs and is involved in the formation of higher-order protein complexes (Riechmann *et al*., 1996;Honma and Goto, 2001;Smaczniak *et al*., 2012).

Tetrameric protein complexes of MIKC-type proteins, according to the Floral Quartet Model, specify the identity of different floral organs. The different quartets probably function as transcription factors of the DNA of the target genes. By activating or repressing these genes, the quartets control the development of the floral organs. For example, AP3, PI, AG and SEP MIKC-type proteins form a quartet that controls the development of stamens and AP3, PI, AP1 and SEP proteins form a quartet that controls the development of petals (Theissen and Saedler, 2001;Smaczniak *et al*., 2012).

Duplications, followed by sub-functionalization, have been suggested to lead to several homologous and paralogous lineages in different plant groups outside the core eudicots. For example the B-class genes, which are

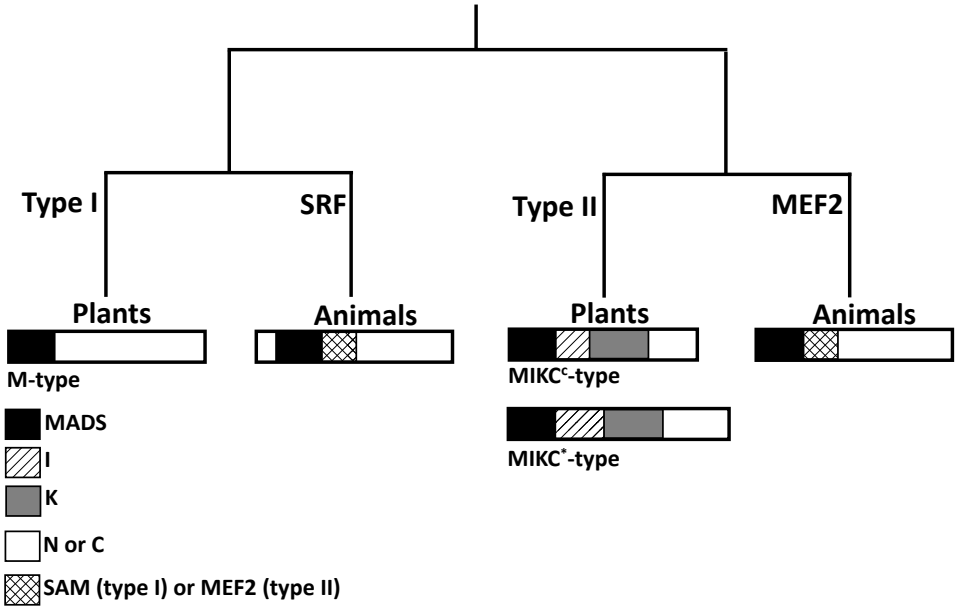



**Figure 3. Domain structures of type I and II MADS-box genes in plants and animals.** Adapted from (Nam *et al*., 2004).

highly conserved in members of the core eudicots, including the model species Arabidopsis, Antirrhinum and Petunia (Angenent *et al*., 1992;Jack *et al*., 1992;Kramer, 1998) were subjected to several duplication events during plant evolution. Before the origin of the angiosperms, gene duplications not only created the *AP3* and *PI* lineages, which are present in all angiosperms, but occurred also in the other MADS-box gene classes (Litt and Irish, 2003;Kramer *et al*., 2004). Duplications, followed by sub-functionalization, are suggested to have led to several homologous and paralogous lineages in different plant groups outside the core eudicots. Duplications of an ancestral MADS-box B-class gene were found to be responsible for the creation of euAP3 and TM6 lineages in core eudicots and B-sister lineages in the gymnosperm *Gnetum gnemon* (Becker *et al*., 2002;Gioppato and Dornelas, 2018) and monocots (Chang *et al*., 2010;Yang *et al*., 2012).

In the orchid lineage, additional duplications have occurred in the B-class lineage, resulting in more (sub-functionalized) genes that may be in part responsible for the enormous flower diversity in the orchid family. The class B MADS-box genes are central to the specification of petal and stamen identity. Most eudicots have two B-class genes, *AP3/DEF-like* and *PI/GLO-like*. However, gene duplication events in orchids have generated several paralogs, in particular of AP3-like genes. In case of *AP3*, a duplication event first gave rise to the AP3A and AP3B clades. Further duplication resulted in four sub-clades, namely AP3B1 (clade 1), AP3B2 (clade 2), AP3A1 (clade 3) and AP3A2 (clade 4). These four sub-clades are found in different orchid subfamilies. In addition to the ABCDE model, the "Orchid Code" and "Homeotic Orchid Tepal (HOT) model" were proposed, both describing the expression of AP3

genes during orchid perianth formation. According to these models, B-class genes are expressed in floral whorl two (petals), following the ABCDE model, but extended to whorl 1 (sepals) when the morphology of the sepals and petals is more or less similar, with expression of a lip-specific copy restricted to this organ (Mondragon-Palomino and Theissen, 2008;Pan *et al*., 2011).

## *Erycina pusilla* as emergent model system for orchid evo-devo research

Arabidopsis (Brassicaceae, Eudicots) is by far the most popular flowering plant model system and widely studied already for over a century. It was the first species of which the genome was fully sequenced because this is relatively small (The Arabidopsis Genome, 2000). Discoveries in this species have proven to be widely applicable to many other plant species. However, based on our knowledge of Arabidopsis only not all plant developmental processes can be understood. The flowers of the monocot orchid species for instance, are very different from the Arabidopsis flowers. Several modifications had to be made to the ABCDE model in the form of the "Orchid code" and "HOT" models as discussed above, and new models discussed and proposed in this PhD thesis.

During a visit in 2006 to Elena Kramer of Harvard University in the United States, my co-promoter Barbara Gravendeel decided to develop a model system for evo-devo research on orchids. Various orchid species were carefully considered for this. What makes a plant a good plant model system? Based on the Field Guide to Plant Model Systems by Chang *et al*. (2016), different properties should be taken into account. For laboratory use, small sized plants, which are easy to culture, with a short generation time and high fecundity are preferential. The capability of self-fertilization for maintenance, being susceptibility to genetic manipulations such as crossing and mutagenesis by for instance UV-irradiation, chemicals, a small-sized diploid genome (preferably fully sequenced and annotated) and the ability to manipulate gene function are more intrinsic properties of a plant species to make it suitable as model system. When laboratory procedures are standardized for e.g. gene transformation, more research institutions will use the model and develop community properties, such as stock centers with genetic strains, reporter gene constructs but also databases with gene annotations, sequences which can be downloaded, and bio-informatics tools (e.g. Blast, gene search, GO and KEGG pathways, chromosome map tools).

Most orchids have long life cycles (~3–5 years), large chromosome numbers and complex genomes, which make functional studies difficult. The final choice therefore fell on the meso-American twig epiphytic species *Erycina pusilla* (Epidendroideae, Oncidiinae), which is easy to maintain and propagate *in vitro*. This species has a low diploid chromosome number (2n, n=6); a relatively small sized genome (1C = 1.5 pg), a short juvenile phase (less than a year from seed to flowering stage) and can complete its life cycle *in vitro* (Chase *et al*., 2005;Felix and

Guerra, 2012;Dirks-Mulder *et al*., 2017). An on-line transcriptome database for *E. pusilla* is available (Chao *et al*., 2017), as well as a protocol for transformation with Agrobacterium, although the efficiency is still low and published only once (Lee *et al*., 2015). Several labs in the world are now using this emergent orchid model, for instance to study MADS-box gene evolution (Lin *et al*., 2016) and the genetic basis of crassulacean acid metabolism (Heyduk *et al*., 2018). In my PhD project, I tried to answer the question 'Why are orchids so diverse?' by carrying out evo-devo research with *E. pusilla*. Below, I will summarize the main results.

## Aim and outline of this PhD thesis

The goal of this PhD project is to gain more insight into the evolutionary development of orchid flowers and fruits by studying the emergent orchid model plant *E. pusilla* with a combination of micro- and macromorphological, molecular and phylogenetic techniques. Research on orchid flowers is described in **chapters 2 and 3.** In **chapter 2** the formation of the lip, a highly modified petal present in most orchid flowers, is described (Gravendeel and Dirks-Mulder, 2015) in line with work from Hsu *et al.* (2015), who proposed the "P-code" model, we found that two different developmental gene complexes are involved in either sepal/petal or lip formation. In **chapter 3** (Dirks-Mulder *et al*., 2017) the evolutionary origin is investigated of three other highly specialized orchid floral organs of *E. pusilla*: the median petaloid sepal, the callus on the lip, and the stelidia along the gynostemium. We discovered that these organs are derived from a sepal, a stamen that gained petal identity, and stamens that became staminodes, respectively. The "Oncidiinae" model was proposed, explaining the duplications, diversifying selection and changes in spatial expression of different MADS-box genes that shaped the sepals, petals and lip, enabling the rewardless flowers of *E. pusilla* to mimic an unrelated rewarding flower for pollinator attraction.

Once an orchid flower is pollinated, the inferior ovary, which is composed of three carpels, develops into a fruit. In **chapter 4** (Dirks-Mulder *et al*., 2019) this process is described in detail for *E. pusilla* from pollination of the flower up to dehiscence of the capsule. Morphological analyses were also carried out on fruits of two other orchid species: *Cynorkis fastigiata* and *Epipactis helleborine* to find further support for the "split carpel" model as proposed by Rasmussen and Johansen (2006). The fruit associated MADS-box genes and proteins together with other dehiscence-related genes were analyzed for *E. pusilla* in order to propose a first "orchid fruit developmental protein and gene network" model.

To further study fruit development of *E. pusilla*, in **chapter 5** (unpublished results) transcriptome analyses are presented that were obtained from different developmental phases as defined in **chapter 4**. The final step in fruit development is the shattering of the seeds, which in most cases are dispersed by wind after the fruit dehisces. Lignification of the fruit valves is generally known to be important for fruit dehiscence but is this also the case for orchid fruits? Also in **chapter 5** this topic

was further studied by analyzing the anatomy of ripe fruits of a total of 41 orchid species from all over the world and investigating possible correlations between fruit valve lignification patterns, life form, growth strategy, ecology, fruit orientation, dehiscence type, number of valves and slits, and phylogenetic relationships. In **chapter 6** the results from the preceding chapters are summarized and discussed and implications for future research and implementation of *E. pusilla* as a plant model system are provided.

## References


Alvarez-Buylla, E.R., Pelaz, S., Liljegren, S.J., Gold, S.E., Burgeff, C., Ditta, G.S., Ribas De Pouplana, L., Martinez-Castilla, L., and Yanofsky, M.F. (2000). An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proceedings of the National Academy of Sciences* 97**,** 5328-5333.

Angenent, G.C., Busscher, M., Franken, J., Mol, J.N., and Van Tunen, A.J. (1992). Differential expression of two MADS box genes in wild-type and mutant petunia flowers. *Plant Cell* 4**,** 983-993.

Angenent, G.C., and Colombo, L. (1996). Molecular control of ovule development. *Trends in Plant Science* 1**,** 228-232.

Becker, A., Kaufmann, K., Freialdenhoven, A., Vincent, C., Li, M.A., Saedler, H., and Theissen, G. (2002). A novel MADS-box gene subfamily with a sister-group relationship to class B floral homeotic genes. *Mol Genet Genomics* 266**,** 942-950.

Brown, J.W., and Smith, S.A. (2018). The Past Sure is Tense: On Interpreting Phylogenetic Divergence Time Estimates. *Syst Biol* 67**,** 340-353.

Chang, C., Bowman, J.L., and Meyerowitz, E.M. (2016). Field Guide to Plant Model Systems. *Cell* 167**,** 325-339.

Chang, Y.Y., Kao, N.H., Li, J.Y., Hsu, W.H., Liang, Y.L., Wu, J.W., and Yang, C.H. (2010). Characterization of the possible roles for B class MADS box genes in regulation of perianth formation in orchid. *Plant Physiol* 152**,** 837-853.

Chao, Y.T., Yen, S.H., Yeh, J.H., Chen, W.C., and Shih, M.C. (2017). Orchidstra 2.0-A Transcriptomics Resource for the Orchid Family. *Plant Cell Physiol* 58**,** e9.

Chase, M.W., Hanson, L., Albert, V.A., Whitten, W.M., and Williams, N.H. (2005). Life history evolution and genome size in subtribe Oncidiinae (Orchidaceae). *Ann Bot* 95**,** 191-199.

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353**,** 31-37.

Crane, P.R., Friis, E.M., and Pedersen, K.R. (1995). The origin and early diversification of angiosperms. *Nature* 374**,** 27-33.

Darwin, C., Darwin, F.S., and Seward, A.C. (1903). *More letters of Charles Darwin. A record of his work in a series of hitherto unpublished letters.* London: J. Murray.

Darwin, C.R. (1862). On the various Contrivances by which British and foreign orchids are fertilized by insects, and on the good effects of intercrossing. By Charles Darwin, M.A., F.R.S. London: John Murray. 12mo. 1862. *The Annals and magazine of natural history; zoology, botany, and geology* 10**,** 384-388.

Dirks-Mulder, A., Ahmed, I., Uit Het Broek, M., Krol, L., Menger, N., Snier, J., Van Winzum, A., De Wolf, A., Van't Wout, M., Zeegers, J.J., Butot, R., Heijungs, R., Van Heuven, B.J., Kruizinga, J., Langelaan, R., Smets, E.F., Star, W., Bemer, M., and Gravendeel, B. (2019). Morphological and Molecular Characterization of Orchid Fruit Development. *Front Plant Sci* 10**,** 137.

Dirks-Mulder, A., Butot, R., Van Schaik, P., Wijnands, J.W., Van Den Berg, R., Krol, L., Doebar, S., Van Kooperen, K., De Boer, H., Kramer, E.M., Smets, E.F., Vos, R.A., Vrijdaghs, A., and Gravendeel, B. (2017). Exploring the evolutionary origin of floral organs of Erycina pusilla, an emerging orchid model system. *BMC Evol Biol* 17**,** 89.

Doyle, J.A. (2012). Molecular and Fossil Evidence on the Origin of Angiosperms. *Annual Review of Earth and Planetary Sciences* 40**,** 301-326.

Doyle, J.A. (2014). Recognising angiosperm clades in the Early Cretaceous fossil record. *Historical Biology* 27**,** 414-429.

Felix, L.P., and Guerra, M. (2012). Chromosome analysis in Psygmorchis pusilla (L.) Dodson & Dressier: the smallest chromosome number known in Orchidaceae. *Caryologia* 52**,** 165-168.

Gioppato, H.A., and Dornelas, M.C. (2018). When Bs Are Better than As: the Relationship between B-Class MADS-Box Gene Duplications and the Diversification of Perianth Morphology. *Tropical Plant Biology* 12**,** 1-11.

Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Iles, W.J., Clements, M.A., Arroyo, M.T., Leebens-Mack, J., Endara, L., Kriebel, R., Neubig, K.M., Whitten, W.M., Williams, N.H., and Cameron, K.M. (2015). Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc Biol Sci* 282.

Gravendeel, B., Smithson, A., Slik, F.J., and Schuiteman, A. (2004). Epiphytism and pollinator specialization: drivers for orchid diversity? *Philos Trans R Soc Lond B Biol Sci* 359**,** 1523-1535.

Henschel, K., Kofuji, R., Hasebe, M., Saedler, H., Munster, T., and Theissen, G. (2002). Two ancient classes of MIKC-type MADS-box genes are present in the moss Physcomitrella patens. *Mol Biol Evol* 19**,** 801-814.

Heyduk, K., Hwang, M., Albert, V., Silvera, K., Lan, T., Farr, K., Chang, T.H., Chan, M.T., Winter, K., and Leebens-Mack, J. (2018). Altered Gene Regulatory Networks Are Associated With the Transition From C3 to Crassulacean Acid Metabolism in Erycina (Oncidiinae: Orchidaceae). *Front Plant Sci* 9**,** 2000.

Honma, T., and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* 409**,** 525-529.

Hsu, H.-F., Hsu, W.-H., Lee, Y.-I., Mao, W.-T., Yang, J.-Y., Li, J.-Y., and Yang, C.-H. (2015). Model for perianth formation in orchids. *Nature Plants* 1**,** 15046.

Jack, T., Brockman, L.L., and Meyerowitz, E.M. (1992). The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is expressed in petals and stamens. *Cell* 68**,** 683-697.

Kramer, E.M. (1998). Molecular Evolution of Genes Controlling Petal and Stamen Development: Duplication and Divergence Within the APETALA3 and PISTILLATA MADS-Box Gene Lineages. *Genetics* 149**,** 765-783.

Kramer, E.M., Jaramillo, M.A., and Di Stilio, V.S. (2004). Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. *Genetics* 166**,** 1011-1023.

Lee, S.-H., Li, C.-W., Liau, C.-H., Chang, P.-Y., Liao, L.-J., Lin, C.-S., and Chan, M.-T. (2015). Establishment of an Agrobacterium-mediated genetic transformation procedure for the experimental model orchid Erycina pusilla. *Plant Cell Tiss Organ Cult (PCTOC)* 120**,** 211-220.

Lin, C.S., Hsu, C.T., Liao, D.C., Chang, W.J., Chou, M.L., Huang, Y.T., Chen, J.J., Ko, S.S., Chan, M.T., and Shih, M.C. (2016). Transcriptome-wide analysis of the MADS-box gene family in the orchid Erycina pusilla. *Plant Biotechnol J* 14**,** 284-298.

Litt, A., and Irish, V.F. (2003). Duplication and diversification in the APETALA1/FRUITFULL floral homeotic gene lineage: implications for the evolution of floral development. *Genetics* 165**,** 821-833.

Magallon, S. (2010). Using fossils to break long branches in molecular dating: a comparison of relaxed clocks applied to the origin of angiosperms. *Syst Biol* 59**,** 384-399.

Magallon, S., Sanchez-Reyes, L.L., and Gomez-Acevedo, S.L. (2019). Thirty clues to the exceptional diversification of flowering plants. *Ann Bot* 123**,** 491-503.

Martinez-Castilla, L.P., and Alvarez-Buylla, E.R. (2003). Adaptive evolution in the Arabidopsis MADS-box gene family inferred from its complete resolved phylogeny. *Proc Natl Acad Sci U S A* 100**,** 13407-13412.

Masiero, S., Colombo, L., Grini, P.E., Schnittger, A., and Kater, M.M. (2011). The emerging importance of type I MADS box transcription factors for plant reproduction. *Plant Cell* 23**,** 865-872.

Mondragon-Palomino, M., and Theissen, G. (2008). MADS about the evolution of orchid flowers. *Trends Plant Sci* 13**,** 51-59.

Moyroud, E., Monniaux, M., Thevenon, E., Dumas, R., Scutt, C.P., Frohlich, M.W., and Parcy, F. (2017). A link between LEAFY and B-gene homologues in Welwitschia mirabilis sheds light on ancestral mechanisms prefiguring floral development. *New Phytol* 216**,** 469-481.

Nam, J., Kim, J., Lee, S., An, G., Ma, H., and Nei, M. (2004). Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms. *Proc Natl Acad Sci U S A* 101**,** 1910-1915.

Pan, Z.J., Cheng, C.C., Tsai, W.C., Chung, M.C., Chen, W.H., Hu, J.M., and Chen, H.H. (2011). The duplicated B-class MADS-box genes display dualistic characters in orchid floral organ identity and growth. *Plant Cell Physiol* 52**,** 1515-1531.

Poinar, G., and Rasmussen, F.N. (2017). Orchids from the past, with a new species in Baltic amber. *Botanical Journal of the Linnean Society* 183**,** 327-333.

Ramirez, S.R., Gravendeel, B., Singer, R.B., Marshall, C.R., and Pierce, N.E. (2007). Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature* 448**,** 1042-1045.

Rasmussen, F.N., and Johansen, B. (2006). Carpology of Orchids. *Selbyana* 27**,** 44-53.

Riechmann, J.L., Krizek, B.A., and Meyerowitz, E.M. (1996). Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. *Proc Natl Acad Sci U S A* 93**,** 4793-4798.

Sauquet, H., Von Balthazar, M., Magallon, S., Doyle, J.A., Endress, P.K., Bailes, E.J., Barroso De Morais, E., Bull-Herenu, K., Carrive, L., Chartier, M., Chomicki, G., Coiro, M., Cornette, R., El Ottra, J.H.L., Epicoco, C., Foster, C.S.P., Jabbour, F., Haevermans, A., Haevermans, T., Hernandez, R., Little, S.A., Lofstrand, S., Luna, J.A., Massoni, J., Nadot, S., Pamperl, S., Prieu, C., Reyes, E., Dos Santos, P., Schoonderwoerd, K.M., Sontag, S., Soulebeau, A., Staedler, Y., Tschan, G.F., Wing-Sze Leung, A., and Schonenberger, J. (2017). The ancestral flower of angiosperms and its early diversification. *Nat Commun* 8**,** 16047.

Silvera, K., Santiago, L.S., Cushman, J.C., and Winter, K. (2009). Crassulacean acid metabolism and epiphytism linked to adaptive radiations in the Orchidaceae. *Plant Physiol* 149**,** 1838-1847.

Smaczniak, C., Immink, R.G., Angenent, G.C., and Kaufmann, K. (2012). Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development* 139**,** 3081-3098.

Smith, S.A., Beaulieu, J.M., and Donoghue, M.J. (2010). An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proc Natl Acad Sci U S A* 107**,** 5897-5902.

Soltis, D.E., Bell, C.D., Kim, S., and Soltis, P.S. (2008). Origin and early evolution of angiosperms. *Ann N Y Acad Sci* 1133**,** 3-25.

The Arabidopsis Genome, I. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* 408**,** 796.

Theissen, G. (2001). Development of floral organ identity: stories from the MADS house. *Curr Opin Plant Biol* 4**,** 75-85.

Theissen, G., and Saedler, H. (2001). Plant biology. Floral quartets. *Nature* 409**,** 469-471.

Vajda, V., Pesquero Fernández, M.D., Villanueva-Amadoz, U., Lehsten, V., and Alcalá, L. (2016). Dietary and environmental implications of Early Cretaceous predatory dinosaur coprolites from Teruel, Spain. *Palaeogeography, Palaeoclimatology, Palaeoecology* 464**,** 134-142.

Yang, X., Wu, F., Lin, X., Du, X., Chong, K., Gramzow, L., Schilling, S., Becker, A., Theissen, G., and Meng, Z. (2012). Live and let die - the B(sister) MADS-box gene OsMADS29 controls the degeneration of cells in maternal tissues during seed development of rice (Oryza sativa). *PLoS One* 7**,** e51435.

# Chapter 2

## Floral development: Lip formation in orchids unraveled

*Nature Plants 1; 15056 (2015)*

Barbara Gravendeel[1,2,3] and Anita Dirks-Mulder[1,2]

[1]Endless Forms group, Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands

[2]Faculty of Science and Technology, University of Applied Sciences Leiden, Zernikedreef 11, 2333 CK Leiden, The Netherlands

[3]Institute of Biology Leiden, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands

Most orchid flowers have an enlarged median petal, the 'lip', which plays a crucial role in attracting pollinators. The existence and appearance of this organ is due to the presence of specific protein complexes involved in floral development, which are differentially expressed in orchid species with more or less pronounced lips. Attracting pollinators for reproduction is a major challenge for plant species. As plants generally cannot move, they evolved flowers to attract pollinators by employing dazzling colors, alluring odors and tactile cues.

Flowers are composed of several whorls. The outer whorl is made up of sepals that generally serve as protection, while the second whorl consists of petals involved in pollinator attraction. Sepals and petals together form the perianth enfolding the male and female reproductive organs. Orchids are unusual in that their floral morphology includes petal-like sepals, brightly colored and ornamented, containing special surface ornaments to attract insects. Furthermore, the median petal is transformed into a lip that acts as the main attractive organ (**Figure 1**). Hsu *et al.* (2015) described how competition between two protein complexes determines the development of this peculiar lip organ. This greatly extends our understanding of the mechanisms leading to the diverse forms of orchid flowers.

**Figure 1. Flower of *Ophrys splendida* with *Andrena* male.** The bee clings to the lip, which mimics a female bee in appearance, fragrance and physical touch. The petals and sepals are bent backwards, ensuring maximal contact between the yellow pollen packages on the bee's head and the female reproductive part of the flower that it attempts to mate with [Photo by Jean Claessens].

The only subfamily of orchids without a pronounced lip is the most basal and least diverse lineage. Therefore the evolution of a lip clearly triggered speciation and can be seen as a key innovation. To dissect the molecular basis of lip development, studies have been conducted on the MADS-box genes involved in promoting petal identity, which revealed a number of duplication events that occurred after orchids started to diversify, more than 60 million years ago (Ramirez *et al*., 2007). These duplication events appear to have generated copies of MADS-box genes with new functions. Hsu *et al.* (2015) comprehensively examined the expression of all known MADS-box genes of the A, B and E classes in the flowers of certain orchid species. They found that

different copies of duplicated MADS-box genes (*AP3* in the B-class and *AGL6* in the E-class) showed different tissue-specific or tissue-biased expression (Hsu *et al*., 2015).

Based on this expression pattern, Hsu *et al.* (2015) proposed the 'Perianth code' to link gene expression with petal/lip identity. According to the perianth code, two protein complexes, the 'L' complex and the 'SP' complex, each produced by four MADS-box genes, compete to promote the formation of lip and petals, respectively. Fluorescence resonance energy transfer (FRET) analyses support the existence of the two complexes. The relative expression of the two determines to what extent, in some orchids, the lip is more or less differentiated from the sepals and petals.

Hsu *et al.*(2015) validated the perianth code in peloric floral mutants. Such a mutant of *Oncidium* 'Gower Ramsey' exhibits complete transformation of petals into lip-like structures. As predicted, it lost the expression of the SP complex and only expressed the L complex in the petals. Similarly, the peloric mutants displaying partial petal-to-lip conversion showed reduced SP expression but elevated L complex expression in the petals. The researchers further tested the perianth code in *Oncidium* species with different types of perianth conversion. Species with lip-like petals or petal-like lips showed co-existence of the SP and L units or absence of both complexes in petals or lips. Hence, the antagonism between the two complexes successfully explains the diverse flower forms across multiple orchid species and cultivars.

Expression analyses only provide part of the evidence for this hypothesis. To substantiate the function of the perianth code, Hsu *et al.* (2015) manipulated the balance between the two complexes using virus induced gene silencing (VIGS), a technique employing agrobacterium strains containing a modified virus for silencing specific genes. The *OAGL6-2* gene encoding a component of the L complex was down regulated by VIGS in *Oncidium* and *Phalaenopsis* orchids, resulting in the conversion of lips to sepal/petal structures in both scenarios. The resulting phenotypes represent solid functional support for the perianth code.

The ABCDE model of floral development describes interactions between specific MADS-box proteins determining the identities of flower organs (Coen and Meyerowitz, 1991). One protein complex consisting of class A and class E MADS-box proteins specifies sepals, whereas another composed of A, B and E class proteins controls petals. According to the floral quartet model, a combination of four proteins forms whorl-specific complexes (Theissen and Saedler, 2001). These complexes function by either activating or silencing target genes during floral development. Both models apply to many flowering plant families. However, for the highly specialized sepals, petals and lip in the first and second whorl of most orchid flowers, the perianth code model (Hsu *et al*., 2015) proves to be a better analogy.

Hsu *et al.* (2015) have shown that two protein complexes produced by A/E and B class genes determine which organs develop in the orchid perianth. When the balance between both complexes is shifted by partially silencing the one involved in lip formation, only sepal- and petal-like organs develop. This experiment nicely shows that during orchid evolution two copies of duplicated developmental genes acquired a new function by specifically promoting lip identity.

In 1877, Charles Darwin published an extensive overview of highly specialized organs in orchid flowers and speculated about their role in maximizing pollination success. It is now assumed that the evolution of these organs provided a strong selection advantage allowing one-third of all orchid species, distributed over many unrelated clades, to be fertilized without offering any food reward to their pollinators (Cozzolino and Widmer, 2005). This makes attraction very challenging as pollinators quickly learn to avoid such flowers. The discovery of the protein complexes involved in lip formation in orchid flowers is intriguing, and provides an important step towards fully understanding how orchids continue to lure pollinators despite the lack of a material benefit. Unraveling the genetic basis of other highly specialized floral organs will undoubtedly provide further insights into the contribution of duplicated and neo-functionalized developmental genes to the evolutionary arms race between orchids and their pollinators.

## References

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353**,** 31-37.

Cozzolino, S., and Widmer, A. (2005). Orchid diversity: an evolutionary consequence of deception? *Trends in Ecology & Evolution* 20**,** 487-494.

Darwin, C. (1877). *The various contrivances by which orchids are fertilised by insects.* London: John Murray.

Hsu, H.-F., Hsu, W.-H., Lee, Y.-I., Mao, W.-T., Yang, J.-Y., Li, J.-Y., and Yang, C.-H. (2015). Model for perianth formation in orchids. *Nature Plants* 1**,** 15046.

Ramirez, S.R., Gravendeel, B., Singer, R.B., Marshall, C.R., and Pierce, N.E. (2007). Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature* 448**,** 1042-1045.

Theissen, G., and Saedler, H. (2001). Plant biology. Floral quartets. *Nature* 409**,** 469-471.

# Chapter 3

## Exploring the evolutionary origin of floral organs of *Erycina pusilla*, an emerging orchid model system

*BMC Evolutionary Biology 17, 89 (2017)*

Anita Dirks-Mulder[1,2], Roland Butôt[1], Peter van Schaik[2], Jan Willem P.M. Wijnands[2], Roel van den Berg[2], Louie Krol[2], Sadhana Doebar[2], Kelly van Kooperen[2], Hugo de Boer[1,7,8], Elena M. Kramer[3], Erik F. Smets[1,6], Rutger A. Vos[1,4], Alexander Vrijdaghs[6] & Barbara Gravendeel[1,2,5]

[1]Endless Forms group, Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands

[2]Faculty of Science and Technology, University of Applied Sciences Leiden, Zernikedreef 11, 2333 CK Leiden, The Netherlands

[3]Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Ave, MA 02138, Cambridge, USA

[4]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

[5]Institute Biology Leiden, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands

[6]Ecology, Evolution and Biodiversity Conservation cluster, KU Leuven, Kasteelpark Arenberg 31, 3001 Leuven, Belgium

[7]The Natural History Museum, University of Oslo, P.O. Box 1172 Blindern, 0318 Oslo, Norway

[8]Department of Organismal Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, SE-75236, Sweden

## Abstract

Thousands of flowering plant species attract pollinators without offering rewards, but the evolution of this deceit is poorly understood. Rewardless flowers of the orchid *Erycina pusilla* have an enlarged median sepal and incised median petal ('lip') to attract oil-collecting bees. These bees also forage on similar looking but rewarding Malpighiaceae flowers that have five unequally sized petals and gland-carrying sepals. The lip of *E. pusilla* has a 'callus' that, together with winged 'stelidia', mimics these glands. Different hypotheses exist about the evolutionary origin of the median sepal, callus and stelidia of orchid flowers.

The evolutionary origin of these organs was investigated using a combination of morphological, molecular and phylogenetic techniques to a developmental series of floral buds of *E. pusilla*. The vascular bundle of the median sepal indicates it is a first whorl organ but its convex epidermal cells reflect convergence of petaloid features. Expression of *AGL6 EpMADS4* and *APETALA3 EpMADS14* is low in the median sepal, possibly correlating with its petaloid appearance. A vascular bundle indicating second whorl derivation leads to the lip. *AGL6 EpMADS5* and *APETALA3 EpMADS13* are most highly expressed in lip and callus, consistent with current models for lip identity. Six vascular bundles, indicating a stamen-derived origin, lead to the callus, stelidia and stamen. *AGAMOUS* is not expressed in the callus, consistent with its sterilization. Out of three copies of *AGAMOUS* and four copies of *SEPALLATA*, *EpMADS22* and *EpMADS6* are most highly expressed in the stamen. Another copy of *AGAMOUS*, *EpMADS20*, and the single copy of *SEEDSTICK*, *EpMADS23*, are most highly expressed in the stelidia; suggesting *EpMADS22* may be required for fertile stamens.

The median sepal, callus and stelidia of *E. pusilla* appear to be derived from a sepal, a stamen that gained petal identity, and stamens, respectively. Duplications, diversifying selection and changes in spatial expression of different MADS-box genes shaped these organs, enabling the rewardless flowers of *E. pusilla* to mimic an unrelated rewarding flower for pollinator attraction. These genetic changes are not incorporated in current models and urge for a rethinking of the evolution of deceptive flowers.

**Keywords**
Deceptive pollination, floral development, MADS-box genes, mimicry, vascular bundles

## Introduction

Flowering plants interact with a wide range of other organisms including pollinators. Pollinators can either receive nectar, oil, pollen or shelter in return for pollen transfer in a rewarding relationship, or nothing at all in a deceptive relationship (Cho *et al*., 1999). One of the deceptive strategies is mimicry, defined as the close resemblance of one living organism, 'the mimic', to another, 'the model', leading to misidentification by a third organism, 'the operator'. Essential for mimicry is the production of a false signal (visual, olfactory and/or tactile) that is used to mislead the operator, resulting in a gain in fitness of the mimic (Cho *et al*., 1999). Mimicry in plants generally serves the purpose of attraction of pollinators to facilitate fertilization. In these cases, an unrewarding plant species mimics traits typical for co-flowering models, such as a specific floral shape, coloration, and presence of nectar guides, glands, trichomes or spurs. In this way, pollinators, that are unable to distinguish the two types of flowers from each other, are fooled (Cho *et al*., 1999;Roy and Widmer, 1999). Despite the fact that deceptive pollination evolved in thousands of plant species, most notably orchids (Ackerman *et al*., 2011), the mechanisms by which this deceit evolved are still poorly understood.

Flowers are the main attractors of the majority of angiosperms to gain attention of pollinators. The outer first whorl of a flower is usually made up of sepals that generally serve as protection covering the other floral parts until anthesis. The outer second whorl consists of often-showy petals mainly involved in pollinator attraction. The sepals and petals together enfold the male and female reproductive organs in the inner floral whorls. Over the past decades, evolutionary developmental (evo-devo) studies have yielded many new insights in the role of duplication and neo-functionalization of developmental genes in floral diversification and the evolution of sepals, petals and male and female reproductive organs. These studies helped redefine the evolutionary origin of such organs (Preston *et al*., 2011).

Theoretically, an orchid flower can be considered to consist of five whorls of floral organs. Three sepals and three petals are present in the outer two whorls. Three external and three internal stamens and three carpels are present in the three inner whorls (**Figure 5a**). Studies of the genetic plant model species *Arabidopsis thaliana* have shown that genes only associated with petals in *A. thaliana* are also expressed in the first floral whorl of petaloid monocots including orchids. Expression of these genes in the first whorl of petaloid monocots plays an important role in the similarity of sepals and petals in lilies, gingers and orchids (Kanno *et al*., 2003;Nakamura *et al*., 2005;Kanno *et al*., 2007). From an evolutionary perspective, retention of expression of genes associated with petals in the outer floral whorl is considered an ancestral character for angiosperms (Soltis *et al*., 2007). In orchid flowers, the median petal, or 'lip', is often enlarged and ornamented with a wart-like structure, or 'callus'. The lip mostly functions as main attractor and landing platform for pollinators. Many hypotheses have been put forward about the evolutionary origin of the lip and its ornaments (Endress, 2016). Hsu *et al*. (2015) showed that the lip is homologous with true petals but gained an additional function possibly due to the duplication of a complex of modified developmental genes that gained novel expression domains.

A stamen usually consists of a filament and an anther where the pollen are produced. Many lineages in plant families such as buttercups, orchids, penstemons and witch-hazels, not only have fertile stamens but also rudimentary, sterile or abortive stamen-like structures. These structures are generally called staminodes and are often positioned between the fertile stamens and carpels, although they can also occur in other positions (Decraene and Smets, 2001). Multiple hypotheses exist about the function of the morphologically very diverse staminodes. In *Aquilegia*, staminodes play a role in protecting the early developing fruits, as they usually remain present after pollination long after the other organs have abscised (Kramer *et al*., 2007). In other plant genera, staminodes are assumed to mediate pollination. Comparative gene expression and silencing studies showed that staminode identity in *Aquilegia* evolved from a pre-existing stamen identity program. Of the genes involved, one lineage duplicated and one paralog became primarily expressed in the staminodia (Decraene and Smets, 2001;Kramer *et al*., 2007).

Characteristic for orchids is that the male and female reproductive organs are incorporated in a so-called 'gynostemium'. This structure is thought to result from a fusion of a maximum of six fertile to (partly) sterile stamens and parts of the pistil, in particular the style and stigma. It is a complex organ and the evolutionary origin of its different parts is not yet clear (Rudall and Bateman, 2002;Rudall *et al*., 2013;Endress, 2016). During the evolution of the orchids over the past 100 million years a reduction in the number of fertile stamens and fusion with the carpels occurred (McKnight and Shippen, 2004;Ramirez *et al*., 2007;Givnish *et al*., 2015). Six fertile stamens, positioned in floral whorls three and four, are commonly present in the closest relatives of the orchids in Asparagales. In the Apostasioideae, the earliest diverging of the five subfamilies of orchids, the number of fertile stamens is reduced to three in the genus *Neuwiedia*, one in floral whorl three and two in whorl four. In the genus *Apostasia,* a staminode develops in floral whorl three or nothing resulting in two fertile stamens (Kocyan and Endress, 2001). In subfamily Cypripedioideae only two fertile stamens are present. A further reduction into a single fertile stamen in floral whorl three evolved in subfamilies Vanilloideae, Orchidoideae and Epidendroideae (Rudall and Bateman, 2002). Since the two subfamilies with either three or two fertile stamens are the least diverse, reduction to a single fertile stamen may have contributed to species diversification. The sterile stamens have evolved into many other structures. In the majority of the Epidendroid orchids with a single fertile stamen, the mature gynostemium evolved appendages projecting to the front or side, clearly differentiating from broadened or flattened tissue at the base, that help pollinators to position themselves in the correct way to remove or deposit pollinia, which ensures pollination. The shapes of these appendages differ greatly and different terms are used to describe them, e.g. column wings or 'stelidia' (Vermeulen, 1959;Kurzweil, 1987;Kurzweil and Kocyan, 2002). The oldest hypothesis postulates that the stelidia are remnants of male reproductive tissue (Brown and Nees von Esenbeck, 1827;Darwin, 1877) and following this hypothesis, stelidia are interpreted as vestiges of the lateral stamens of the third and fourth floral whorls (Swamy, 1948).

**Current models explaining floral organ development**

The genetic basis of floral organ formation can be explained with various genetic models of MADS-box transcription factors. The core eudicot 'ABCDE model' included the A-class gene *APETALA1* (*AP1*), B-class genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*), C-class gene *AGAMOUS* (*AG*), D-class gene *SEEDSTICK* (*STK*) and E-class gene *SEPALLATA* (*SEP*). This model has been revised for the monocots to reflect two key differences: (i) there are no *AP1* orthologs outside the core eudicots so *FRUITFULL* (*FUL*)-like genes are the closest homologs, and (ii) many monocots have entirely petaloid perianths. Class A+B+E genes specify petaloid sepals, A+B+E control petals, B+C+E determine stamens, C+E specify carpels, and D+E are necessary for ovule development (Coen and Meyerowitz, 1991;Theissen, 2001;Theissen and Saedler, 2001) (**Figure 1a**). As in the core eudicots, these genetic combinations are thought to function as protein complexes, as proposed by Theissen and Saedler (2001) in the now well accepted 'floral quartet model' (**Figure 1b**). For the highly specialized flowers of most orchid lineages, further elaborations have been proposed, including the 'orchid code'(Mondragon-Palomino and Theissen, 2009;2011), 'Homeotic Orchid Tepal' (HOT) model (Pan *et al*., 2011) and 'Perianth code' (P-code) (Hsu *et al*., 2015).

The orchid code and HOT model (**Figure 1c**) postulate that the four *AP3* lineages in orchids have experienced sub- and neo-functionalization to give rise to distinct petal and lip identity programs. In addition to original MADS-box genes incorporated in the ABCDE model, several *AGAMOUS-LIKE-6* (*AGL6*) gene copies were recently found to play an important role in orchid flower formation. According to the P-code model (**Figure 1d**), there are two MADS-box protein complexes active in orchid flowers, one consisting of a set of *AP3/AGL6/PI* copies, specific for sepal/petal formation, and one consisting of another set of *AP3/AGL6/PI* copies, specific for the formation of the lip. When the ratio of these two complexes is skewed towards the latter, the lip is large. When the ratio is skewed towards the former, intermediate lip-structures are formed (Hsu *et al*., 2015). The P-code model has been functionally validated for wild type *Oncidium* and *Phalaenopsis*, and also for *Oncidium* peloric mutants, in which the two petals are lip-like. The P-code model was also validated in orchids from other subfamilies than the Epidendroideae, to which *Oncidium* and *Phalaenopsis* belong, i.e. Cypripedioideae, Orchidoideae and Vanilloideae, and used to detect gene expression profiles in species with intermediate lip formation (Hsu *et al*., 2015).

***Erycina pusilla* as an emergent orchid model: current resources and terminology**

MADS-box genes have now been identified for several commercially important orchid genera (e.g. *Cymbidium*, *Dendrobium, Oncidium* and *Phalaenopsis*) (Pan *et al*., 2011;Su *et al*., 2013;Cai *et al*., 2015) but long life cycles, large chromosome numbers and complex genomes of these genera hamper functional studies. DNA-mediated transformation can be used to study the function of orchid genes and *E. pusilla*, with its relatively short life cycle, functions as an emergent orchid model species for such studies (Lee *et al*., 2014;Lin *et al*., 2016).

*Erycina pusilla* belongs to the Oncidiinae, which is a highly diverse subtribe of meso- and south-American epiphytic orchids in subfamily Epidendroideae (Neubig

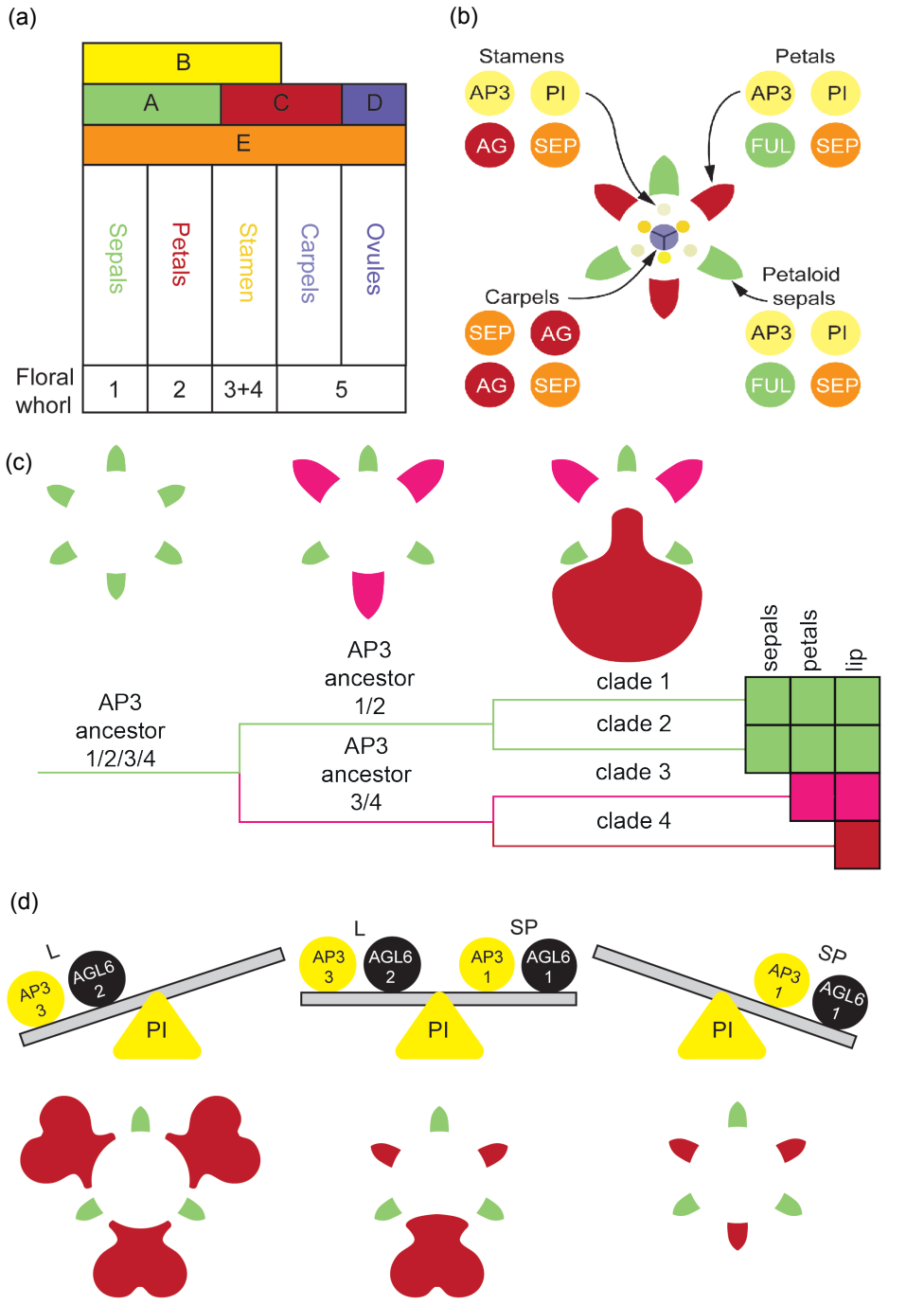



**Figure 1. Current models explaining floral organ development. (a)** ABCDE model of floral development in petaloid monocots. **(b)** Floral quartet model. **(c)** Orchid code and HOT model. **(d)** Perianth code model [Illustrations by Bas Blankevoort].

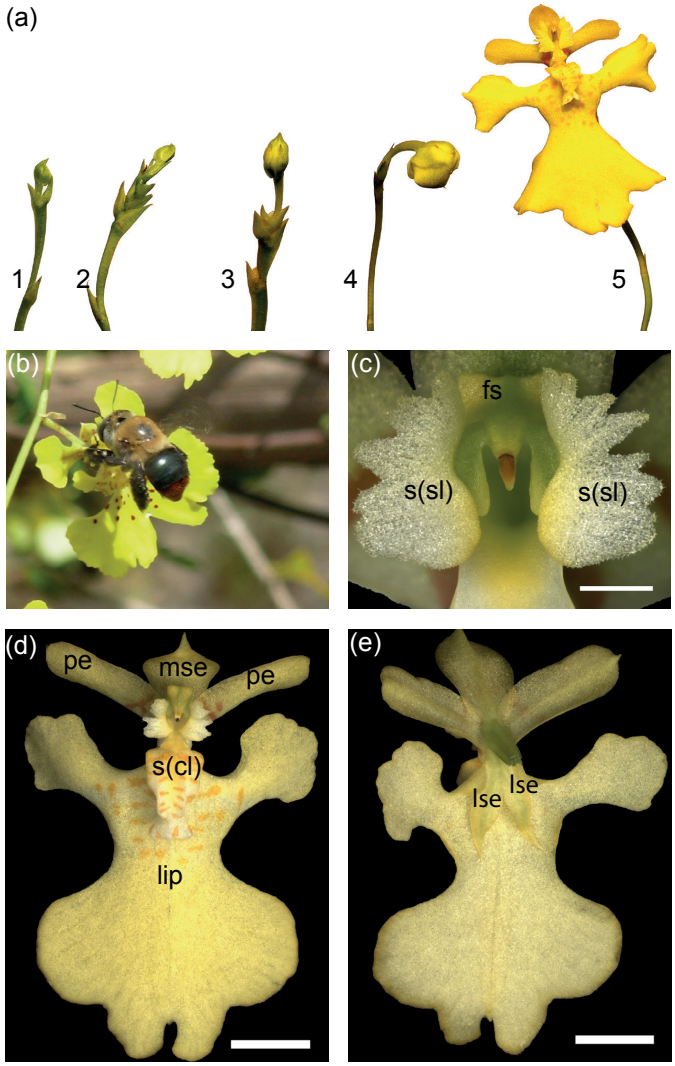



**Figure 2. General overview of *E. pusilla* flowers, pollinator and floral parts**. **(a)** Five floral stages of *E. pusilla* [Photo by Rogier van Vugt]. **(b)** A female *Centris poecila* bee pollinating a flower of *Tolumnia guibertiana*, a close relative of *E. pusilla*, in Cuba [Photo by Angel Vale], showing the function of the stelidia and callus in freshly opened flowers of these orchids, i.e. attraction and providing a holdfast for the pollinator. **(c)** Frontal view of fully developed stelidia. **(d)** Adaxial side (with respect to the floral axis) of a flower. (e) Abaxial side (with respect to the floral axis). Abbreviations: s(cl) = callus; lse = lateral sepal; mse = median sepal; pe = petal; s(sl) = stelidium; fs = fertile stamen.

*et al*., 2012). It is a rapidly growing orchid species with a low chromosome number (n = 6) and a, for orchids, relatively small sized diploid genome of 1.475 Gb (Chase *et al*., 2005;Felix and Guerra, 2012). It can be grown from seed to flowering stage in less than a year (Lee *et al*., 2014;Lin *et al*., 2016) and plantlets can be grown without mycorrhizae in test tubes. Flowers develop in a few days in which five distinct floral developmental stages can be observed (**Figure 2a**). The species produces deceptive

flowers that are self-compatible but incapable of spontaneous self-pollination.

Oil-collecting *Centris* bees are the main pollinators (Pridgeon *et al*., 2009). The lateral sepals of *E. pusilla* are small and green. The median sepal is larger and more colorful than the lateral sepals. The lip is the largest part of the flower and very different in shape compared to the lateral petals and sepals. On the basal part of the lip or 'hypochile', a callus is present that guides pollinators towards the stamen and stigma to either remove or deposit pollinia effectively. The gynostemium is enveloped on both sides by two large, wing-shaped structures that we further refer to as stelidia. During floral visits, *Centris* bees cling to these stelidia and the callus with their forelegs while searching for oils (**Figure 2b**). In *E. pusilla* however, these bees are fooled because the flowers employ food deception by Batesian mimicry by resembling flowers of rewarding species of the unrelated Malpighiaceae (Pridgeon *et al*., 2009;Vale *et al*., 2011;Papadopulos *et al*., 2013). Flowers of this family have five clawed petals that are often unequal in size. The sepals carry oil glands. It is generally assumed that the enlarged median sepal, incised lip, callus and stelidia of Oncidiinae evolved to mimic the shape of the petals and oil glands of rewarding flowers of Malpighiaceae (**Figure 2b-d** and **Figure 3**) in order to attract oil-collecting bees for pollination (Carmona-Díaz and García-Franco, 2008;Pridgeon *et al*., 2009;Neubig *et al*., 2012;Papadopulos *et al*., 2013).

*Agrobacterium*-mediated genetic transformation was recently developed for *E. pusilla* (Lee *et al*., 2014) and knockdown of genes is currently being optimized. It is expected that the entire genome will have been analyzed using a combination of next-generation sequencing techniques within the following years. Furthermore, transcriptome data of *E. pusilla* are included in the Orchidstra database (Su *et al*., 2013). Twenty-eight MADS-box genes from *E. pusilla* have been identified thus far including the most important floral developmental ones (Lin *et al*., 2016). These resources make *E. pusilla* an ideal orchid model for evo-devo studies. Lin *et al*. (2016) published expression data of MADS-box genes isolated from sepals, petals, lip, column and ovary of flowers of *E. pusilla* after anthesis together with a basic phenetic gene lineage analysis.

In this study, we employed a combination of micro-, macro-morphological, molecular and phylogenetic techniques to assess the evolutionary origin of the

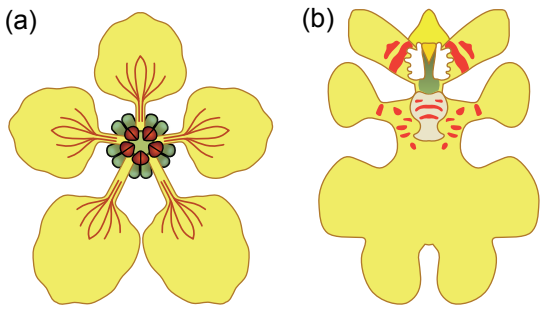



**Figure 3. Graphical representation of a flower** belonging to **(a)** Malpigiaceae and **(b)** Oncidiinae [Illustrations by Bas Blankevoort].

median sepal, callus and stelidia of the flowers of *E. pusilla*. To accomplish this goal, we investigated early and late floral developmental stages with scanning electron microscopy (SEM), light microscopy (LM), 3D-Xray microscopy (micro-CT) and expression (RT-qPCR) of MADS-box genes belonging to six different lineages. In addition, we investigated gene duplication and putative neo-functionalization as indicated by inferred episodes of diversifying selection. Our aim was to test the hypotheses that the median sepal, callus and stelidia are derived from sepals, petals and stamens, respectively, to unravel the genetic basis of the evolution of deceptive flowers.



## Material and methods

**Plant material and growth conditions**

A more than 15 year old inbred line of *E. pusilla* originally collected in Surinam was grown in climate rooms under controlled conditions (7.00 h – 23.00 h light regime), at a temperature of 20 °C and a relative humidity of 50%. The orchids were cultured *in vitro* under sterile conditions on Phytamax orchid medium with charcoal and banana powder (Sigma-Aldrich) mixed with 4 g/L Gelrite™ (Duchefa) culture medium. Pollinia of flowers from different plants were placed on each other's stigma after which ovaries developed into fruits. After 18-22 weeks, seeds were ripe and sown into containers with sterile fresh nutrient culture medium. The seeds developed into a new *E. pusilla* flowering plant within 20 weeks.

**Fixation for micromorphology**

Flowers and flower buds were fixed with standard formalin-aceto-alcohol (FAA: absolute ethanol, 90%; glacial acetic acid, 5%, formalin; 5% acetic acid) for one hour under vacuum pressure at room temperature and for 16 hours at 4 °C on a rotating platform. They were washed once and stored in 70% ethanol until further use.

**Scanning Electron Microscopy (SEM)**

Floral buds at different developmental stages were dissected in 70% ethanol under a Wild M3 stereo-microscope (Leica Microsystems AG, Wetzlar, Germany) equipped with a cold-light source (Schott KL1500; Schott-Fostec LLC, Auburn, New York, USA). Subsequently, the material was washed with 70% ethanol and then placed in a mixture (1:1) of 70% ethanol and DMM (dimethoxymethane) for five minutes for dehydration. The material was then transferred to 100% DMM for 20 minutes and critical point dried using liquid $CO_2$ with a Leica EM CPD300 critical point dryer (Leica Microsystems, Wetzlar Germany). The dried samples were mounted on aluminium stubs using Leit-C carbon cement or double-sided carbon tape and coated with Platina-Palladium with a Quorum Q150TS sputtercoater (Quorum Technologies, Laughton, East Sussex, UK). Images were obtained with a JEOL JSM-7600F Field Emission Scanning Electron Microscope (JEOL Ltd., Tokyo, Japan).

For the images presented in **Figure 4**, fixed floral buds were critical point dried using liquid $CO_2$ with a CPD 030 critical point dryer (BAL-TEC AG, Balzers, Lichtenstein) and

coated with gold with a SPI-ModuleTM Sputter Coater (SPI Supplies, West-Chester, Pennsylvania, USA). Scanning electron microscope (SEM) images were obtained with a Jeol JSM-6360 (JEOL Ltd., Tokyo) at the Laboratory of Plant Conservation and Population Biology (KU Leuven, Belgium).

**X-ray micro-computed tomography (micro-CT)**
Fully grown flowers were infiltrated with 1% phosphotungstic acid (PTA) in 70% ethanol for seven days in order to increase the contrast (Staedler *et al*., 2013). The PTA solution was changed every 1-2 days. The flowers were embedded in 1% low melting point agarose (Promega) prior to scanning. The scans were performed on a Zeiss Xradia 510 Versa 3D X-ray with a Sealed transmission 30-160 kV, max 10 W x-ray sources. Scanning was performed using the following settings: acceleration voltage/power 40 kV/3 W; source current 75 µA; exposure time 2 s; picture per sample 3201; camera binning 2; optical magnification 4 x, with a pixel size of 3.5 µm. The total exposure time was approximately 3,2 hours. 3D images were stacked and processed with Avizo 3D software version 8.1.

**RNA extraction**
For organ dissection, floral buds of *E. pusilla* were collected from floral stages 2 and 4 (**Figure 2a**). The earliest floral stage to dissect the different flower parts was at floral stage 2. The lateral sepals, median sepal, petals, lip, callus, stamen and the remaining part of the gynostemium with stelidia but excluding the ovary were dissected (**Figure 2c-e**) and collected in individual tubes and immediately frozen on dry ice and stored at –80 °C until RNA extraction. Total RNA was extracted from seven different floral organs of *E. pusilla* using the RNeasy Plant Mini Kit (QIAGEN), following the manufacturer's protocol. A maximum of 100 mg plant material was placed in a 2.2 ml micro centrifuge tube with 7 mm glass bead. The TissueLyser II (QIAGEN) was used to grind the plant material. The amount of RNA was measured using the NanoVue Plus™ (GE Healthcare Life Sciences) and its integrity was assessed on an Agilent 2100 Bioanalyzer using the Plant RNA nano protocol. RNA samples with an RNA Integrity Number (RIN) < 7 were discarded. RNA was stored at -80 °C until further use. Extracted RNA was treated with DNase I, Amp Grade (Invitrogen 1U/µl) to digest single- and double-stranded DNA following the manufacturer's protocol.

**cDNA synthesis**
cDNA was synthesized with up to 1 µg of DNase-treated RNA using iScript™ cDNA Synthesis Kit (Bio-Rad Laboratories) following the manufacturer's protocol. A reaction mixture was prepared by addition of 1 µg of RNA, 4 µl 5x iScript reaction mix, 1 µl iScript reverse transcriptase to nuclease-free water up to a total volume of 20 µl. The reaction mixture was incubated at 25 °C for 5 minutes, 42 °C for 30 minutes and 85 °C for 5 minutes using a C1000 Touch™ thermal cycler machine (Bio-Rad). During this reaction, a positive control (CTRL) and no reverse transcriptase (NRT) control were included.

**Primer design**

DNA sequences were downloaded from NCBI Genbank and Orchidstra (http://orchidstra.abrc.sinica.edu.tw). For the MADS-box genes primers were designed on the C-terminal of the DNA sequences to avoid cross –amplification. Beacon Designer™ (Premier Biosoft, www.oligoarchitect.com) software was used to design primers (**Tables S1-S2**). All primer pairs were screened for their specificity against the Orchidstra database and in a gradient PCR reaction. The reaction mixture (25 µl) contained: 2.5 ng cDNA, 0.2 µM of each primer, 0.1 mM dNTP's and 0.6 U *Taq* DNA polymerase (QIAGEN) in 1x Coral Load Buffer (QIAGEN). The amplification protocol was as follows: initial denaturation step of 5 min 94°C followed by 40 cycles of [20 s 94 °C, 20 s <55-65> °C, 20 s 72 °C], one final amplification step of 7 min 72 °C and ∞ 15 °C. Based on the results of the gradient PCR, the annealing temperature was set to 61.3 °C for the Quantitative Real-time PCR as this value gave the best results. Only when a specific product was detected was the primer pair used for subsequent quantification.

**Reference genes and Quantitative Real-time PCR**

Experimental and computational analyses with LinRegPCR (http://www.hartfaalcentrum.nl, v2015.1) (Ruijter *et al*., 2009;Tuomi *et al*., 2010), indicate that *E. pusilla Ubiquitin-2*, *Actin*, and *F-box* were stably expressed in the tissues of interest and these genes were chosen as reference genes for the expression assay. Expression of all MADS-box genes was normalized to the geometric mean of these three reference genes.

Quantitative real-time PCR was performed using the CFX384 Touch Real-Time PCR system (Bio-Rad Laboratories). The assays were performed using the iQ™ SYBR® Green Supermix (Bio-Rad Laboratories). The reaction mixture (7 µl) contained: 1x iQ™ SYBR® Green Supermix, 0.2 µM of each primer, 1 ng cDNA template from a specific floral organ (biological triplicate reactions) for each target gene and floral organ for two sets of isolated RNA (six reactions in total). All reactions were performed in Hard-Shell® Thin-Wall 384-Well Skirted PCR Plates (Bio-Rad Laboratories). For each amplicon group, a positive control was included (=CTRL, flower buds from floral stage 1 to 4), a negative control (=NTC, reaction mixture without cDNA) and a no reverse transcriptase treated sample (= NRT, control sample during the cDNA synthesis). For all the qPCR reactions, the amplification protocol was as follows: initial denaturation of 5 min 95 °C followed by; 20 s 95 °C; 30 s 61.3 °C; 30 s 72 °C; plate read, for 50 cycles; then followed by a melting curve analysis of 5 s, 65 °C to 95 °C with steps of 0.2 °C to confirm single amplified products (**Figure S2**).

**Normalization, data analysis and statistical analysis**

The non-baseline corrected data were exported from the Bio-Rad CFX Manager™ (v3.1) to a spreadsheet. Quantification Amplification results (QAR) were used for analysis with LinRegPCR (v2015.1, dr. J.M. Ruijter). The calculated $N_0$-values represented the starting concentration of a sample in fluorescence units. Removal of between-run variation in the multi-plate qPCR experiments was done using Factor qPCR© (v2015.0) (Ruijter *et al*., 2006;2015). Geometric means of the corrected $N_0$-

values were calculated from the six samples together, i.e. two biological and three technical replicates. GraphPad Prism version 7.00 (www.graphpad.com) was used to perform a Two-Way ANOVA with Sidak's multiple comparison test to calculate significant differences between the two floral stages 2 and 4, and graphed with Standard Error of Measurement (SEM) error bars. Tukey's multiple comparisons test was used to compare the means between the floral organs. Variation for the two biological replicates was assessed by tests in triplicate.

**Phylogenetic analyses**

Nucleotide sequences of floral developmental genes were downloaded from NCBI GenBank® (**Table S1**) and separate data sets were constructed for MADS-box gene classes *FUL-, AP3-, PI-, AG-, STK-, SEP-* and *AGL6-like*. For each gene class, protein-guided codon alignments were constructed by first performing multiple sequence alignments of the protein translations using MAFFT v.7.245 (with the algorithm most suited for proteins with multiple conserved domains, E-INS-I or "oldgenafpair" for backward compatibility), with a maximum of 1,000 iterations (Katoh and Standley, 2013) and then reconciling the nucleotide sequences with their aligned protein translations.

Gene trees were inferred from the codon alignments using PhyML v3.0_360-500M (Guindon *et al*., 2010) under a GTR+G+I model with 6 rate classes and with base frequencies, proportion of invariant sites, and γ-shape parameter α estimated using maximum likelihood. Optimal topologies were selected from results obtained by traversing tree space with both nearest neighbor interchange (NNI) and subtree prune and regraft (SPR) branch swap algorithms, ie. PhyML's "BEST"option. Support values for nodes were computed using approximate likelihood ratio tests (SH-like aLRT, (Anisimova and Gascuel, 2006)).

To infer where on the gene trees duplications may have occurred the GSDI algorithm (Zmasek and Eddy, 2001) was used as implemented in forester V1.038 (https://sites.google.com/site/cmzmasek/home/software/forester). Fully resolved species trees for GSDI testing were constructed based on the current understanding of the phylogeny of the species under study (**Figure S4**).

Lastly, to detect lineage-specific excesses of non-synonymous substitutions, BranchSiteREL (Kosakovsky Pond *et al*., 2011) analyses were performed as implemented in HyPhy (Pond *et al*., 2005) on the Datamonkey (http://datamonkey.org) cluster.

## Results

**Ontogeny, macro- and micromorphology of flowers of *E. pusilla***

Floral ontogeny in *E. pusilla* can be divided into two main phases: early and late. Early ontogeny starts from floral initiation (floral stage 1) up to the three-carpel-apex stage (floral stage 2) and late ontogeny starts from the three-carpel-apex stage (floral stage 2) until anthesis (floral stages 3,4 and 5, **Figure 2a**) (Kull and Arditti, 2013).

The inflorescence of *E. pusilla* is branched and multiple flowers develop in succession (**Figure 2a**). Up to floral stage 1, the perianth is formed following a classic

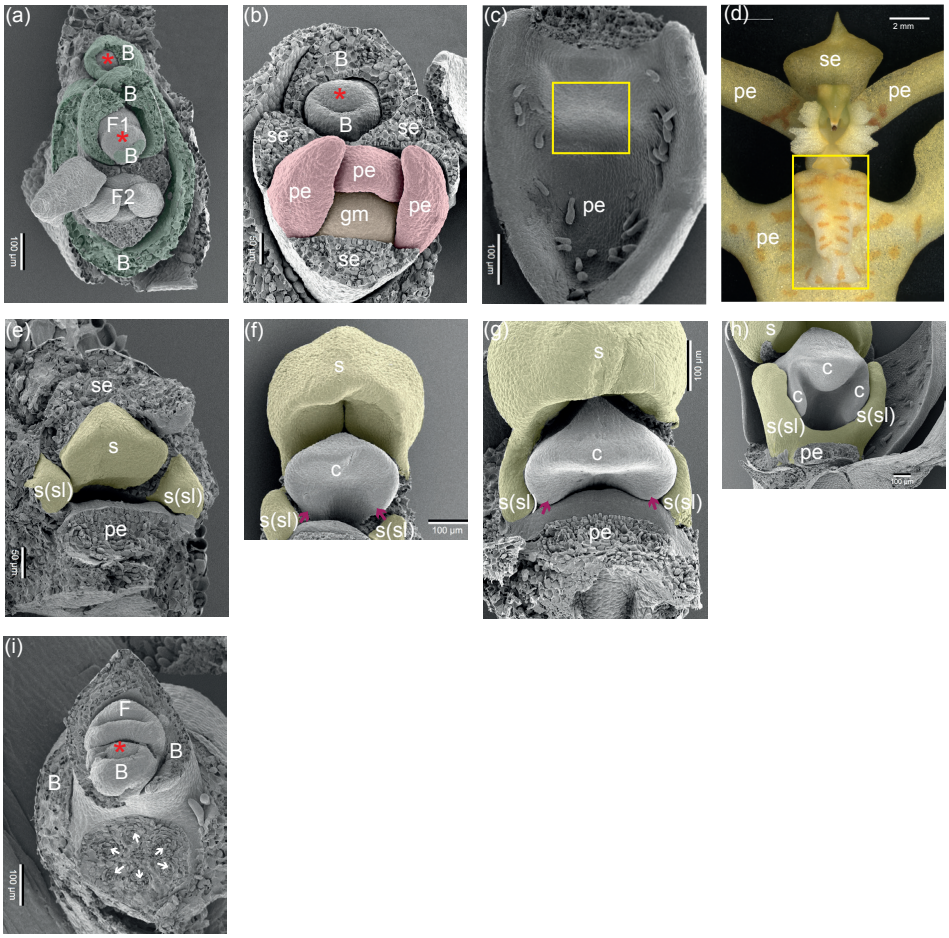



**Figure 4. Developing inflorescence of *E. pusilla*. (a)** Apical view of a young developing inflorescence. A central meristem is present and below it two flowers are visible, each subtended by a bract. The distal flower (F1) is primordial and the next flower (F2) is somewhat more developed. **(b)** Apical view of a developing flower in an early developmental stage. The scars of the three removed sepals are visible, two are adaxially (lateral sepals) and one is abaxially (median sepal) situated. More central in the flower, two abaxial-lateral petals and one adaxial developing petal (lip) are present. Most central in the flower is the primordium of the gynostemium. **(c-d)** Developing adaxial petal (lip) with callus (boxed). **(e-h)** Successive stages of the development of the gynostemium with the developing fertile stamen central and stelidia laterally. In [e], the scar of the removed abaxial sepal is visible. Below the fertile stamen, the scar of the adaxial petal (lip) can be seen. In between the fertile stamen and the adaxial petal (lip), the stigmatic cavity is present. In [f and g], the two adaxial (lateral) carpels are visible (arrowed). In [h], the abaxial carpel is incorporated in the stigmatic cavity. **(i)** Apical view of an inflorescence axis with a removed developing flower. In the upper half of the micrograph, the apex of the axis is visible as well as a flower at very early developmental stage, subtended by a bract. In the lower half, in the scar of the removed developing flower, six vascular bundles are visible (arrowed).

Abbreviations: Red asterisk = apical meristem; B = bract; F = flower (primordium); c = carpel; gm = gynostemium; pe = petal; se = sepal; s = fertile stamen; s(sl) = stelidium. Color codes: dark green = bract; red = petals; orange = gynostemium; yellow = androecium.

monocot developmental pattern (**Figure 5a**) (Rudall and Bateman, 2004) in which the sepals are among the first organs to become visible, followed by the petals. The position of the two abaxial petals is slightly shifted laterally (**Figure 4a**). Stamen and carpel primordial are not visible in the course of the early phase, but instead a single massive primordium is present from which the gynostemium will develop (**Figure 4b**).

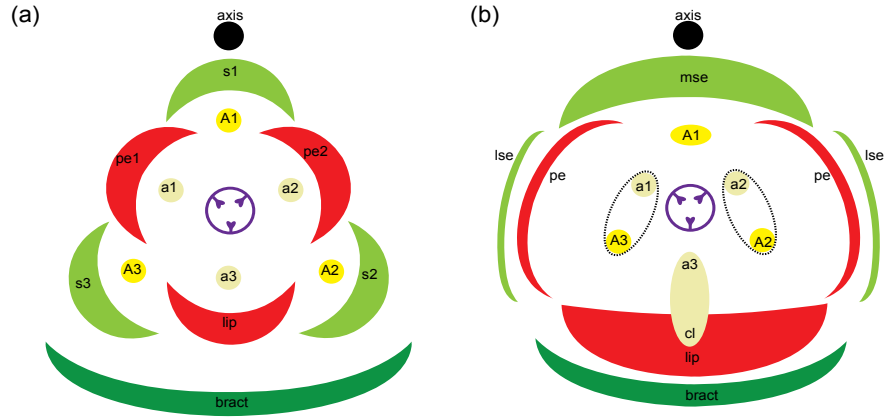



**Figure 5. Floral diagrams**. (**a)** A typical monocot flower. **(b)** A resupinate flower of *E. pusilla*. Abbreviations: $s_{1-3}$ = sepals; $p_{1-3}$ = petals; $A_{1-3}$ = anther in outer floral whorl; $a_{1-3}$ = anther in inner floral whorl; lse = lateral sepal; mse = median sepal; pe = petal; cl = callus. Color codes: black interrupted = stelidia and callus on lip; purple = gynoecium [Illustrations by Erik-Jan Bosch].

On the hypochile of the lip a callus is formed from floral stage 2 onwards (**Figure 4c-d**). The fertile stamen differentiates after floral stage 1. The stelidia appear at each side of the gynostemium (**Figure 4e-h**) from where they elongate and start forming wing-like appendices (**Figure 2e**). The abaxial carpel is incorporated in the stigmatic cavity, which forms a compound structure with the fertile stamen (**Figure 4h**). The three-carpel-apex stage is clearly visible in floral stage 2. At this stage the six staminal vascular bundles can also be observed just above the inferior ovary (**Figure 4i**). In floral stage 3, no new organs are formed, but in floral stage 4 (**Figure 2a**) the mature flower becomes resupinate (**Figure 5b**). The terms adaxial and abaxial are used here to indicate the position of the distinct floral parts with respect to the inflorescence axis (**Figure 4a-b**), thereby taking the position of the primordia of the floral organs as a reference. For example, with respect to the inflorescence axis, the lip is the adaxial petal, which by resupination becomes the lowermost part of the flower.

Using micro-CT scanning, vascular bundles were observed in a fully-grown floral stage 5 flower (**Figure 6a-f** and **Movie S1**). In the inferior ovary six vascular bundles could be discerned, indicated in purple. Three of these vascular bundles, indicated in green, run to the adaxial (median) sepal and abaxial (lateral) sepals, respectively. Three main groups of vascular bundles, indicated in red, run towards the petals including the lip, where they split up. Four vascular bundles (indicated in yellow) are present; one bundle, already split into two at the base, runs to the fertile stamen, where it splits up further towards the two pollinia (**Figure 6a-e**); two vascular bundles, originated from two pairs, run up into the stelidia (**Figure 6b-c; e-f**)

and one vascular bundle runs all the way up into the callus of the lip (**Figure 6b; e-f**). When following the yellow vascular bundles downwards, they connect in a plexus situated on top of the inferior ovary with the rest of the vascular system of the flower.

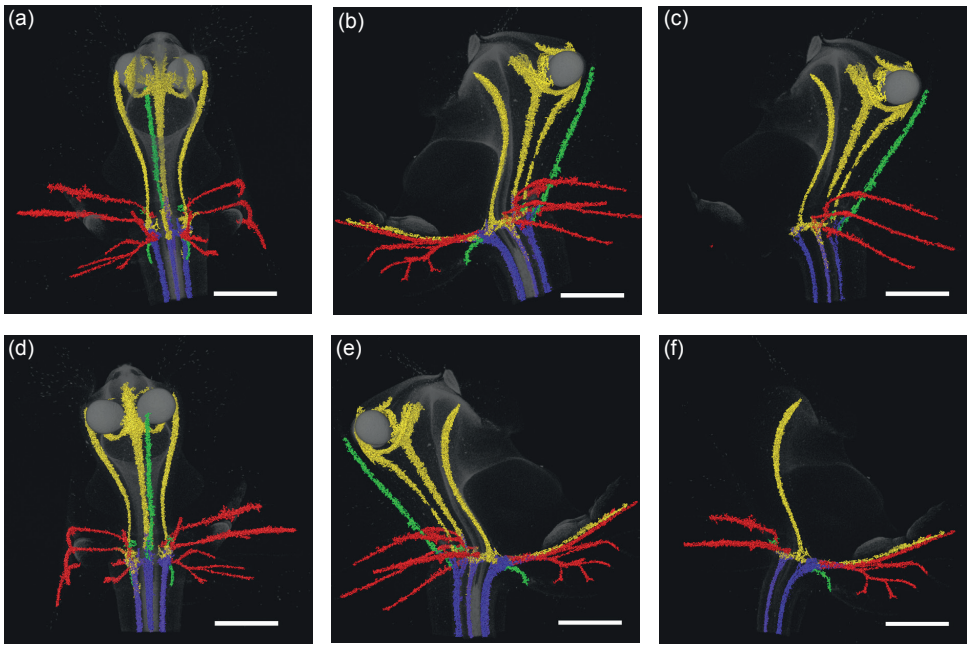



**Figure 6. Vascular bundle patterns of *E. pusilla.* (a)** Frontal view of a 3D X-ray macroscopical reconstruction of the vascular bundle patterns in a mature flower of *E. pusilla.* **(b)** Successive clockwise turn of 45°. **(c)** Simplified version of [b]. **(d)** Successive clockwise turn of 90°. **(e)** Successive clockwise turn of 135°. **(f)** Simplified version of [e]. Color codes: green = vascular bundles in sepals; red = vascular bundles in petals; purple = vascular bundles in gynoecium; yellow = vascular bundles in androecium. Scale bar = 1 mm.

Throughout late ontogeny, epidermal cells in all floral organs remained relatively undifferentiated and only expanded in size. Epidermal cells on the abaxial side of floral organs were mostly similar to the cells on the adaxial side, but more convex shaped (**Figure S1**). Epidermal cells of the lateral sepals were irregular, flattened and rectangular shaped and longitudinally orientated from the base to the apex (**Figure 7a-c**). Epidermal cells of the median sepal, as well as of the petals and the lip, develop from irregularly flattened shaped cells at floral stage 2, to a more convex shape in floral stage 5 (**Figure 7d-l**). Epidermal cells of the callus develop from convex shaped cells in floral stage 2 to cells with a more conical shape in floral stage 5 (**Figure 7m-o**). Epidermal cells of the stelidia become convex shaped during floral stage 2 and develop papillae on their apices during floral stage 5 (**Figure 7p-r**).

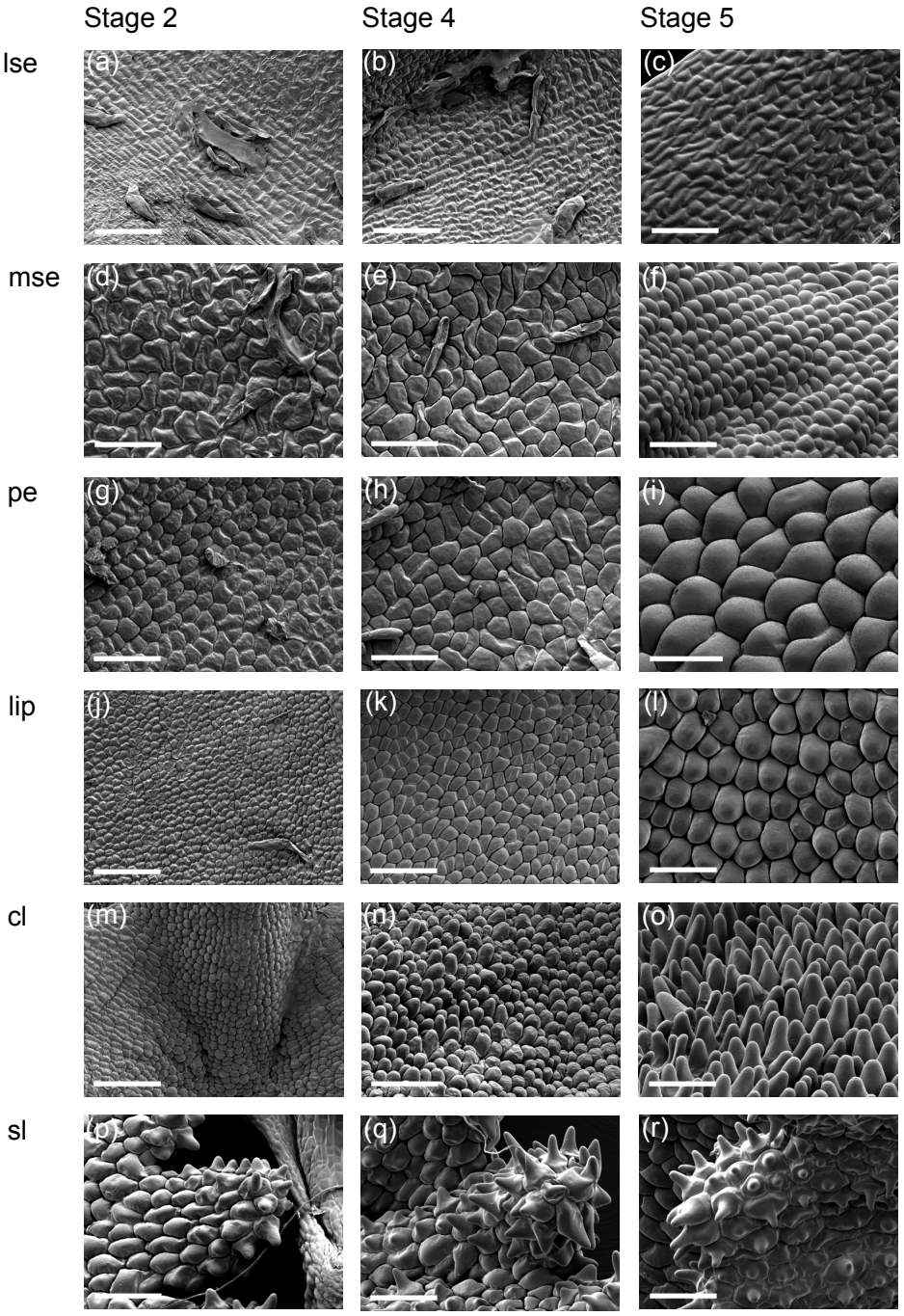



**Figure 7. Micromorphology of the epidermal cells on the adaxial side of a flower of** *E. pusilla***.** The three columns represent, from left to right, floral stage 2, 4 and 5 of the floral organs. Epidermal cells of (a-c) lateral sepal, (d-f) median sepal, (g-i) petal, (j-l) lip, (m-o) callus on lip and (p-r) stelidia. Scale bar = 100 µm. Abbreviations: lse = lateral sepal; mse = median sepal; pe = petal; cl = callus; sl = stelidia

## Duplications, diversifying evolution and expression of eighteen MADS- box genes in selected floral organs of *E. pusilla* in two developmental stages

### *FUL-, SEP-* and *AGL6-like* genes

The closest homologs of the *Arabidopsis* A class gene *APETALA1* in *E. pusilla* are the three *FUL*-like genes copies *EpMADS10*, *11* and *12*. Our phylogenetic analyses reconstructed three orchid clades of *FUL*-like genes, containing the three copies present in the genome of *E. pusilla* (**Figure S5a**), which was consistent with previous studies (Acri-Nunes-Miranda and Mondragon-Palomino, 2014). Diversifying selection was detected along the branch following the gene duplication leading to *EpMADS10*. The three *FUL*-like gene copies were expressed in all floral organs of *E. pusilla* but at low levels only (**Figure S3**). During development, expression generally decreased in most floral organs for *EpMADS10* and *11* whereas it generally increased for the majority of floral organs for *EpMADS12* (**Figure S3** and **Table S3**).

Four *SEP*-like orchid clades were retrieved (**Figure S5f**), encompassing the four copies of *E. pusilla*, consistent with previous studies (Acri-Nunes-Miranda and Mondragon-Palomino, 2014;Pan *et al*., 2014). The branch leading to the duplication that gave rise to *EpMADS6* and *EpMADS7* shows evidence of diversifying selection. *EpMADS6, 7, 8* and *9* were expressed in all floral organs at varying levels. *EpMADS6* was mainly expressed in the fertile stamen, a statistically significant difference as compared to the other six floral organs (**Figure S3** and **Table S3**).

Three *AGL6* orchid clades, also found by (Hsu *et al*., 2015) were retrieved, containing the three different copies present in the *E. pusilla* genome (**Figure S5g**). Evidence for a moderate degree of diversifying selection could be detected on the branch leading to *EpMADS4.* The three different copies of *AGL6*-genes were not expressed in all floral organs and the level of expression also varied. *EpMADS3* was most highly expressed in the sepals and petals. *EpMADS4* was more highly expressed in the lateral sepals as compared with the median sepal, petals and lip. *EpMADS5* was mainly expressed in the lip and callus (**Figure 8**).

### *AP3-like* and *PI-like* genes

Initial phylogenetic analyses reconstructed the main duplication between the *AP3* and *PI* genes also found in many other studies (Mondragon-Palomino *et al*., 2009;Pan *et al*., 2011;Hsu *et al*., 2015) so two separate gene trees were retrieved for each lineage (**Figure S5b-c**). Four orchid *AP3*-clades and three *PI*-clades were identified in these analyses. The three copies of *AP3* and a single copy of *PI* present in the genome of *E. pusilla* were placed in *AP3*-clades 1,2 and 3 and *PI*-clade 2, respectively. No evidence for diversifying selection could be detected along the branches leading to the *PI*-clade containing *EpMADS16* but evidence for diversifying selection along the branch in the *AP3*-1 clade encompassing *EpMADS15* was found. *AP3*-like gene copy *EpMADS14* was most highly expressed in the lateral sepals than in the median sepal, lip and callus. *AP3*-like gene copy *EpMADS13* was more highly expressed in the lip and callus than in the sepals and petals (**Figure 8**). The *PI*-like gene *EpMADS16* was more highly expressed in the first four floral whorls in both floral stages (**Figure 8** and **Figure 9**).

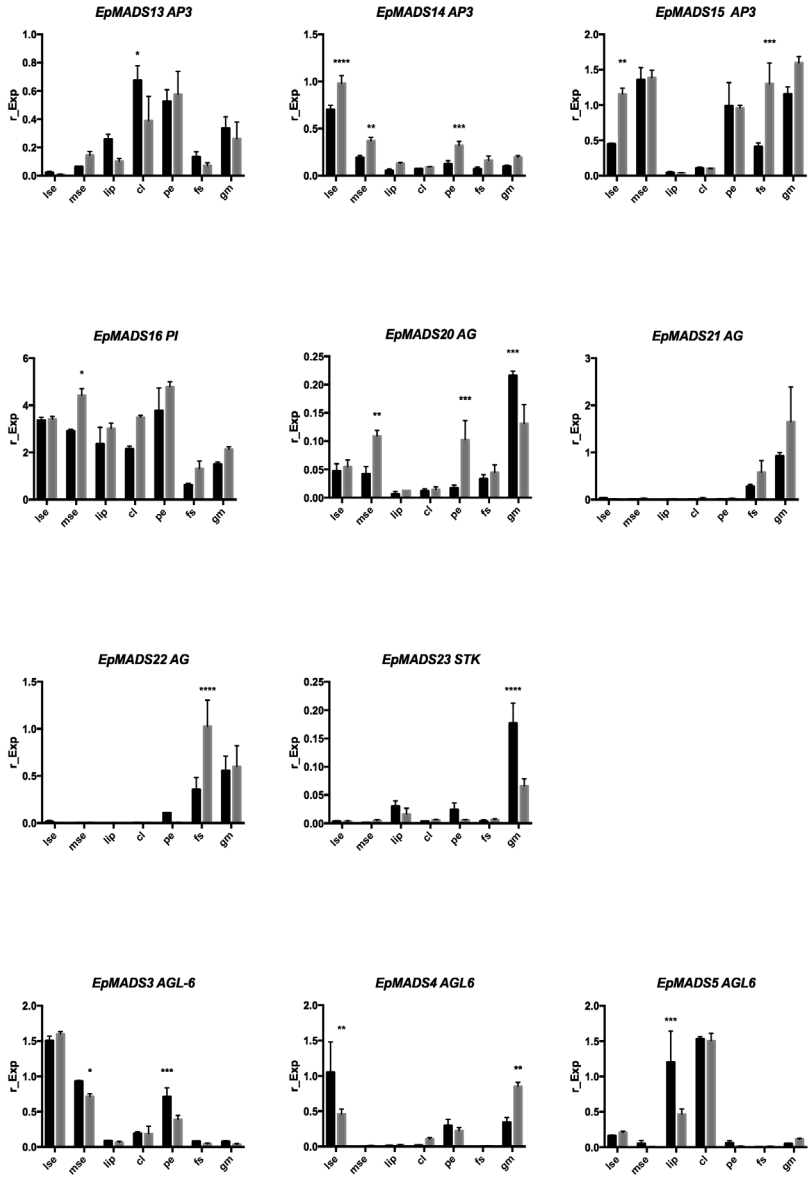



**Figure 8. Floral organ specific expression levels of selected MADS-box gene copies in *E. pusilla*.** *AP3* (top row), *PI* (second row), *AG* (second and third row), *STK* (second row), *ALG6* (third row). RNA was extracted from seven different floral organs during two stages of development of *E.* pusilla and used for cDNA synthesis.

Expression of the MADS-box genes was normalized to the geometric mean of three reference genes *Actin*, *UBI2* and *Fbox*. Each column shows the relative expression of 20 floral organs in two cDNA pools (10 floral organs per isolation), both tested in triplicate. Abbreviations: lse = lateral sepal; mse = median sepal; cl = callus; pe = petal; fs = fertile stamen; gm = gynostemium. Dark grey = floral stage 2 and light grey = floral stage 4. Y-axis: relative gene expression. The error bars represent the Standard Error of Mean. P-value style: GP: >0.05 (ns), <0.05 (*), <0.01 (**), <0.001 (***), <0.0001 (****).

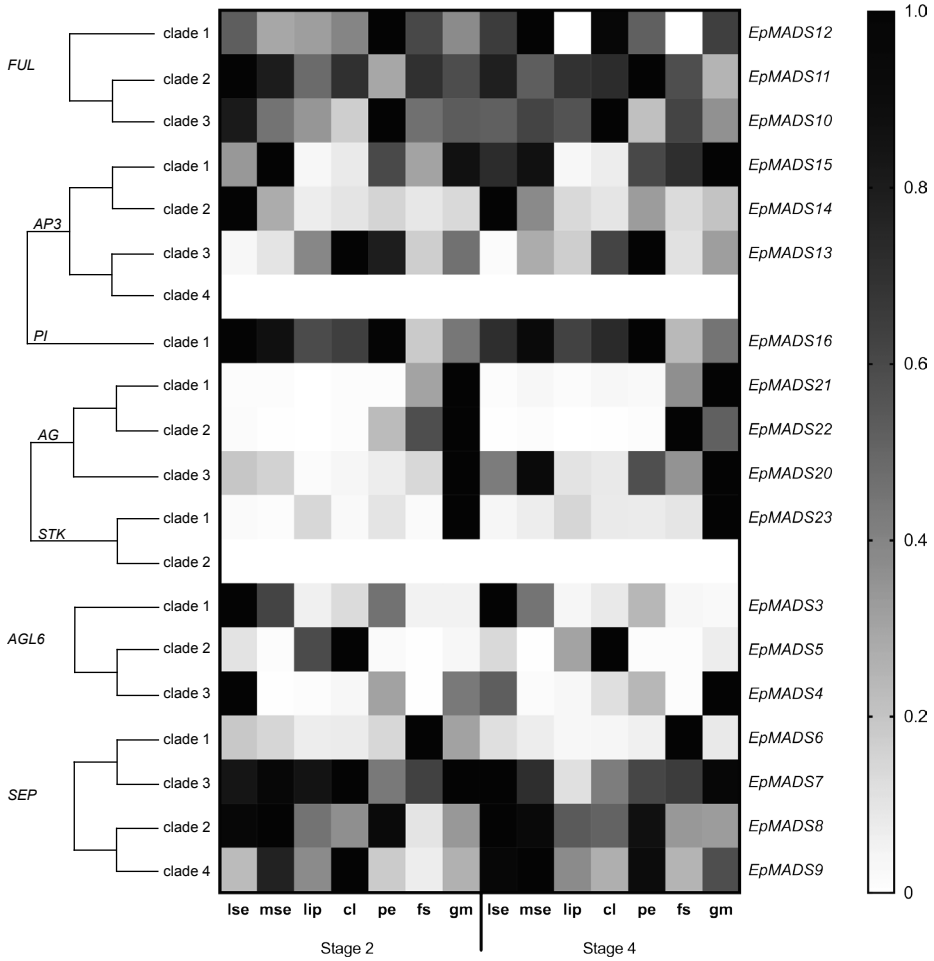



**Figure 9**. **Heat map representation of MADS-box gene expression in *E. pusilla.*** The *FUL-, AP3-, PI-, AG-, STK-, SEP-* and *ALG6-* like copies were retrieved from different gene lineage clades during two stages of floral development.

Expression of the MADS-box genes was normalized to the geometric mean of three reference genes *Actin*, *UBI2* and *Fbox*. The relative gene expression was normalized with the CTRL sample (= flower buds from floral stages 1-4). The scales for each gene and developmental stage are independent of each other and set to 1 for the highest value. Abbreviations: lse = lateral sepal; mse = median sepal; cl = callus; pe = petal; fs = fertile stamen; gm = gynostemium

### *AG*- and *STK-like* genes

Three orchid *AG*-clades and two *STK*-clades were identified in the phylogenetic analyses (**Figure S5d-e**). *EpMADS20, 21 and 22* were placed in *AG*-clades 3, 1 and 2, respectively, and *EpMADS23* was placed in *STK*-clade 1, as also found by Lin *et al*. (2016). No evidence for diversifying selection in the branches supporting the three orchid *AG*-clades and *STK*-clade containing copies present in the genome of *E. pusilla* could be detected. *AG*-like gene copy *EpMADS20* was most highly expressed in the stelidia, whereas *EpMADS22* was most highly expressed in the stamen as

compared with all other floral organs analyzed (**Figure 8**). No expression of *AG*-like genes could be detected in the callus. *STK*-like gene copy *EpMADS23* was most highly expressed in the stelidia as compared with all other floral organs analyzed (**Figure 8** and **Figure 9**).

## Discussion

**Homology of the median sepal of *Erycina pusilla***

The floral ontogenetic observations and vascularization patterns indicate that the median sepal is derived from the first floral whorl. In contrast, the presence of convex epidermal cells suggests a petaloid origin (Whitney *et al*., 2011). The *AGL6* and *AP3* copies *EpMADS3* and *EpMADS15,* members of the sepal/petal-complex of the P-code model, were most highly expressed in the median sepal, lateral sepal and petal. A possible correlation between expression and petaloidy was found for *AGL6* and *AP3* copies *EpMADS4* and *EpMADS14.* These two genes were lowly expressed in the median sepal, lip and petal as compared with the lateral sepal. Additional functional studies are needed to show whether loss of function of *EpMADS4* and *EpMADS14* is linked to sepal morphology in *E. pusilla* and other species that also possess a petaloid median sepal. The *AGL6* gene copy *EpMADS4* copy showed evidence of diversifying evolution. Lin *et al*. (2016) Identified fifteen motifs in the MIKC-type MADS-box proteins of *E. pusilla*. Two differences can be noticed within the K-region and C-terminal-region of *AP3* and *AGL6* genes of *E. pusilla*: (i) *AP3 EpMADS14* is missing motif 11, while the other B-class genes all contain motif 11. *AGL6 EpMADS4* also contains motif 11, while the other *AGL6* gene copies lack this motif; (ii) *AGL6 EpMADS4* is missing motif 6 whereas all the other *AGL6* gene copies contain motif 6. The differences found may contribute to the morphological differences between the median and lateral sepals *of E. pusilla*.

**Homology of the lip and callus of *Erycina pusilla***

The convex shaped epidermal cells on the lip and conical shaped epidermal cells on the callus are indicative of a petaloid function (Whitney *et al*., 2011). The *FUL*-like gene copy *EpMADS12*, *AP3*-like *EpMADS13* and *AGL6*-like *EpMADS5* are most highly expressed in lip and callus, further confirming a lip identity based on the ABCDE, floral quartet and P-code models, that dictate joint expression of A, B, E and *AGL6*-like genes in the petals and lip, respectively. According to these models, B, C and E class genes should be expressed in stamens but no evidence of expression of C class genes was found in the lip or callus. Notwithstanding, multiple lines of evidence support the possible staminal origin of the callus. First of all, the ontogeny and function of the lip of *E. pusilla* are very different as compared with the ontogeny and function of the callus. The lip is formed from floral stage 1 onwards, mainly acts as a long distance attraction and functions as a soft landing platform for pollinating bees. The callus is formed from floral stage 2 onwards and functions as short distance attraction by offering a sturdy holdfast to pollinators. This is in line with Carlquist (1969), who states that different vascularization patterns are driven by different functional needs. Many Oncidiinae have a callus on the lip and in some of these

species, the callus produces oil, making the functions of the lip and the callus even more distinct. Flowers with an oil-producing callus evolved twice in unrelated clades from species with non-rewarding flowers according to the molecular phylogeny of the Oncidiinae as presented in Genera Orchidacearum by (Pridgeon *et al*., 2009). One of the two rewarding clades, i.e. the one containing the genus *Gomesa*, is the sister group of the *Erycina* clade, showing that changes between an oil-producing and a non-rewarding callus occur quite easily in this group of orchids. This suggests that evolution towards oil production is correlated with increased venation as also stated by Carlquist (1969). We argue, however, that the venation in the callus is not only driven by functional needs but that the venation pattern is also informative regarding the evolutionary origin of the callus, as the callus of *E. pusilla* is connected with only one of the six original staminal bundles, physically distinct from the two adjacent vascular bundles leading to the lip. We consider this indicative of a possible staminal origin of the callus because of the occasional appearance of an infertile staminodial structure at this particular position, the inner adaxial stamen (a3), in teratologous orchid flowers (Bateman and Rudall, 2006). Terata of monandrous orchids with both stelidia carrying an additional anther on their tip next to the anther on the apex of the gynostemium, such as *Bulbophyllum triandrum* and *Prostechea cochleata* var. *triandrum*, are commonly seen as support for a staminal origin of stelidia. Similarly, mutants in *Dactylorhiza*, for instance, with a staminodial structure on their lip (Bateman and Rudall, 2006) could be interpreted as support for a staminal origin of the callus. Alternatively, these phenotypes could be caused by ectopic C gene expression that is transforming petal into stamen tissue. Homeotic transformation is not necessarily indicative of derivation. According to Carlquist (1969) data from teratology are therefore not useful for studying the evolution of flowers. This publication was written at a time that experimental mutants could not yet be made though. Ongoing work on B- and C- class homeotic mutants in the established plant models *Arabidopsis, Antirrhinum* and *Petunia* shows how much can be gained from teratology. We hope that these mutants can be created in emerging orchid models such as *E. pusilla* in the future to provide more evidence for the evolutionary origin of the callus on the lip.

**Homology of the stamen and stelidia of *Erycina pusilla***

Five vascular bundles, indicating a stamen-derived origin, lead to the stamen and stelidia. Our observations concur with those of Swamy (1948) who showed that the ovary is traversed by multiple vascular bundles in monandrous orchids. He visualized 'compound' bundles of staminal origin in the ovary of a species of *Dendrobium* and discovered vascularizing bundles in the stelidia. In several other plant families, e.g. Brassicaceae (*Arabidopsis*), Commelinaceae (*Tradescantia*), and Cyperaceae (*Cyperus*), it has been shown that vascular bundles of different organs originate in the developing organs and grow towards the stele rather than being branched from the stele (Endress and Steiner-Gafner, 1996;Pizzolato, 2006;Scarpella *et al*., 2006;Reynders, 2012). Based on **Figure 6** and **Movie S1**, we hypothesize that especially the staminal vascular bundles are connected in a similar way to the rest of the vascular system. Of the three copies of *AG* and four copies of *SEP*, *EpMADS22*

and *EpMADS6* were found to be highest expressed in the stamen. Another copy of *AG EpMADS20*, and the single copy of *STK*, *EpMADS23*, were found to be most highly expressed in the stelidia, suggesting that *EpMADS23* expression may be correlated with sterility.

**Implications for current floral models**

The ABCDE, orchid code, HOT and P-code models do not explain the morphological difference between median and lateral sepals as present in orchid species such as *E. pusilla*. Our results show that a differentiation between the sepaloid lateral sepals and petaloid median sepal of *E. pusilla* is correlated with a significant reduction of expression of *AP3*-like *EpMADS14* and *ALG6*-like *EpMADS4* in all petaloid organs (**Figure 10a**).

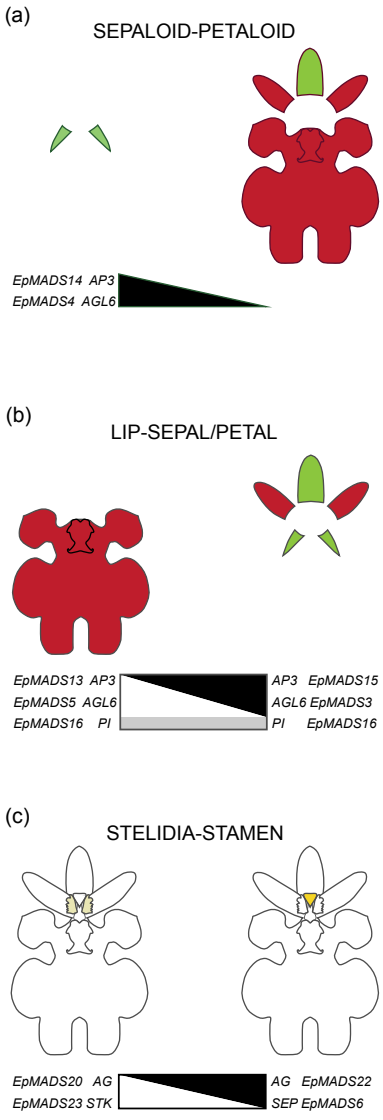



**Figure 10. Summary of expression of MADS-box genes involved in the differentiation of selected floral organs of *E. pusilla*. (a)** Expression of *EpMADS4/14* (in black) correlating with a sepaloid-petaloid identity is high in the lateral sepals (left side) but low in the remainder of the perianth (right side), **(b)** Expression of the lip complex *EpMADS5/13/16* (in white/grey)) correlating with a lip identity is high in in the lip and callus (left side) but low in the remainder of the perianth (right side). Expression of the sepal/petal-complex *EpMADS3/15/16* (in black/grey) correlating with a sepal and petal identity is low in the lip (left side) but high in the sepals and petals (right side), **(c)** Expression of *EpMADS20/23* (in white) correlating with a stelidia-stamen identity is high in the stelidia (left side) but low in the stamen (right side). Expression of *EpMADS6/22* (in black) is low in the stelidia (left side) but high in the stamen (right side). [Illustrations by Erik-Jan Bosch].

The P-code model explains the development of the lip of *E. pusilla* as the SP-complex (*AP3*- like *EpMADS15/AGL6*-like *EpMADS3/PI*-like *EpMADS16)* was found to be most highly expressed in the sepals and petals, whereas the L-complex (*AP3*-like *EpMADS13/AGL6*-like *EpMADS5/PI*-like *EpMADS16)* was found to be most highly expressed in the lip (**Figure 10b**). However, the model does not yet account for the development of the callus and the high expression of *AGL6*-like *EpMADS5* in this particular organ. To incorporate all new evidence found for the evolution and development of first and second floral whorl organs, we propose an Oncidiinae model (**Figure 11**), summarizing the gene expression data presented in this study for *E. pusilla* and earlier studies carried out on *Oncidium* Gower Ramsey (Hsu *et al*., 2015).

All four MADS-box B class gene copies were found to be expressed in the fertile stamen of *E. pusilla*. In addition, *AG*-like *EpMADS22* and *SEP*-like *EpMADS6* were most highly expressed in this floral organ, confirming a stamen identity as predicted by the ABCDE model. The high expression of *AG*-like *EpMADS20* and *STK*-like *EpMADS23* in the stelidia cannot be explained with the ABCDE model. All current orchid floral models only describe evolution and development of the first and second whorl floral organs. We found evidence for differential gene expression in organs in the third and fourth floral whorl, i.e. the stamen and stelidia (**Figure 10c**), and this argues for the development of additional models.

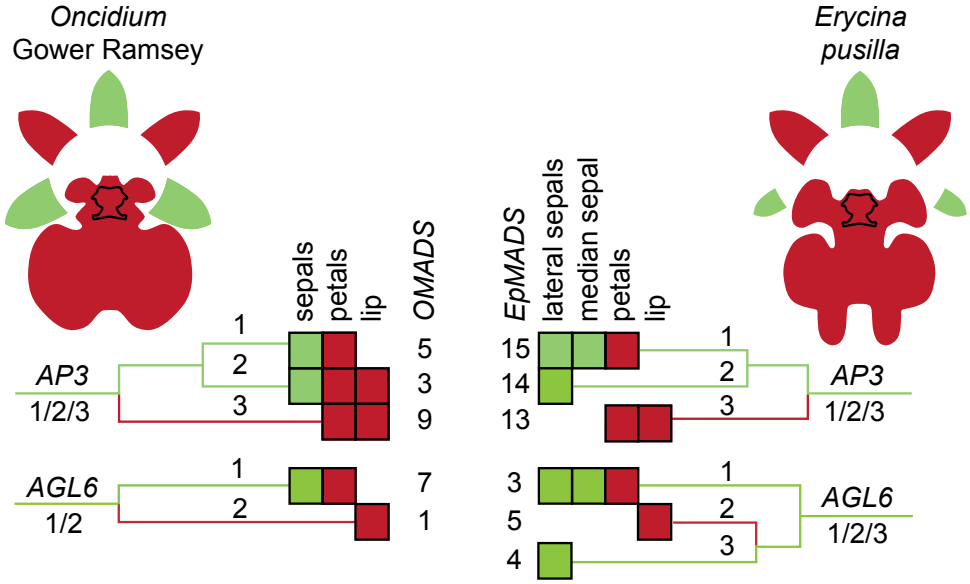



**Figure 11. Oncidiinae model** summarizing expression of MADS-box genes involved in the differentiation of the perianth of *Oncidium* Gower Ramsey (left) and *E. pusilla* (right). Clade 1 *AP3*-like *OMADS5* and *EpMADS15* and clade *AGL6*-like genes *OMADS7* and *EpMADS3* are expressed in the sepals and petals of both species. Clade 2 *AP3*-like *OMADS3* is expressed in the entire perianth of *O.* Gower Ramsey whereas *EpMADS14* is only expressed in the lateral sepals of *E. pusilla*. Clade 2 *AGL6*-like genes *OMADS1* and *EpMADS5* are expressed in the lip only of both species. Clade 3 *AP3*-like *OMADS9* and *EpMADS13* are expressed in the petals and lip of both species. Clade 3 *AGL6*-like gene *EpMADS4* is only expressed in the lateral sepals of *E. pusilla*. [Illustrations by Erik-Jan Bosch].

**Conclusions**

After examining vascularization, macro- and micromorphology, gene duplications, diversifying evolution and expression of different MADS-box genes in selected floral organs in two developmental stages, it can be concluded that: (i) the median sepal obtained a petal-identity, thus representing a particular character state of the character 'sepal', (ii) that the lip was derived from a petal but the callus from a stamen that gained petal identity, and (iii) the stelidia evolved from stamens. Duplications, diversifying selection and changes in spatial expressions of *AP3 EpMADS14* and *AGL6 EpMADS4* may have contributed to an increase of petaloidy of the median sepal. The same can be applied to *AP3 EpMADS13* and *AGL6 EpMADS5* in the lip and callus. Differential expression of *AG* copies *EpMADS20* and *EpMADS22*, *STK* copy *EpMADS23* and *SEP* copy *EpMADS6* appear to be associated with the evolution of the stamen and stelidia, respectively.

The evolutionary origin of the median sepal, callus and stelidia of *E. pusilla* cannot be explained with any of the currently existing floral developmental models. Therefore, new models, like our Oncidiinae model, need to be developed to summarize MADS-box gene expression in more complex floral organs. Such models need validation by functional analyses. The genetic mechanisms discovered in this study ultimately contributed to the evolution of a deceptive orchid flower mimicking the morphologies of rewarding Malpighiaceae flowers. This mimicry enabled flowers of *E. pusilla,* and many other species in the highly diverse Oncidiinae, to successfully attract *Centris* bees for pollination, often, as is the case for *E. pusilla*, without offering a reward. Pollination by deceit is one of the most striking adaptations of orchids to pollinators. It is estimated that approximately a third of all orchid species employ deceit pollination, and that food mimicry is the most common type. Deceptive pollination is hypothesized to be correlated with species diversification as subtle changes in floral morphology can attract different pollinators and eventually lead to reproductive isolation. It was recently discovered that deceptive pollination augmented orchid diversity, not by accelerating speciation but by adding more species at roughly the same rate through time (Givnish *et al*., 2015). Ongoing research on the genomics of *E. pusilla* and other emergent plant models will shed more light on the role that key developmental genes played in the evolution of deceptive flowers.

**Supplementary material**

If not published in this thesis further supplementary material for this chapter can be found online at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5364718/

**Table S1.** List of sequences used in the alignments and phylogenetic analyses.

**Table S2.** Transcript primer sequences and amplicon characteristics used for quantitative real-time PCR validation of the expression profiles of eighteen MADS-box transcripts following MIQE guidelines (Bustin *et al*., 2009).

**Table S3.** Difference in MADS-box gene expression between floral organs; variance analysis of measures

using Tukey multicomparisons test. P-value style: GP: >0.05 (ns), <0.05 (*), <0.01 (**), <0.001 (***), <0.0001 (****). Abbreviations: lse = lateral sepal, mse = median sepal, cl = callus, pe = petal, fs = fertile stamen and gm = gynostemium.

**Figure S1.** Scanning electron micrographs of epidermal cells on the abaxial side of an *E. pusilla* flower. The three columns represent, from left to right, stage 2, 4 and 5 floral organs. Epidermal cells of (a-c) lateral sepal, (d-f) median sepal, (g-i) petal and (j-l) lip. Scale bar = 100 µm. Abbreviations: lse = lateral sepal; mse = median sepal; pe = petal.



**Figure S2**. Melting curve analysis of all primer pairs used in this study performed at the end of the PCR cycles to confirm the specificity of primer annealing.

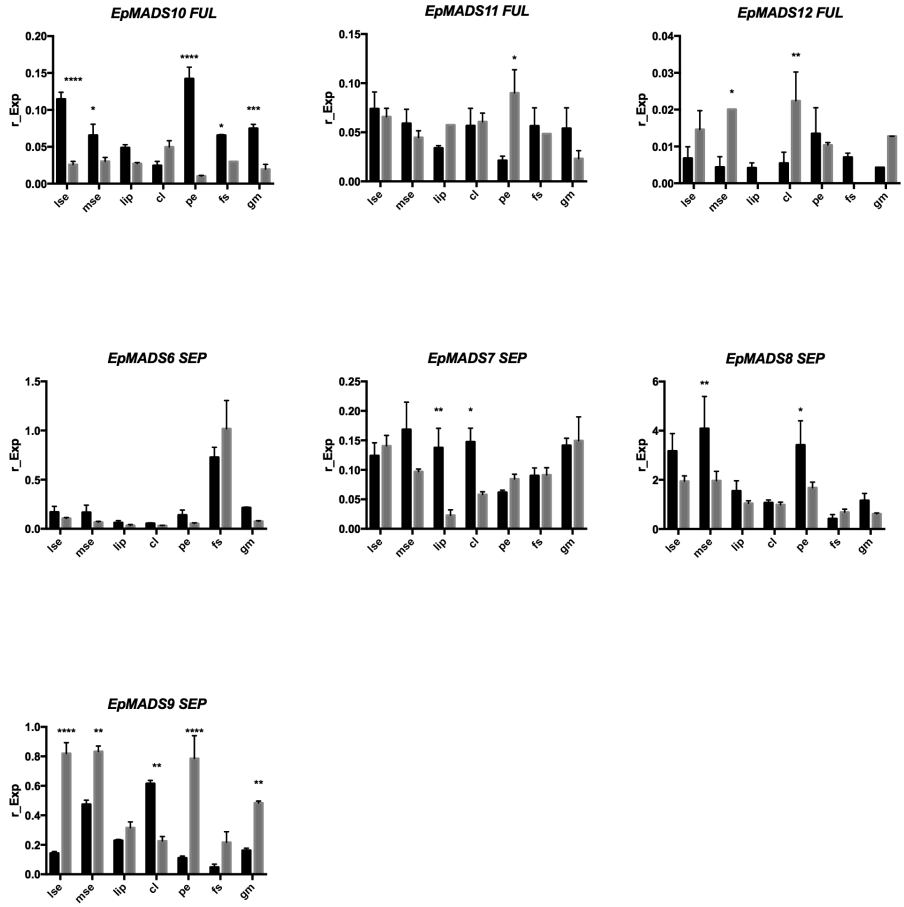



**Figure S3**. **Floral organ specific expression levels of FUL EpMADS10, EpMADS11 and EpMADS12 and SEP EpMADS6, EpMADS7, EpMADS8 and EpMADS9.** RNA was extracted from seven different floral organs during two stages of development of *E. pusilla* and used for cDNA synthesis. Expression of the MADS-box genes was normalized to the geometric mean of three reference genes *Actin*, *UBI2* and *Fbox*. Each column shows the relative expression of 20 floral organs in two cDNA pools (10 floral organs per isolation), both tested in triplicate. Abbreviations: lse = lateral sepal; mse = median sepal; cl = callus; pe = petal; fs = fertile stamen; gm = gynostemium. Dark grey = floral stage 2 and light grey = floral stage 4. Y-axis: relative gene expression. The error bars represent the Standard Error of Mean. P-value style: GP: >0.05 (ns), <0.05 (*), <0.01 (**), <0.001 (***), <0.0001 (****).

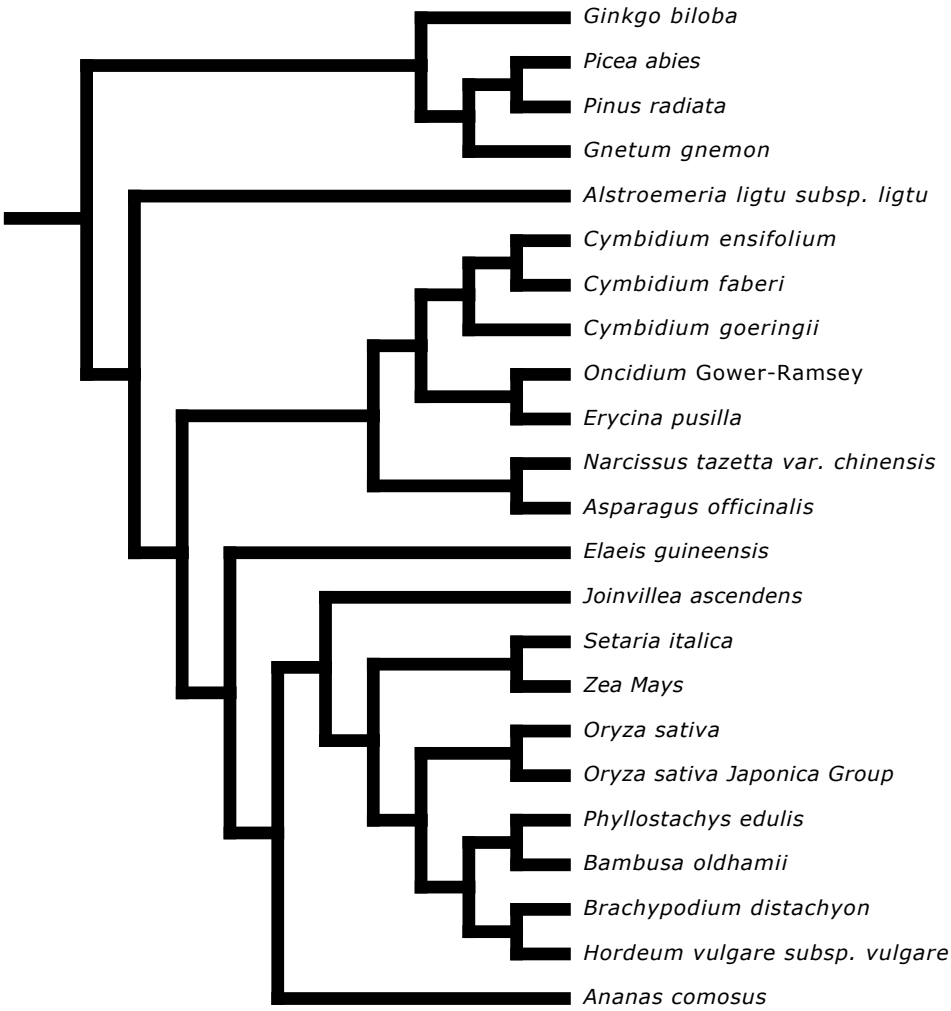



**Figure S4**. **Species phylogeny** compiled based on Topik *et al*. (2005), Biswal *et al*. (2013), Takamiya *et al*. (2014) and Chase *et al*. (2015) for **(a)** *FUL*-, (**b)** *AP3*-, (**c)** *PI*- (**d)** *AG*- and *STK*-, (**e**) *SEP*- and **(f)** ***AGL6-like*** **MADS-box gene lineage trees** (data for *FUL, AP3, PI, AG/STK* and *SEP* not shown in this thesis).

**Movie S1.** Animation of the 3D visualization as depicted in **Figure 6**.

## Author's contributions

ADM and BG designed the gene expression study and KvK, PvS and LK collected the expression data. RAV carried out the phylogenetic analyses with help of JWW. RvdB and SD collected the anatomical and micro-CT data. AV collected the electron microscope data and helped with the interpretation of the floral ontogeny. RB assisted with plant breeding. All authors contributed to the writing of the manuscript.

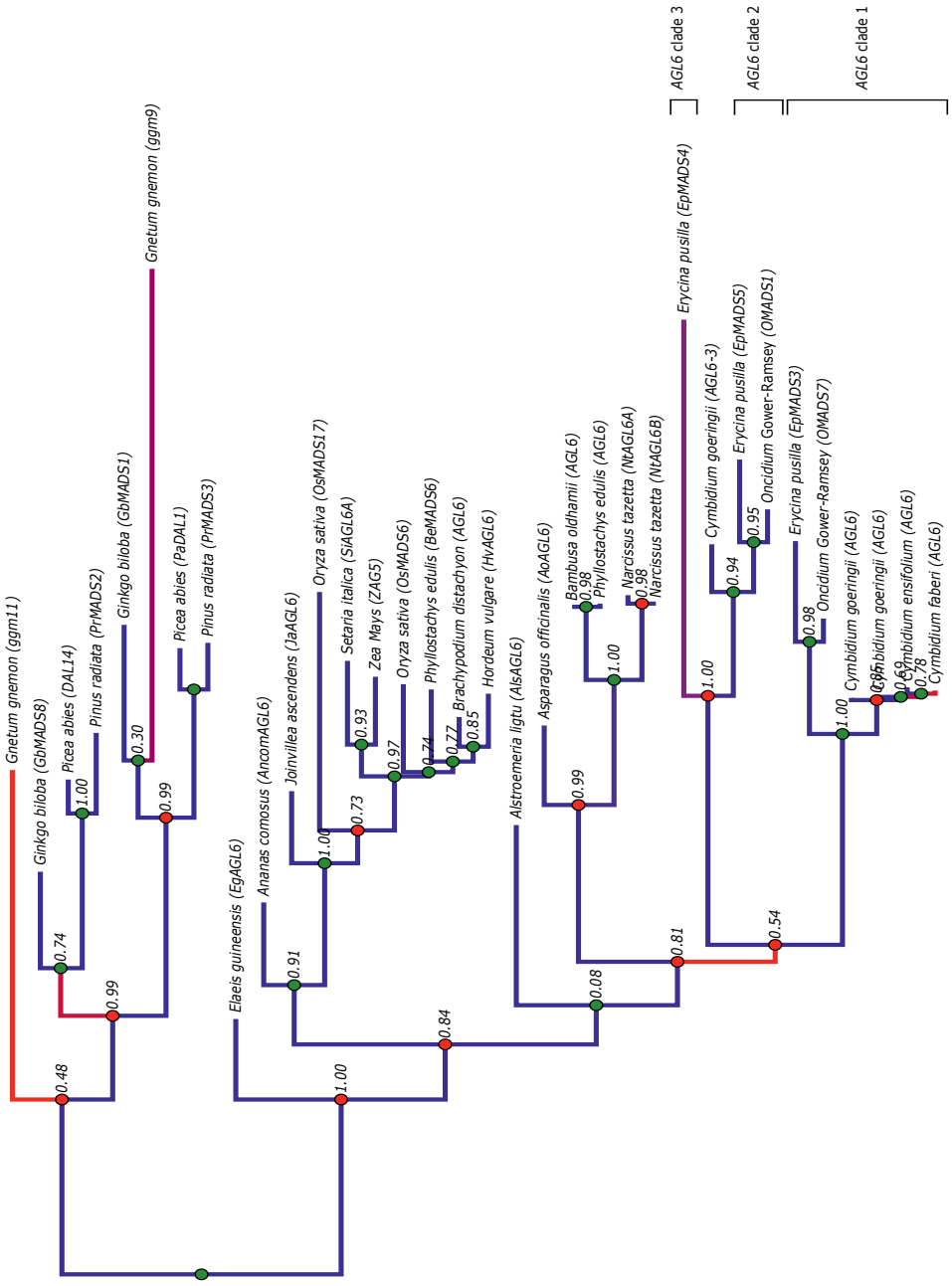



**Figure S5**. **MADS-box gene lineage trees**. **(a)** *FUL-*, **(b)** *AP3-*, **(c)** *PI-*, **(d)** *AG-*, **(e)** *STK-*, **(f)** *SEP-* and **(g)** ***AGL6*-like** trees. Color codes: green node=speciation event: red node=duplication event. Branches are colored along a gradient between blue and red, in proportion to the value of omega (dN/dS) for the third (i.e. the highest) rate class in the BranchSiteREL analysis. Hence, blue and red branches may be interpreted as suggesting, respectively, stabilizing and diversifying selection. Purple branches implicate a moderate level of diversifying selection (data for *FUL, AP3, PI, AG/STK* and *SEP* not shown in this thesis).

**Funding**
This work was supported by grant 023.003.015 from the Netherlands Organization for Scientific Research (NWO) to AD and a Fulbright grant to BG.

**Competing interests**
The authors declare that they have no competing interests.

**Acknowledgements**
We thank Johan Keus for the culturing of *E. pusilla,* Bas Blankevoort, Erik-Jan Bosch, Rogier van Vugt (Hortus botanicus Leiden) and Angel Vale for the illustrations and photographs, Pieter van der Velden (LUMC), Stef Janson and Jan M. de Ruijter (UvA) for their support and input with the qPCR, Anneke de Wolf for support with the SEM, Øyvind Hammer for help with the Zeiss X radia and Marcel Lombaerts and Jan Oliehoek for help with the construction of the DNA sequence alignments.

**Availability of Data and Materials**
Scripts and alignments can be found at https://zenodo.org/record/44533#.VpPCP5PhCCR.

## References

Ackerman, J.D., Cuevas, A.A., and Hof, D. (2011). Are deception-pollinated species more variable than those offering a reward? *Plant Systematics and Evolution* 293**,** 91-99.

Acri-Nunes-Miranda, R., and Mondragon-Palomino, M. (2014). Expression of paralogous SEP-, FUL-, AG- and STK-like MADS-box genes in wild-type and peloric Phalaenopsis flowers. *Front Plant Sci* 5**,** 76.

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. *Syst Biol* 55**,** 539-552.

Bateman, R.M., and Rudall, P.J. (2006). The good, the bad, and the ugly: Using naturally occuring terata to distinguish the possible from the impossible in orchid floral evolution. *Aliso* 22**,** 481-496.

Biswal, D.K., Marbaniang, J.V., and Tandon, P. (2013). Age estimation for Asian Cymbidium (Orchidaceae: Epidendroideae) with implementation of fossil data calibration using molecular markers (ITS2 & matK) implying phylogeographic inference. *PeerJ PrePrints* 1**,** e94v91.

Brown, R., and Nees Von Esenbeck, C.G.D. (1827). *Prodromus florae Novae Hollandiae et Insulae Van-Diemen : exhibens characteres plantarum.* Norimbergae: Sumtibus L. Schrag.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., and Wittwer, C.T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55**,** 611-622.

Cai, J., Liu, X., Vanneste, K., Proost, S., Tsai, W.-C., Liu, K.-W., Chen, L.-J., He, Y., Xu, Q., Bian, C., Zheng, Z., Sun, F., Liu, W., Hsiao, Y.-Y., Pan, Z.-J., Hsu, C.-C., Yang, Y.-P., Hsu, Y.-C., Chuang, Y.-C., Dievart, A., Dufayard, J.-F., Xu, X., Wang, J.-Y., Wang, J., Xiao, X.-J., Zhao, X.-M., Du, R., Zhang, G.-Q., Wang, M., Su, Y.-Y., Xie, G.-C., Liu, G.-H., Li, L.-Q., Huang, L.-Q., Luo, Y.-B., Chen, H.-H., De Peer, Y.V., and Liu, Z.-J. (2015). The genome sequence of the orchid Phalaenopsis equestris. *Nature Genetics* 47**,** 65-72.

Carlquist, S. (1969). Toward acceptable evolutionary interpretations of floral anatomy. *Phytomorphology* 19**,** 332-362.

Carmona-Díaz, G., and García-Franco, J.G. (2008). Reproductive success in the Mexican rewardless Oncidium cosymbephorum (Orchidaceae) facilitated by the oil-rewarding Malpighia glabra (Malpighiaceae). *Plant Ecology* 203**,** 253-261.

Chase, M.W., Cameron, K.M., Freudenstein, J.V., Pridgeon, A.M., Salazar, G., Berg Van Den, C., and Schuiteman, A. (2015). An updated classification of Orchidaceae. *Botanical Journal of the Linnean Society* 177**,** 151-174.

Chase, M.W., Hanson, L., Albert, V.A., Whitten, W.M., and Williams, N.H. (2005). Life history evolution and genome size in subtribe Oncidiinae (Orchidaceae). *Ann Bot* 95**,** 191-199.

Cho, S., Jang, S., Chae, S., Chung, K.M., Moon, Y.-H., An, G., and Jang, S.K. (1999). Analysis of the C-terminal region of Arabidopsis thaliana APETALA1 as a transcription activation domain. *Plant Molecular Biology* 40**,** 419-429.

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature* 353**,** 31-37.

Darwin, C. (1877). *The various contrivances by which orchids are fertilised by insects.* London: John Murray.

Decraene, L.P.R., and Smets, E.F. (2001). Staminodes: Their Morphological and Evolutionary Significance. *Botanical Review* 67**,** 351-402.

Endress, P.K. (2016). Development and evolution of extreme synorganization in angiosperm flowers and diversity: a comparison of Apocynaceae and Orchidaceae. *Ann Bot* 117**,** 749-767.

Endress, P.K., and Steiner-Gafner, B. (1996). *Diversity and Evolutionary Biology of Tropical Flowers.* Cambridge University Press.

Felix, L.P., and Guerra, M. (2012). Chromosome analysis in Psygmorchis pusilla (L.) Dodson & Dressier: the smallest chromosome number known in Orchidaceae. *Caryologia* 52**,** 165-168.

Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Iles, W.J., Clements, M.A., Arroyo, M.T., Leebens-Mack, J., Endara, L., Kriebel, R., Neubig, K.M., Whitten, W.M., Williams, N.H., and Cameron, K.M. (2015). Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc Biol Sci* 282.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59**,** 307-321.

Hsu, H.-F., Hsu, W.-H., Lee, Y.-I., Mao, W.-T., Yang, J.-Y., Li, J.-Y., and Yang, C.-H. (2015). Model for perianth formation in orchids. *Nature Plants* 1**,** 15046.

Kanno, A., Nakada, M., Akita, Y., and Hirai, M. (2007). Class B gene expression and the modified ABC model in nongrass monocots. *ScientificWorldJournal* 7**,** 268-279.

Kanno, A., Saeki, H., Kameya, T., Saedler, H., and Theissen, G. (2003). Heterotopic expression of class B floral homeotic genes supports a modified ABC model for tulip (Tulipa gesneriana). *Plant Mol Biol* 52**,** 831-841.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30**,** 772-780.

Kocyan, A., and Endress, P.K. (2001). Floral structure and development of Apostasia and Neuwiedia (Apostasioideae) and their relationships to other Orchidaceae. *International Journal of Plant Sciences* 162**,** 847-867.

Kosakovsky Pond, S.L., Murrell, B., Fourment, M., Frost, S.D., Delport, W., and Scheffler, K. (2011). A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol* 28**,** 3033-3043.

Kramer, E.M., Holappa, L., Gould, B., Jaramillo, M.A., Setnikov, D., and Santiago, P.M. (2007). Elaboration of B gene function to include the identity of novel floral organs in the lower eudicot Aquilegia. *Plant Cell* 19**,** 750-766.

Kull, T., and Arditti, J. (2013). *Orchid Biology VIII: Reviews and Perspectives.* Springer Netherlands.

Kurzweil, H. (1987). Developmental studies in orchid flowers I: Epidendroid and vandoid species. *Nordic Journal of Botany* 7**,** 427-442.

Kurzweil, H., and Kocyan, A. (2002). "Ontogeny of Orchid Flowers," in *Orchid Biology VIII: Reviews and*

*Perspectives* eds. T. Kull & J. Arditti (Leiden: Springer The Netherlands).

Lee, S.-H., Li, C.-W., Liau, C.-H., Chang, P.-Y., Liao, L.-J., Lin, C.-S., and Chan, M.-T. (2014). Establishment of an Agrobacterium-mediated genetic transformation procedure for the experimental model orchid Erycina pusilla. *Plant Cell, Tissue and Organ Culture (PCTOC)* 120**,** 211-220.

Lin, C.S., Hsu, C.T., Liao, D.C., Chang, W.J., Chou, M.L., Huang, Y.T., Chen, J.J., Ko, S.S., Chan, M.T., and Shih, M.C. (2016). Transcriptome-wide analysis of the MADS-box gene family in the orchid Erycina pusilla. *Plant Biotechnol J* 14**,** 284-298.

Mcknight, T., and Shippen, D. (2004). Plant Telomer Biology. *The Plant cell* 16**,** 794-803.

Mondragon-Palomino, M., Hiese, L., Harter, A., Koch, M.A., and Theissen, G. (2009). Positive selection and ancient duplications in the evolution of class B floral homeotic genes of orchids and grasses. *BMC Evol Biol* 9**,** 81.

Mondragon-Palomino, M., and Theissen, G. (2009). Why are orchid flowers so diverse? Reduction of evolutionary constraints by paralogues of class B floral homeotic genes. *Ann Bot* 104**,** 583-594.

Mondragon-Palomino, M., and Theissen, G. (2011). Conserved differential expression of paralogous DEFICIENS- and GLOBOSA-like MADS-box genes in the flowers of Orchidaceae: refining the 'orchid code'. *Plant J* 66**,** 1008-1019.

Nakamura, T., Fukuda, T., Nakano, M., Hasebe, M., Kameya, T., and Kanno, A. (2005). The modified ABC model explains the development of the petaloid perianth of Agapanthus praecox ssp. orientalis (Agapanthaceae) flowers. *Plant Mol Biol* 58**,** 435-445.

Neubig, K.M., Whitten, W.M., Williams, N.H., Blanco, M.A., Endara, L., Burleigh, J.G., Silvera, K., Cushman, J.C., and Chase, M.W. (2012). Generic recircumscriptions of Oncidiinae (Orchidaceae: Cymbidieae) based on maximum likelihood analysis of combined DNA datasets. *Botanical Journal of the Linnean Society* 168**,** 117-146.

Pan, Z.J., Chen, Y.Y., Du, J.S., Chen, Y.Y., Chung, M.C., Tsai, W.C., Wang, C.N., and Chen, H.H. (2014). Flower development of Phalaenopsis orchid involves functionally divergent SEPALLATA-like genes. *New Phytol* 202**,** 1024-1042.

Pan, Z.J., Cheng, C.C., Tsai, W.C., Chung, M.C., Chen, W.H., Hu, J.M., and Chen, H.H. (2011). The duplicated B-class MADS-box genes display dualistic characters in orchid floral organ identity and growth. *Plant Cell Physiol* 52**,** 1515-1531.

Papadopulos, A.S., Powell, M.P., Pupulin, F., Warner, J., Hawkins, J.A., Salamin, N., Chittka, L., Williams, N.H., Whitten, W.M., Loader, D., Valente, L.M., Chase, M.W., and Savolainen, V. (2013). Convergent evolution of floral signals underlies the success of Neotropical orchids. *Proc Biol Sci* 280**,** 20130960.

Pizzolato, Thompson d. (2006). Procambial Initiation for the Vascular System in the Shoot of Tradescantia zebrina (Commelinaceae). *International Journal of Plant Sciences* 167**,** 59-81.

Pond, S.L., Frost, S.D., and Muse, S.V. (2005). HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21**,** 676-679.

Preston, J.C., Hileman, L.C., and Cubas, P. (2011). Reduce, reuse, and recycle: developmental evolution of trait diversification. *Am J Bot* 98**,** 397-403.

Pridgeon, A., Cribb, P., Chase, M., and Rasmussen, F. (2009). *Genera Orchidacearum. Epidendroideae (Part Two).* Oxford, United Kingdom: Oxford University Press.

Ramirez, S.R., Gravendeel, B., Singer, R.B., Marshall, C.R., and Pierce, N.E. (2007). Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature* 448**,** 1042-1045.

Reynders, M. (2012). Gynoecial anatomy and development in Cyperoideae (Cyperaceae, Poales): congenital fusion of carpels facilitates evolutionary modifications in pistil structure. *Plant Ecology and Evolution* 145**,** 96-125.

Roy, B.A., and Widmer, A. (1999). Floral mimicry: a fascinating yet poorly understood phenomenon. *Trends Plant Sci* 4**,** 325-330.

Rudall, P.J., and Bateman, R.M. (2002). Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots. *Biol Rev Camb Philos Soc* 77**,** 403-441.

Rudall, P.J., and Bateman, R.M. (2004). Evolution of zygomorphy in monocot flowers: iterative patterns and developmental constraints. *New Phytologist* 162**,** 25-44.

Rudall, P.J., Perl, C.D., and Bateman, R.M. (2013). Organ homologies in orchid flowers re-interpreted using the Musk Orchid as a model. *PeerJ* 1**,** e26.

Ruijter, J.M., Ramakers, C., Hoogaars, W.M., Karlen, Y., Bakker, O., Van Den Hoff, M.J., and Moorman, A.F. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* 37**,** e45.

Ruijter, J.M., Ruiz Villalba, A., Hellemans, J., Untergasser, A., and Van Den Hoff, M.J. (2015). Removal of between-run variation in a multi-plate qPCR experiment. *Biomol Detect Quantif* 5**,** 10-14.

Ruijter, J.M., Thygesen, H.H., Schoneveld, O.J., Das, A.T., Berkhout, B., and Lamers, W.H. (2006). Factor correction as a tool to eliminate between-session variation in replicate experiments: application to molecular biology and retrovirology. *Retrovirology* 3**,** 2.

Scarpella, E., Marcos, D., Friml, J., and Berleth, T. (2006). Control of leaf vascular patterning by polar auxin transport. *Genes Dev* 20**,** 1015-1027.

Soltis, D.E., Chanderbali, A.S., Kim, S., Buzgo, M., and Soltis, P.S. (2007). The ABC model and its applicability to basal angiosperms. *Ann Bot* 100**,** 155-163.

Staedler, Y.M., Masson, D., and Schonenberger, J. (2013). Plant tissues in 3D via X-ray tomography: simple contrasting methods allow high resolution imaging. *PLoS One* 8**,** e75295.

Su, C.L., Chao, Y.T., Yen, S.H., Chen, C.Y., Chen, W.C., Chang, Y.C., and Shih, M.C. (2013). Orchidstra: an integrated orchid functional genomics database. *Plant Cell Physiol* 54**,** e11.

Swamy, B.G.L. (1948). Vascular anatomy of orchid flowers *Botanical Museum Leaflets Harvard University* 13**,** 61-95.

Takamiya, T., Wongsawad, P., Sathapattayanon, A., Tajima, N., Suzuki, S., Kitamura, S., Shioda, N., Handa, T., Kitanaka, S., Iijima, H., and Yukawa, T. (2014). Molecular phylogenetics and character evolution of morphologically diverse groups, Dendrobium section Dendrobium and allies. *AoB Plants* 6**,** plu045.

Theissen, G. (2001). Development of floral organ identity: stories from the MADS house. *Curr Opin Plant Biol* 4**,** 75-85.

Theissen, G., and Saedler, H. (2001). Plant biology. Floral quartets. *Nature* 409**,** 469-471.

Topik, H., Yukawa, T., and Ito, M. (2005). Molecular phylogenetics of subtribe Aeridinae (Orchidaceae): insights from plastid matK and nuclear ribosomal ITS sequences. *J Plant Res* 118**,** 271-284.

Tuomi, J.M., Voorbraak, F., Jones, D.L., and Ruijter, J.M. (2010). Bias in the Cq value observed with hydrolysis probe based quantitative PCR can be corrected with the estimated PCR efficiency value. *Methods* 50**,** 313-322.

Vale, A., Navarro, L., Rojas, D., and Alvarez, J.C. (2011). Breeding system and pollination by mimicry of the orchid Tolumnia guibertiana in Western Cuba. *Plant Species Biol* 26.

Vermeulen, P. (1959). The Different Structure of the Rostellum in Ophrydeae and Neottieae. *Acta Botanica Neerlandica* 8**,** 338-355.

Whitney, H.M., Bennett, K.M., Dorling, M., Sandbach, L., Prince, D., Chittka, L., and Glover, B.J. (2011). Why do so many petals have conical epidermal cells? *Ann Bot* 108**,** 609-616.

Zmasek, C.M., and Eddy, S.R. (2001). A simple algorithm to infer gene duplication and speciation events on a gene tree. *Bioinformatics* 17**,** 821-828.

# Chapter 4

## Morphological and molecular characterization of orchid fruit development

*Frontiers in Plant Science 10, 137 (2019)*

Anita Dirks-Mulder[1,2], Israa Ahmed[2], Mark uit het Broek[2], Louie Krol[2], Nino Menger[2], Jasmijn Snier[2], Anne van Winzum[2], Anneke de Wolf[2], Martijn van 't Wout[2], Jamie J. Zeegers[2], Roland Butôt[1], Reinout Heijungs[3,4], Bertie Joan van Heuven[1], Jakob Kruizinga[5], Rob Langelaan[1], Erik F. Smets[1,6,7], Wim Star[1], Marian Bemer[8] & Barbara Gravendeel[1,2,6]

[1]Endless Forms group, Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands

[2]Faculty of Science and Technology, University of Applied Sciences Leiden, Zernikedreef 11, 2333 CK Leiden, The Netherlands

[3]Department of Econometrics and Operations Research, School of Business and Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

[4]Institute of Environmental Sciences (CML), Leiden University, Einsteinweg 2, 2333 CC, Leiden, The Netherlands

[5]Hortus botanicus, Leiden University, Rapenburg 73, 2311 GJ Leiden, The Netherlands

[6]Institute of Biology Leiden, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands

[7]Ecology, Evolution and Biodiversity Conservation cluster, KU Leuven, Kasteelpark Arenberg 31, 3001 Leuven, Belgium

[8]Department of Plant Sciences, Laboratory of Molecular Biology, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

## Abstract

Efficient seed dispersal in flowering plants is enabled by the development of fruits, which can be either dehiscent or indehiscent. Dehiscent fruits open at maturity to shatter the seeds, while indehiscent fruits do not open and the seeds are dispersed in various ways. The diversity in fruit morphology and seed shattering mechanisms is enormous within the flowering plants. How these different fruit types develop and which molecular networks are driving fruit diversification is still largely unknown, despite progress in eudicot model species. The orchid family, known for its astonishing floral diversity, displays a huge variation in fruit dehiscence types, which have been poorly investigated. We undertook a combined approach to understand fruit morphology and dehiscence in different orchid species to get more insight into the molecular network that underlies orchid fruit development.

We describe fruit development in detail for the epiphytic orchid species *Erycina pusilla* and compare it to two terrestrial orchid species: *Cynorkis fastigiata* and *Epipactis helleborine*. Our anatomical analysis provides further evidence for the split carpel model, which explains the presence of three fertile and three sterile valves in most orchid species. Interesting differences were observed in the lignification patterns of the dehiscence zones. While *C. fastigiata* and *E. helleborine* develop a lignified layer at the valve boundaries, *E. pusilla* fruits did not lignify at these boundaries, but formed a cuticle-like layer instead. We characterized orthologs of fruit-associated MADS-domain transcription factors and of the Arabidopsis dehiscence-related genes *INDEHISCENT (IND)/HECATE 3 (HEC3)*, *REPLUMLESS (RPL)* and *SPATULA (SPT)/ALCATRAZ (ALC)* in *E. pusilla* and found that the key players of the eudicot fruit regulatory network appear well conserved in monocots. Protein-protein interaction studies revealed that MADS-domain complexes comprised of FRUITFULL (FUL), SEPALLATA (SEP) and AGAMOUS (AG) /SHATTERPROOF (SHP) orthologs can also be formed in *E. pusilla*, and that the expression of *HEC3, RPL* and *SPT* can be associated with dehiscence zone development similar to Arabidopsis. Our expression analysis also indicates differences, however, which may underlie fruit divergence.

**Keywords**
Cuticle layer, *Erycina pusilla*, fruit-gene and protein network, lignification, MADS-box genes, fruit ontogeny

## Introduction

Developmental mechanisms driving fruit diversification are still poorly understood, despite progress in the study of fruit formation in model plant species such as Arabidopsis (*Arabidopsis thaliana* (L.) Heyhn) and tomato (*Solanum lycopersicum* L.) (Gu *et al*., 1998;Ferrandiz *et al*., 1999;Vrebalov *et al*., 2009;Pabon-Mora and Litt, 2011). Research in monocots so far has focused mainly on cereal species such as rice, maize and wheat, all of which have relatively simple indehiscent fruits that consist of a one-layered pericarp. This leaves a big gap in the knowledge about evolution and development of fruits of other monocots, especially of the orchid family. Orchids are known for their spectacular floral diversity, but they also exhibit a large variety of fruit morphologies and dehiscence types (Brown, 1831;Beer, 1863;Dressler, 1993;Rasmussen and Johansen, 2006).

Orchid fruits are very diverse in size and shape, but almost all share the same basic pattern and variation results from specific differentiation and development of the carpels. Orchid flowers are epigynous with an inferior ovary composed of three fused carpels containing many tiny ovules. After pollination the inferior ovary further develops into a dehiscent or indehiscent fruit. The three fused carpels develop into six valves: three fertile valves with a placenta, bearing the ovules, and three sterile valves. The origin and nature of these valves has been debated since the beginning of the nineteenth century. Rasmussen and Johansen (2006) presented the “split-carpel model” of the orchidaceous ovary, giving an explanation of the hexamerous pattern. According to this model, a typical orchid ovary consists of three sterile valves (located at the sepal bases) and three fertile valves (located at the petal bases), each consisting of two carpel-halves (**Figure 1**).

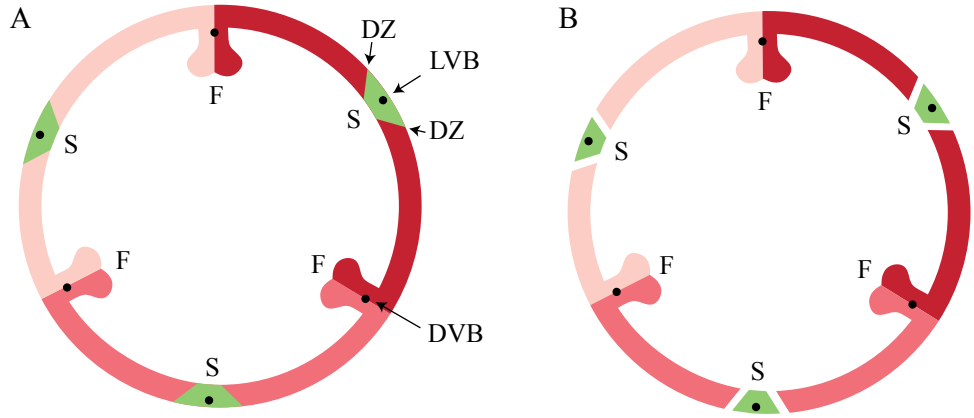



**Figure 1. Tricarpellate orchidaceous fruit according to Brown (1831) and Rasmussen and Johansen (2006).** (**A**) Immature fruit, (**B**) Mature fruit. Abbreviations: DVB = dorsal vascular bundle; DZ = dehiscence zone; F = fertile valve; LVB = lateral vascular bundle; S = sterile valve. Color codes: red = fertile valve; green = sterile valve. [Illustrations by Erik-Jan Bosch].

According to Horowitz (1901), the main difference in morphology of the fertile and sterile valves is the size and/or number of cells. Cells of sterile valves do not become much larger during fruit maturation, whereas cells of

fertile valves expand considerably. This would be in agreement with the split carpel model, where the fertile valves are the actual carpels, while the sterile valves are structures containing the mid-nerves descending from the sepals.

Almost all orchids have dehiscent dry fruits, with a few exceptions, such as the berries of *Neuwiedia zollingeri* Rchb. f. (Kocyan and Endress, 2001) and fleshy pods of *Vanilla pompona* Schiede (Pridgeon *et al*., 1999). Fruit development of dry dehiscent orchid fruits has been studied in *Oncidium flexuosum* Sims (Mayer *et al*., 2011), which led to the identification of a special layer of cells involved in fruit dehiscence. However, a detailed description of dehiscent fruits from other orchid species is lacking, as well as any molecular data about the genes that underlie fruit development in orchids.

MADS-box genes have been shown to play an important role in fruit development, maturation and ripening in several angiosperm species, among which the dry fruit species Arabidopsis and the fleshy fruit species tomato. However, whether there is a conserved regulatory network operating at the base of dry and fleshy fruit development is still unclear. In dry dehiscent fruits of Arabidopsis, the MADS-domain proteins AGAMOUS (AG), SHATTERPROOF 1/2 (SHP1/2) and FRUITFULL (FUL) are essential for carpel formation (AG), as well as for fruit development and dehiscence (SHP and FUL) (Gu *et al*., 1998;Ferrandiz *et al*., 2000;Ferrandiz and Fourquin, 2014). FUL represses *SHP1/2* expression in the valves of the fruit, which ensures proper dehiscence zone development (Ferrandiz *et al*., 2000). SHP1/2 in their turn are expressed in the valve margins, where they activate the expression of *INDEHISCENT (IND)* and *ALCATRAZ (ALC)*, which are required for separation of the valves and the formation of a lignified cell layer initiating this separation (Rajani and Sundaresan, 2001;Liljegren *et al*., 2004). *REPLUMLESS (RPL)* is expressed in the replum at the other side of the valve margin and controls the development of the Arabidopsis replum by repression of SHP1/2 (Roeder *et al*., 2003). Both RPL and FUL are necessary for the proper development of a functional dehiscence zone in Arabidopsis fruits by repressing *SHP* expression in the valve margins (Ballester and Ferrandiz, 2017). In addition, the MADS-domain factors SEPALLATA 1-3 (SEP1-3), which promote higher-order complex formation, are also highly active in the Arabidopsis fruit and can interact with AG, SHP and FUL (de Folter *et al*., 2004;2005). Orthologs of these MADS-domain factors have been shown to play important roles in fruit development and ripening in tomato (Vrebalov *et al*., 2009;Bemer *et al*., 2012;Ferrandiz and Fourquin, 2014). Thus, complexes consisting of homologs of the MADS-domain proteins AG, SHP, FUL and/or SEP seem to be generally important for fruit development in the eudicots. Whether the same MADS-domain factors play a role in orchid fruit development is still unclear, but a recent study from Lin *et al*. (2016) revealed that there are several MADS-box genes expressed in mature *Erycina pusilla* (L.) N.H. Williams & M.W. Chase fruits, pointing to a role for MADS-box genes in orchid fruit development as well. There is no data available yet about the presence and activity of homologs of the downstream target genes *IND/HECATE3* (*HEC3*), *SPATULA (SPT)/ALC* and *RPL* in orchids. Homologs of these Arabidopsis genes have also been found to be expressed in fruits of different Solanaceae species, suggesting that their role may be more broadly conserved.

To increase our knowledge of fruit anatomy in different orchid species and of the molecular gene regulatory network that underlies fruit development in orchids, we undertook a combined approach, in which we performed a detailed anatomical and molecular analysis of fruit development of the orchid species *E. pusilla,* which has dry dehiscent fruits*. Erycina pusilla* belongs to the subfamily Epidendroideae and subtribe Oncidiinae, and is a fast growing, small sized epiphytic orchid species occurring in the wild in South America with a relatively short life cycle. It develops from seed to flowering stage in less than a year and is an upcoming model system for orchid research (Pan *et al*., 2012;Chou *et al*., 2013;Lin *et al*., 2013;Lee *et al*., 2015;Lin *et al*., 2016;Dirks-Mulder *et al*., 2017). To expand the study of fruit divergence in the orchid family, we compared development and dehiscence of *E. pusilla* fruits with those of fruits from the terrestrial species *Epipactis helleborine* (L.) Crantz (subfamily Epidendroideae, tribe Neottieae from Europe, Asia and North-Africa) and *Cynorkis fastigiata* Thouars (subfamily Orchidoideae, tribe Orchideae from Madagascar and surrounding islands).

To investigate whether the regulatory network underlying fruit development in eudicots could to some extent be conserved in orchids, we investigated the fruit-expressed MADS-box genes in *E. pusilla* by performing detailed expression analysis and determining the protein-protein interactions of the proteins encoded by these genes. In addition, we performed expression analysis of close homologs of other well-known Arabidopsis fruit genes to investigate to what extent the genetic network driving fruit patterning and lignification of Arabidopsis corresponds to that of *E. pusilla*.

## Material and methods

### Plant material

A more than 20-year-old inbred line of *E. pusilla* originally collected in Surinam was grown in climate rooms under controlled conditions (7.00h – 19.00h light regime), at a temperature of 22 °C and a relative humidity of 50%. The orchids were cultured *in vitro* under sterile conditions on Phytamax™ orchid medium with charcoal and banana powder (Sigma-Aldrich) with 4g/L Gelrite™ (Duchefa) culture medium. Pollination was conducted manually by placing the pollinia of flowers on each other's stigma. The seeds were ripe after 14-16 weeks and subsequently sown into containers (Duchefa) using sterile fresh culture medium. Fruits were collected from this laboratory strain of *E. pusilla* at 0, 1, 3, 5 days after pollination (DAP) and 1, 2, 3, 4, 8 and 12 weeks after pollination (WAP). Ripe seeds were collected from open fruits, after 16 WAP.

Fruits of *E. helleborine* and *C. fastigiata* were collected at different developmental stages in the Hortus botanicus (Leiden, The Netherlands). Different developmental stages were determined by assessing the relative degradation of the floral remains and the size of the fruits.

### Fixation for micromorphology

Fruits were fixed with standard formalin aceto-alcohol (FAA: 50% ethanol; 5% glacial

acetic acid; 5% formalin (Sigma-Aldrich and Boom)) for one hour under vacuum conditions at room temperature. They were placed on a rotating platform for 16 hours (at room temperature). The fruits were washed once in 70% ethanol and subsequently stored in 70% ethanol at room temperature.

**London Resin (LR) – polyhydroxy-aromatic acrylic resin – White embedding**
Fruits, stored in 70% ethanol, were cut off transversally and rinsed in absolute ethanol for one hour. Subsequently, the fruits were incubated in the following solutions: eight hours in 3:1 absolute ethanol in LR White (SPI supplies, Pennsylvania); overnight in 2:1 absolute ethanol in LR White overnight; eight hours in 1:1 absolute ethanol in LR White; overnight in 1:2 absolute ethanol in LR White; eight hours in 1:3 absolute ethanol in LR White and lastly overnight in LR White. Gelatin capsules (Electron Microscopy Sciences) were filled with the tissue and fresh LR White and placed in an oven at 60 °C for 48 hours.

**LR White sectioning**
LR White resin embedded samples were sectioned using a Leica RM2265 microtome (Leica Biosystems, Germany). The samples were trimmed until the tissue of interest was reached. Using a tungsten knife (Leica), at a 4°angle, sections of 5 μm thickness were obtained. The sections were placed in a drop of 40% acetone on a microscope slide. The slides were placed on a hot plate at 70 °C for at least one hour, after which they were stained.

**Staining, visualization, valve area and cell layer measurements**
LR white embedded sections were stained for two minutes with a solution of 0.2% Toluidine Blue and 0.2% Borax in distilled water, rinsed with distilled water, placed on a hot plate at 50-60 °C for 20 seconds and mounted with Entellan mounting medium (Merck-Millipore). The slides were scanned using Bright field and Z-stacking on a 2D Scanning Panoramic Viewer 250 (LUMC, Leiden, The Netherlands). Scanned slides were viewed and analyzed with Case Viewer software (3DHISTECH). Areas of individual fertile- and sterile valves of the fruits were measured using Case Viewer software (3DHISTECH). For the reliability of the valve area measurements, the valve areas of cross-sections of the fruits were used as the metric for fruit size. These areas are quite robust against distortions and angle under which the anatomical slides were made, and highly reproducible. For *E. pusilla* the areas of the fertile and sterile valves were determined using the perimeter of these valves in fruits of 0 DAP, 5 DAP, 7 DAP, 2 WAP, 5 WAP, 8 WAP, 11 WAP, 15 WAP and 16 WAP. Per time point, at least three fruits obtained of the same inbred laboratory strain were used. Of these fruits, six sections, three fertile and sterile valves per section, were measured. The number of cell layers of the fruit walls of all three orchid species was determined from at least 4 different fruits. Counts were performed in quarto for 3-4 slides per fruit. Cell number was determined in the same sections for 6-9 valves per developmental stage. Cells in the vascular bundles and placental tissues were not included.

Handmade cross sections of *E. pusilla*, *E. helleborine* and *C. fastigiata* fruits, stored in 70% ethanol, were stained with 1% phloroglucinol (Sigma-Aldrich) in 96%

ethanol for one hour. The cross sections were subsequently washed with 25%-(v/v) hydrochloric acid (HCl) (Sigma-Aldrich) and immediately examined under a Binocular microscope (Zeiss SteREO Discovery.V12).

**X-ray micro-computed tomography (micro-CT)**
Fruits were infiltrated with 1% phosphotungstic acid (Brunschwig) in 70% ethanol for three to four days, where PTA solution was refreshed daily. Scans were performed on a Zeiss Xradia 510 Versa 3D. Data were stacked and processed with Dragonfly Pro 2.0 (Object Research Systems, Montreal Canada).



**Scanning Electron Microscopy (SEM)**
Fruits were dehydrated twice for 20 minutes in 90% ethanol and twice for 20 minutes in absolute ethanol. The fruits were dried using liquid carbon dioxide ($CO_2$) with a Leica EM CPD300 critical point dryer (Leica Microsystems, Wetzlar Germany). Dried fruits were then placed on a stub with Leit-C conductive carbon cement (Neubauer) and spray-coated with 20nm of Platinum/Palladium in a Quorum Q150TS sputter-coater. Fruits were observed with a JEOL JSM-7600F field emission scanning electron microscope.

**Transmission Electron Microscopy (TEM)**
Dehiscence zones of *E. pusilla* fruits were cut and fixed with Karnovsky fixative (2% formaldehyde and 2.5% gluteraldehyde) for 3 hours on a rotating platform at 4° C and 2 hours post-fixed with 1% osmium tetroxide ($OsO_4$) in the dark, both in 0.1 M sodium cacodylate buffer (pH 7.2). Samples were stained and dehydrated with a 1% uranyl acetate replacement (UAR) (Electron Microscopy Sciences) in 30% ethanol and dehydrated in an ascending 1% UAR ethanol series of 50-70-96 for 10 minutes each, twice with absolute ethanol for 20 minutes and once with acetonitrile for 20 minutes each, all on a rotating platform. They were embedded in epoxy-resin (48% EMbed-812, 21% dodecenyl succinic anhydride, 29% methyl-5-norbornene-2,3-dicarboxylic anhydride (Electron Microscopy Sciences) and 2% benzyl dimethylamine (Agar Scientific)) through a graded series of epoxy-resin:acetonitrile; 1:2, 1:1 both for one hour, 1:1 overnight, 2:1 for one hour and 100% epoxy-resin for 3 hours. The submerged samples were placed in a vacuum for 20 minutes. Epoxy-resin was placed in molds and placed under vacuum for 20 minutes after which the samples were polymerized in an oven for 48 hours at 60 °C. Ultra-thin sections of 70 nm were cut with Leica Ultracut-S (Leica Co. Ltd) and directly mounted on copper grids (G2010-Cu, Electron Microscopy Sciences). The grids were rinsed in triple distilled water for 20 minutes and stained with 4% UAR for 20 minutes in the dark. The grids were subsequently rinsed 3 times with distilled water for 30 seconds. They were stained a second time with lead citrate (Electron Microscopy Sciences) according to Reynolds (1963). Images were made using the JEM-1400plus transmission electron microscope (JEOL Ltd).

**RNA extraction, cDNA synthesis and Quantitative real-time PCR**
Total-RNA was extracted from two different pools of fruits and seeds of the same inbred laboratory strain of *E. pusilla* using the RNeasy Plant Mini Kit.

Two biological replicates were used in this study of *E. pusilla*, as the variation in expression of developmental genes between individuals is negligible. Extracted RNA was treated with DNase I, Amp Grade (Invitrogen 1U/µl) to digest single- and double-stranded DNA following the manufacturer's protocol.

cDNA was synthesized with up to 1 µg of DNase-treated RNA using iScript™ cDNA Synthesis Kit (Bio-Rad Laboratories) following the manufacturer's protocol. A positive control (CTRL) and a no reverse transcriptase (NRT) control were included.

Beacon Designer™ (Premier Biosoft, www.oligoarchitect.com) software was used to design primers (**Table S3**). Quantitative real-time PCR was performed using the CFX384 Touch Real-Time PCR system (Bio-Rad Laboratories) and iQ™ SYBR® Green Supermix (Bio-Rad Laboratories). The reaction mixture contained 1x iQ™ SYBR® Green Supermix, 0.2 µM of each primer, 1 ng cDNA template (triplicate reactions) for each target gene and from a fruit time-point for two sets of isolated RNA (six reactions in total). For each amplicon group, a positive control was included (=CTRL, RNA extracted from *E. pusilla* flower buds), a negative control (=NTC, reaction mixture without cDNA) and a no reverse transcriptase treated sample (=NRT, control sample during the cDNA synthesis). For all the qPCR reactions, the amplification protocol was as follows: initial denaturation of 5 min 95 °C followed by 20 s 95 °C; 30 s 61 °C; 30 s 72 °C; plate read, for 50 cycles; followed by a melting curve analysis of 5 s, 65 °C to 95 °C with steps of 0.2 °C to confirm single amplified products. Quantification Amplification results (QAR) were used for analysis with LinRegPCR (v2017.0, dr. J.M. Ruijter) (Ruijter *et al*., 2006; 2015).

**Yeast two-hybrid analysis (Y2H)**

A yeast two-hybrid screening was performed as described by De Folter *et al*. (2005;2011). Full-length clones were used for the construction of the yeast two-hybrid vectors (**Table S1**).

**Protein alignment and phylogenetic analysis**

Nucleotide sequences of *SPT/ALC, IND/HEC3* and *POUNDFOOLISH (PNF)/RPL* genes were downloaded from NCBI GenBank (www.ncbi.nlm.nih.gov), OneKP (https://sites.google.com/a/ualberta.ca/onekp) and Phytozome (https://phytozome.jgi.doe.gov). Most of the orchid nucleotide sequences were downloaded from Orchidstra (orchidstra2.abrc.sinica.edu.tw) and belong to orthologous group ORGP07662 for the predicted *SPT* genes, ORGP11571 for the predicted *HEC3* genes (both Pfam ID00010, HLH) and ORGP08194 for the predicted *RPL* genes (Pfam ID07526, POX and PF05920, homeobox_KN). A multiple sequence alignment was performed using the ClustalW alignment tool within Geneious v7.1.5 (www.geneious.com), based on translated nucleotides), taking into account protein domains and amino acid motifs that have been reported as conserved for the three gene lineages by Pabon-Mora *et al*. (2014). Regions that did not align were removed prior to further analysis. For the visualization of the alignments, Bioedit (www.mbio.ncsu.edu/BioEdit/bioedit.html) was used. Phylogenetic trees were generated with the Geneious Tree Builder plug-in using the Maximum likelihood (ML) method with gymnosperm gene lineages as out-group based on Pabon-Mora *et al*. (2014). Numbers above the branches

represent bootstrap support values from 100 replicates.

## Results

### Description of *Erycina pusilla* fruit development

To obtain more insight into the development of orchid fruits we documented changes in anatomy and morphology during fruit maturation of *Erycina pusilla*, which develops a dry dehiscent capsule, a very common fruit type for the orchid family (Dressler, 1993). Fruit development starts around one day after pollination (1 DAP) and the fruit reaches its final size at 16 weeks after pollination (16 WAP) **(Figure S1**). From 16 WAP onwards, the valves of the *E. pusilla* fruit slowly separate longitudinally, starting from the apex, but remaining fused at the base. The fruits usually open along three of the six fusion zones and split into three wide seed-bearing fertile valves whereas the narrow sterile valves remain connected to one of the fertile valves (**Figure S1H)**.



To examine the development of the different fruit tissues and the dehiscence zone in more detail, we investigated the *E. pusilla* fruits using light microscopy (LM), Scanning Electron Microscopy (SEM), Transmission Electron Microscopy (TEM) and X-ray micro computed tomography (micro-CT). **Figure 2** shows sections of the different fruit stages observed with LM. At 0 DAP the ovary consists of three sterile and three fertile valves, which form broad protrusions. The fertile valves contain two zones with small cells that will develop into the placenta (**Figure 2A**). In the first week after pollination, the fruits slowly expand while the ovules start to develop from the placenta of the fertile valves (**Figure 2B-C**). The valves each consist of three major fruit tissue layers: exocarp, mesocarp and endocarp. The outer exocarp is single layered; the mesocarp, or middle part of the fruit wall, has multiple cell layers consisting of parenchyma cells and the innermost endocarp consists of a few cell layers from which the placenta differentiates in the fertile valves (**Figure 2D**). At 2 WAP, specific cell files can be clearly observed at both sides of the sterile valve, marking the initiation of three V-shaped dehiscence zones. These zones emerge in the first week after pollination at the boundaries of the sterile and fertile valves and consist of a layer of small cells (**Figure 2D-E**). In orchids, the signal to initiate fruit maturation is provided by pollination rather than fertilization. Fertilization of the ovules inside the ovary takes place much later; this can vary from a few days up to months, depending on the circumstances and orchid species (Arditti, 1992;Chen and Fang, 2016;Chen *et al*., 2016;Fang *et al*., 2016). For *E. pusilla* six bundles of pollen tubes, at each side of the placenta, develop around 2 WAP and increase in diameter until 4 WAP, when fertilization probably takes place and seed development starts (**Figure 2D-E**). The pollen tube bundles begin to shrink at 6 WAP and have completely disappeared at 11 WAP (**Figure 2F-I**), indicating that the ovules have been fertilized. At 16 WAP, the fruit opens where the endocarp borders the sterile valves. Opening progresses along the dehiscence zones at the boundaries of the fertile and sterile valves by rupture of the cells (**Figure 2J-K**). The main morphological changes observed during the development of the fruit of *E. pusilla* are summarized in **Table 1**.

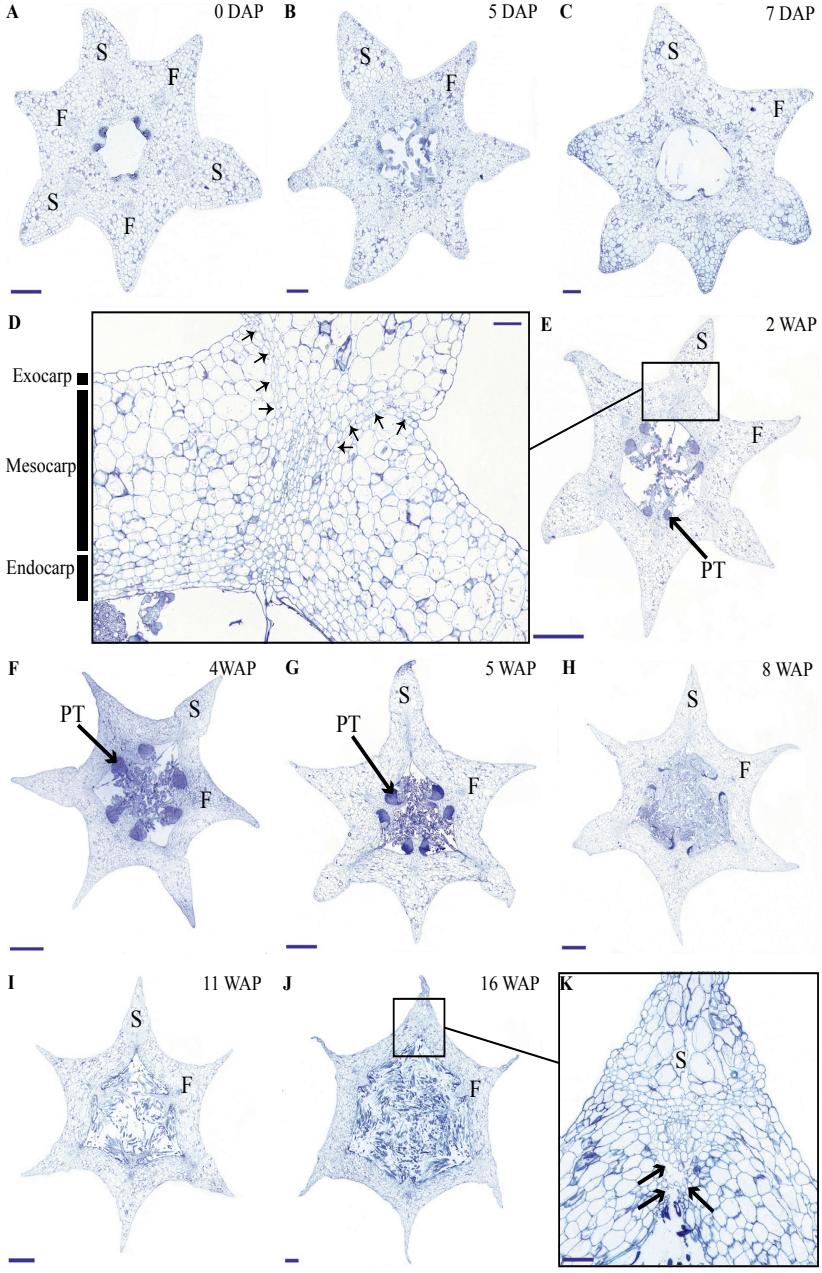



**Figure 2. Time-line of developing *E. pusilla* fruit cross sections, embedded in LR White and stained with toluidine blue. (A)** 0 DAP. **(B)** 5 DAP. **(C)** 7 DAP. **(D)** Magnified part of the sterile valve at 2 WAP. Arrows indicate the dehiscence zone. Black boxes the exo-, meso- and endocarp layer. **(E)** 2 WAP. **(F)** 4 WAP. **(G)** 5 WAP. **(H)** 8 WAP. **(I)** 11 WAP. **(J)** 16 WAP. **(K)** Magnified part of the sterile valve at 16 WAP. Black arrows indicate the dehiscence zone. Abbreviations: DAP = days after pollination; WAP = weeks after pollination; F = fertile valve; S = sterile valve; PT = pollen tube. Scale bar (A-C, K) = 0.2 mm, (D) = 0.1 mm, (E-I) = 1 mm, (J) = 0.5 mm.

**Table 1. The main morphological changes of *E. pusilla* fruits observed during development.**

| Time (days/weeks) | Main morphological changes |
|---|---|
| 0 DAP – 7 DAP | Elongation of the fruit<br>Cell division in the sterile and fertile valves<br>Trichome development |
| 2 WAP – 5 WAP | Increase of the volume of the fruit<br>Cell division and growth in the fertile and sterile valves<br>Development of six pollen tube bundles<br>Formation of dehiscence zones<br>Thickening of trichome walls |
| 6 WAP – 11 WAP | Increase of the volume of the cells in the sterile and fertile valves<br>Shrinking of the pollen tube bundles<br>Development of dehiscence zones<br>Lignification of the trichomes |
| 12 WAP – 16 WAP | Increase of the volume of the cells in the sterile and fertile valves<br>Disappearance of pollen tube bundles<br>Lignification of the endocarp<br>Dehiscence of the fruit |

Abbreviations: DAP = days after pollination, WAP = weeks after pollination.



Micro-CT was used to visualize internal structures of entire *E. pusilla* fruits during development. Fruits older than 4 WAP were full of seeds, which made it difficult to visualize other internal structures. In these fruits, many umbilical cords (funiculi) could be detected. The funiculi were connected to the three main dorsal vascular bundles of the fertile valves and the developing ovules in the placenta regions (Movie S1). According to Österberg (1883) and Rasmussen and Johansen (2006), fertile valves are located at the bases of the petals and sterile valves at the sepal bases. By following the vascular bundles in a micro-CT scan of an *E. pusilla* fruit of 5 DAP from the base upward to the wilted floral organs at the apex, we found further support for this hypothesis (**Figure 3**). The six main vascular bundles in the fruit connect in a plexus (**Figure 3**, indicated in purple) situated between the base of the fruit and the apex with wilted floral organs. The dorsal vascular bundles are connected to a petal (orange circles, **Figure 3A, C-D**) and the lateral vascular bundles to a sepal (green circles, **Figure 3B-D**).

Cross-sections of fruits in different developmental stages of the terrestrial orchid species *C. fastigiata* and *E. helleborine* were compared with those of the epiphytic orchid species *E. pusilla* (**Figure 2** and **Figure S3**). The number of cell layers is very constant throughout development for all three orchid species (**Table S2**). The fruit wall of *E. pusilla* consists of 13-19 cell layers at the narrowest parts of the fertile valve. The terrestrial species have fruits walls consisting of less cell layers, with 6-9 layers for *C. fastigiata* and 7-11 for *E. helleborine*

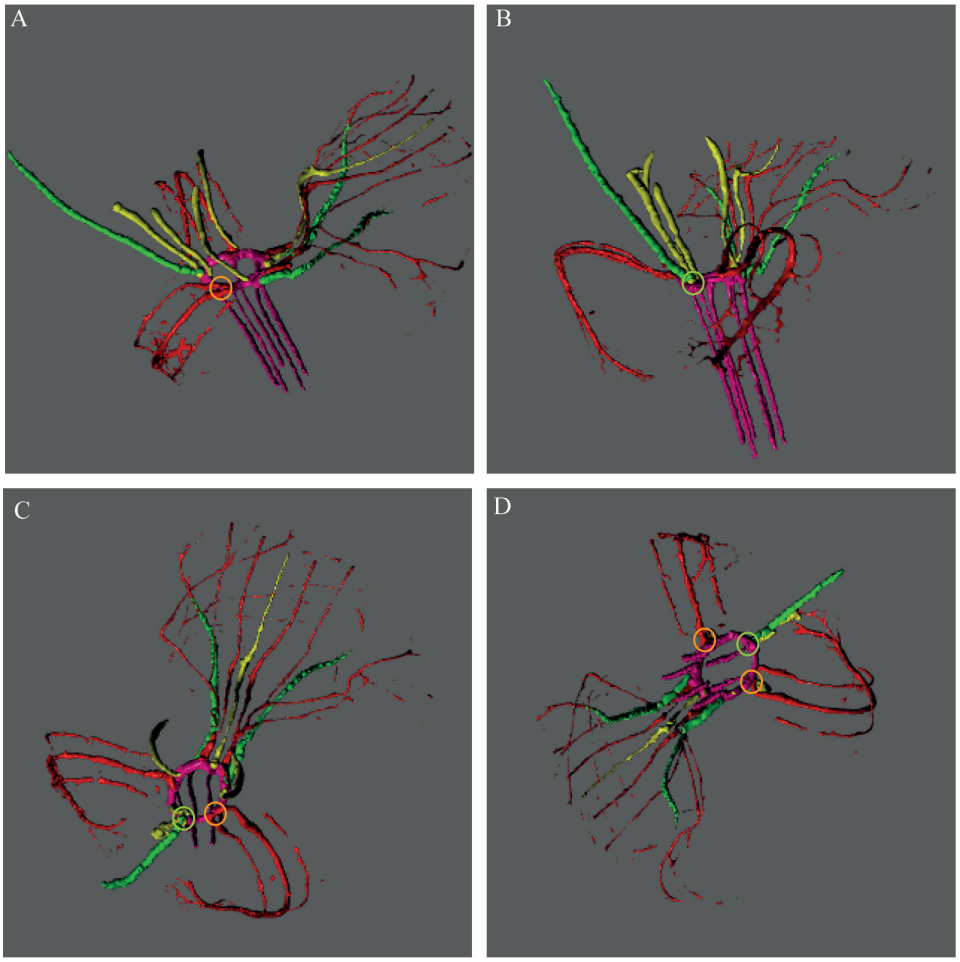



**Figure 3. Vascular bundle patterns in an *E. pusilla* fruit and wilted flower at 5 DAP visualized with a micro-CT reconstruction and depicted at different angles. (A)** Lateral view of the fruit and wilted flower from the right-hand side. **(B)** Posterior view of the fruit and wilted flower. **(C)** Apical view of the fruit with the wilted labellum projected upwards. **(D)** Inferior view of the fruit with the wilted labellum projected downwards. Color codes: green = vascular bundles in sepals; red = vascular bundles in petals; pink = plexus and vascular bundles in fruit; yellow = vascular bundles in stamens, stelidia and callus. Orange circle = connection of a dorsal vascular bundle with a petal. Green circle = connection of a lateral vascular bundle with a sepal.

(**Table S2**), both measured in the fertile valves, which is in agreement with the observations made by Beer (1863) that terrestrial species have thinner fruit walls.

While the number of cell layers in the fruit wall is stable, anticlinal cell divisions are responsible for initially expansion of the *E. pusilla* fruit. **Figure S4A** and **B** show that the cells in the fertile valves divide frequently until 5 DAP, after which the cell division rate slows down. At 2 WAP, only occasional cell division events were still observed and the maximum number of cells in the fertile valves is reached around 4 WAP. Cell expansion in the fertile valve is initiated between 5 and 7 DAP

(**Figure S4C**) and continues to contribute to fruit growth until 12-16 WAP (**Figure S4C** and **Figure 2F–J**), after which fruit size decreases again due to dehydration. Thus, the fertile valves of *E. pusilla* grow mainly by cell division between 0 and 5 DAP, by a combination of cell division and cell expansion between 5 DAP and 2 WAP, and by cell expansion after 2 WAP. In contrast to the fertile valves, only a limited number of cell divisions was observed in the sterile valves between 0 and 5 DAP (**Figure S4A**), after which cell division was terminated completely. Growth of the sterile valves between 0 and 5 DAP was mainly caused by cell expansion (**Figure S4C**), which was also responsible for further growth in the stages thereafter. However, the fertile valves appeared to expand to a much larger extent than the sterile valves. To investigate this in more detail, we determined growth of the different valves in fruits of *E. pusilla* by measuring the perimeter of the fertile and sterile valves during development and calculating the total area per valve ( **A-D**). Although the spread of the area measurements is high during the first two weeks of fruit development ( **D**), a clear increase of the fertile valve-area and a relative decrease of the sterile valve-area can be seen towards fruit maturation, confirming the hypothesis of Horowitz (1901) that increase in fruit volume of *E. pusilla* during development is mainly caused by expansion of the fertile valves.



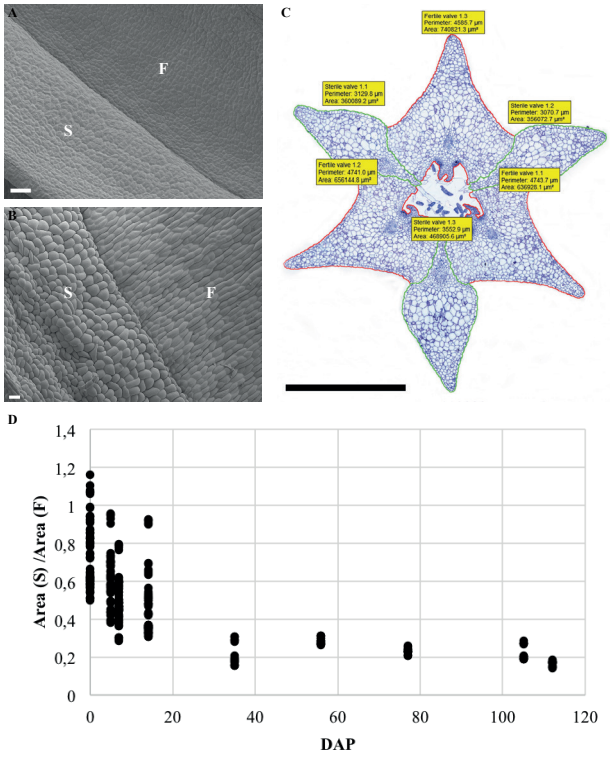



**Figure 4. Ratio and normalized area of the fertile and sterile valve of *E. pusilla* fruits during development from 0 to 112 DAP**. **(A)** SEM image of external surface of 1 WAP. **(B)** SEM image of external surface of 16 WAP. **(C)** Perimeter and area measurement of a fruit cross-section at 5 WAP. **(D)** Spread of the ratio area (S)/area (F). Each dot represents one individual fruit. Abbreviations: F = fertile valve; S = sterile valve. DAP = days after pollination. Color codes: green = sterile valve; red = fertile valve. Scale bar (A and B) = 100 µm, (C) = 1 mm.

**Dehiscence zone development in fruits of different orchid species**

The most characteristic aspect of the development of dry dehiscent fruits is the formation of the dehiscence zone. In dry fruits of Arabidopsis, specific cell files in the valve margin become lignified towards maturation and the fruit splits at a cell file adjacent to this lignified layer, called the separation layer (Ferrandiz *et al*., 1999;2002;Liljegren *et al*., 2004). The innermost endocarp layer also becomes lignified in Arabidopsis. In the Solanaceae, lignification occurs in dehiscent dry fruit species as well, but the extent and location is variable in different species (Pabon-Mora and Litt, 2011). To investigate lignification of *E. pusilla* fruits, we stained sections with phloroglucinol at different developmental stages. Lignin in the cell walls of the vascular bundles in the valves was stained with phloroglucinol. Until 6 WAP there was no sign of lignification anywhere besides the vascular bundles. Staining was observed in the trichomes, which appeared at 5 DAP in the endocarp on each side of the fertile valves, but became lignified around 8 WAP. The innermost endocarp cell layer also became lignified, which was evident around 10 WAP (**Figure 5A-D**). During the entire development of the fruit there was no sign of lignification of the dehiscence zones at the boundaries of the fertile and sterile valves of *E. pusilla*. This lignification pattern is similar to the one observed for the fruits of the epiphytic species *O. flexuosum* Sims (Mayer *et al*., 2011).

To investigate dehiscence zone development in dry dehiscent fruits of other orchid species, we also performed phloroglucinol staining of fruits from the terrestrial species *E. helleborine* and *C. fastigiata*. Interestingly, different patterns were observed for these species. The fertile valves of *E. helleborine* became lignified during development along the entire endocarp layer (**Figure 5G-H**), whereas lignification was absent in the fertile valves of *C. fastigiata* (**Figure 5I-J**). Lignification of the dehiscence zones of *E. helleborine* was not evident, but staining of three ripe fruits did reveal some lignification at the sites of separation between the sterile and fertile valves (**Figure 5H**). By contrast, the entire V-shaped dehiscence zone, including the small innermost endocarp cell layer of the sterile valves, became lignified in *C. fastigiata* towards maturation of the fruit (**Figure 5J**).

While no lignification of the dehiscence zones was observed in the fruits of *E. pusilla*, ultra-structural observations revealed the formation of an electron-dense, possible cuticular lipid-layer (**Figure 6**). This layer developed from the exocarp starting from an invagination between the fertile and the sterile valve (**Figure 6A**) and expanded during fruit development towards the endocarp (**Figure 6B-C).** At higher magnification polysaccharide fibers were visible (**Figure 6D-E**), known to be often present in cuticle layers (**Figure 6F**) (Fernandez *et al*., 2016). In micro-CT sections taken from 3D-models of *E. pusilla* fruits in different developmental stages, the vascular bundles of each of the six valves were clearly visible (**Figure 7**). Interestingly, additional stained cells could be observed between each fertile and sterile valve. These files have a similar location as the developing layer of small cells observed with LM (**Figure 2D**) and the lipid-layer observed with TEM (**Figure 6B-C**). The differential contrast of these cells indicates that their chemical content differs from that of their neighboring cells, and thus strengthens the idea that cells at the sterile-fertile valve border produce a yet uncharacterized

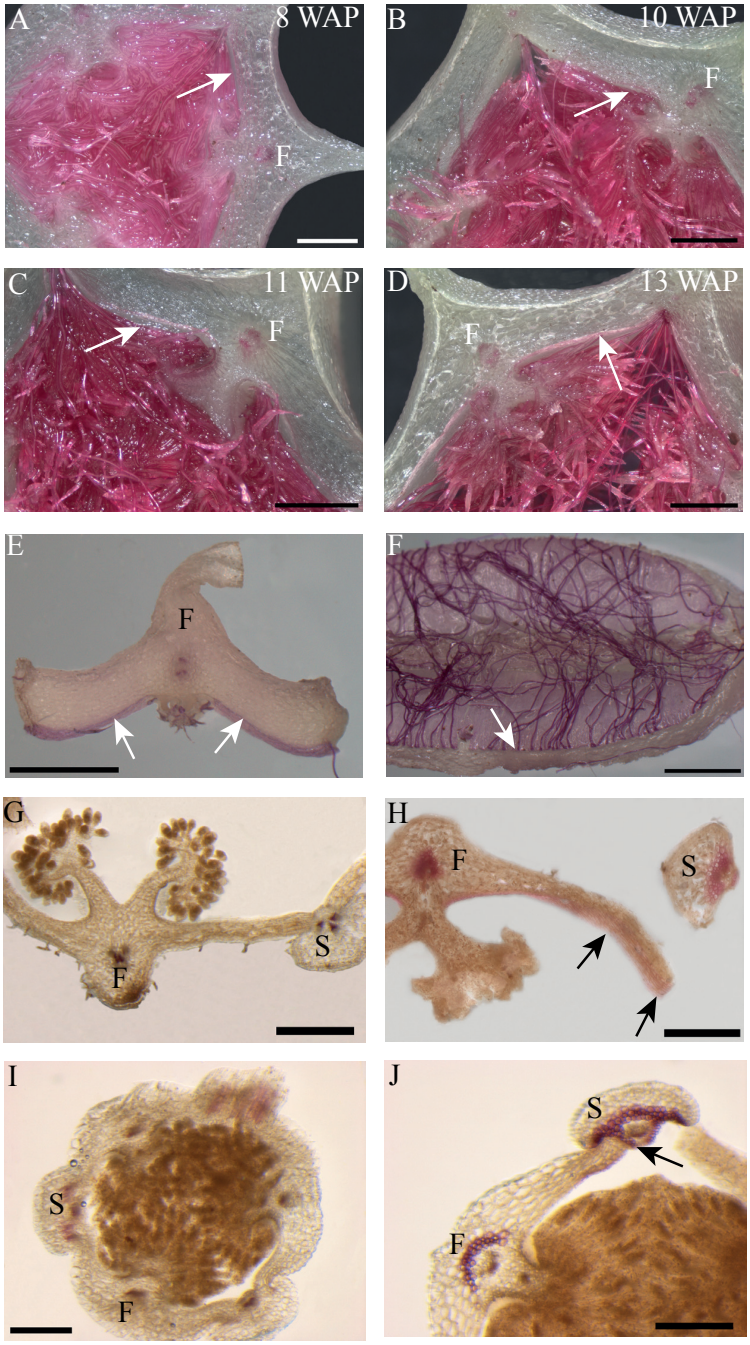



**Figure 5. Phloroglucinol staining of fruit cross-sections.** *Erycina pusilla*: **(A)** 8 WAP. **(B)** 10 WAP. **(C)** 11 WAP. **(D)** 13 WAP. **(E-F)** Dehisced fertile fruit valve. *Epipactis helleborine*: **(G)** Unripe and indehisced fruit. **(H)** Ripe and dehisced fruit. *Cynorkis fastigiata*: **(I)** Unripe and indehisced fruit. **(J)** Ripe and dehisced fruit. Abbreviations: F = Fertile valve; S = sterile valve. White and black arrows indicate lignified endocarp. Scale bar (A-D, G and H) = 1 mm. (E and F) = 2 mm. (I and J) = 0.5 mm.

substance instead of lignin prior to dehiscence.

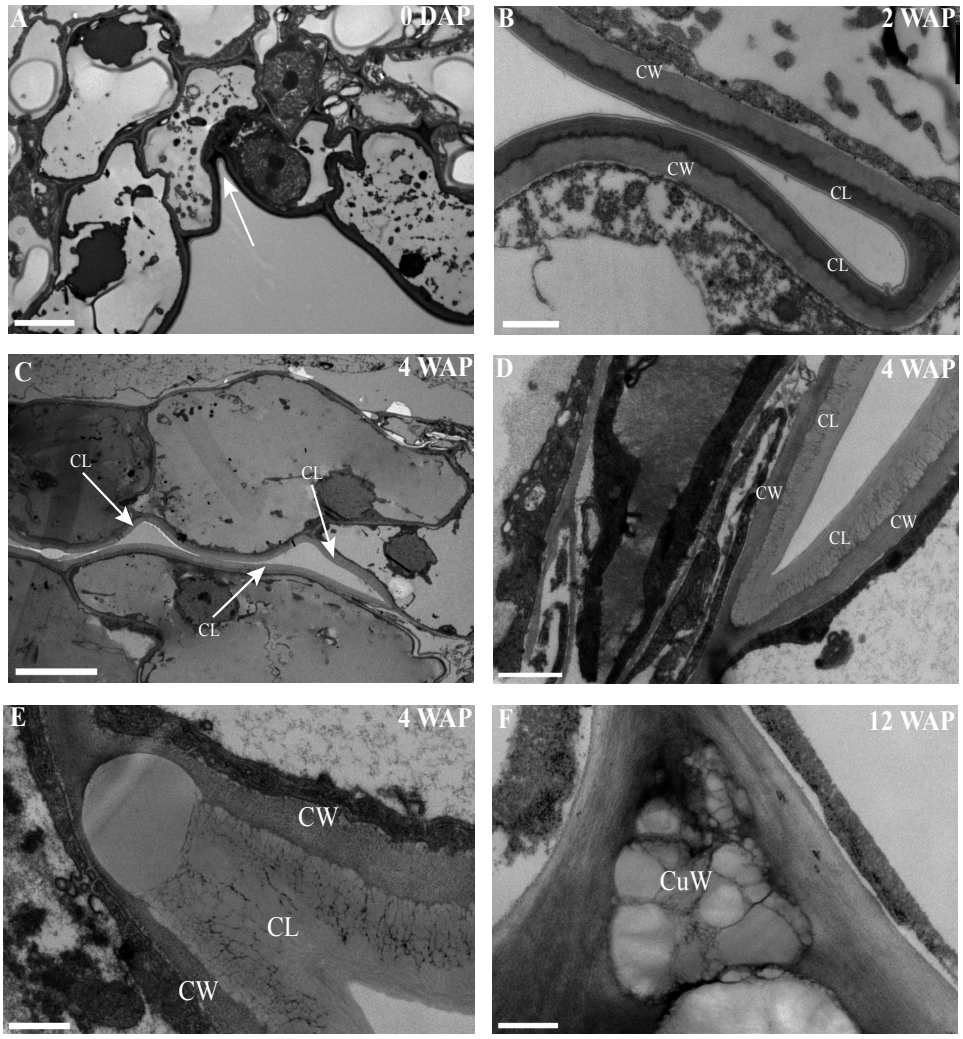



**Figure 6. An electron-dense, cuticular lipid layer developing in the dehiscence zone of *E. pusilla* fruits visualized with TEM. (A)** 0 DAP, white arrow pointing to the incision between a fertile and a sterile valve. Cuticle-like layer in the dehiscence zone **(B)** 2 WAP. **(C)** 4 WAP. White arrows point out the direction of the developing cuticle-like layer in the dehiscence zone. **(D-E)** Detailed image at 4 WAP. **(F)** 12 WAP. Abbreviations: CW = cell wall, CL = cuticular-like layer, CuW = cuticular-like wax. Scale bar (A) = 5 µm, (B and D) = 1 µm, (C) = 10 µm, (D and E) = 500 nm.

## Gene expression changes during orchid fruit development

To obtain a better understanding of the molecular mechanisms driving development and dehiscence of *E. pusilla* fruits, a detailed expression study was performed on MADS-box genes known to be expressed in fruits of this orchid species (Lin *et al.*, 2016) together with homologs of two bHLH-like genes (*HEC3* and *SPT*) and the

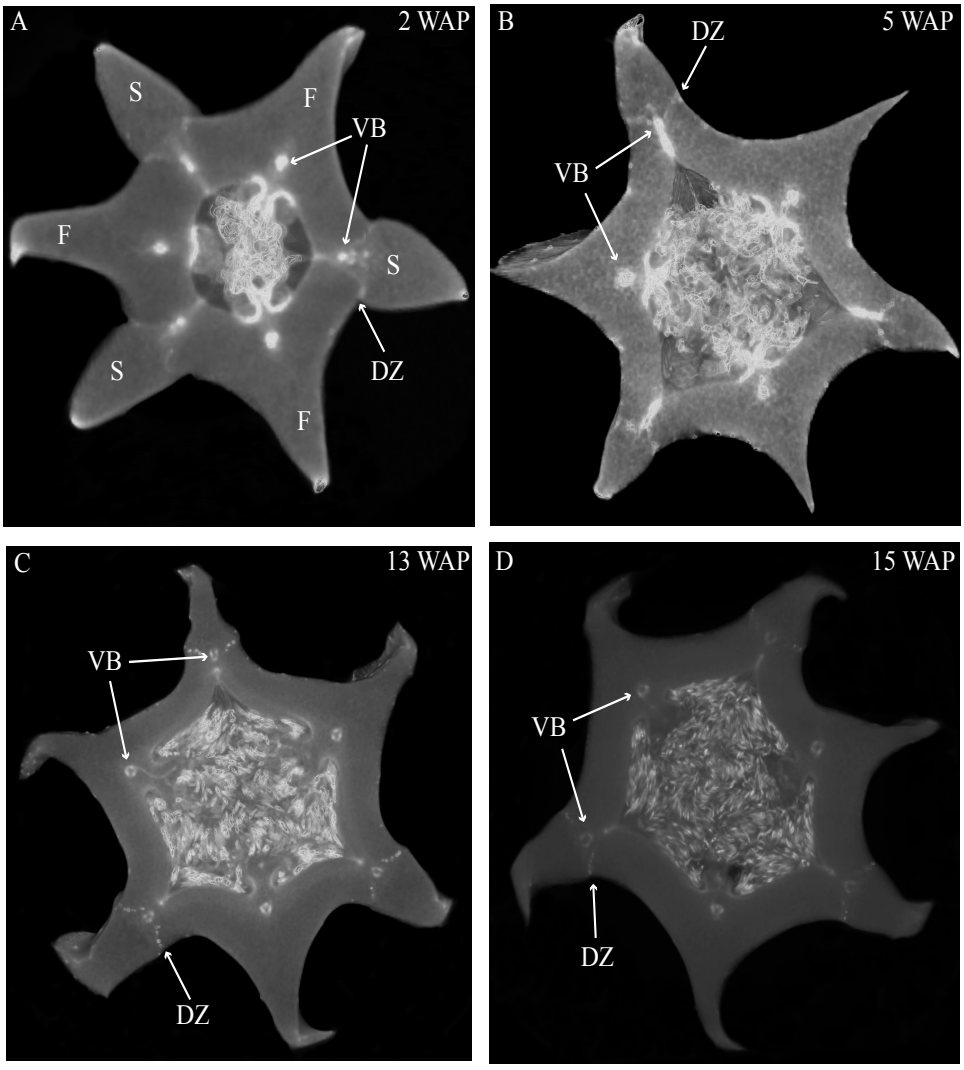



**Figure 7**. **Cross sections of 3D reconstructions of micro-CT scans of *E. pusilla* fruits stained with PTA. (A)** 2 WAP. **(B)** 5 WAP. **(C)** 13 WAP. **(D)** 15 WAP. Long arrows indicate one vascular bundle of a fertile or sterile valve. Short arrows indicate the location of the dehiscence zone between a fertile and a sterile valve. Abbreviations: VB = vascular bundle, F = fertile valve, S = sterile valve. (No scale bars can be included for 3D images).

homeodomain transcription factor (*RPL*), which are part of the Arabidopsis fruit gene regulatory network (Ferrandiz *et al*., 2000;Ferrandiz, 2002;Girin *et al*., 2011). Eleven different developmental stages of *E. pusilla* fruits were used, as well as a separate sample of seeds from mature fruits, to determine in which phases of fruit development the different genes may be functioning. To confirm whether orchid *RPL, HEC3* and *SPT* sequences, downloaded from different databases were the correct orthologs, they were aligned according to Pabon-Mora *et al*. (2014). With these alignments, ML analyses were performed. The orchid protein

sequences downloaded all aligned well with other SPT (**Figure S6** and **Table S5**), HEC3 (**Figure S7** and **Table S6**), and RPL (**Figure S8** and **Table S7**) homologs. Our ML analysis yielded trees with considerable up to high (62-99%) BS support for the orchid clades of SPT (**Figure S9**), HEC3 (**Figure S10**) and RPL (**Figure S11**).

The *AGL6-like* gene copy *EpMADS3* and the *SEP-like* genes *EpMADS8* and *-9* were expressed throughout fruit development (**Figure 8**), in line with their putative role in the formation of various higher order complexes. At 16 WAP, *EpMADS8* was clearly more highly expressed than in the other fruit stages (**Table S4**). The other MADS-box genes showed more specific expression patterns. The *AP1/FUL-like* genes *EpMADS10, -11 and -12* all showed a major increase of expression between 0 and 5 DAP, while their expression could hardly be detected after this stage. This expression pattern was also observed for the *AP3-like* genes *EpMADS14* and *EpMADS15*, the *AG-like* genes *EpMADS20* and *EpMADS22* and the *STK-like* gene *EpMADS23*. The *AG* homolog *EpMADS21* displayed a different expression pattern, with substantial expression until 4 WAP, while also having a peak at 5 DAP. Both *EpMADS21* and *EpMADS22* were still expressed around 16 WAP, when dehiscence was initiated. Interestingly, the *SHORT VEGETATIVE PHASE* homolog *EpMADS18* was also expressed in the fruit, with the highest expression at 16 WAP, indicating that it may contribute to the initiation of dehiscence. Overall, the *E. pusilla fruit*-expressed MADS-box genes showed predominantly high expression during fruit patterning, but *EpMADS18* and *EpMADS8* were also clearly expressed at 16 WAP, when the *AG* homologs *EpMADS21* and *EPMADS22* also still exhibited expression (**Figure 8** and **Table S4**).

The expression patterns of the *SVP* homolog *EpMADS18* and *EpRPL* were quite similar: expression between 0 and 5 DAP, a decrease in expression between 1 to 12 WAP and a significant increase in fruits at 16 WAP. The *bHLH-like* gene *SPT* showed a similar expression pattern, although it was also moderately expressed between 5 DAP and 12 WAP. This indicates that these genes may both be involved in fruit patterning as well as in the initiation of dehiscence. The bHLH-like gene *HEC3* was mainly expressed after 4 WAP with a significant increase at 16 WAP. These expression data suggest that all three homologs of Arabidopsis fruit specification genes are also important for the development of orchid fruits **(Figure 8-9** and **Table S4)**.

**MADS-box protein-protein interaction during orchid fruit development**

To investigate whether the fruit regulatory network present in Arabidopsis (Dinneny *et al*., 2005) and tomato (Bemer *et al*., 2012) could be conserved in orchids, we performed yeast two-hybrid experiments with fruit-expressed MADS-domain proteins from *E. pusilla*. Full-length proteins were fused to the GAL4 binding domain or activation domain and all baits and preys were screened against each other. The auto-activation test revealed that EpMADS8*,* EpMADS21 and EpMADS22 exhibited auto-activation, and these proteins were therefore only tested as prey (fused to the activation domain). The results of the yeast two-hybrid analysis are summarized in **Table 2. Figure S5** shows the growth of the spotted yeast colonies.

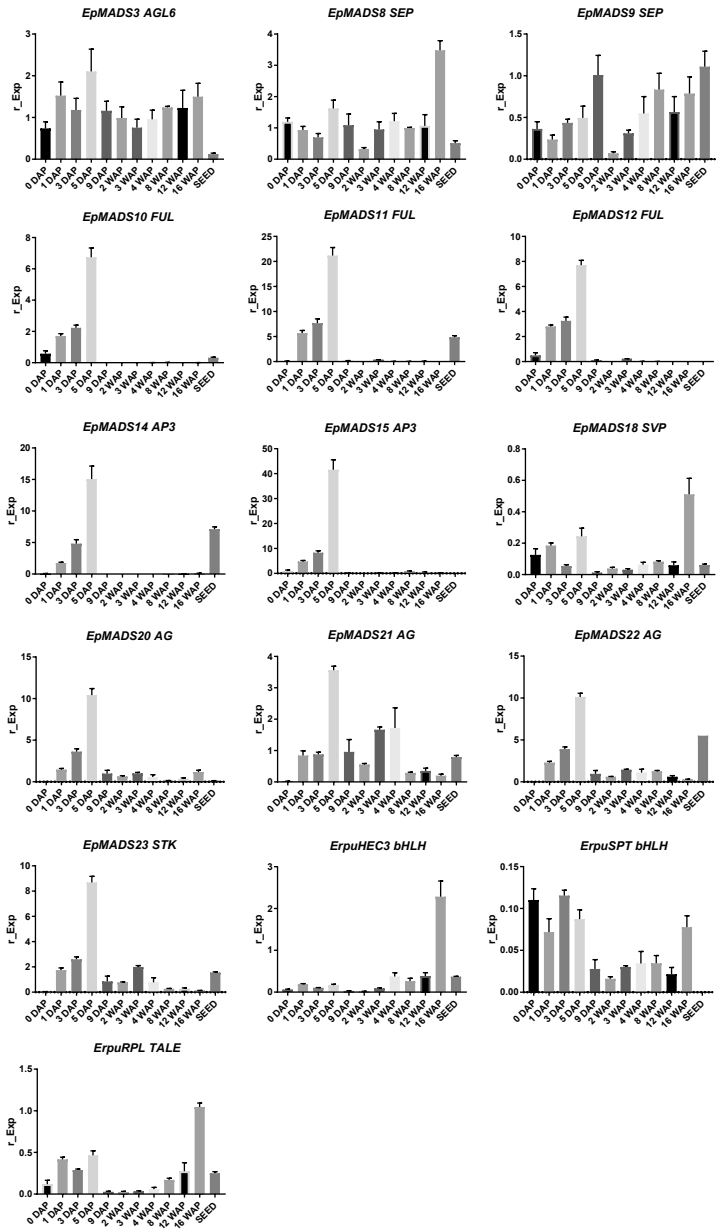



**Figure 8. Fruit specific expression patterns of selected MADS-box gene copies in *E. pusilla* of *AGL6*, *SEP*, *AP1/FUL, AP3, SVP, AG*, *STK* and three fruit specific genes *SPT, HEC3 and RPL*.** Each graph shows the relative expression during twelve stages of development. Expression of the genes was normalized to the geometric mean of three reference genes *Actin, UBI2* and *Fbox*. Each column shows the relative expression of two cDNA pools of different fruits of the same inbred laboratory strain, both tested in triplicate. Abbreviations: DAP = Days After Pollination, WAP = Weeks after Pollination. Y-axis: relative gene expression. The error bars represent the Standard Error of Mean. P-value style: GP: >0.05 (ns), <0.05 (*), <0.01 (**), <0.001 (***), <0.0001 (****).

**Table 2**. **Yeast two-hybrid screening of interactions between MADS-box proteins of *E. pusilla* and Arabidopsis FUL (AtFUL).**

| Arabidopsis homolog | *E. pusilla* | EpMADS3 | EpMADS8 | EpMADS9 | EpMADS10 | EpMADS11 | EpMADS12 | EpMADS14 | EpMADS15 | EpMADS18 | EpMADS20 | EpMADS21 | EpMADS22 | EpMADS23 | AtFUL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AGL6 | EpMADS3 | | | | | | | | | | | | | | |
| SEP | EpMADS8 | | X | | | | | | | | | | | | |
| SEP | EpMADS9 | | | | | | | | | | | | | | |
| AP1 | EpMADS10 | | | | | | | | | | | | | | |
| FUL | EpMADS11 | | | | | | | | | | | | | | |
| FUL | EpMADS12 | | | | | | | | | | | | | | |
| AP3 | EpMADS14 | | | | | | | | | | | | | | |
| AP3 | EpMADS15 | | | | | | | | | | | | | | |
| SVP | EpMADS18 | | | | | | | | | | | | | | |
| AG | EpMADS20 | | | | | | | | | | | | | | |
| AG | EpMADS21 | | | | | | | | | | | X | | | |
| AG | EpMADS22 | | | | | | | | | | | | X | | |
| STK | EpMADS23 | | | | | | | | | | | | | | |
| AtFUL | AtFUL | | | | | | | | | | | | | | |

Light grey: Interaction in one direction only (either with AD or BD). Dark grey: Interaction in both directions. White: No interactions. X: Not analyzed.

The AP1/FUL protein EpMADS11 showed a similar interaction pattern compared to the Arabidopsis FUL-protein and could also form homo-dimers like AtFUL. EpMADS10 did not interact at all. EpMADS18, an SVP-like protein, formed homo-dimers and bound to most of the proteins used in this study. No interaction was found between members of the AG- and STK clades. Two AP3 proteins, EpMADS14 and -15 did not interact at all with the MADS-box proteins used in this study.

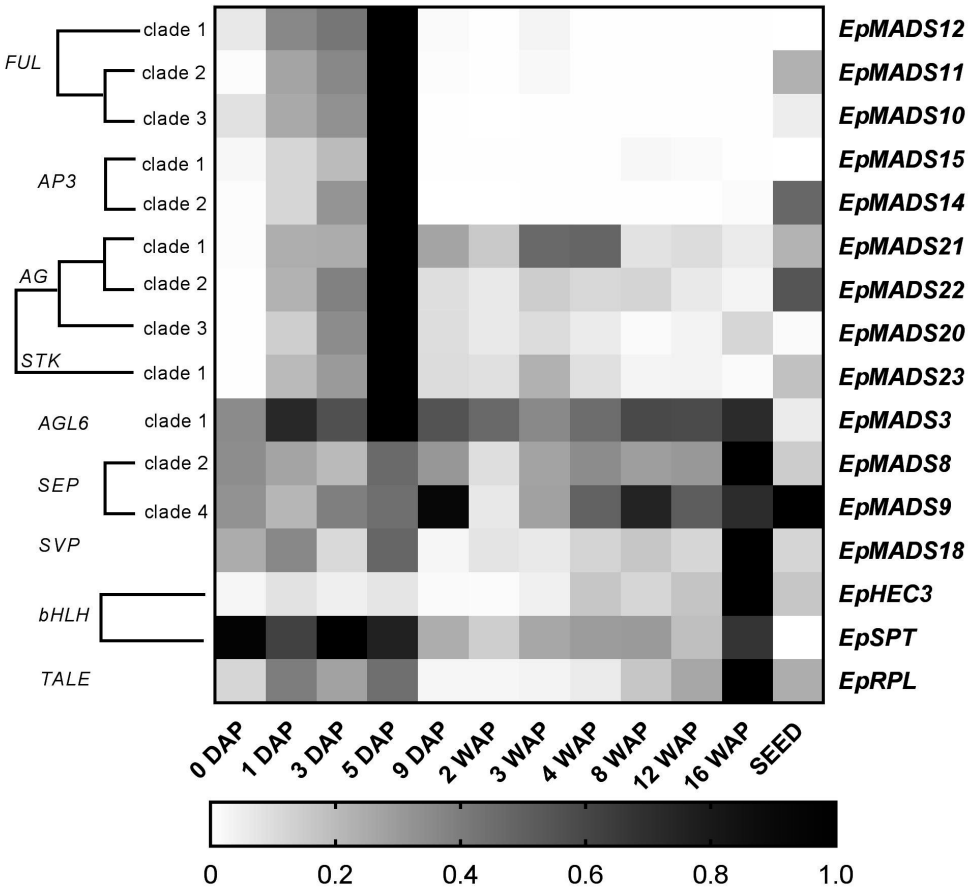



**Figure 9. Heat map representation of expression of developmental genes in *E. pusilla* fruits and ripe seeds.** The *FUL-, AP3-, AG-, STK-, AGL6-, SEP-, SVP-, bHLH-* and *TALE-like* copies were retrieved from different gene lineage clades during eleven stages of fruit development. Expression of the genes was normalized to the geometric mean of the reference genes *Actin, Fbox* and *UBI2*. The scale for each gene was set to 1 for the highest value. Abbreviations: DAP = Days after pollination, WAP = Weeks after pollination.

## Discussion

**Conservation and divergence of fruit anatomy within the orchid family**
According to Beer (1863), epiphytic and terrestrial orchid fruits are thought to develop differently regarding fruit wall thickness. We found that both fruits of the terrestrial orchid species *C. fastigiata* and *E. helleborine* do indeed have relatively thinner walls than fruits of the epiphytic orchid species *E. pusilla* **(Figure 2** and **S3)**. Fruits of terrestrial and epiphytic orchids differ in the number of cell layers in the exocarp, resulting in thin-walled terrestrial fruits and thick-walled epiphytic fruits, possibly to protect the latter against the much more fluctuating moisture and UV radiation levels present in tree canopies as compared with the forest floor.

Several models of fruit dehiscence have been described for angiosperms

with parietal fruits: (i) loculicidal dehiscence: each locule splits at the middle of each carpel while the septa remain intact (ii) septicidal dehiscence: each septum, bordering a locule, splits in two and (iii) septifragal dehiscence: each placental region with its adjacent valves breaks away from the sterile (septal) region. The orchid species we investigated all showed septifragal dehiscence, whereby each carpel consisted of two halves of fertile valves and one sterile valve (**Figure 1**) as described by Brown (1831).

During the entire development of the fruit there was no sign of lignification of the dehiscence zone at the boundaries of the fertile and sterile valves of E. pusilla. Oncidium flexuosum, another epiphytic member of the Oncidiinae, showed the same lignification pattern as *E. pusilla* and the valves of this orchid also separated longitudinally along a dehiscence zone consisting of small cells (Mayer *et al*., 2011). Using osmium tetroxide as a secondary fixative, not only enhanced the contrast but it also increased the retention of lipids in the tissue (Hayat, 1970). With a two-step staining procedure with UAR followed by lead citrate (Reynolds, 1963) in TEM, we could visualize the development of a cuticular lipid layer in E. pusilla fruits. Cross sections of 3D reconstructions of micro-CT scans of E. pusilla fruits, stained with phosphotungstic acid (PTA), a negative staining agent for lipids (Melchior *et al*., 1980;Bello *et al*., 2010), also revealed a white-stained layer at the dehiscence zones of the valves of *E. pusilla* (**Figure 7**). On the other hand, in both terrestrial species a layer of lignified cells can be seen at the valve margins (**Figure 5H** and **5J**).

The major components of plant cuticles are cutin and cuticular wax (Yeats and Rose, 2013). Both are primarily composed of fatty acid derivates. Renault *et al*. (2017) described a lignin-related biochemical pathway in mosses, which is responsible for the formation of cuticles. These authors found that cutin and lignin have the same evolutionary origin whereby the enzyme CYP98 (from the family of cytochromes P450) plays an important role in either the production of lignin in seed plants and in the development of a phenol-enriched cuticle in mosses. If indeed there is a common ancestor of lignin, cutin and suberin (wax) polymers, this suggests that orchids use different strategies for the dehiscence of their fruits. Cutin synthase plays an important role in the cutin biosynthesis pathway and belongs to the GDSL-lipase family, which is widely represented in orchid transcriptomes. We found more than 75 possible hits for *E. pusilla* in the Orchidstra database (orchidstra2.abrc.sinica.edu.tw). Which of these proteins are responsible for the formation of the cuticular lipid layer at the dehiscence zone of *E. pusilla* fruits is still unknown and deserves further investigation.

**Fruit molecular networks appear to be partly conserved between eudicots and orchids**

We show in this study that orchid homologs of well-known Arabidopsis fruit development genes are also expressed during orchid fruit development, and that the encoded MADS-domain transcription factors are able to form dimeric complexes with a similar composition as the Arabidopsis complexes. However, there were also distinct differences. For example, two *AP3*-clade members were found to be expressed during orchid fruit development. This might be correlated with the fact that floral remains stay attached to the developing fruit in orchids, whereas in tomato,

pepper and thale cress these remains fall off after fruit maturation has been initiated.

The *FUL* homologs *EpMADS11* and *EpMADS12* are expressed during fruit patterning, suggesting that they may regulate the initiation and specification of the fruit in the cell division stage. Remarkably, both genes are not expressed during later stages when the dehiscence zone is specified. This suggests that this mechanism is regulated without the contribution of *FUL-like* genes, which would be different from the situation in Arabidopsis.

In line with protein-protein interaction data from Arabidopsis (de Folter *et al*., 2005) and tomato (Leseberg *et al*., 2008), the FUL-homologs EpMADS11 and EpMADS12 are both able to interact with the AG homologs EpMADS21 and EpMADS22. EpMADS11 also interacts with the third C-class gene EpMADS20. Because the split of the C-class genes *AG* and *SHP* occurred at the base of the core Eudicots (Kramer *et al*., 2004;Zahn *et al*., 2006), it is not possible to distinguish the three *E. pusilla* C-class genes based on their sequence identity. However, distinct expression of *EpMADS21* and *EpMADS22* in the seeds suggests that these genes could function as SHP homologs. In addition to interacting with C-class proteins, the FUL ortholog EpMADS11 can also interact with all three SEP/AGL6 homologs, while EpMADS12 interacts with the SEP homolog EpMADS8 only. Interestingly, *EpMADS10*, which belongs to the *AP1/FUL* clade and is highly expressed in fruits at 5 DAP, does not interact with any other fruit-expressed MADS proteins. A closer inspection of the EpMADS10 protein sequence unveiled a deletion of a 9 amino acids domain at the border of the I- and K-domains. Because both regions are essential for the dimerization of MADS-domain proteins (van Dijk *et al*., 2010), this deletion is a probable cause for the lack of interactions found (**Table 2** and **Figure S5**). We also could not detect any interactions for the AP3 homologs EpMADS14 and -15, which form obligate heterodimers with PISTILLATA (PI) homologs (Bartlett *et al*., 2016) in most species. The *AP3* homolog *EpMADS13* and *PI* homolog *EpMADS16* are hardly expressed in orchid fruits (Lin *et al*., 2016). It is good to note however, that both PI and AP3 did not show any interactions in Arabidopsis yeast two-hybrid studies (de Folter *et al*., 2005) despite ubiquitous proof that they are interacting *in planta*, suggesting that false negative data may easily be obtained for these proteins. Based on the combination of our expression and interaction data, it is very likely that an AG/SHP-SEP-FUL regulatory module is also acting during *E. pusilla* fruit development. The FUL paralogs EpMADS11 and EpMADS12 can both interact with SEP and AG/SHP homologs, but EpMADS12 exhibits a more specific interaction pattern and may have sub-functionalized after duplication. To compare the interaction capacity of the *E. pusilla* FUL homologs with that of Arabidopsis FUL (AtFUL), we also tested AtFUL against the fruit-expressed *E. pusilla* MADS-domain proteins and found that the interaction profile of AtFUL was highly similar to that of EpMADS11, indicating conservation of protein structure and function.

Orchids disperse their seeds by different ways using either dehiscent or indehiscent fruits and by various vectors such as wind, water or animals. Here we propose a fruit developmental network for the dry dehiscent fruit of the orchid species *E. pusilla* based on the expression data and the protein-protein interactions data obtained in this study **(Figure 10)**.

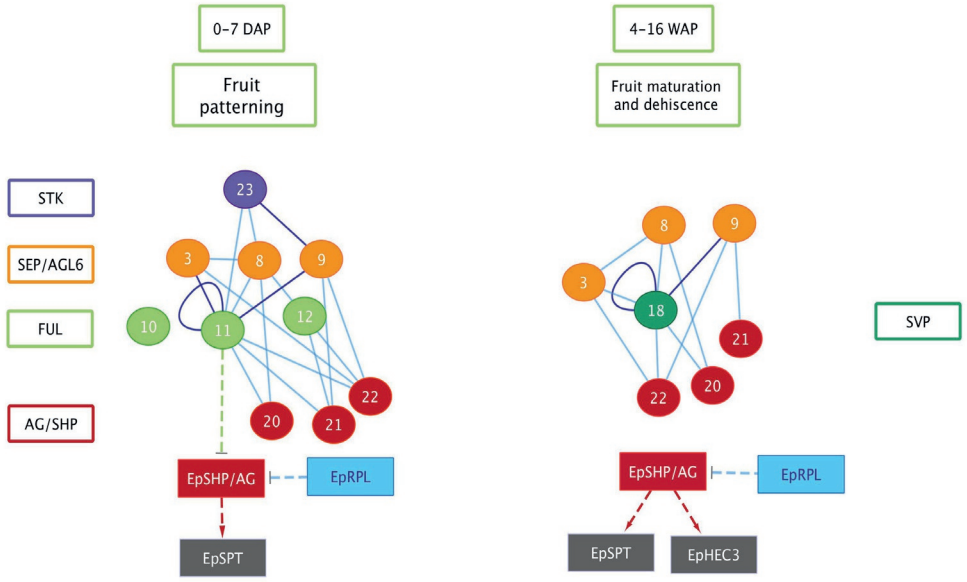



**Figure 10. Orchid fruit developmental protein and gene network for *E. pusilla*.** Abbreviations: Circles: MADS-box proteins; rectangles: genes; Solid lines: validated protein–protein interactions (blue: one direction, purple: both directions); dashed arrows: putative activation interactions; dashed T-bars: putative repression.

Our model is based on the fruit core genetic regulatory network from Arabidopsis (Ferrandiz *et al*., 2000;2002;Roeder *et al*., 2003). Based on our results, there is a clear difference between the regulators involved in fruit patterning and fruit maturation and dehiscence in *E. pusilla*. During fruit patterning, FUL-like proteins interact with AG/SHP-like proteins and SEP-like proteins, thereby possibly regulating downstream targets that are also involved in Arabidopsis fruit development, such as SPT. *RPL* is also moderately expressed in this stage and may suppress the transcription of *AG/SHP* similar to Arabidopsis. During maturation and dehiscence, the MADS-box genes are less active, except for the E-class homologs *EpMADS8* and *EpMADS9* and the SVP homolog *EpMADS18*. In addition, homologs of the Arabidopsis dehiscence-zone specifiers *SPT*, *HEC3* and *RPL* are all highly expressed, indicating that they perform a more prominent role in the last stage of fruit development, when the dehiscence zone is formed. *HEC3* and *RPL* are only expressed during late fruit development, which is different from the Arabidopsis fruit regulatory network. RPL may suppress the expression of *HEC3* and *SPT* in the orchid fruit regulatory network.

In conclusion, the expression and protein-protein interaction data that we present here suggests that orchid fruit development may be regulated via a similar regulatory network as Arabidopsis fruit development. However, more functional data are required to validate our orchid fruit regulatory model, such as knockout studies of fruit-expressed transcription factors using for instance CRISPR/Cas9 mutagenesis, in combination with *in situ* hybridization studies for higher spatial resolution. These need to be performed in *E. pusilla* and other orchid species with different dehiscence patterns, to determine a general regulatory network for orchid

fruits as well as the main molecular factors that are responsible for divergence in dehiscence zone formation.

## Supplementary material

If not published in this thesis further supplementary material for this chapter can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019. 00137/full#supplementary-material.



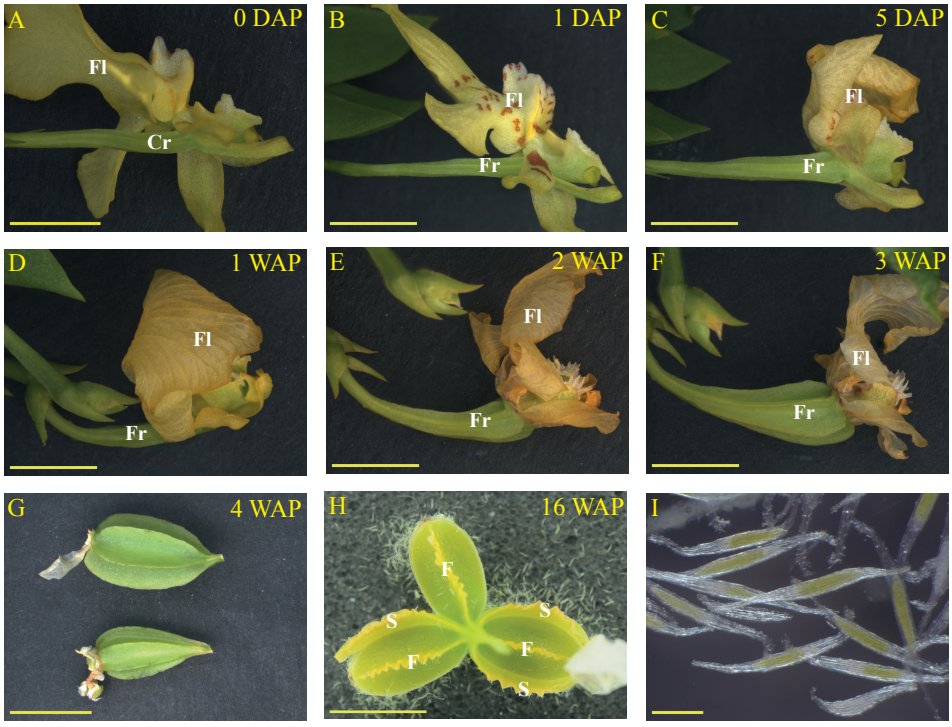



**Figure S1**. **Development of *E. pusilla* fruits from unfertilized ovary to mature fruit with ripe seeds.** **(A)** Ovary. **(B)** Fruit 1 DAP. **(C)** Fruit 5 DAP. **(D)** Fruit 1 WAP. **(E)** Fruit 2 WAP. **(F)** Fruit 3 WAP. **(G)** Fruit 4 WAP. **(H)** Dehisced fruit 16 WAP. **(I)** Ripe seeds of dehisced fruit. Abbreviations: DAP = days after pollination, WAP = weeks after pollination, Fl = flower, Cr = carpel, Fr = fruit, F = fertile valve and S = sterile valve. Scale bars (A–G) = 5 mm, (H) = 10 mm and (I) = 1 mm.

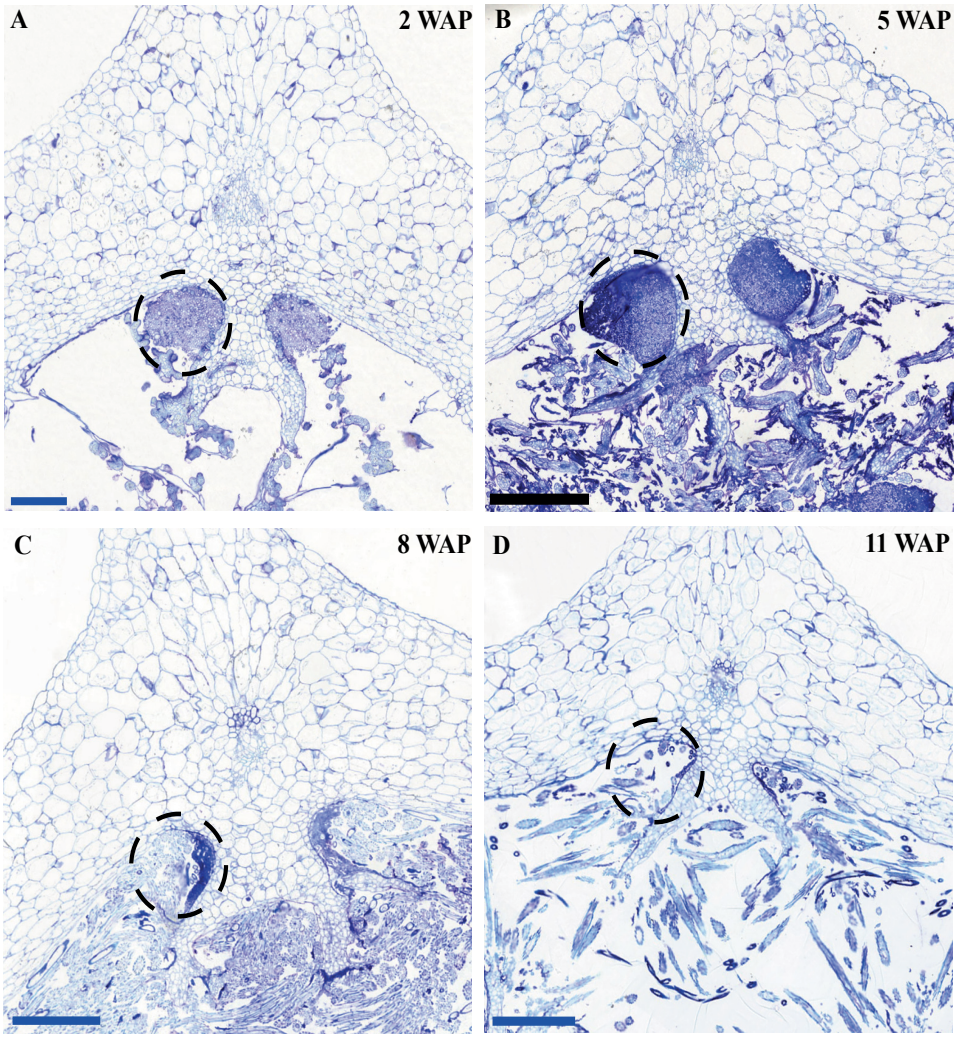



**Figure S2**. **Time-line of developing *E. pusilla* fruit cross-sections of one of the fertile valves. (A)** 2 WAP. **(B)** 5 WAP. **(C)** 8 WAP. **(D)** 11 WAP. Dashed circles indicate the left pollen tube bundle located in the fertile valve. Scale bar (A) = 0.2 mm, (B-D) = 0.5 mm.

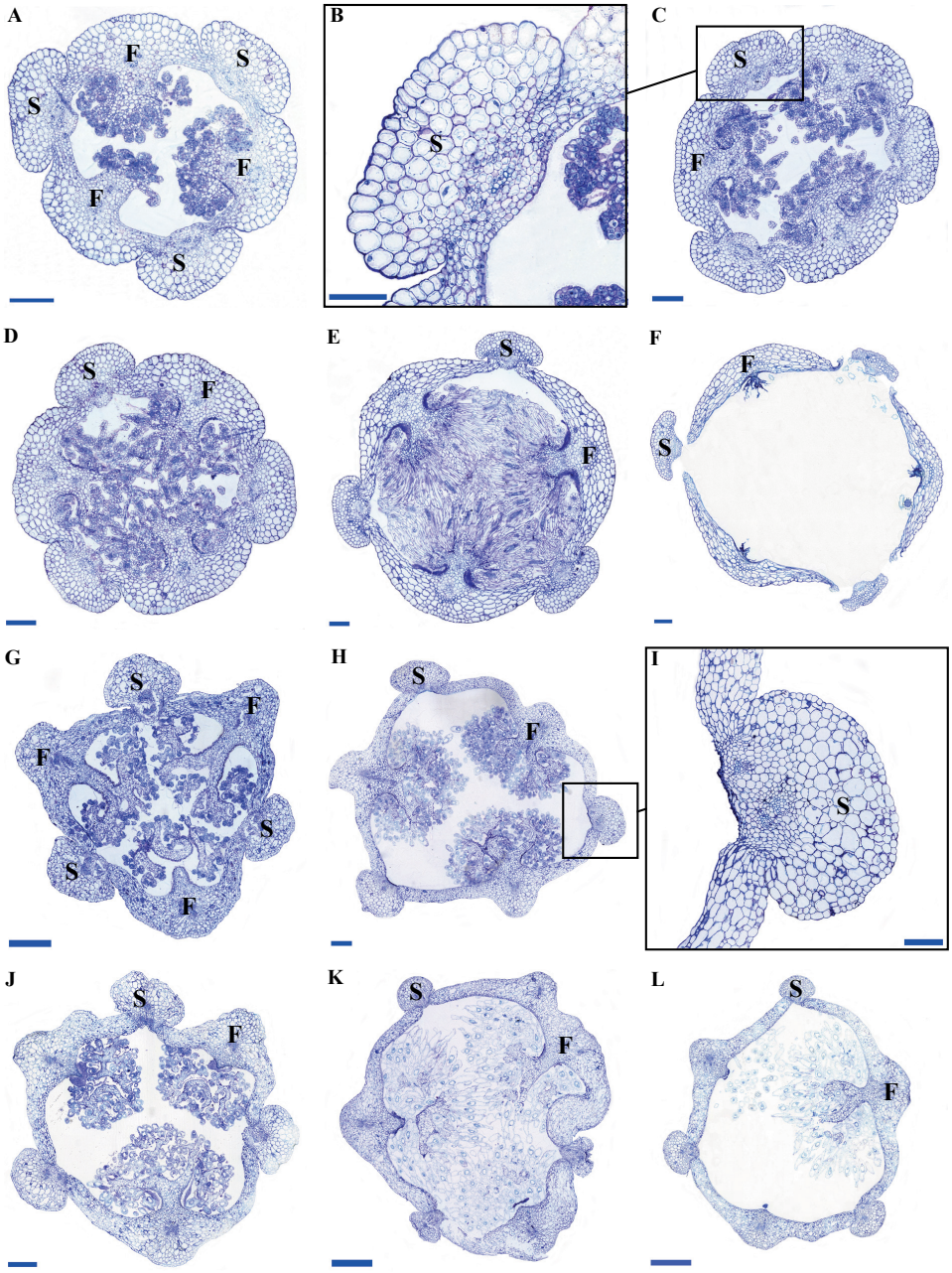



**Figure S3**. **Time-line of developing *Cynorkis fastigiata* and *Epipactis helleborine* fruits, embedded in LR White and stained with toluidine blue. (A-F)** *C. fastigiata*. **(G-I)**. *E. helleborine.* Scale bars (A, C-F, I) = 0.2 mm, (B) = 0.1 mm, (G, H, J) = 0.5 mm. (K, L) = 1 mm. Abbreviations: F = fertile valve; S = sterile valve.

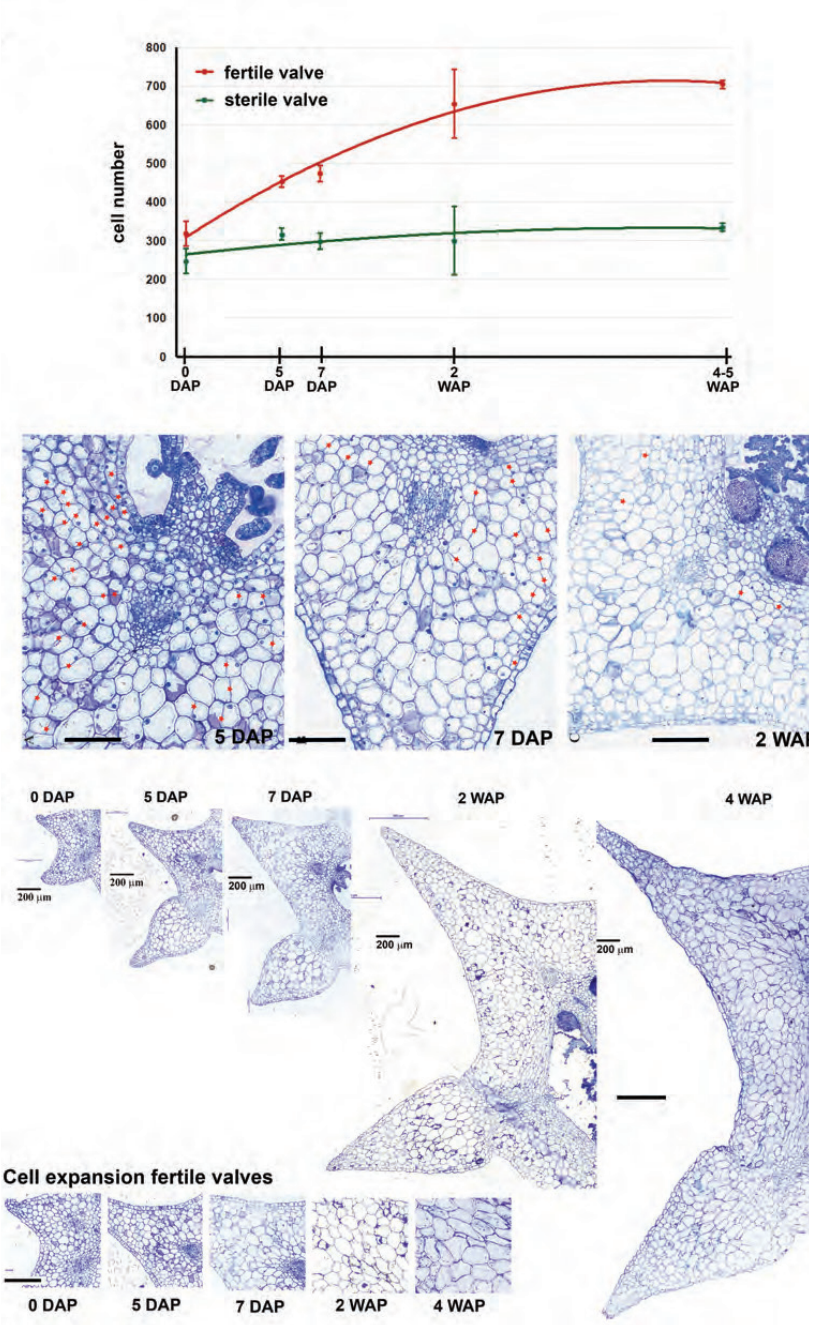



**Figure S4**. **Growth of *E. pusilla* fruits by cell division and cell elongation. (A)** Graph showing the number of cells in different growth stages. Cell number was determined in cross sections of fruits at different growth stages from at least 6 sterile and 6 fertile valves per growth stage. A trend line was drawn through the different data points. The error bars depict the SE. **(B)** Details of fertile valve sections showing the cell divisions with red asterisks. **(C)** The size of sterile and fertile valves in different growth stages. The detailed pictures (bottom) show the increase in cell size. Scale bar (A-C) = 0.2 mm.

**Figure S5. Results of the yeast-two hybrid assay.** Thirteen *E. pusilla* MADS-box proteins and one Arabidopsis protein (AtFUL) were screened against each other. After mating, the diploid yeast were grown for five days on SD medium lacking Leu, Trp, and His, supplemented with 5 mM 3-amino-1,2,4 triazole. Growth indicates an interaction between bait and prey.

**Figure S6**. **Alignment of the bHLH domain of SPATULA/ALCATRAZ proteins based on Pabon-Mora *et al*. (2014) extended with orchid gene lineages.** The bHLH was drawn based on (Toledo-Ortiz *et al*., 2003) and corresponds with positions K359-Q410. Within the bHLH domain, black arrows indicate positions E13, R16, L27, K39, L56, which are conserved in all bHLH plant and animal genes. The H9 and R17 positions (red arrows) show amino acids that provide the SPT/ALC proteins with G-box (CACGTG) binding activity. Black dashed boxes: N-flank and C-flank showing the conserved motif LQLQVQ.



**Figure S7. Alignment of the bHLH domain of HECATE3/INDEHISCENT proteins based on Pabon-Mora *et al*. (2014) extended with orchid gene lineages.** The bHLH was drawn based on (Toledo-Ortiz *et al*., 2003) and corresponds with positions N462-L515. Left black dashed box: N-flank of the bHLH domain: HEC domain (Kay *et al*., 2013) including domain 17 (Pires and Dolan, 2010). Right black dashed box: poorly conserved C-flank.

**Figure S8. Alignment of the BELL-domain and the Homeo-domain of REPLUMLESS/POUNDFOOLISH (PNF) proteins based on Pabon-Mora *et al*. (2014) extended with orchid gene lineages.** The BELL domain (Smith *et al*., 2002) and the conserved Homeodomain, based on Mukherjee *et al*. (2009) were drawn.

**Figure S9**. **ML tree of the *SPATULA/ALCATRAZ* genes in seed plants**. Branch colors denote the following taxa: Persian green = Gymnosperms, Blue = Basal angiosperms, Middle washed yellow = Monocots, Green = Basal eudicots, Purple = Core eudicots, Red = Brassicaceae. Bootstrap values are placed above the nodes.

**Figure S10. ML tree of the *HECATE3/INDEHISCENT* genes in seed plants.** Branch colors denote the following taxa: Persian green = Gymnosperms, Blue = Basal angiosperms, Middle washed yellow = Monocots, Green = Basal eudicots, Purple = Core eudicots, Red = Brassicaceae. Bootstrap values are placed above the nodes.

**Figure S11. ML tree of the *REPLUMLESS/POUNDFOOLISH* genes in seed plants.** Branch colors denote the following taxa: Persian green = Gymnosperms, Blue = Basal angiosperms, Middle washed yellow = Monocots, Green = Basal eudicots, Purple = Core eudicots, Red = Brassicaceae. Bootstrap values are placed above the nodes.

**Movie S1. 3D X-ray macroscopic reconstruction of a 3 WAP *E. pusilla* fruit.** Umbilical cords (funiculi) can be detected between the vascular bundles of the fertile valves and the placenta regions. (No scale bar can be included for a 3D movie). Abbreviations: WAP = week after pollination.

**Table S1. AttB-primers used for the creation of inserts for Gateway cloning.**

**Table S2. Number of cell layers of *Erycina pusilla*, *Epipactis helleborine* and *Cynorkis fastigiata* fruits during development.** Abbreviations: DAP = days after pollination, WAP = weeks after pollination, Cf = *Cynorkis fastigiata*, Eh = *Epipactis helleborine*.

**Table S3**. **Transcript primer sequences and amplicon characteristics used for quantitative real-time PCR validation of the expression profiles of different transcripts, following MIQE guidelines (Bustin *et al*., 2009).** The sequences listed here were downloaded from NCBI GenBank (www.ncbi.nlm.nih.gov), Orchidstra (orchidstra2.abrc.sinica.edu.tw) and our own fruit transcriptome dataset (see chapter 5).

**Table S4. Difference in MADS-box gene expression between developmental stages of the fruit of *E. pusilla* as calculated using a variance analysis of measures using a Tukey multi-comparisons test.** P-value style: GP: >0.05 (ns), <0.05 (*), <0.01 (**), <0.001 (***), <0.0001 (****). No value = No expression. Abbreviations: DAP = days after pollination, WAP = weeks after pollination.

**Table S5. Accession numbers of *SPT/ALC bHLH* transcription factor sequences used in the alignment.** The Orchidaceae subfamilies are provided in parentheses.

**Table S6. Accession numbers of *IND/HEC3* sequences used in the alignments and phylogenetic analyses.** The Orchidaceae subfamilies are provided in parentheses.

**Table S7. Accession numbers of *RPL/PNF* sequences used in the alignments and phylogenetic analyses.** The Orchidaceae subfamilies are provided in parentheses.

### Author contributions

AD, ES, MB and BG designed the research. AD, IA, LK, NM, JS, AW, MW, JZ, RB, BJH, JK, RL, WS and MB performed the research. AD, IA, MB, LK, JS, AW, MW, JZ, RH, AW, RH and MB analyzed the data. AD, ES, MB and BG wrote the paper.


### Funding

This study was financially supported by grant 023.003.015 from the Netherlands Organization for Scientific Research (NWO) to AD.


### Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.


### Acknowledgements

We thank Elza Duim, Marcel Eurlings, Dirk van der Marel and Kees van den Berg for their technical support, Karoly Szuhai for his support at the Microscopy Core Facility of the Cell and Chemical Biology Department, Leiden University Medical Center (The Netherlands), Dick Groenenberg and Nemi Dorst for their help with the phylogenetic analyses and Steve Donovan and Philipp Schlüter for critical reading and helpful comments to improve the manuscript.


## References


Arditti, J. (1992). *Fundamentals of orchid biology.* New York: John Wiley & Sons.

Ballester, P., and Ferrandiz, C. (2017). Shattering fruits: variations on a dehiscent theme. *Curr Opin Plant Biol* 35**,** 68-75.

Bartlett, M., Thompson, B., Brabazon, H., Del Gizzi, R., Zhang, T., and Whipple, C. (2016). Evolutionary Dynamics of Floral Homeotic Transcription Factor Protein-Protein Interactions. *Mol Biol Evol* 33**,** 1486-1501.

Beer, J.G. (1863). *Beiträge zur morphologie und biologie der familie der orchideen.* Wien: Carl Gerold's Sohn.

Bello, V., Mattei, G., Mazzoldi, P., Vivenza, N., Gasco, P., Idee, J.M., Robic, C., and Borsella, E. (2010). Transmission electron microscopy of lipid vesicles for drug delivery: comparison between positive and negative staining. *Microsc Microanal* 16**,** 456-461.

Bemer, M., Karlova, R., Ballester, A.R., Tikunov, Y.M., Bovy, A.G., Wolters-Arts, M., Rossetto Pde, B., Angenent, G.C., and De Maagd, R.A. (2012). The tomato FRUITFULL homologs TDR4/FUL1 and MBP7/FUL2 regulate ethylene-independent aspects of fruit ripening. *Plant Cell* 24**,** 4437-4451.

Brown, R. (1831). *Observations on the organs and mode of fecundation in Orchideae and Asclepiadeae.* London : Printed by Richard Taylor, Red lion court, Fleet Street.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., and Wittwer, C.T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55**,** 611-622.

Chen, J.C., and Fang, S.C. (2016). The long pollen tube journey and in vitro pollen germination of Phalaenopsis orchids. *Plant Reprod* 29**,** 179-188.

Chen, J.C., Wei, M.J., and Fang, S.C. (2016). Expression analysis of fertilization/early embryogenesis-associated genes in Phalaenopsis orchids. *Plant Signal Behav* 11**,** e1237331.

Chou, M.L., Shih, M.C., Chan, M.T., Liao, S.Y., Hsu, C.T., Haung, Y.T., Chen, J.J., Liao, D.C., Wu, F.H., and Lin, C.S. (2013). Global transcriptome analysis and identification of a CONSTANS-like gene family in the orchid Erycina pusilla. *Planta* 237**,** 1425-1441.

De Folter, S., Busscher, J., Colombo, L., Losa, A., and Angenent, G.C. (2004). Transcript profiling of transcription factor genes during silique development in Arabidopsis. *Plant Mol Biol* 56**,** 351-366.

De Folter, S., Immink, R.G., Kieffer, M., Parenicova, L., Henz, S.R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M.M., Davies, B., and Angenent, G.C. (2005). Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. *Plant Cell* 17**,** 1424-1433.

De Folter, S., and Immink, R.G.H. (2011). "Yeast Protein–Protein Interaction Assays and Screens," in *Plant Transcription Factors: Methods and Protocols,* eds. L. Yuan & S.E. Perry. (Totowa, NJ: Humana Press), 145-165.

Dinneny, J.R., Weigel, D., and Yanofsky, M.F. (2005). A genetic framework for fruit patterning in Arabidopsis thaliana. *Development* 132**,** 4687-4696.

Dirks-Mulder, A., Butot, R., Van Schaik, P., Wijnands, J.W., Van Den Berg, R., Krol, L., Doebar, S., Van Kooperen, K., De Boer, H., Kramer, E.M., Smets, E.F., Vos, R.A., Vrijdaghs, A., and Gravendeel, B. (2017). Exploring the evolutionary origin of floral organs of Erycina pusilla, an emerging orchid model system. *BMC Evol Biol* 17**,** 89.

Dressler, R.L. (1993). *Phylogeny and Classification of the Orchid Family.* Cambridge, UK: Cambridge University Press.

Fang, S.C., Chen, J.C., and Wei, M.J. (2016). Protocorms and Protocorm-Like Bodies Are Molecularly Distinct from Zygotic Embryonic Tissues in Phalaenopsis aphrodite. *Plant Physiol* 171**,** 2682-2700.

Fernandez, V., Guzman-Delgado, P., Graca, J., Santos, S., and Gil, L. (2016). Cuticle Structure in Relation to Chemical Composition: Re-assessing the Prevailing Model. *Front Plant Sci* 7**,** 427.

Ferrandiz, C. (2002). Regulation of fruit dehiscence in Arabidopsis. *J Exp Bot* 53**,** 2031-2038.

Ferrandiz, C., and Fourquin, C. (2014). Role of the FUL-SHP network in the evolution of fruit morphology and function. *J Exp Bot* 65**,** 4505-4513.

Ferrandiz, C., Liljegren, S.J., and Yanofsky, M.F. (2000). Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. *Science* 289**,** 436-438.

Ferrandiz, C., Pelaz, S., and Yanofsky, M.F. (1999). Control of carpel and fruit development in Arabidopsis. *Annu Rev Biochem* 68**,** 321-354.

Girin, T., Paicu, T., Stephenson, P., Fuentes, S., Korner, E., O'brien, M., Sorefan, K., Wood, T.A., Balanza,

V., Ferrandiz, C., Smyth, D.R., and Ostergaard, L. (2011). INDEHISCENT and SPATULA interact to specify carpel and valve margin tissue and thus promote seed dispersal in Arabidopsis. *Plant Cell* 23**,** 3641-3653.

Gu, Q., Ferrandiz, C., Yanofsky, M.F., and Martienssen, R. (1998). The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. *Development* 125**,** 1509-1517.

Hayat, M.A. (1970). *Principles and techniques of electron microscopy: biological applications.* New York: Van Nostrand Reinhold Co.

Horowitz, A. (1901). Ueber den anatomischen Bau und das Aufspringen der Orchideen frűchte. *Beihefte zum Botanischen Centralblatt* 11**,** 486-521.

Kay, P., Groszmann, M., Ross, J.J., Parish, R.W., and Swain, S.M. (2013). Modifications of a conserved regulatory network involving INDEHISCENT controls multiple aspects of reproductive tissue development in Arabidopsis. *New Phytol* 197**,** 73-87.

Kocyan, A., and Endress, P. (2001). Floral Structure And Development of Apostasia and Neuwiedia (Apostasioideae) and their Relationships to other Orchidaceae. *International Journal of Plant Sciences* 162**,** 847-867.

Kramer, E.M., Jaramillo, M.A., and Di Stilio, V.S. (2004). Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. *Genetics* 166**,** 1011-1023.

Lee, S.-H., Li, C.-W., Liau, C.-H., Chang, P.-Y., Liao, L.-J., Lin, C.-S., and Chan, M.-T. (2015). Establishment of an Agrobacterium-mediated genetic transformation procedure for the experimental model orchid Erycina pusilla. *Plant Cell Tiss Organ Cult (PCTOC)* 120**,** 211-220.

Leseberg, C.H., Eissler, C.L., Wang, X., Johns, M.A., Duvall, M.R., and Mao, L. (2008). Interaction study of MADS-domain proteins in tomato. *J Exp Bot* 59**,** 2253-2265.

Liljegren, S.J., Roeder, A.H., Kempin, S.A., Gremski, K., Ostergaard, L., Guimil, S., Reyes, D.K., and Yanofsky, M.F. (2004). Control of fruit patterning in Arabidopsis by INDEHISCENT. *Cell* 116**,** 843-853.

Lin, C.S., Chen, J.J., Huang, Y.T., Hsu, C.T., Lu, H.C., Chou, M.L., Chen, L.C., Ou, C.I., Liao, D.C., Yeh, Y.Y., Chang, S.B., Shen, S.C., Wu, F.H., Shih, M.C., and Chan, M.T. (2013). Catalog of Erycina pusilla miRNA and categorization of reproductive phase-related miRNAs and their target gene families. *Plant Mol Biol* 82**,** 193-204.

Lin, C.S., Hsu, C.T., Liao, D.C., Chang, W.J., Chou, M.L., Huang, Y.T., Chen, J.J., Ko, S.S., Chan, M.T., and Shih, M.C. (2016). Transcriptome-wide analysis of the MADS-box gene family in the orchid Erycina pusilla. *Plant Biotechnol J* 14**,** 284-298.

Mayer, J.L.S., Carmello-Guerreiro, S.M., and Appezzato-Da-Glória, B. (2011). Anatomical development of the pericarp and seed of Oncidium flexuosum Sims (ORCHIDACEAE). *Flora - Morphology, Distribution, Functional Ecology of Plants* 206**,** 601-609.

Melchior, V., Hollingshead, C.J., and Cahoon, M.E. (1980). Stacking in lipid vesicle-tubulin mixtures is an artifact of negative staining. *The Journal of Cell Biology* 86**,** 881-884.

Mukherjee, K., Brocchieri, L., and Burglin, T.R. (2009). A comprehensive classification and evolutionary analysis of plant homeobox genes. *Mol Biol Evol* 26**,** 2775-2794.

Österberg, J.A. (1883). Bidrag till Kännedomen of pericarpiets anatomi och kärlsträngförloppet i blomman hos Orchideerna. Öfvers. Förh. Kongl. Svenska Vetensk.-Akad. 3**,** 47-62

Pabon-Mora, N., and Litt, A. (2011). Comparative anatomical and developmental analysis of dry and fleshy fruits of Solanaceae. *Am J Bot* 98**,** 1415-1436.

Pabon-Mora, N., Wong, G.K., and Ambrose, B.A. (2014). Evolution of fruit development genes in flowering plants. *Front Plant Sci* 5**,** 300.

Pan, I.C., Liao, D.C., Wu, F.H., Daniell, H., Singh, N.D., Chang, C., Shih, M.C., Chan, M.T., and Lin, C.S. (2012). Complete chloroplast genome sequence of an orchid model plant candidate: Erycina pusilla apply in tropical Oncidium breeding. *PLoS One* 7**,** e34738.

Pires, N., and Dolan, L. (2010). Origin and diversification of basic-helix-loop-helix proteins in plants. *Mol Biol Evol* 27**,** 862-874.

Pridgeon, A., Cribb, P., Chase, M.W., and Rasmussen, F.N. (1999). *Genera Orchidacearum Volume 3: Orchidoideae (Part 2) Vanilloideae.* Oxford: Oxford Universitty Press.

Rajani, S., and Sundaresan, V. (2001). The Arabidopsis myc/bHLH gene ALCATRAZ enables cell separation in fruit dehiscence. *Curr Biol* 11**,** 1914-1922.

Rasmussen, F.N., and Johansen, B. (2006). Carpology of Orchids. *Selbyana* 27**,** 44-53.

Renault, H., Alber, A., Horst, N.A., Basilio Lopes, A., Fich, E.A., Kriegshauser, L., Wiedemann, G., Ullmann, P., Herrgott, L., Erhardt, M., Pineau, E., Ehlting, J., Schmitt, M., Rose, J.K.C., Reski, R., and Werck-Reichhart, D. (2017). A phenol-enriched cuticle is ancestral to lignin evolution in land plants. *Nature Communications* 8**,** 14713.

Reynolds, E.S. (1963). The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *J Cell Biol* 17**,** 208-212.

Roeder, A.H., Ferrandiz, C., and Yanofsky, M.F. (2003). The role of the REPLUMLESS homeodomain protein in patterning the Arabidopsis fruit. *Curr Biol* 13**,** 1630-1635.

Ruijter, J.M., Ruiz Villalba, A., Hellemans, J., Untergasser, A., and Van Den Hoff, M.J. (2015). Removal of between-run variation in a multi-plate qPCR experiment. *Biomol Detect Quantif* 5**,** 10-14.

Ruijter, J.M., Thygesen, H.H., Schoneveld, O.J., Das, A.T., Berkhout, B., and Lamers, W.H. (2006). Factor correction as a tool to eliminate between-session variation in replicate experiments: application to molecular biology and retrovirology. *Retrovirology* 3**,** 2.

Smith, H.M., Boschke, I., and Hake, S. (2002). Selective interaction of plant homeodomain proteins mediates high DNA-binding affinity. *Proc Natl Acad Sci U S A* 99**,** 9579-9584.

Toledo-Ortiz, G., Huq, E., and Quail, P.H. (2003). The Arabidopsis basic/helix-loop-helix transcription factor family. *Plant Cell* 15**,** 1749-1770.

Van Dijk, A.D., Morabito, G., Fiers, M., Van Ham, R.C., Angenent, G.C., and Immink, R.G. (2010). Sequence motifs in MADS transcription factors responsible for specificity and diversification of protein-protein interaction. *PLoS Comput Biol* 6**,** e1001017.

Vrebalov, J., Pan, I.L., Arroyo, A.J., Mcquinn, R., Chung, M., Poole, M., Rose, J., Seymour, G., Grandillo, S., Giovannoni, J., and Irish, V.F. (2009). Fleshy fruit expansion and ripening are regulated by the Tomato SHATTERPROOF gene TAGL1. *Plant Cell* 21**,** 3041-3062.

Yeats, T.H., and Rose, J.K. (2013). The formation and function of plant cuticles. *Plant Physiol* 163**,** 5-20.

Zahn, L.M., Leebens-Mack, J.H., Arrington, J.M., Hu, Y., Landherr, L.L., Depamphilis, C.W., Becker, A., Theissen, G., and Ma, H. (2006). Conservation and divergence in the AGAMOUS subfamily of MADS-box genes: evidence of independent sub- and neofunctionalization events. *Evol Dev* 8**,** 30-45.

# Chapter 5

## Transcriptome and ancestral character state analysis of orchid fruits

Anita Dirks-Mulder[1,2], Israa Ahmed[2], Marian Bemer[3], Mark uit het Broek[2], Nemi Dorst[2], Jasmijn Snier[2], Anne van Winzum[2], Martijn van 't Wout[2], Stephen Pieterman[2], Floyd Wittink[2], Jamie Zeegers[2], Bertie Joan van Heuven[1], Jaco Kruizinga[4], Erik F. Smets[1,5,6], Diego Bogarin[1] & Barbara Gravendeel[1,2,6]

[1]Endless Forms group, Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands

[2]Faculty of Science and Technology, University of Applied Sciences Leiden, Zernikedreef 11, 2333 CK Leiden, The Netherlands

[3]Department of Plant Sciences, Laboratory of Molecular Biology, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[4]Hortus botanicus, Leiden University, Rapenburg 73, 2311 GJ Leiden, The Netherlands

[5]Ecology, Evolution and Biodiversity Conservation cluster, KU Leuven, Kasteelpark Arenberg 31, 3001 Leuven, Belgium

[6]Institute of Biology Leiden, Leiden University, Sylviusweg 72, 2333 BE Leiden, The Netherlands

## Abstract

The orchid family is known for its vast floral diversity. Orchid fruits are highly diverse as well but much less studied. The first aim of this study is to detect genes and gene networks involved in fruit development of *Erycina pusilla* so targeting a much wider set than only the MADS-box genes discussed in the previous chapter. Secondly, we investigated possible correlated evolution of a total of 12 binary traits of orchid fruits from 41 different species in all five subfamilies on a combined nrITS, *mat*K and *rbc*L based phylogeny.

Next-generation sequencing of RNA isolated from entire fruits of *E. pusilla* harvested from different developmental stages, was used to generate a full fruit transcriptome. Fresh and spirit conserved mature fruits were collected, hand cut and stained with phloroglucinol to visualize lignification. A reversible-jump Markov chain Monte Carlo search was employed to search among possible models to describe joint evolution of fruit traits on a phylogeny of the orchid family.

Preliminary results from the orchid fruit transcriptomes analyzed, indicate that genes related to pollen tube formation, seed development and lignification are differentially expressed during development. A striking variation was found in fruit type, direction, ripening period, number of slits and lignification patterns. The latter varied from lignification of the exocarp, various tissue types such as valves or dehiscence zones in the endocarp, to no lignification at all except for the trichomes. The posterior distributions calculated indicate that an epiphytic or lianaceous habit and fruit ripening period of more than four months have clearly coevolved in orchids. There is also strong evidence for the hypothesis that an epiphytic or lianaceous habit co-evolved with a smaller number of opening slits of orchid fruits. Similarly, pendant orchid fruits co-evolved with a preference for growing at intermediate to high temperatures. Strong support was also found for the hypothesis that pendant orchid fruits co-evolved with an epiphytic or lianaceous habit. Lastly, lignification of the valves evolved in both pendant and erect orchid fruits. Future comparative transcriptome analyses including more species may reveal which developmental genes drive the high morphological diversity of orchid fruits.

Key words: Bayesian analyses - Coevolution - Endocarp - Exocarp - Fruit dehiscence - Lignification - OmicsBox - RNAseq

## Introduction

Pollination of flowers is a key regulatory event in plant reproduction. When an orchid flower is pollinated, the perianth withers, the stigma closes, and the ovary starts to grow. The ovary of orchids is inferior and made up of three carpels. Unlike most flowering plants, pollination and fertilization are separated in orchids by relatively long periods of time in which the oocyte and embryo sac develop (Dirks-Mulder *et al*., 2019). Once a flower is pollinated, pollen tubes grow from the stigma into the gynostemium towards the ovary, the placenta starts proliferating and ovules are formed. Fertilization of orchid ovules can occur days to months after pollination. Finally, the inferior ovary develops into a fruit, containing up to millions of seeds.

There are two different main types of orchid fruits, fleshy and dry (Dressler, 1993). Most orchids produce dehiscent capsules, which split open along the midline of each carpel when the capsule is fully ripe. Just a few orchids species produce fleshy berries or capsules, which can be found in several genera such as *Neuwiedia* (Apostasioideae), *Palmorchis* (Epidendroideae), *Cyrtosia* and *Vanilla* (Vanilloideae) (Dressler, 1981;1993;Kocyan and Endress, 2001). Indehiscent capsules can be found in *Dictyophyllaria* and *Vanilla* (Vanilloideae) and *Selenipedium* (Cypripedioideae) (Stern *et al*., 2014).

Fleshy orchid fruits are generally indehiscent, except for a few exceptions such as *Vanilla trigonocarpa*, while dry fruits can be either dehiscent or indehiscent. Dehiscent fruits open and release the seeds to be spread by wind, water or animals, while indehiscent fruits have to decay or be eaten and digested by animals in order to spread their seeds. Dehiscence can be defined as the completion of the reproductive cycle, in which the fruit will open and release seeds. This process is achieved by a coordinated cell separation (Mayer *et al*., 2011).

Although orchid fruits are very diverse in size and shape, they all have the same basic structure, which is the result of the differentiation and development of the inferior carpels. The carpels develop into six different valves, three fertile valves, with a placenta, and three sterile valves, not connected to any seed tissue. When a capsule starts to dehisce, the splitting of the valves occurs along the midline of a carpel, between the sterile and the fertile valves. The fruit maximally splits into three wide fertile valves and three narrow sterile ones (Dirks-Mulder *et al*., 2019). The number of dehiscent splits can vary from one, two, three to six splits (Beer, 1863;Horowitz, 1901;Dressler, 1981;1993). In **chapter 4**, we discovered that at the dehiscence zone of orchid fruits either a cuticle layer develops, such as for instance in *Erycina pusilla*, or that valves become lignified during fruit development, such as in *Epipactis helleborine* (both Epidendroideae) and *Cynorkis fastigiate* (Orchidaceae). Lignin is a very common biopolymer found in vegetables, fruits and the secondary cell walls of plants. In Arabidopsis, lignification of the valve margin and endocarp is necessary for creating a mechanical tension in the fruit to open (dehisce) and shatter its seeds (Di Vittori *et al*., 2019).

To gain more insight into the genes regulating these processes, we performed RNA-seq analysis of *E. pusilla* fruits from different stages: 0 and 5 days after pollination (DAP) and 4 and 8 weeks after pollination (WAP). In **chapter 4,**

we divided fruit development of *E. pusilla* in four developmental stages based on morphology: (I) 0 DAP - 1 WAP, (II) 2 WAP - 5 WAP, (III) 6 WAP - 11 WAP and (IV) 12 WAP - 16 WAP. In the first stage, cell division takes place in the fertile and sterile valves and the trichomes start to develop. We therefore expected genes involved in these processes to be up-regulated. In the second stage, when the fruit starts to elongate, six bundles of pollen tubes are formed. We therefore expected an upregulation of expression of genes involved in pollen tube formation and cell elongation in this stage.

During the third stage, when all the trichomes become lignified, the dehiscence zone is developing, and embryonic development takes place. We expected genes involved in all these processes to show a change in expression here.

Secondly, to investigate which fruit traits evolved together and which evolved independently of each other, we expanded our analyses to ripe fruits from additional species belonging to all five subfamilies of the orchids and carried out an ancestral character state study for twelve different fruit characters (Camus *et al*., 1921;Ziegenspeck, 1936). We hypothesized that lignification could be an adaptation for fast fruit dehiscence and therefore expected it to have evolved especially in orchids coping with relatively short seasons suitable for seed dispersal. To test this hypothesis, we studied lignin formation in fruits of orchid species from five subfamilies (Apostasioideae, Cypripedioideae, Epidendroideae, Orchidoideae and Vanilloideae) with different growth forms (either epiphytic, terrestrial or lianaceous). A reversible-jump Markov chain Monte Carlo search was employed to search among possible models to describe joint evolution on a phylogeny. According to this model, transitions among combinations of states result from two binary valuables. Posterior distributions were calculated to estimate statistical support for hypotheses of correlated evolution.

## Materials and methods

### Plant material

A more than 20-year-old inbred line of *Erycina pusilla* originally collected in Surinam was grown in climate rooms under controlled conditions (7.00h – 19.00h light regime), at a temperature of 22 °C and a relative humidity of 50%. The orchids were cultured in vitro under sterile conditions on Phytamax™ orchid medium with charcoal and banana powder (Sigma-Aldrich) with 4g/L Gelrite™ (Duchefa) culture medium. Pollination was conducted manually by placing the pollinia of flowers on each other's stigma to produce fruits and seeds.

### Total RNA extraction and RNA-sequencing

Carpels and fruits were collected from *E. pusilla* at 0 and 5 days after pollination (DAP) and 4 and 8 weeks after pollination (WAP), frozen in liquid nitrogen and stored at -80 °C until they were used for RNA extraction. There were three replicates for each sample. Total RNA was extracted using the RNeasy Plant Mini Kit (QIAGEN) and DNase I (QIAGEN) treated, following the manufacturer's protocol. A maximum of 100 mg plant material was placed in a 2.2 ml micro centrifuge tube with 7 mm

glass beads. The TissueLyser II (QIAGEN) was used to grind the plant material. The amount of RNA was measured using the NanoVue Plus™ (GE Healthcare Life Sciences) and its integrity was assessed on a 2100 Bioanalyzer (Agilent) using the Plant RNA nano protocol. RNA was stored at -80 °C until further use.

Messenger-RNA selection, library preparation and strand-specific sequencing of in total twelve RNA pools (0 DAP, 5 DAP, 4 WAP and 8 WAP x 3 biological replicates) was performed at BGI to generate 150-bp paired-end (PE) reads following the Illumina TruSeq Stranded sample preparation protocol. Summarized, mRNA was isolated using oligo(dT)-attached magnetic beads and fragmented. Then cDNA was synthesized using random hexamer primers, purified, end-repaired, poly-A tailed, adaptor ligated and loaded on an Illumina Hiseq4000™ sequencer (Illumina, United States) for sequencing. The raw reads were filtered with internal software at BGI for low-quality, adaptor-polluted and high content of unknown base(N) reads in order to obtain clean reads. The cleaned reads (in fastq format) were downloaded and stored on a High-End Workstation (**Table S1**) for further analysis.



**RNA sequence analysis**

Two different methods were used for the RNAseq analysis, both implemented in OmicsBox (v1.0.34) (OmicsBox;Conesa *et al*., 2005). First of all, a d*e novo* assembly for whole-transcriptome construction was carried out using the clean reads from BGI. The reads were assembled with Trinity (Grabherr *et al*., 2011;Haas *et al*., 2013), creating a sequence table containing the assembled transcripts sequences. Using local Blast, the assembled sequences were blasted against a local database consisting of all orchid sequences "txid4747" [Organism] in the NCBI database. BLAST hits were then functionally annotated using Blast2GO (Conesa *et al*., 2005;Gotz *et al*., 2008) and Gene Ontology (GO) terms (Ashburner *et al*., 2000) were assigned to transcripts with BLAST hits to understand gene function in the developing fruits.

RNA-Seq by Expectation-Maximization (RSEM) was used for alignment-based abundance estimation using Bowtie (Li and Dewey, 2011;Langmead and Salzberg, 2012). Since the assembled transcriptome represented all the transcripts found in the fruits of all four time-points, the sequence reads from each sample were aligned to the assembled transcriptome to differentiate transcripts unique to one stage from those present throughout fruit development. The read counts for each transcript were analyzed using edgeR (Robinson *et al*., 2010) for differential expression analysis, and the logarithmic fold change and false detection rate (FDR) were calculated for each transcript. The results of differential expression and GO annotation were combined for Gene Ontology enrichment using GOseq, which identifies GO terms enriched as a result of differential expression between samples.

Secondly, the *Phalaenopsis equestris* reference genome (Cai *et al*., 2015) together with its annotation were downloaded from Plaza 4.0 (Van Bel *et al*., 2018). The *E. pusilla* Illumina reads were mapped to the *P. equestris* reference genome using STAR (Spliced Transcript Alignment to a Reference) (Dobin *et al*., 2013)/, changing the default parameters to a minimal match of 0.05 in order to obtain >75% of the total reads aligned. The output BAM files (annotated files) were used to create a Count Table for every time point, based on HTSeq package (Anders *et al*.,

2015). A Single Series Time-Course Expression Analysis was performed using default settings (Conesa *et al*., 2006) in order to obtain expression profiles, either clustered or by gene.

**Fruit material**

A total of 41 ripe fruits from different orchid species and, as outgroup, *Narcissus* and *Hypoxis* fruits were freshly collected in the Hortus botanicus (Leiden, The Netherlands), Lankester Botanical Garden (Cartago, Costa Rica), Utrecht Botanic Gardens (The Netherlands), Bochum Botanical Garden (Germany) and from private collections of orchid breeders throughout Europe. The ripe fruits were fixed and stored in 70% ethanol at room temperature.

**Lignin staining and visualization**

Lignin formation in fruits was visualized using phloroglucinol staining. Handmade cross sections of fruits were cleared with 100% lactic acid (Merck) for one hour at 60 °C, stained with 1% phloroglucinol (Merck) in 96% ethanol for one hour and 25%-(v/v) hydrochloric acid (HCl) (Merck) for 2-5 minutes until a clear staining was visible and immediately examined under a Binocular microscope (Zeiss SteREO Discovery. V12). Between each step the sections were washed three times with demineralized water.

**Alignments and phylogenetic analyses**

Representative *mat*K, *nr*ITS and *rbc*L sequences were downloaded from NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/) including 117 accessions of 50 species covering every subfamily of the Orchidaceae (**Table S2**). In addition, we included two accessions of *Hypoxis curtissii* Rose (Hypoxidaceae) and two of *Narcissus bulbocodium* L. (Amaryllidaceae) as outgroups due to the close relationship of these families with orchids within the monocots (Givnish *et al*., 2015). We aligned and trimmed the matrices of each marker in Geneious® R9 (Biomatters Ltd., Auckland, New Zealand (Kearse *et al*., 2012) using MAFFT (Multiple Alignment using Fast Fourier Transform). The concatenated dataset (nrITS+*mat*K+*rbc*L) was built with Sequence Matrix v100.0 (Vaidya *et al*., 2011). When sequences for a specific marker were not available, they were included as missing data. When a species analyzed morphologically was not available in NCBI GenBank, it was substituted in the DNA concatenated matrix by a DNA sequence of a close relative.

To obtain ultrametric trees for the ancestral state reconstructions, we estimated the divergence times in BEAST v.1.8.2 using the CIPRES Science Gateway (Miller *et al*., 2010). The following MCMC parameters were set to $20\times10^6$ generations and a sampling frequency of 1000 yielding 20,001 trees per run. The substitution model selected was GTR, estimated with 4 gamma categories, estimated lognormal relaxed clock (uncorrelated) and the Yule Process (Y) speciation tree model. A normal prior distribution of 105.06 (±2.5 standard deviations) Ma was assigned to the root node of the Orchidaceae and 94.75 (±3.5 standard deviations) Ma to the node containing all Orchidaceae members. These secondary calibrations were obtained from various dating studies of the Orchidaceae (Ramirez *et al*., 2007;Chomicki *et*

*al*., 2015;Givnish *et al*., 2015;Poinar and Rasmussen, 2017). The convergence of independent runs and the MCMC parameters (ESS values >200) were inspected in Tracer v.1.6. Finally, a maximum clade credibility (MCC) tree was obtained with a 10% burnin using TreeAnnotator v.1.8.2. Resulting trees and the 95% highest posterior density (HPD) estimations were visualized in FigTree v1.4.3 (Rambaut, 2014) and manipulated with R programming language (R Core Team, 2018) under R Studio (Gandrud, 2015) using the packages APE, ggtree and phytools (Paradis *et al*., 2004;Revell, 2012;Yu *et al*., 2016).

**Ancestral State Reconstruction**

Ancestral state reconstructions (ASRs) were assessed with ML and stochastic character mapping (SCM) using ultrametric trees. For the ML approach we tested several models: equal rates (ER), symmetrical (SYM) and all rates different (ARD) with the re-rooting method of Yang *et al*. (1995) and the function ACE implemented in the R (R Development Core Team, 2018)packages APE, ggtree and phytools (Paradis *et al*., 2004;Revell, 2012;Yu *et al*., 2016). A likelihood ratio test, comparing the log-likelihoods among models, was used to select the best-fitting model. For the SCM analysis we performed 100 replicates on 100 randomly selected trees (10,000 mapped trees) derived from the BEAST analysis. These trees were randomly selected using the R function *samples.trees* (http://coleoguy.blogspot.de/ 2012/09/ randomly-sampling-trees.html). Results of transitions and the proportion of time spent in each state were obtained with the functions *make.simmap* and *describe. simmap* (Bollback, 2006;Revell, 2012). The mean probabilities retrieved at each node were plotted with phytools on the MCC tree for each character analyzed.



**Correlated evolution between traits**

We tested correlations among traits using the program BayesTraits V3.0.1 (Pagel, 1994;Pagel and Cunningham, 1999;Pagel and Meade, 2006) by performing 120 comparisons between each pair of the 12 characters of fruits assessed. We tested the two models under Bayesian Inference: a dependent model which allows correlation among traits against the independence model with no correlation among traits (correlations set to zero). To account for phylogenetic uncertainty, we used a set of 1000 trees (randomly selected with the R function *samples.trees* as described above) from the post burnin sample of the 20,001 ultrametric trees obtained from the time-calibrated BEAST analysis. The MCMC parameters of each model were set to 1,010,000 iterations, sample period of 1,000, burnin of 10,000, auto tune rate deviation and stepping stones 100 10,000 with reversible jump hyper-prior exponential. The BayesTraits outputs files were analyzed in R with the BayesTraits wrapper (btw) by Randi H Griffin (http://rgriff23.github.io/projects/ btw.html) and other functions from btrtools and BTprocessR (https://github.com/ hferg). The MCMC stationarity of parameters (ESS values >200) and convergence of chains were checked with the R package coda (Plummer *et al*., 2006) and the function *mcmcPlots* of BTprocessR. Using the log marginal likelihoods obtained from the outputs, we estimated a log Bayes factor (logBF) for the dependent model and the independent model with: logBF = 2*(SS dependent model (complex model)

– SS independent model (simple model)). We interpreted the logBF as suggested in the BayesTraitsV3 manual (http://www.evolution.rdg.ac.uk/BayesTraitsV3.0.1/Files/BayesTraitsV3.Manual.pdf): <2: weak evidence; >2: positive evidence; 5-10: strong evidence and >10: very strong evidence.

## Results

**Transcriptome sequencing of *Erycina pusilla* fruits**

To discover more about genes involved in orchid fruit formation, total RNA was extracted of *E. pusilla* fruits from different stages: 0 and 5 days after pollination (DAP) and 4 and 8 weeks after pollination (WAP). The RNA was sequenced at the Beijing Genomics Institute (BGI) using an Illumina Hiseq 4000 high-throughput sequencing platform.

Because there is no *E. pusilla* reference genome available, we performed a *de novo* assembly. With the clean reads obtained from BGI, Trinity (Grabherr *et al*., 2011) produced a set of 501,110 non-redundant putative transcripts and 260,320 unigenes with an average length of 788 bp and an N50 value of 1,357 bp. The length of transcripts ranged from 201 to 47,179 bp with a GC content of 36.9%.

The assembled 260,320 unigenes were annotated by comparison to the orchid proteins in the NCBI's non-redundant (NR) protein database. A Principal Component Analysis (PCA) was conducted, resulting in an MDS Plot, showing that not all three independent biological replicates of each sample had a good reproducibility (**Figure S1A**). The two deviating samples were excluded from further analysis (**Figure S1B**). Through BLASTx, transcripts (43.8% of the transcriptome) could be annotated with a description based on sequence homology. These transcripts were functionally annotated using Gene Ontology (GO) terms to understand the role of the genes in the tissue analyzed. Of the 219,296 transcripts annotated using BLASTx, 196,859 (89.8%) could be assigned to GO terms.

To identify differentially expressed genes (DEGs) between the different fruit time points, we determined relative expression levels using DEseq and found an increasing number of both up- and down-regulated features (**Figure 1**).

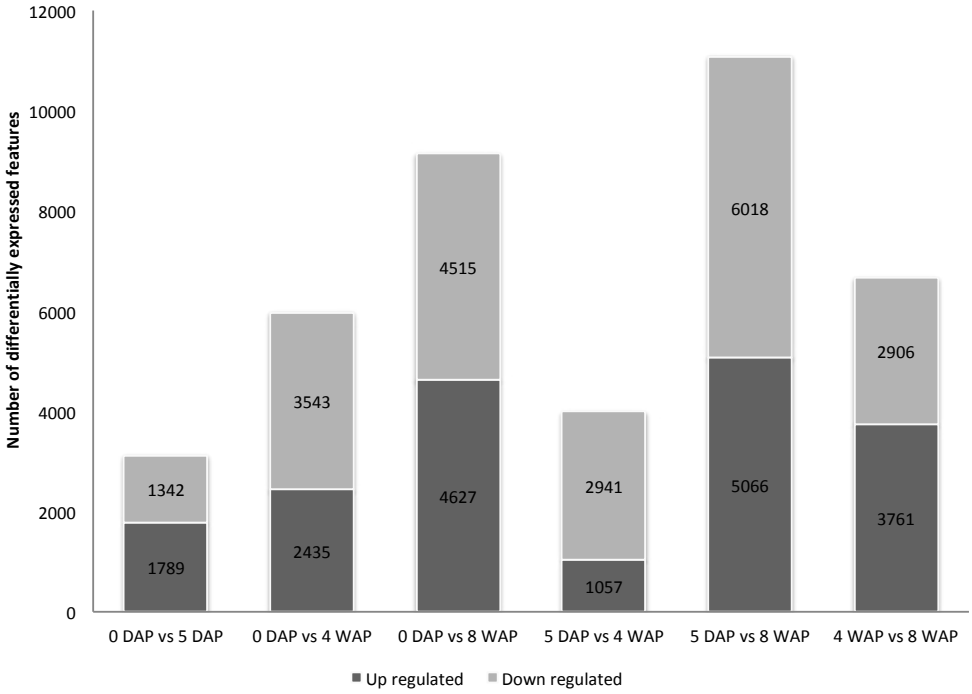



**Figure 1**: **Changes in gene expression profiles of *E. pusilla* fruits sampled at 0 DAP, 5 DAP, 4 WAP and 8 WAP as expressed in the number of features**. Abbreviations: DAP = days after pollination, WAP = weeks after pollination.

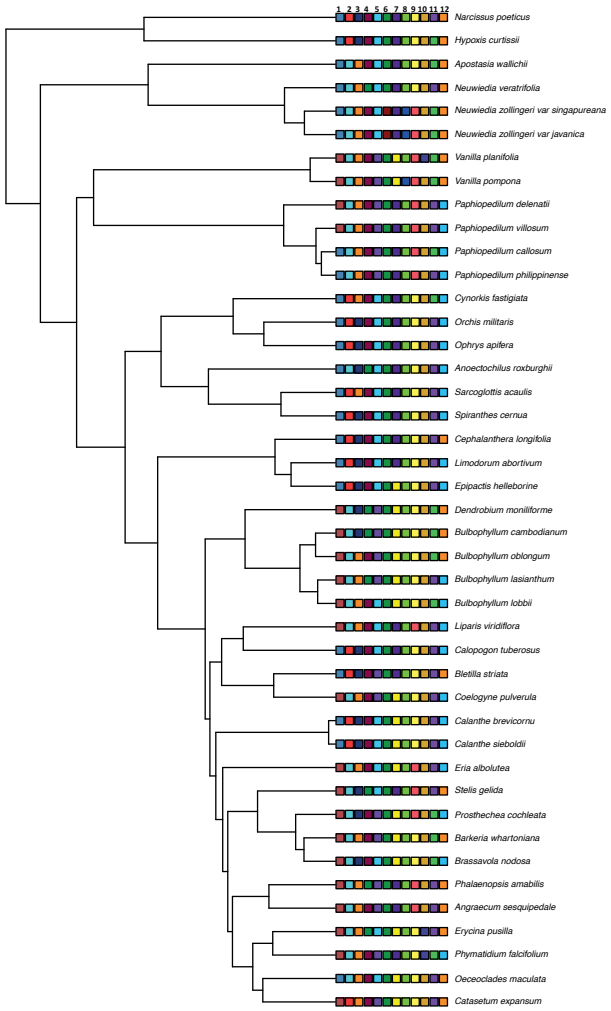
For *E. pusilla* we were not able to detect gene-networks involved in fruit development with a *de novo* assembled transcriptome approach, because of high redundancy in the raw contigs. Using CD-HIT (Li and Godzik, 2006)d to remove redundant contigs from *de novo* contig assemblies didn't improve much.

With the availability of an annotated reference genome of the orchid *Phalaenopsis equestris,* we decided to map the reads against this genome (Cai *et al*., 2015;Van Bel *et al*., 2018).

After aligning the *E. pusilla* reads against the *P. equestris* genome, >75% of the reads could be aligned. A Count Table (BAM+GFF) for gene level quantification was created in OmixcBox and the Counts per Category were visualized showing the number of reads of each input file sorted by different categories. For every input file, more than 10,000 read counts could be assigned to a feature (**Figure S2**), of a total number of 29,431 features.

For the Time Course Expression analyses, we performed a Single Series Time Course, with Time factor Group. The results were visualized using K-Means Clustering, resulting in 5,984 significant features divided over 9 clusters (**Table S3**). Genes belonging to Cluster 2 decreased their expression after pollination, for cluster 5 expression increased at 4 WAP and for cluster 7-9 expression increased at 8 WAP. Genes in these clusters belonging to GO terms involved in fruit and seed development, pollen tube development and lignin biosynthesis were further

analyzed in order to identify gene-networks involved in orchid fruit development (**Figure S3** and **Table S4**).




| Number | Character | State 0 | State 1 |
|---|---|---|---|
| 1 | Life form | Terrestrial | Epiphytic/Liana |
| 2 | Strategy | Evergreen | Deciduous |
| 3 | Minimum temperature requirement | Cool/Intermediate (<15 °C) | High (> 15 °C) |
| 4 | Production | All year round | Seasonal |
| 5 | Ripening period (months) | ≤ 4 | > 4 |
| 6 | Type | Berry | Capsule |
| 7 | Direction | Erect | Pendant |
| 8 | Dehiscence type | Indehiscent | Dehiscent |
| 9 | Number of slits in mature fruit | <3 | ≥ 3 |
| 10 | United valves | United at the base and apex | United at the base |
| 11 | Valves | No | Yes |
| 12 | Dehiscence zones | No | Yes |

**Figure 2**: **Consensus tree used for the ancestral state analyses of 41 orchid species and two non-orchid species (used as outgroup) plotted together with a character matrix score for 12 characters.** Almost all branches are supported by 90% bootstrap support or higher.

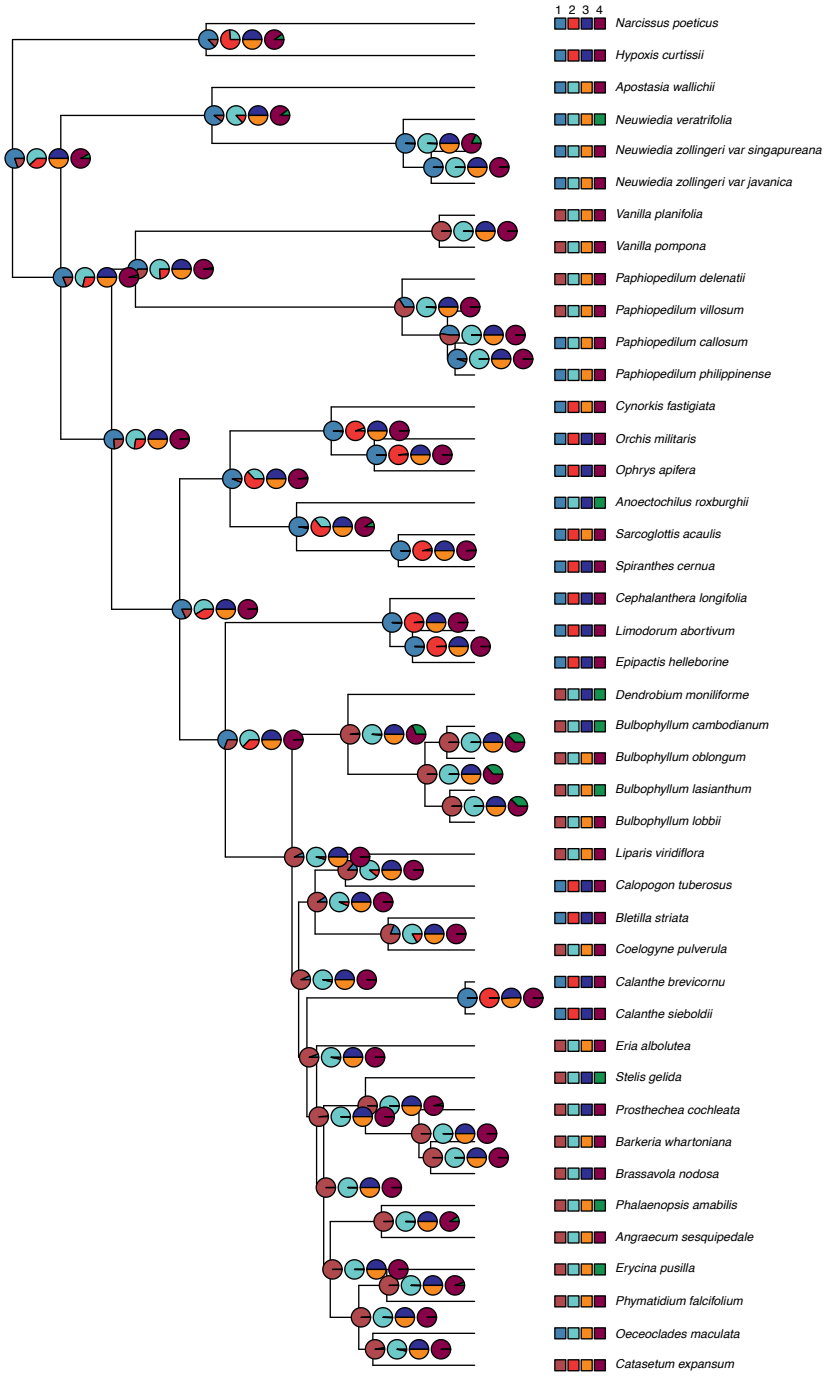



**Figure 3**: **Ancestral state reconstructions of selected characters from stochastic mapping analysis based on joint sampling (10,000 mapped trees).** Posterior probabilities (pie charts) are mapped in a random stochastic character map. Each tree represents four of the twelve characters represented in **Figure 2**.

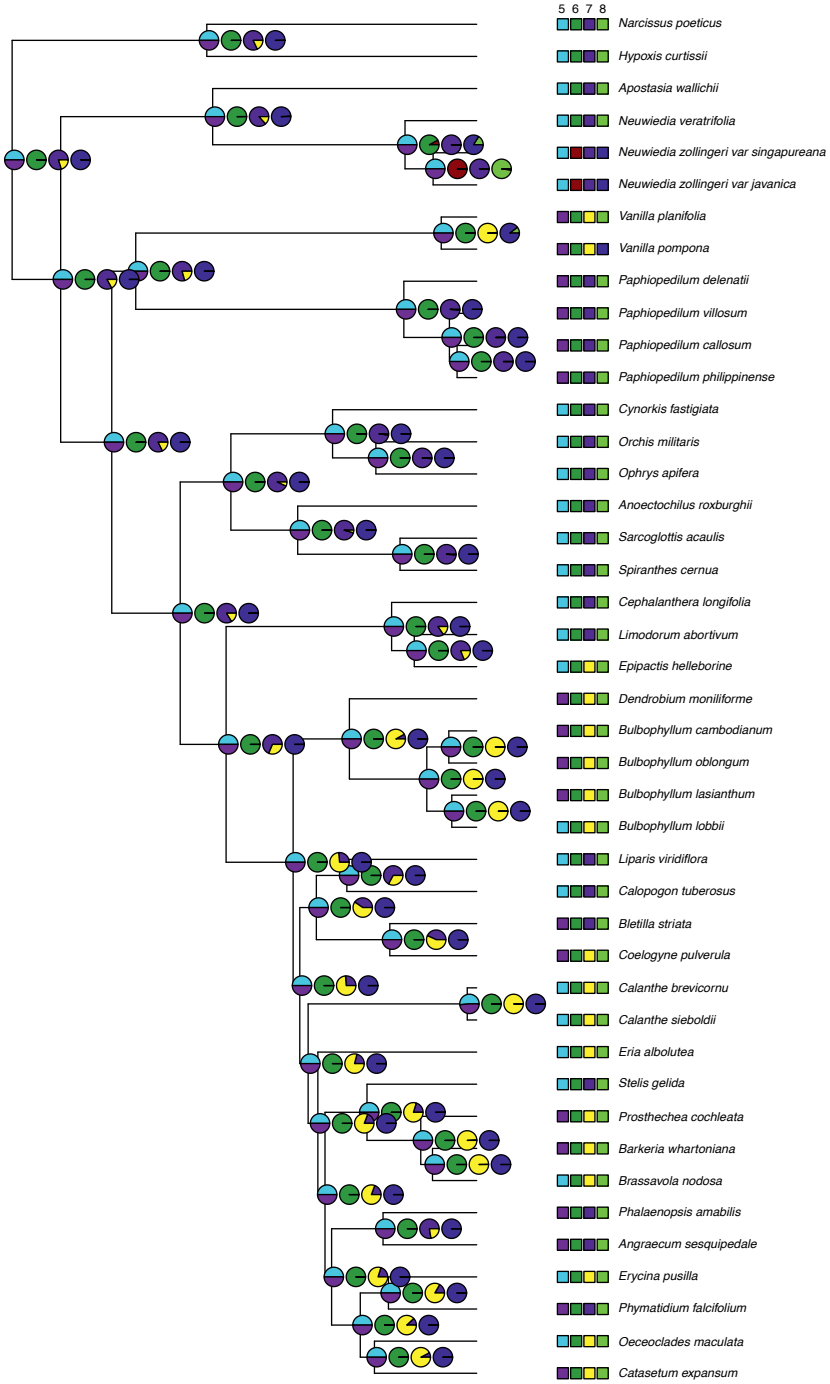



**Figure 3 continued**: Ancestral state reconstructions of selected characters from stochastic mapping analysis based on joint sampling (10,000 mapped trees).

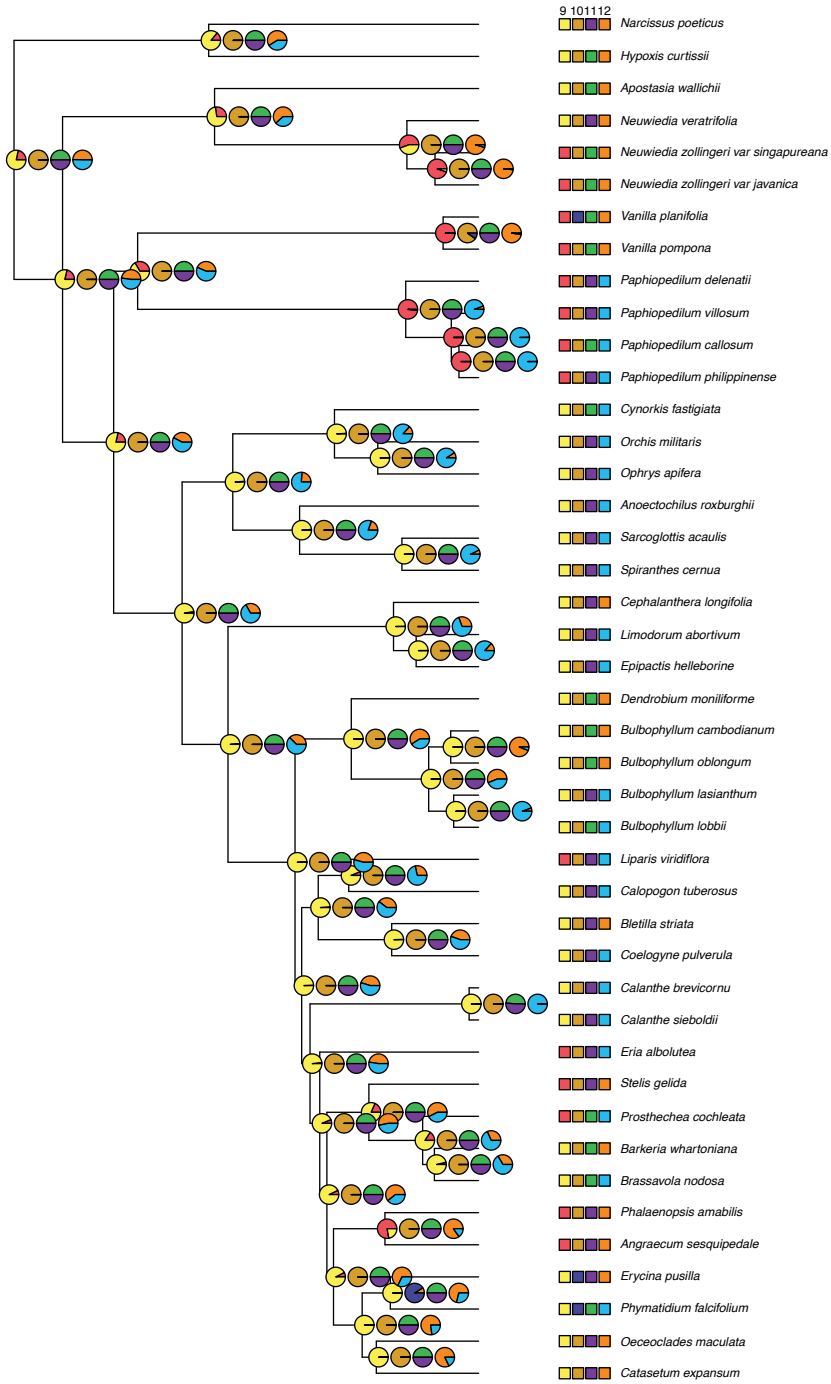



**Figure 3 continued**: Ancestral state reconstructions of selected characters from stochastic mapping analysis based on joint sampling (10,000 mapped trees).

**Ancestral character state analyses**

A total of 12 different characters, all binary states, could be scored based on the variation among the species studied as observed from literature, personal observations in the greenhouse, and after staining with phloroglucinol (**Figure S4, S5 and S6**). These twelve characters were plotted on a consensus of the orchid phylogeny (**Figure 2**) and further analyzed with ancestral state analyses.

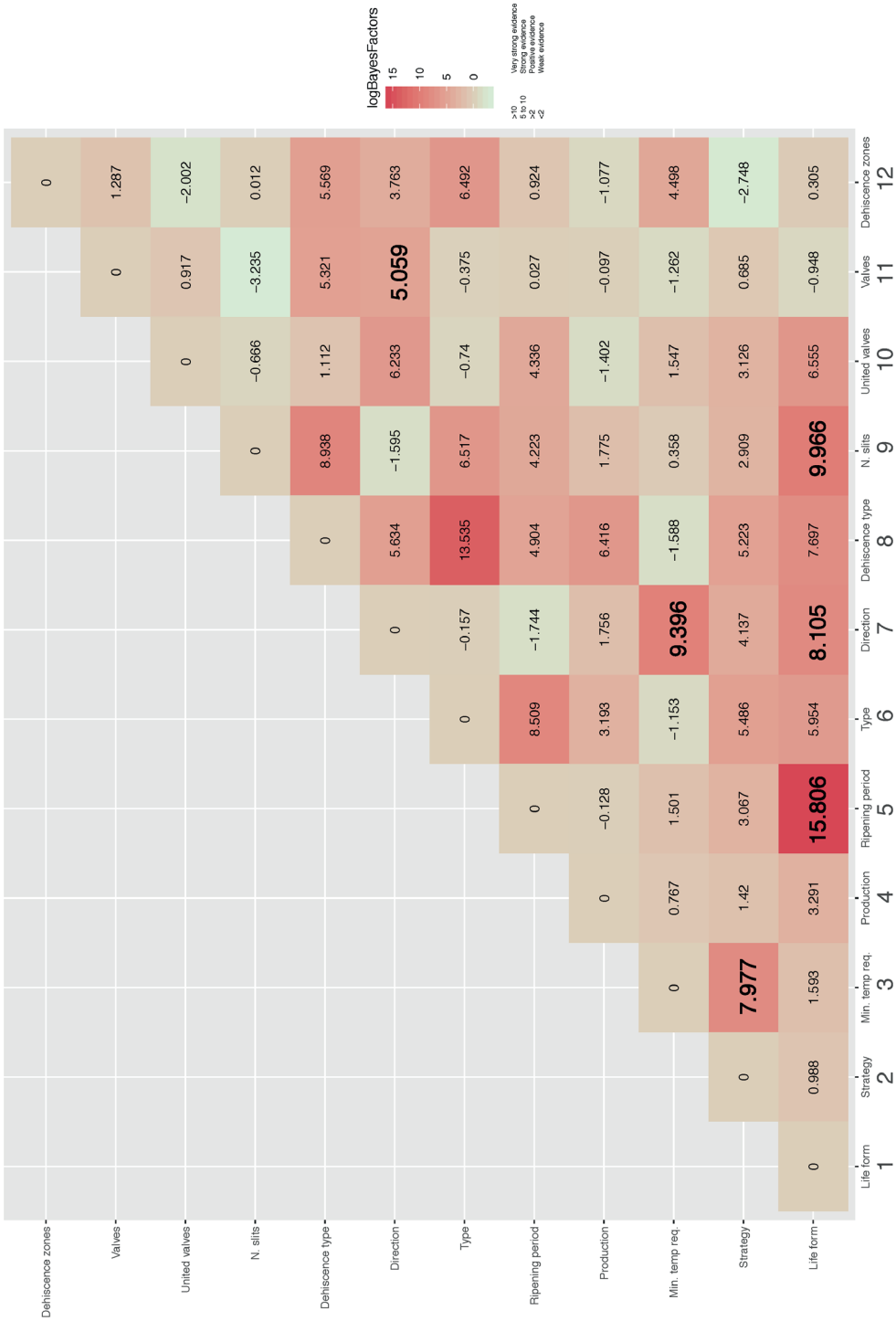
The ASR were based on the ER (equal rate) model, which performed consistently better than the SYM (Symmetrical rate) and the ARD (all rates different) models. The results of the Bayesian analyses are summarized in **Figure 3** showing that for character 1 terrestrial, character 7 erect and for character 9 ≥ 3 slits are the most ancestral states.

Possible concerted evolution among the different traits was further investigated using reversible jump Markov Chain Monte Carlo searches. For several character combinations, models of dependent evolution received high statistical support. Of the 66 possible correlations (**Figure 4**), we found two correlations with very strong support (LogBayes Factor >10), 19 with strong support (LogBayes Factor 5 to 10), 11 with some support (LogBayes Factor >2) and 34 with weak support (LogBayes Factor <2).

Because of the fact that only two indehiscent berries were included (character 6), and almost all included capsules dehisced (character 8), correlations found among these characters were excluded. Correlations found for fruits which are united at the base or base and apex (character 10) were also excluded due to too small sampling size, resulting in five comparisons with either very strong or strong evidence of correlation, as depicted in **Figure 5**.

As can be seen in **Figure 5**, there is strong evidence for the hypothesis that an epiphytic or lianaceous habit co-evolved with a relatively longer orchid fruit ripening time (**Figure 5A**). There is also evidence for the hypothesis that an epiphytic or lianaceous habit co-evolved with a smaller number of opening slits of orchid fruits (**Figure 5B**). Pendant orchid fruits seem to co-evolved with a preference for growing at intermediate to high temperature, although the interpretation of these two characters is ambiguous (**Figure 5C**). Strong support was also found for the hypothesis that pendant orchid fruits co-evolved with an epiphytic or lianaceous habit (**Figure 5D**). Lastly, orchids that have a deciduous life cycle are more common in cool to intermediate temperatures (**Figure 5E**), which makes sense due to the presence of more strictly defined seasons in temperate regions as compared with the tropics. No direction for valve lignification and the direction of orchid fruits was found (**Figure 5F**).

**Figure 4 (next page)**: **Heatmap of the correlations between each character.** LogBayes factor >10 Very strong evidence (warm color), 5 to10 strong evidence, >2 positive evidence and <2 weak evidence (cold color).


logBayesFactors
15
10
5
0
>10
>5 to 10
>2
<2
Very strong evidence
Strong evidence
Positive evidence
Weak evidence
Life form 1
Strategy 2
Min. temp req. 3
Production 4
Ripening period 5
Type 6
Direction 7
Dehiscence type 8
N. slits 9
United valves 10
Valves 11
Dehiscence zones 12
Life form
Strategy
Min. temp req.
Production
Ripening period
Type
Direction
Dehiscence type
N. slits
United valves
Valves
Dehiscence zones
0
0.988
1.593
3.291
15.806
5.954
8.105
7.697
6.555
–0.948
0.305
0
7.977
1.42
3.067
5.486
4.137
5.223
3.126
0.685
–2.748
0
0.767
1.501
–1.153
9.396
–1.588
1.547
–1.262
4.498
0
–0.128
3.193
1.756
6.416
1.775
–1.402
–0.097
–1.077
0
8.509
–1.744
4.904
4.223
4.336
0.027
0.924
0
–0.157
13.535
6.517
–0.74
–0.375
6.492
0
5.634
–1.595
6.233
5.059
3.763
0
8.938
1.112
5.321
5.569
0
–0.666
–3.235
0.012
0
–2.002
0
0.917
1.287
0

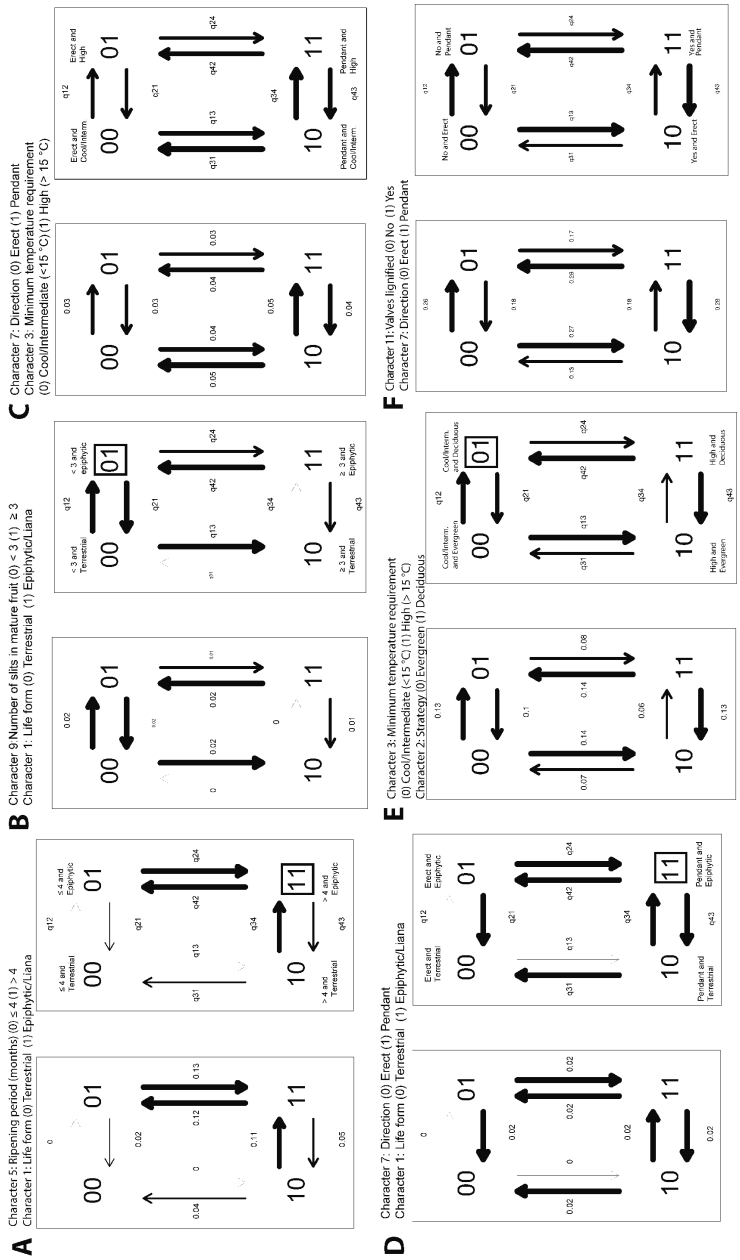



**Figure 5**: **Ancestral state reconstructions of selected morphological orchid fruit characters from stochastic mapping analyses based on joint sampling (10,000 mapped trees).** Arrows represent transitions between states and numbers represent the estimated number of evolutionary changes with proportion in parenthesis and the time spent in each state. Posterior probabilities (pie charts) are mapped in a random stochastic character map. (A) Character 5 vs. 1, (B) Character 9 vs. 1, (C) Character 7 vs. 3, (D) Character 7 vs. 1, (E) Character 3 vs. 2 and (F) Character 11 vs. 7. Character combinations indicated with a square are interpreted as having co-evolved.

## Discussion

**Transcriptome analyses of *Erycina pusilla* fruits**

In contrast to orchid floral transcriptomes, not many orchid fruit transcriptomes have been published yet. We therefore generated transcriptomes of fruits of *E. pusilla* sampled at different developmental stages in order to discover more about gene networks involved in orchid fruit development. We used two different strategies for data analysis.

First of all, the candidate unigenes that most likely had the longest ORFs (Open Reading Frames) or same Trinity cluster (e.g. DNXXXX_c0_g1) were identified and then filtered by their fragments per kilobase per million mapped base pairs of sequenced (FPKM) values. The high number of unigenes retrieved could not be analyzed with the computer facilities available though. After assembly, we were therefore not able to select the longest transcript from each locus as a unigene for subsequent annotation and thus could not reduce redundancy and potential assembly errors for all the unigenes. The *de novo* assembled unigenes and related groups therefore did not provide very specific information about possible gene networks involved in orchid fruit formation.

Aligning the Ilumina reads to an annotated reference genome, in our case *P. equestris*, downloaded from Plaza4.0 (Van Bel *et al*., 2018) turned out to be more effective as more than 75% of the *E. pusilla* reads could be aligned to this reference. With a Time-Course Expression Analysis, we could detect genes differentially expressed during development of the fruits of *E. pusilla*. In particular the genes involved in the lignin pathway (**Figure S3C** and **Table S4**), which probably also include genes involved in suberin and cutin biosynthesis (Taylor-Teeples *et al*., 2015), are interesting for further analysis. Expression differences in these genes may explain the variation observed in character state 12 (**Figure 2** and **3**), including the presence of a peculiar waxy layer in *E. pusilla* fruits (**Chapter 4**), compared to different lignification patterns in other fruits.

Aligning reads against a reference genome turned out to be more promising and a first analysis revealed putative candidate genes involved in seed formation, pollen tube development and lignification of orchid fruits.

**Ancestral character state analysis of orchid fruits**

A first important finding of our study was that orchid fruit orientation (erect or pendant), opening (dehiscent or indehiscent) and lignification of the valves are phylogenetically informative within the orchid family. The pendant fruits in this study evolved from an erect one and indehiscent fruits evolved from dehiscent fruits multiple times. Non-lignified fruits seem the most ancestral, from this state lignified endocarp evolved first, followed by lignified exocarp.

A second important finding was that the orchid fruits that we investigated thus far showed that lignification in the dehiscence zone had a positive correlation with temperature. It is not clear yet if lignification of the dehiscence zone is most prominent at a high temperature or intermediate to cool temperature.

Thirdly, the majority of the species with erect orchid fruits displayed lignification

in the dehiscence zone. The hypothesis that lignification could be an adaptation to fast fruit development implies this character state to be present in orchids coping with relatively short seasons suitable for seed dispersal, cannot yet be answered with the current dataset. Expanding the species sampling might shed more light on this.

For the ancestral character state analysis, we did not discriminate between valve lignification in either the endocarp or exocarp layer, but we clearly saw that only in *Bletilla striata*, *Bulbophyllum lasianthum* and *Coelogyne dayana*, all species belonging to subfamily Epidendroideae, the exocarp layer becomes lignified during fruit development. Lignification of the endocarp occurred in a single species of the Apostasioideae, none of the species of the Vanilloideae, the majority of the species sampled of the Cypripedioideae and Epidendroideae, and in all species sampled from the Orchidoideae (**Table S5)**. It remains to be seen if these trends remain standing once the sampling is expanded. Based on the Genera Orchidacearum series edited by Pridgeon *et al*., of the two genera in Apostasioideae, both were sampled (100%), of the 15 genera in Vanilloideae, one was sampled (6.6%), of the 5 genera in Cypripedioideae, one was sampled (20%), of the 200 genera in the Orchidoideae, 6 were sampled (3%) and of the 650 genera in the Epidendroideae, 21 were sampled (3.2%). Increase in sampling size should be particularly focused on genera with both terrestrial and epiphytic species such as *Cymbidium* and *Malaxis*, and fruits of the enigmatic underground flowering genus *Rhizanthella*.

Another observation is that some orchid fruits do dehisce and shatter their seeds, while not showing any sign of lignification in either the dehiscence zone or valves (e.g. *Apostasia wallichii*, *Barkeria scandens*, *Bulbophyllum phalaenopsis*, *Dendrobium* sp. and *Vanilla planifolia*). Whether these fruits develop a dehiscence zone consisting of small cells similar to *E. pusilla* and *Oncidium flexuosum* (Mayer *et al*., 2011;Dirks-Mulder *et al*., 2019) remains to be investigated. Staining tissue sections of fruits of these orchid species with toluidine blue O at different developmental stages, as has been done for *E. pusilla*, *Epipactis hellborine* and *Cynorkis fastigiata*, and staining entire fruits with phosphotungtic acid (PTA) together with X-ray computed microtomography (micro-CT) is needed to visualize the possible presence of cuticle-like layers.

For the ancestral character state analysis of orchid fruits we can thus far conclude that an epiphytic or lianaceous habit and fruit ripening period of more than four months clearly coevolved in orchids, similar to an epiphytic or lianaceous habit with a smaller number of opening slits of orchid fruits, a pendant orientation of orchid fruits with a preference for growing at intermediate to high temperatures, a pendant orientation of orchid fruits with an epiphytic or lianaceous habit and lignification of the valves in both pendant and erect orchid fruits. Future comparative transcriptome analyses of orchid fruits may reveal which developmental genes drive this morphological diversity.

## Supplementary figures and tables

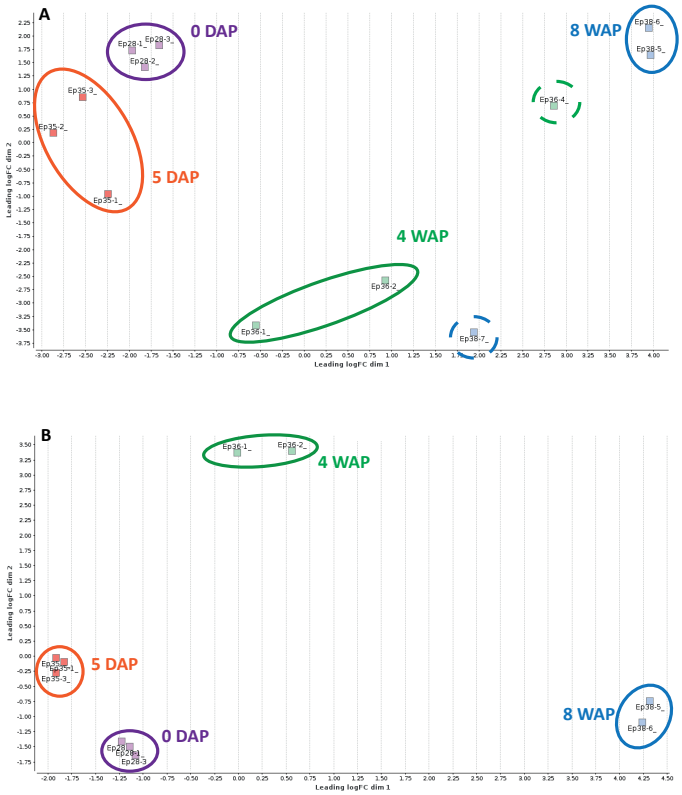



**Figure S1: MDS plot resulting from the RNAseq assembly.** (A) All samples included. (B) Two deviating samples excluded. DAP, days after pollination; WAP, weeks after pollination.

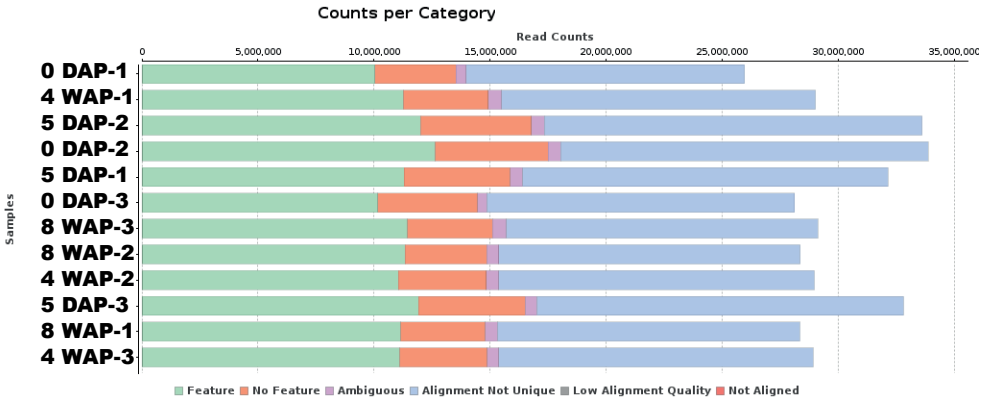



**Figure S2**: **Read counts per category split up over the different fruits samples analyzed resulting from the RNAseq alignment against the *P. equestris* genome.**

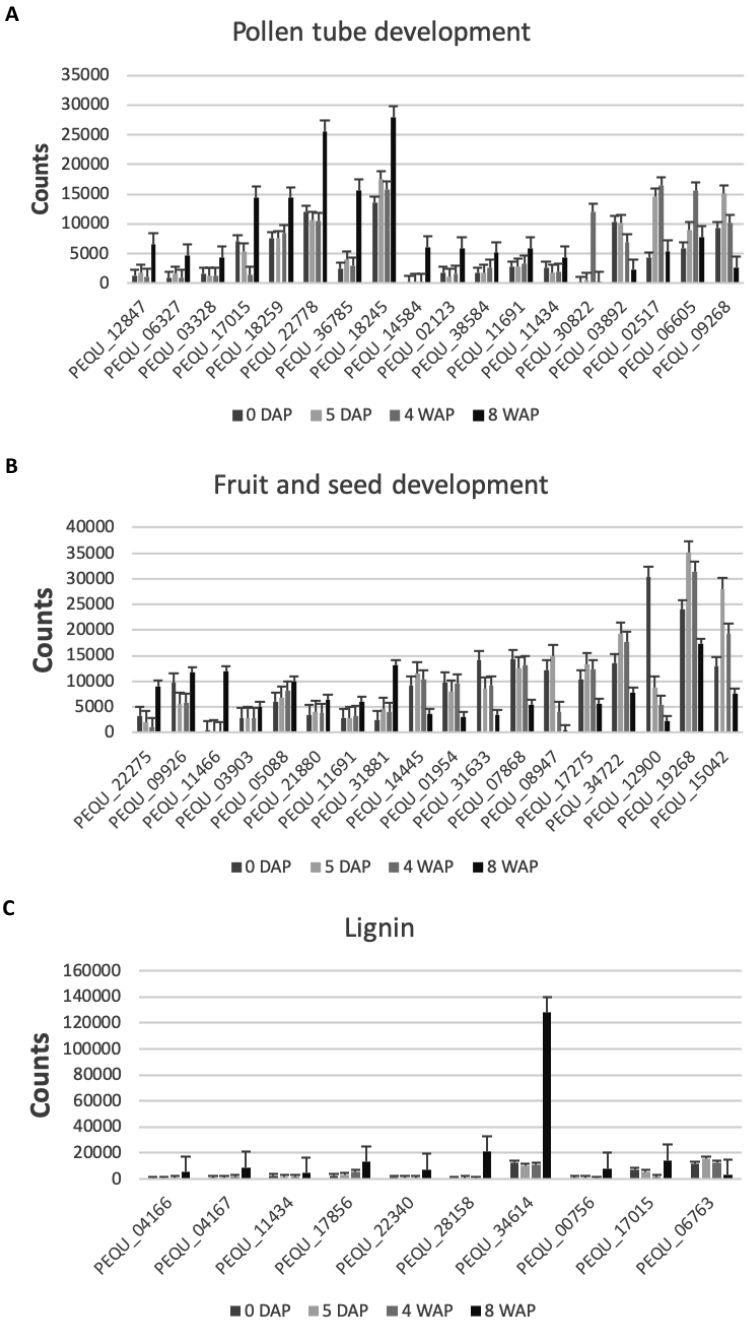



**Figure S3**: **Fruit specific expression patterns of genes involved in (A) Pollen tube development, (B) Fruit and seed development and (C) Lignification.** Each graph shows the counts of the *E. pusilla* reads aligned against the homologous *P. equestris* gene. DAP, days after pollination; WAP, weeks after pollination. The error bars represent the standard deviation.

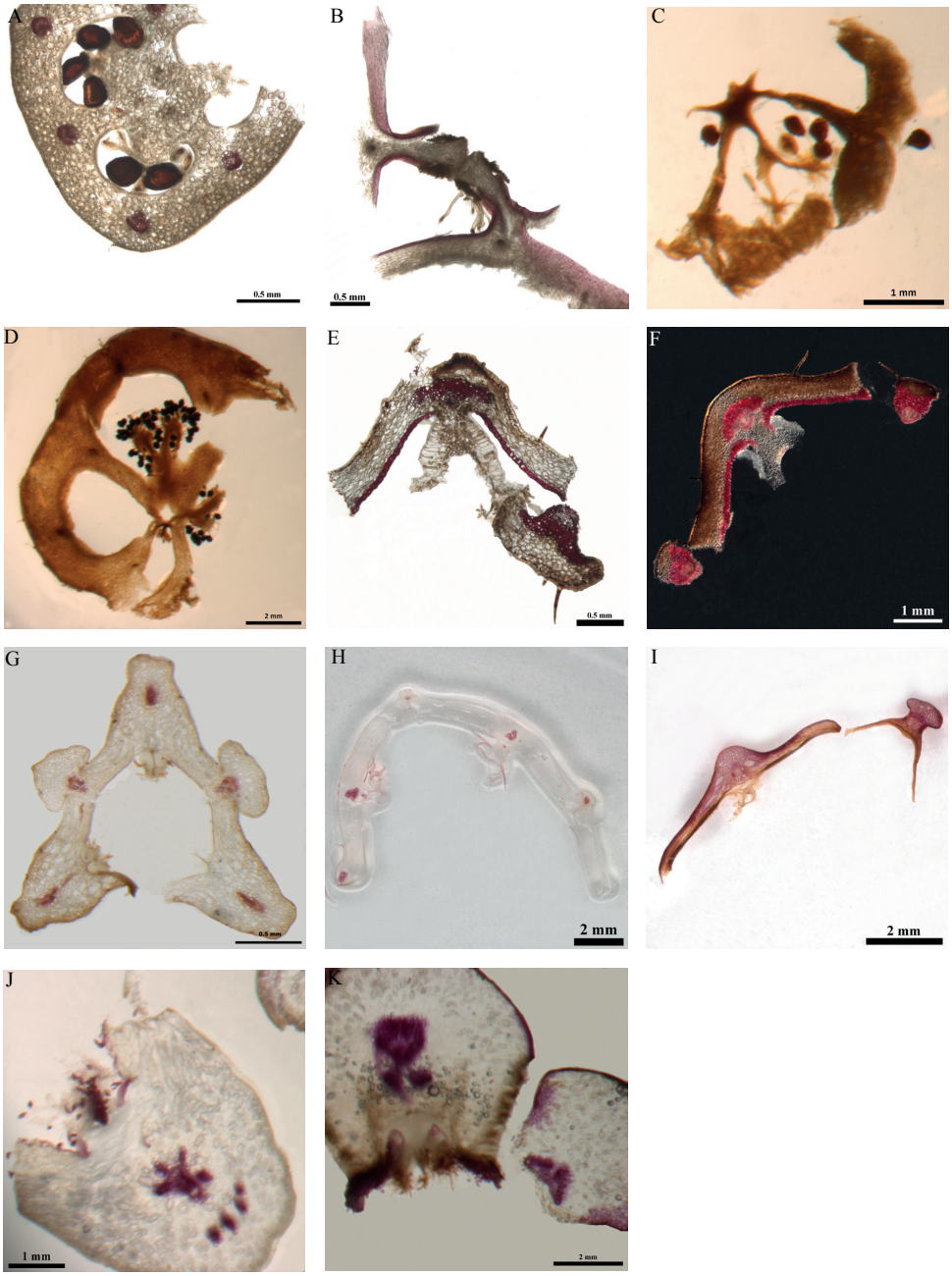



**Figure S4**: **Phloroglucinol staining of fruit cross-sections. (A)** *Apostasia wallichii*, **(B)** *Neuwiedia veratrifolia*, **(C)** *Neuwiedia zollingeri* var. *javanica*, **(D)** *Neuwiedia zollingeri* var. *singapureana*, **(E)** *Paphiopedilum delenatii*, **(F)** *Paphiopedilum philippinensis*, **(G)** *Angraceum*, **(H)** *Barkeria scandens*, **(I)** *Bletilla striata*, **(J)** *Bulbophylum grandiflorum*, **(K)** *Bulbophylum lasianthu*m. Scale bars: 0.5 mm (A, B, E, G), 1 mm (C, F, J) and 2 mm (D, H, I, K).

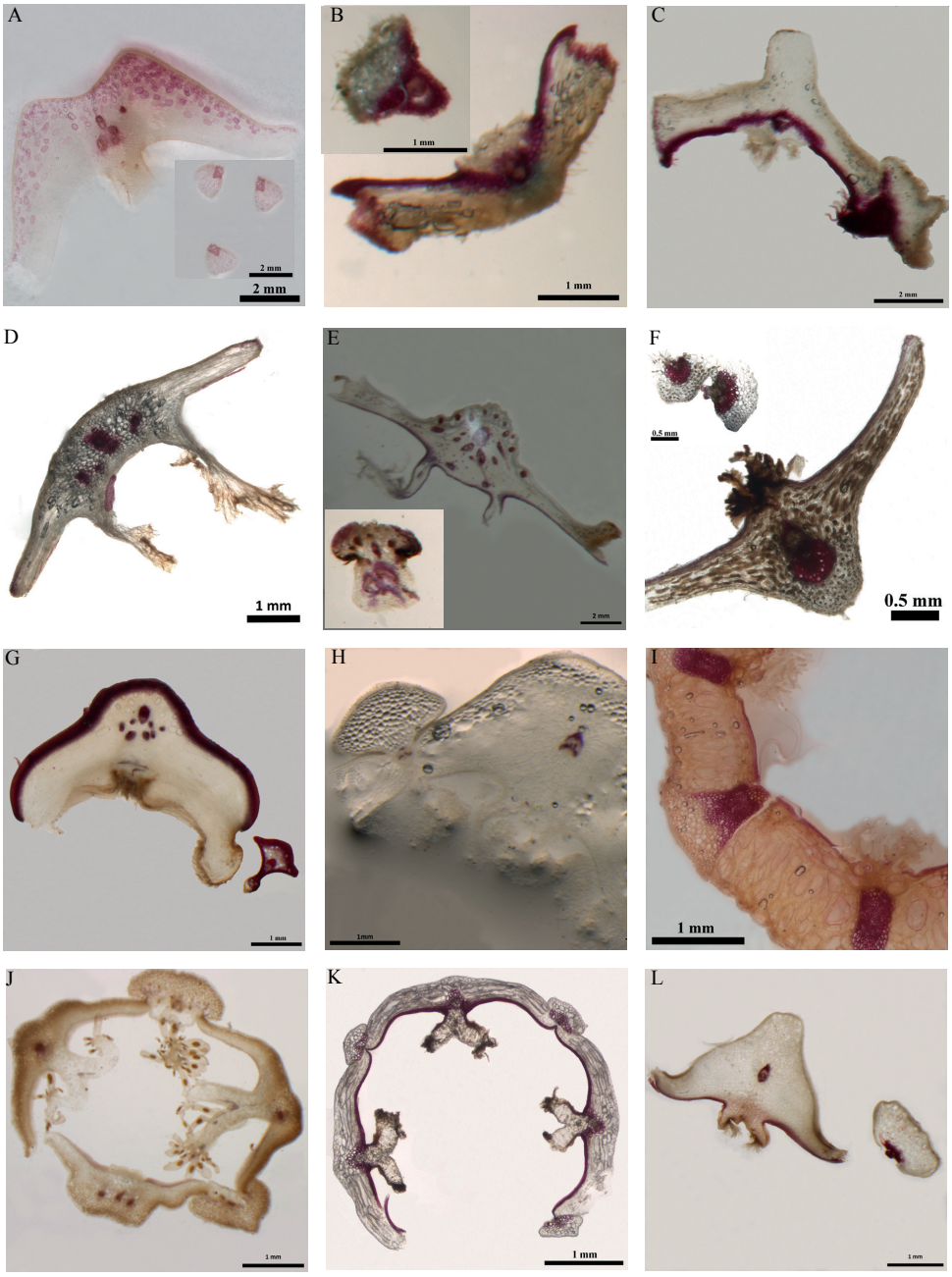



**Figure S5**: **Phloroglucinol staining of fruit cross-sections. (A)** *Brassavola nodosa*, **(B)** *Calanthe brevicornu*, **(C)** *Calanthe sieboldii*, **(D)** *Calopogon tuberosus*, **(E)** *Catasetum planiceps*, **(F)** *Cephalanthera longifolia*, **(G)** *Coelogyne dayana*, **(H)** *Dendrobium*, **(I)** *Eria albutea*, **(J)** *Limodorum abortivum*, **(K)** *Liparis viridiflora*, **(L)** *Oeceoclades maculat*a. Scale bars: 0.5 mm (F), 1 mm (B, D, G-L) and 2 mm (A, C, E).


**Figure S6**: **Phloroglucinol staining of fruit cross-sections**. **(A)** *Phalaenopsis*, **(B)** *Phymatidium delicatulum*, **(C)** *Pleurothallis*, **(D)** *Encyclia*, **(E)** *Anoectochilus*, **(F)** *Ophrys*, **(G)** *Orchis militaris*, **(H-I)** *Sarcoglottis*, **(J)** *Spiranthes cernua*, **(K)** *Vanilla planifolia*, **(L)** *Vanilla pompon*a. Scale bars: 0.5 mm (D-G, J), 1 mm (B, C) and 2 mm (A, H, I, K, L). Arrows (B): lignified dehiscence zone.

**Table S1. Recommended computer specifications to run OmicsBox.**

| Product | Description |
|---|---|
| Memory (RAM) | 64 GB Kingston DDR4 2666MHz ECC-registered (2 x 32 GB) |
| Drives | Storage drive: 2 TB WD Blue™ 3D NAND 2,5" SSD, (max 560 MB/sR \| 530 MB/sW)<br>SSD-station: 2 TB SAMSUNG 970 EVO M.2, PCIe NVMe (max 3500 MB/R, 2500 MB/W) |
| Processor (CPU) | Intel® Xeon® W-2155 10 processor cores (3,3 GHz, 4,5 GHz Turbo, 13,75M Cache) |
| Motherboard | ASUS® WS C422 PRO/SE (DDR4 RDIMM, 6 Gb/s, CrossFireX/SLI) |
| Graphics Card | 24 GB NVIDIA TITAN RTX - HDMI, 3x DP - RTX VR Ready |
| Processor cooling | Corsair H55 Hydro-serie power CPU-cooling |

**Table S2: Accession numbers of *rbc*L, *mat*K and nrITS sequences used in the phylogenetic analysis for the character state analysis.**

| | NCBI GenBank accession number | | |
|---|---|---|---|
| **Species/Phylogenetic marker** | ***rbc*L** | ***mat*K** | **nrITS** |
| *Acianthera_ochreata* | AF518038 | AY008458 | AF366934 |
| *Angraecum_sesquipedale* | AF074106 | AF263621 | KX669265 |
| *Anoectochilus_roxburghii* | | KY966708 | KY966417 |
| *Anoectochilus roxburghii* | | | KR815829 |
| *Apostasia_wallichii* | HM640552 | KC172547 | AY557228 |
| *Barkeria_whartoniana* | | FJ238568 | AF260170 |
| *Bletilla_striata* | AF074114 | EU490679 | AF273334 |
| *Brassavola_nodosa* | | JQ771572 | AF260219 |
| *Bulbophyllum_cambodianum* | KM924495 | KM924448 | KM924472 |
| *Bulbophyllum_lasianthum* | JF428026 | JF305818 | JF428116 |
| *Bulbophyllum_lobbii* | AF074115 | AY368395 | AF521074 |
| *Bulbophyllum_oblongum* | KM924498 | KM924475 | KM924452 |
| *Calanthe_brevicornu* | KF852738 | | |
| *Calanthe_sieboldii* | KF296674 | | AY882613 |
| *Calanthe brevicornu* | | KF852693 | |
| *Calopogon_tuberosus* | AF264161 | AF263635 | AF273372 |
| *Catasetum_expansum* | AF074121 | KF660300 | KU295260 |
| *Cephalanthera_longifolia* | FJ454875 | KF262096 | AY146447 |
| *Coelogyne_dayana* | matK | AY003879 | AF281126 |
| *Coelogyne_pulverula* | KU877826 | KU877846 | |
| *Cynorkis_fastigiata* | AY381117 | MF350008 | MF944264 |
| *Dendrobium_catenatum* | | AB847714 | KJ881390 |

| Species/Phylogenetic marker | NCBI GenBank accession number | | |
|---|---|---|---|
| | *rbc*L | *mat*K | nrITS |
| *Dendrobium_moniliforme* | KJ187385 | AB847816 | |
| *Epipactis_helleborine* | Z73707 | EU490692 | AY154383 |
| *Erycina_pusilla* | NC_018114: 54886-56328 | JN598952 | AF350538 |
| *Hypoxis_curtissii* | KJ773578 | KJ772842 | |
| *Limodorum_abortivum* | JX051376 | | AY351378 |
| *Liparis_viridiflora* | matK | AY907174 | AY907107 |
| *Narcissus bulbocodium* | KY992380 | JX464569 | |
| *Neuwiedia_veratrifolia* | AF074200 | KC172553 | AY557227 |
| *Neuwiedia_zollingeri_var._javanica* | matK | KC172554 | AY557226 |
| *Neuwiedia_zollingeri_var._singapureana* | LC199503: 57921-59384 | LC086542 | KY966622 |
| *Oeceoclades_maculata* | JQ593044 | LN831626 | KF318917 |
| *Ophrys_apifera* | AJ542396 | HE858501 | |
| *Ophrys_sitiaca* | | | AM711736 |
| *Ophrys_sphegodes* | AP018717: 52206-53639 | | AY699974 |
| *Orchis_militaris* | KF997273 | KF997352 | AY699977 |
| *Paphiopedilum_callosum* | KP311864 | KC692129 | JQ929308 |
| *Paphiopedilum_delenatii* | KX264996 | AY368379 | JX088548 |
| *Paphiopedilum_philippinense* | KX755546 | KX755566 | JQ929341 |
| *Paphiopedilum_villosum* | KX755529 | KX755549 | GU993851 |
| *Phalaenopsis_amabilis* | matK | EU256323 | AB217571 |
| *Phymatidium delicatulum* | nrITS | KR709309 | KT709688 |
| *Prosthechea_cochleata* | KJ773788 | KJ773041 | AY008545 |
| *Sarcoglottis_acaulis* | AJ542424 | AJ543928 | AJ539500 |
| *Spiranthes_cernua* | AF074229 | KM213805 | AF301444 |
| *Vanilla_planifolia* | AF074242 | AF263687 | AF391786 |
| *Vanilla_pompona* | NC_036809 | | EU498164 |

**Table S3: Summary of time course expression results.**

| Data overview | Features |
|---|---|
| Total number | 29,431 |
| Number after filtering | 27,171 |
| Identified differentially expressed | 12,410 |
| Significant | 5,984 |
| Cluster 1 | 214 |
| Cluster 2 | 8 |
| Cluster 3 | 4,484 |
| Cluster 4 | 1 |
| Cluster 5 | 35 |
| Cluster 6 | 1,118 |
| Cluster 7 | 4 |
| Cluster 8 | 100 |
| Cluster 9 | 20 |

**Table S4**: **Heatmap representation of expression profiles of genes involved in pollen tube development, fruit and seed development, and lignin biosynthesis.** The scale for each gene was set to the highest green value. DAP = days after pollination; WAP = weeks after pollination. Colors: No color: PEQU from cluster 7-8-9, orange: PEQU from cluster 2, green: PEQU from cluster 5.

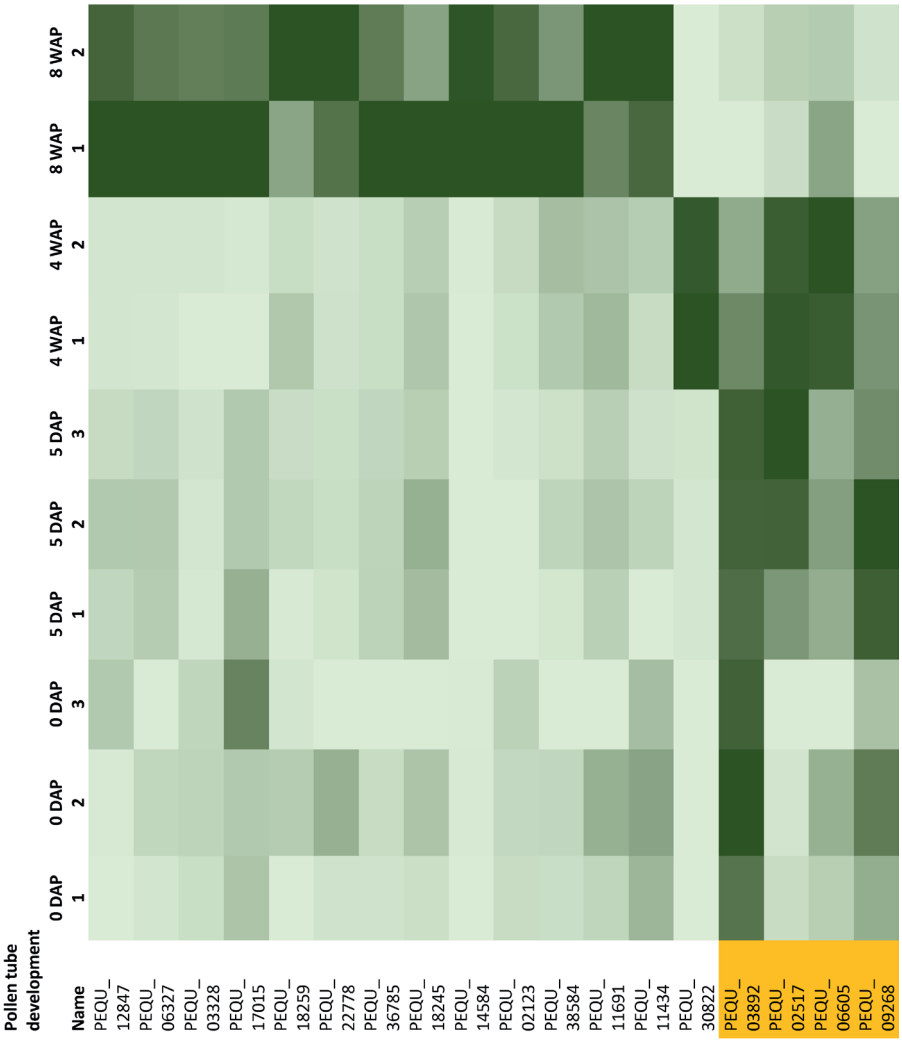


**Table S4 continued**: Heatmap representation of expression profiles of genes involved in pollen tube development, fruit and seed development, and lignin biosynthesis.

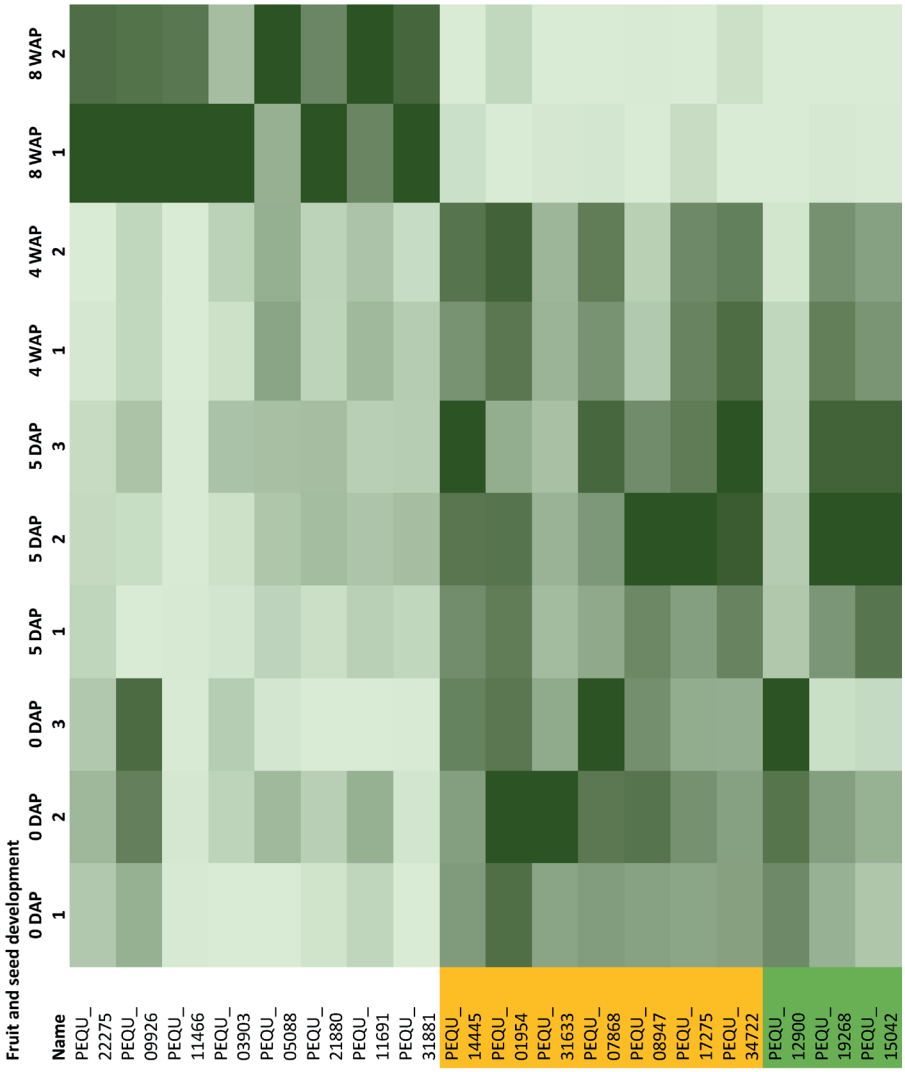


**Table S4 continued**: Heatmap representation of expression profiles of genes involved in pollen tube development, fruit and seed development, and lignin biosynthesis.

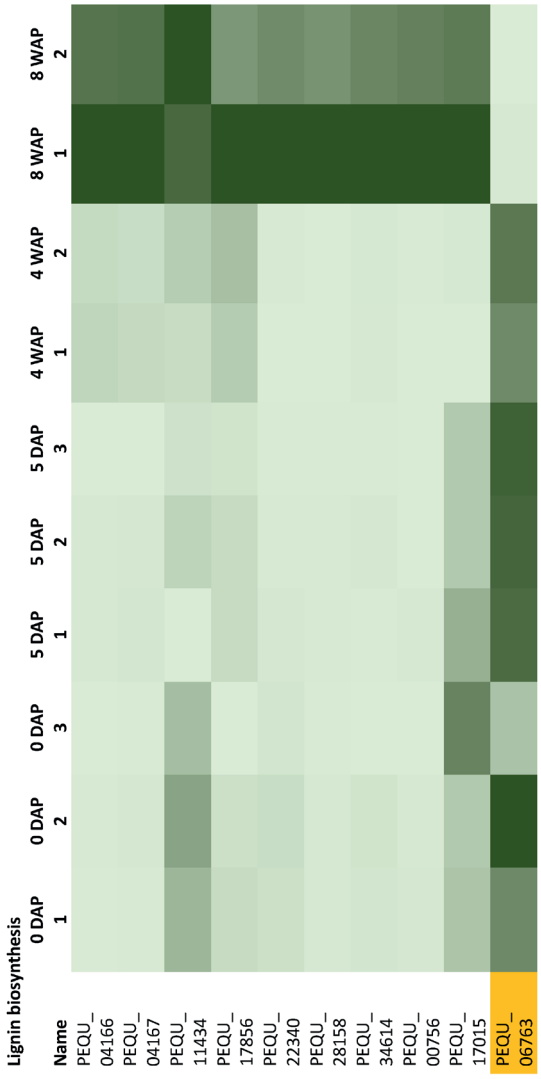


**Table S5**: **Lignification of the dehiscence zone (DZ) and the valves of 41 orchid species and two non-orchid species.**

| Subfamily | Name | Lignification of dehisence zone | Lignification of valves in endocarp/exocarp |
|---|---|---|---|
| Apostasioideae | *Apostasia wallichii* | No | No |
| Apostasioideae | *Neuwiedia veratrifolia* | No | Endocarp |
| Apostasioideae | *Neuwiedia zollingeri var. javanica* | No | No |
| Apostasioideae | *Neuwiedia zollingeri var. singapureana* | No | No |
| Cypripedioideae | *Paphiopedilum callosum* | Yes | No |
| Cypripedioideae | *Paphiopedilum delenatii* | Yes | Endocarp |
| Cypripedioideae | *Paphiopedilum philippinense* | Yes | Endocarp |
| Cypripedioideae | *Paphiopedilum villosum* | Yes | Endocarp |
| Epidendroideae | *Angraecum sp.* | No | Endocarp |
| Epidendroideae | *Barkeria scandens* | No | No |
| Epidendroideae | *Bletilla striata* | No | Exocarp |
| Epidendroideae | *Bulbophyllum grandiflorum* | No | No |
| Epidendroideae | *Bulbophyllum lasianthum* | Yes | Exocarp |
| Epidendroideae | *Bulbophyllum oreonastes* | Yes | No |
| Epidendroideae | *Bulbophyllum phalaenopsis* | No | No |
| Epidendroideae | *Brassavola nodosa* | Yes | No |
| Epidendroideae | *Calanthe brevicornu* | Yes | Endocarp |
| Epidendroideae | *Calanthe sieboldii* | Yes | Endocarp |
| Epidendroideae | *Calopogon tuberosus* | Yes | Endocarp |
| Epidendroideae | *Catasetum planiceps* | No | Endocarp |
| Epidendroideae | *Cephalanthera longifolia* | No | Endocarp |
| Epidendroideae | *Coelogyne dayana (C. pulverula)* | Yes | Exocarp |
| Epidendroideae | *Dendrobium sp. section Crinivera* | No | No |
| Epidendroideae | *Epipactis helleborine* | Yes | Endocarp |
| Epidendroideae | *Eria albolutea* | Yes | Endocarp |
| Epidendroideae | *Erycina pusilla* | No | Endocarp |
| Epidendroideae | *Limodorum abortivum* | Yes | Endocarp |
| Epidendroideae | *Liparis viridiflora* | Yes | Endocarp |
| Epidendroideae | *Oeceoclades maculata* | No | Endocarp |
| Epidendroideae | *Phaleanopsis hybrid* | No | Endocarp |
| Epidendroideae | *Phymatidium delicatulum* | Yes | No |
| Epidendroideae | *Pleurothallis sp.* | No | Endocarp |
| Epidendroideae | *Prosthechea cochleata (= Encyclia)* | Yes | No |
| Orchidoideae | *Anoectochilus papuanus* | Yes | Endocarp |
| Orchidoideae | *Cynorkis fastigiata* | Yes | No |
| Orchidoideae | *Ophrys sp.* | Yes | Endocarp |
| Orchidoideae | *Sarcoglottis sp.* | Yes | Endocarp |
| Orchidoideae | *Orchis militaris* | Yes | Endocarp |
| Orchidoideae | *Spiranthes cernua* | Yes | Endocarp |
| Vanilloideae | *Vanilla planifolia* | No | No |
| Vanilloideae | *Vanilla pompona* | No | No |
| Amaryllidoideae | *Narcissus* | No | Endocarp |
| Hypoxidaceae | *Hypoxis angustifolia* | No | No |

## Acknowledgements

We thank Alexander Kocyan, Elaine Lopes Pereira Nunes, Rogier van Vugt, Gert-Jan de Waard, and Anneke Wagner for providing us with fruit material.

## References

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31**,** 166-169.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics* 25**,** 25-29.

Beer, J.G. (1863). *Beiträge zur morphologie und biologie der familie der orchideen.* Wien: Carl Gerold's Sohn.

Bollback, J.P. (2006). SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics* 7**,** 88.

Cai, J., Liu, X., Vanneste, K., Proost, S., Tsai, W.-C., Liu, K.-W., Chen, L.-J., He, Y., Xu, Q., Bian, C., Zheng, Z., Sun, F., Liu, W., Hsiao, Y.-Y., Pan, Z.-J., Hsu, C.-C., Yang, Y.-P., Hsu, Y.-C., Chuang, Y.-C., Dievart, A., Dufayard, J.-F., Xu, X., Wang, J.-Y., Wang, J., Xiao, X.-J., Zhao, X.-M., Du, R., Zhang, G.-Q., Wang, M., Su, Y.-Y., Xie, G.-C., Liu, G.-H., Li, L.-Q., Huang, L.-Q., Luo, Y.-B., Chen, H.-H., De Peer, Y.V., and Liu, Z.-J. (2015). The genome sequence of the orchid Phalaenopsis equestris. *Nature Genetics* 47**,** 65-72.

Camus, E.G., Lecomte, H., and Camus, A.E. (1921). *Iconographie des orchidées d'Europe et du bassin Méditerranéen.* Paris: P. Lechavalier.

Chomicki, G., Bidel, L.P., Ming, F., Coiro, M., Zhang, X., Wang, Y., Baissac, Y., Jay-Allemand, C., and Renner, S.S. (2015). The velamen protects photosynthetic orchid roots against UV-B damage, and a large dated phylogeny implies multiple gains and losses of this function during the Cenozoic. *New Phytol* 205**,** 1330-1341.

Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21**,** 3674-3676.

Conesa, A., Nueda, M.J., Ferrer, A., and Talon, M. (2006). maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 22**,** 1096-1102.

Di Vittori, V., Gioia, T., Rodriguez, M., Bellucci, E., Bitocchi, E., Nanni, L., Attene, G., Rau, D., and Papa, R. (2019). Convergent Evolution of the Seed Shattering Trait. *Genes (Basel)* 10**,** 68.

Dirks-Mulder, A., Ahmed, I., Uit Het Broek, M., Krol, L., Menger, N., Snier, J., Van Winzum, A., De Wolf, A., Van't Wout, M., Zeegers, J.J., Butot, R., Heijungs, R., Van Heuven, B.J., Kruizinga, J., Langelaan, R., Smets, E.F., Star, W., Bemer, M., and Gravendeel, B. (2019). Morphological and Molecular Characterization of Orchid Fruit Development. *Front Plant Sci* 10**,** 137.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29**,** 15-21.

Dressler, R.L. (1981). *The orchids: natural history and classification.* Cambridge: Harvard University Press.

Dressler, R.L. (1993). *Phylogeny and Classification of the Orchid Family.* Cambridge, UK: Cambridge University Press.

Gandrud, C. (2015). *Reproducible Research with R and R Studio.* New York: Chapman and Hall/CRC.

Givnish, T.J., Spalink, D., Ames, M., Lyon, S.P., Hunter, S.J., Zuluaga, A., Iles, W.J., Clements, M.A., Arroyo, M.T., Leebens-Mack, J., Endara, L., Kriebel, R., Neubig, K.M., Whitten, W.M., Williams, N.H.,

and Cameron, K.M. (2015). Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc Biol Sci* 282.

Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J., and Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36**,** 3420-3435.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* 29**,** 644-652.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., Macmanes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., and Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols* 8**,** 1494-1512.

Horowitz, A. (1901). Ueber den anatomischen Bau und das Aufspringen der Orchideen frűchte. *Beihefte zum Botanischen Centralblatt* 11**,** 486-521.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., and Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28**,** 1647-1649.

Kocyan, A., and Endress, P. (2001). Floral Structure And Development of Apostasia and Neuwiedia (Apostasioideae) and their Relationships to other Orchidaceae. *International Journal of Plant Sciences* 162**,** 847-867.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9**,** 357-359.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12**,** 323.

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22**,** 1658-1659.

Mayer, J.L.S., Carmello-Guerreiro, S.M., and Appezzato-Da-Glória, B. (2011). Anatomical development of the pericarp and seed of Oncidium flexuosum Sims (ORCHIDACEAE). *Flora - Morphology, Distribution, Functional Ecology of Plants* 206**,** 601-609.

Miller, M.A., Pfeiffer, W., and Schwartz, T. (Year). "Creating the CIPRES Science Gateway for inference of large phylogenetic trees", in: *2010 Gateway Computing Environments Workshop (GCE)*), 1-8.

Omicsbox Bioinformatics made easy. BioBam Bioinformatics.

Pagel, M. (1994). Detecting Correlated Evolution on Phylogenies: A General Method for the Comparative Analysis of Discrete Characters. *Proceedings: Biological Sciences* 255**,** 37-45.

Pagel, M., and Cunningham, C. (1999). The Maximum Likelihood Approach to Reconstructing Ancestral Character States of Discrete Characters on Phylogenies. *Systematic Biology* 48**,** 612-622.

Pagel, M., and Meade, A. (2006). Bayesian analysis of correlated evolution of discrete characters by reversible-jump Markov chain Monte Carlo. *Am Nat* 167**,** 808-825.

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20**,** 289-290.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R news* 6**,** 7-11.

Poinar, G., and Rasmussen, F.N. (2017). Orchids from the past, with a new species in Baltic amber. *Botanical Journal of the Linnean Society* 183**,** 327-333.

Pridgeon, A., Cribb, P., Chase, M.W., and Rasmussen, F.N. (1999a). *Genera Orchidacearum Volume 3: Orchidoideae (Part 2) Vanilloideae.* Oxford: Oxford Universitty Press.

Pridgeon, A.M., Cribb, P., Chase, M.W., and Rasmussen, F.N. (2005). *Genera Orchidacearum Volume 4: Epidendroideae.* OUP Oxford.

Pridgeon, A.M., Cribb, P., Chase, M.W., and Rasmussen, F.N. (2009). *Genera Orchidacearum Volume 5: Epidendroideae.* OUP Oxford.

Pridgeon, A.M., Cribb, P.J., Chase, M.W., and Rasmussen, F. (1999b). *Genera Orchidacearum: Volume 1: Apostasioideae and Cypripedioideae.* OUP Oxford.

Pridgeon, A.M., Cribb, P.J., Chase, M.W., and Rasmussen, F.N. (2001). *Genera Orchidacearum: Volume 2. Orchidoideae.* OUP Oxford.

Pridgeon, A.M., Cribb, P.J., Chase, M.W., and Rasmussen, F.N. (2014). *Genera Orchidacearum Volume 6: Epidendroideae.* OUP Oxford.

R Development Core Team (2018). R: A Language and Environment for Statistical Computing}.

Rambaut, A. (2014). *FigTree. v. 1.4. 2: Tree drawing tool.* [Online]. Available: http://tree.bio.ed.ac.uk/software/figtree/ [Accessed].

Ramirez, S.R., Gravendeel, B., Singer, R.B., Marshall, C.R., and Pierce, N.E. (2007). Dating the origin of the Orchidaceae from a fossil orchid with its pollinator. *Nature* 448**,** 1042-1045.

Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3**,** 217-223.

Robinson, M.D., Mccarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* 26**,** 139-140.

Stern, W.L., Gregory, M., and Cutler, D.F. (2014). *Anatomy of the Monocotyledons Volume X: Orchidaceae.* OUP Oxford.

Taylor-Teeples, M., Lin, L., De Lucas, M., Turco, G., Toal, T.W., Gaudinier, A., Young, N.F., Trabucco, G.M., Veling, M.T., Lamothe, R., Handakumbura, P.P., Xiong, G., Wang, C., Corwin, J., Tsoukalas, A., Zhang, L., Ware, D., Pauly, M., Kliebenstein, D.J., Dehesh, K., Tagkopoulos, I., Breton, G., Pruneda-Paz, J.L., Ahnert, S.E., Kay, S.A., Hazen, S.P., and Brady, S.M. (2015). An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature* 517**,** 571-575.

Vaidya, G., Lohman, D.J., and Meier, R. (2011). SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* 27**,** 171-180.

Van Bel, M., Diels, T., Vancaester, E., Kreft, L., Botzki, A., Van De Peer, Y., Coppens, F., and Vandepoele, K. (2018). PLAZA 4.0: an integrative resource for functional, evolutionary and comparative plant genomics. *Nucleic Acids Res* 46**,** D1190-D1196.

Yang, Z., Kumar, S., and Nei, M. (1995). A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics* 141**,** 1641-1650.

Yu, G., K. Smith, D., Zhu, H., Guan, Y., and Tsan-Yuk Lam, T. (2016). *ggtree : an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data.*

Ziegenspeck, H. (1936). *In: von Kirchner O, Loew E, Schroter C, eds. Lebensgeschichte der Blutenpflanzen Mitteleuropas.* Stuttgart, Germany: Eugen Ulmer Verlag.

# Chapter 6

## General discussion and conclusions

## In vitro propagation of *Erycina pusilla*

During this PhD project, several innovative insights were obtained about the genetic basis of the development of floral organs and fruits of the orchid species *Erycina pusilla*. First, though, several basic skills had to be developed at Naturalis Biodiversity Center. One of the first challenges was to establish an efficient *in vitro* propagation protocol for *E. pusilla*, which could easily be implemented in other laboratories as well. We started in 2014 using an existing protocol developed by the Dutch orchid breeder Johan Keus. One of the biggest challenges that we encountered was contamination of the medium. Only after we started to use Sodium dichloroisocyanurate (NaDCC) as sterilant and the antimicrobial and Plant Preservative Mixture (PPM) as antimicrobial substance (Kendon *et al*., 2017), our *in vitro* cultures of *E. pusilla* could be completed from seed to fruiting stage without loss of plant material due to undesirable fungal and bacterial infections.

## Life cycle of *Erycina pusilla*

Once a growing protocol had been established, a second necessity was to characterize the full development of the flowers and fruits of *E. pusilla* by studying the full ontogeny using the microscopy facilities present at Naturalis Biodiversity Center and Leiden University Medical Centre. Floral morphology and ontogeny of *E. pusilla* flowers was unraveled in **chapter 3**, in which the development of *E. pusilla* flowers was subdivided into five stages based on macro-morphological changes from the first emergence of an inflorescence stalk up to a fully open flower. *Erycina pusilla* fruits on the other hand were divided into four stages based on different macro- and micro-morphological changes as described in **chapter 4**.

*Erycina pusilla* is a common twig epiphyte in the meso-American tropics; our laboratory strain came from Surinam. In nature, *Centris* bees pollinate the yellow non-rewarding flowers. During floral visits, these bees cling to the stelidia (wings along the gynostemium) and the callus on the lip with their forelegs while searching for floral oils that are not present. Flowers of *E. pusilla* are mimicking the rewarding flowers of Malpighiaceae, which do have rewarding oil glands. The stelidia were detected to be the remnants of the six stamens, which are still fertile in a few basal orchids with rather primitive pollination syndromes such as *Apostasia wallichii* but reduced to sterile organs in most derived lineages with highly specialized pollination syndromes such as *E. pusilla*.

In the lab, we pollinated flowers of *E. pusilla* ourselves by removing the pollinarium with a forceps and placing the pollinia into the stigmatic cavity of the gynostemium. After a flower was successfully pollinated it wilted after a few days and the stelidia folded together (**Figure 1**). We actually postulate that this process, either active or an effect of senescence once pollination is achieved, protects the stigma with the developing pollen tubes against UV radiation and herbivores.

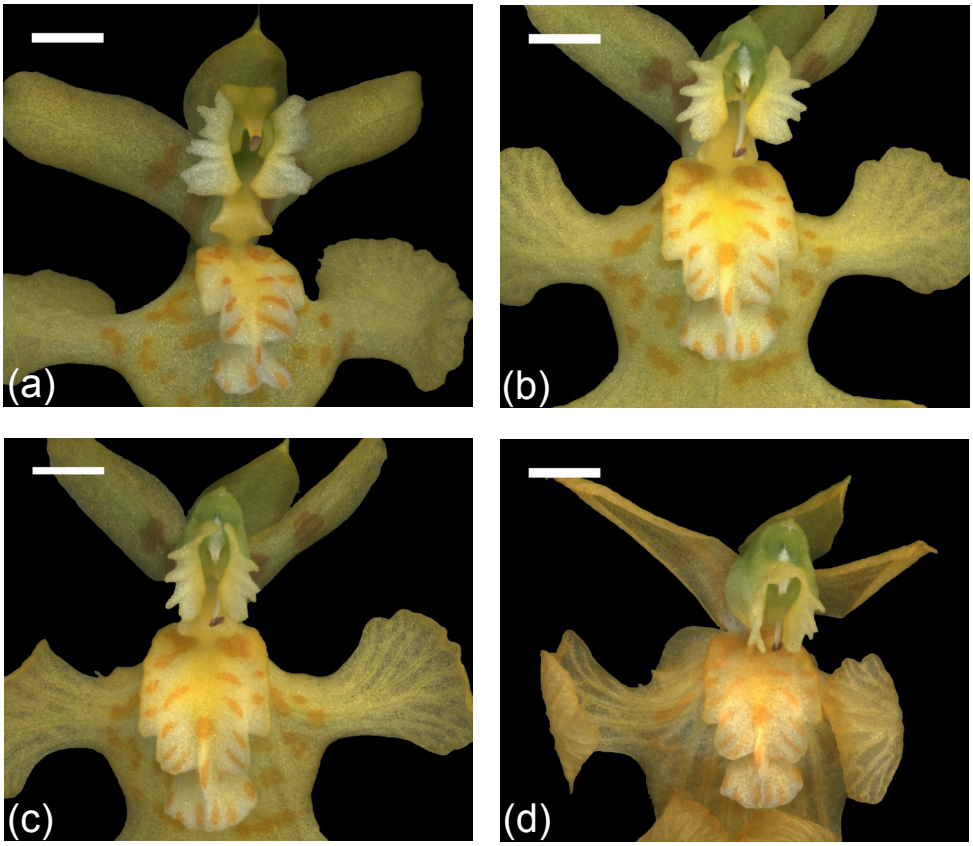



**Figure 1. Macro-morphology of E. pusilla prior to and after pollination. (a)** Fresh flower prior to pollination. **(b)** Flower one day after pollination. **(c)** Flower two days after pollination. **(d)** Flower five days after pollination. Scale bar=2 mm.

Pollen tubes were macroscopically visible in developing fruits until 7-8 weeks after pollination (WAP) suggesting that ovules were fertilized. After 16 WAP the fruit was ripe and opened to release the seeds. The seeds germinated in ~3-4 months from protocorms into protocorm-like bodies (PLBs). Once the first roots and leaves had developed from the PLBs, the sterile plantlets were transferred to individual tubes, in which they developed into full plants with inflorescences in ~ 3-4 months. One aspect that still needs to be optimized for our *E. pusilla* Surinam strain is cryopreservation of seeds, or embryos, for short- and long-term storage. Up until now the strain has been propagated by continuous tissue culture, which is time consuming, expensive and risky. Different attempts were made in the beginning of this PhD project, including dry storage at 4 °C and -20 °C but unfortunately these attempts all failed due to infected plates and lack of germination. Various articles were published over the last few years (Popova *et al*., 2016;Cerna *et al*., 2018;Magrini *et al*., 2018;Schofield *et al*., 2018) with instructions to set up a simple cryopreservation protocol for *E. pusilla* seeds and use triphenyl tetrazolium chloride

(TTC) staining to test seed viability before sowing the seeds for future attempts.

## MADS-box genes involved in flower and fruit development

The shape of orchid flowers is very complicated due to gene duplications and sub-functionalizations that evolved over millions of years. MADS-box genes play an important role in flower development so to find out what the role is of these different genes was, we conducted an expression study of 20 different MADS-box genes isolated from floral organs of early and late developing floral buds. We then constructed gene lineage trees for every MADS-box gene class to investigate to which clades all duplicates belonged and whether they evolved at the same rate or not. Combining these data with micro-morphological data of developing floral organs and epidermal cell structures using SEM, the Oncidiinae model was developed in **chapter 3.** This model describes how different copies of MADS-box genes *AP3* (*EpMADS14*) and *AGL6* (*EpMADS4*) code for the shape of the lateral sepals. When these gene copies are expressed in the perianth, the lateral sepals have a sepaloid appearance. Different developmental MADS-box gene copies determine the various shapes of orchid sepals, lips and stelidia. Following the perianth-code model by Hsu *et al*. (2015), the development of the lip compared with the sepals and petals is based on the interaction of *PI* with two different *AP3* and *AGL6* copies. We found for *E. pusilla* flowers that (i) the Lip-complex is fully expressed in the lip and the callus on the lip, (ii) the SP-complex is fully expressed in the sepals and petals and (iii) *AG* (*EpMADS20*) and *STK* (*EpMADS23*) copies shape the stelidia and a copy of *AG* (*EpMADS22*) and *SEP* (*EpMADS6*) shape the stamen.

We concluded that the enlarged median sepal, incised lip, callus and stelidia of *E. pusilla* evolved to mimic the shape of the petals and oil glands of flowers of Malpighiaceae in order to attract oil-collecting bees for pollination. All these organs evolved for a perfect fit with bodies of specific pollinators to optimize pollination success.

Fruit-associated MADS-box proteins were studied in **chapter 4** by performing a yeast-two-hybrid assay to study protein-protein interactions. Together with an expression study on a series of developing fruits of *E. pusilla*, we characterized orthologs of fruit-associated MADS-domain transcription factors and of the *Arabidopsis thaliana* dehiscence-related genes *INDEHISCENT* (*IND*)/*HECATE3* (*HEC3*), *REPLUMLESS* (*RPL*) and *SPATULA* (*SPT*)/*ALCATRAZ* (*ALC*). We found that the key players of the eudicot fruit regulatory network appear well conserved in monocots. Protein-protein interaction studies revealed that MADS-domain complexes comprised of *FRUITFULL* (*FUL*), *SEPALLATA* (*SEP*) and *AGAMOUS* (*AG*) /*SHATTERPROOF* (*SHP*) orthologs can also be formed in *E. pusilla*, and that the expression of *HEC3*, *RPL* and *SPT* can be associated with dehiscence zone development similar to Arabidopsis. Our gene expression analysis also indicates differences, however, which may underlie fruit divergence.

Massive parallel sequencing of short RNA molecules added important additional insights in the genetic basis of the orchid sepals, petals and lip, as described in the Perianth Code model (Gravendeel and Dirks-Mulder, 2015;Hsu *et al*., 2015)

and Oncidiinae model. It also provided the first glimpses of the genetic basis of other orchid organs in the third and fourth floral whorls such as the stamen and stelidia, and fruit dehiscence. A third challenge was to start up transcriptome analyses. For this, initially, existing data in Orchidstra (Su *et al*., 2013;Chao *et al*., 2017) were mined. This was sufficient for the first expression studies on floral and fruit organs as presented in **chapter 3 and 4**, but we could not find any *E. pusilla HEC3* homologue in this database. This meant that we had to produce our own fruit transcriptomic data.

## Orchid transcriptome analysis

It became apparent that we needed to generate transcriptome data ourselves to be able to detect additional genes involved in fruit development besides the MADS-box genes. This meant that we needed a full suite of tailor-made bioinformatics pipelines. Bioinformaticians of Naturalis Biodiversity Center and many bioinformatics students, staff and lecturers of the University of Applied Sciences Leiden were vital for developing the orchid genomic toolkit for generating a *de novo* transcriptome using Blast2GO (Gotz *et al*., 2008). This bioinformatics platform was used for analysis of the orchid transcriptomes generated. Vital for the annotating process as well was the creation of a custom made Orchid-Blast-Database to run local Blasts. This database turned out to be dominated by floral transcriptomes, though, and did not contain that many genes yet involved in fruit formation. By mapping reads against the assembled and annotated reference genome of *P. equestris*, though, several fruit transcriptomes of *E. pusilla* could be analyzed with various bioinformatics tools as presented in **chapter 5**. This resulted in the identification of a first group of candidate genes involved in seed formation, pollen tube development and lignification, that are either up- or down-regulated during development of the fruits of *E. pusilla.*



Currently, genome sequences are published for a total of seven orchid species: *Dendrobium officinale* and *D. catenatum* (Yan *et al*., 2015;Zhang *et al*., 2016), *Phalaenopsis equestris* and *P. aphrodite* (Cai *et al*., 2015;Chao *et al*., 2018) (all subfamily Epidendroideae), *Apostasia shenzhenica* (subfamily Apostasioideae) (Zhang *et al*., 2017), *Gastrodia elata* (Yuan *et al*., 2018) (subfamily Epidendroideae) and a draft genome of *Vanilla planifolia* (Hu *et al*., 2019) (subfamily Vanilloideae). Comparing these orchid genomes gave important first insights into the existence of new gene families, and how these gene families either expanded or contracted during the evolution of this plant family. This was done by comparing the genome of the rather basal orchid species *A. shenzhenica* with the genomes of the more derived species *Phalaenopsis equestris* and *D. catenatum* (Zhang *et al*., 2017). The authors found gains and losses of certain MADS-box and other genes (e.g. Myb factors) controlling a diverse suite of processes, e.g. the development of the lip and gynostemium, pollinia, and seeds without endosperm.

During this PhD project, the first attempts were made for full genome sequencing and hybrid assembly of the *E. pusilla* genome. We combined second

generation Illumina HiSeq sequencing, creating short (50-100 bp long) reads with high (35x) coverage, and PacBio analysis, creating longer (3-5 kb long) reads, sequenced with low (5x) coverage. Using different assembly methods, a first *de novo* hybrid genome assembly was created, consisting of more than 450,000 scaffolds with only 2% of the total data used. Ongoing developments in third generation sequencing technologies will soon enable retrieval of even higher quality genomes and also transcriptomes by sequencing longer reads that are expected to ultimately encompass full-length transcripts. An important emerging research tool is the MinION Oxford Nanopore DNA/RNA sequencing technology to generate such long (up to 50 kb) sequence reads, which can be used to read entire plant transcriptomes. This technique has already been used for whole genome sequencing of other plant species, also in combination with optical mapping (Belser *et al*., 2018;Deschamps *et al*., 2018). Plant genomes are large, often polyploid, and may contain over 90% of repetitive DNA (Mehrotra and Goyal, 2014) so sequencing the relatively small and diploid orchid genome of *E. pusilla* is feasible. For future studies, I therefore recommend completing the sequencing and assembly of the full genome of *E. pusilla*. Anonymous (short) reads, either genomic or transcriptomic, can then be matched to this reference genome to be able to carry out future gene identification and editing studies.

## Genome editing

For *E. pusilla* to become an established research model for evolutionary, developmental and genetic studies, not only genomic data have to become available but also ways to genetically manipulate plants by e.g. transformation to either overexpress, down-regulate or knock-out genes. Lee *et al*. (2015) published an *Agrobacterium*-mediated genetic transformation protocol for *E. pusilla* using three-month-old protocorms for expressing *Arabidopsis thaliana* genes within 14 months. Unfortunately for every transformation experiment, ca. 2,500 protocorms have to be used and screened for more than ten rounds on selective media. For research purposes this is not a practical approach, although some adjustments can be made. Hsing *et al*. (2016) published a similar method for *Phalaenopsis* orchids using only three selection rounds and then transferring the transformed plants to non-selective media, which speeds up the protocol considerably.

To study gene functions of orchids, genes can either be transiently silenced or inhibited using RNA-interference (RNAi). RNA interference refers to suppression of gene expression of sequence-specific, homologous RNA molecules and is triggered by a double-stranded RNA (dsRNA) mediator to generate small interfering RNA, which then results in sequence-specific RNA degradation.

Different approaches have been used over the last two decades among which Virus Induced Gene Silencing (VIGS). Different *Cymbidium mosaic virus*-based VIGS vectors (pCymMV) have been used to induce gene silencing in orchids of which the pCymMV-Gateway is, in my opinion, the most convenient vector (Lu

*et al*., 2012), because of its Gateway cassette, which is optimal for easy and high-throughput cloning and screening of genes of interest. Before the start of my PhD in 2013, we requested this vector to suppress gene expression in *E. pusilla*. At that time, sequence information for *E. pusilla* was scarce and gene expression unknown so knocking-down genes of interest had to wait until the first expression studies had been carried out. Despite the fact that time did not permit me to use this technique during my PhD study, I advise to give VIGS a go with *E. pusilla* and first use the homologous MADS-box genes that Hsieh *et al*. (2013) used in *Phalaenopsis*, which are *PI* (*EpMADS16*) and *AP3*, clade 2 (*EpMADS14*). Down regulation can be monitored using RT-qPCR and SEM as carried out in **Chapter 3**. The VIGS technique has been used for examining knockdown phenotypes in non-model plants and is less labor intensive and more rapid than stable transformation approaches. Some considerations have to be taken into account, such as the low transformation efficiency and difficulty to assay the specific knockdown of a gene of interest if multiple copies exist.

Delivery of dsRNA in plants can also be achieved by directly rubbing the dsRNA into the plant. For this dsRNA is produced by the RNase III-deficient *E. coli* strain HT115 (DE3) containing the pL4440 plasmid, which carries the gene of interest. With this method, orchid plants were successfully protected against CymMV infection using dsRNA from the viral coat protein (Lau *et al*., 2015a) and the shape of epidermal cells of a *Dendrobium* orchid was changed using R2R3MYB transcription factor dsRNA (Lau *et al*., 2015b). For *E. pusilla* we cloned a *MYB* homologue and *PI*, two genes involved in floral color, in pL4440 and tried to downregulate these genes once by applying the dsRNA directly on young flower buds. No phenotypic effects were observed and due to time constraints and priority given to other research, this part of this PhD project could not be completed. Because this method seems so easy and does show an effect in other orchids, I definitely recommend spending more time on this when studying gene functions of orchid genes. I recommend treating not only young flower buds with dsRNA but also PLBs, use multiple treatments, apply purified dsRNA instead of crude lysate, and use vacuum to penetrate the dsRNA as possible options for *E. pusilla* to knock down gene expression.

Another method to edit genomes is CRISPR/Cas9 and with an increasing number of scientific papers published over the past five years in plants, this seems a very promising technique for editing orchid genomes as well. In 2017, Kui *et al.* published for the first time a method using *Agrobacterium* to deliver a CRISPR/Cas9 construct into *Dendrobium officinale* and showed that insertions, deletions and substitutions for a given gene target could be made. During the writing of this PhD thesis, no other CRISPR/Cas9 study on orchids was published.

### Ancestral character state analysis of orchid fruits

In **chapter 5**, character evolutionary analysis of general orchid fruit traits was carried out to find possible patterns of co-evolution. In **chapter 4**, I showed that *E. pusilla* fruits dehisce without lignification at or near the dehiscence zone. We found evidence for the formation of a cuticle-like layer in the fruits of *E. pusilla*, using TEM and micro-CT scanning, causing the fruit to dehisce. Such a layer is not present in the fruits of two other orchid species, *Epipactis helleborine* and *Cynorkis fastigiata,* where lignification does play a role in fruit dehiscense. When examining dehiscence zone development in these fruits, no cuticle-like layer was observed. Examining more fruits from orchid species from all subfamilies for lignification patterns and combining this information with other characteristics such as ripening time, orientation dehiscence type, and number of slits revealed that an epiphytic or lianaceous habit, longer fruit ripening period and smaller number of opening slits clearly co-evolved in orchids. Similarly, pendant orchid fruits co-evolved with a preference for growing at intermediate to high temperatures and an epiphytic or lianaceous habit. All the methods discussed above now enable addressing fundamental evolutionary questions for any orchid species within a relatively short timeframe. An integrated toolbox can now be used for tracing character evolution to unravel the full genetic basis of the highly specialized organs that make orchids such fascinating subjects for evolutionary studies.

## References

Belser, C., Istace, B., Denis, E., Dubarry, M., Baurens, F.C., Falentin, C., Genete, M., Berrabah, W., Chevre, A.M., Delourme, R., Deniot, G., Denoeud, F., Duffe, P., Engelen, S., Lemainque, A., Manzanares-Dauleux, M., Martin, G., Morice, J., Noel, B., Vekemans, X., D'hont, A., Rousseau-Gueutin, M., Barbe, V., Cruaud, C., Wincker, P., and Aury, J.M. (2018). Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat Plants* 4**,** 879-887.

Cai, J., Liu, X., Vanneste, K., Proost, S., Tsai, W.-C., Liu, K.-W., Chen, L.-J., He, Y., Xu, Q., Bian, C., Zheng, Z., Sun, F., Liu, W., Hsiao, Y.-Y., Pan, Z.-J., Hsu, C.-C., Yang, Y.-P., Hsu, Y.-C., Chuang, Y.-C., Dievart, A., Dufayard, J.-F., Xu, X., Wang, J.-Y., Wang, J., Xiao, X.-J., Zhao, X.-M., Du, R., Zhang, G.-Q., Wang, M., Su, Y.-Y., Xie, G.-C., Liu, G.-H., Li, L.-Q., Huang, L.-Q., Luo, Y.-B., Chen, H.-H., De Peer, Y.V., and Liu, Z.-J. (2015). The genome sequence of the orchid Phalaenopsis equestris. *Nature Genetics* 47**,** 65-72.

Cerna, M., Valdivieso, P., Cella, R., Matyas, B., and Aucapina, C. (2018). Cryopreservation of orchid seeds through rapid and step freezing methods. *F1000Res* 7**,** 209.

Chao, Y.T., Chen, W.C., Chen, C.Y., Ho, H.Y., Yeh, C.H., Kuo, Y.T., Su, C.L., Yen, S.H., Hsueh, H.Y., Yeh, J.H., Hsu, H.L., Tsai, Y.H., Kuo, T.Y., Chang, S.B., Chen, K.Y., and Shih, M.C. (2018). Chromosome-level assembly, genetic and physical mapping of Phalaenopsis aphrodite genome provides new insights into species adaptation and resources for orchid breeding. *Plant Biotechnol J* 16**,** 2027-2041.

Chao, Y.T., Yen, S.H., Yeh, J.H., Chen, W.C., and Shih, M.C. (2017). Orchidstra 2.0-A Transcriptomics Resource for the Orchid Family. *Plant Cell Physiol* 58**,** e9.

Deschamps, S., Zhang, Y., Llaca, V., Ye, L., Sanyal, A., King, M., May, G., and Lin, H. (2018). A chromosome-scale assembly of the sorghum genome using nanopore sequencing and optical mapping.

*Nat Commun* 9**,** 4844.

Gotz, S., Garcia-Gomez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talon, M., Dopazo, J., and Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36**,** 3420-3435.

Gravendeel, B., and Dirks-Mulder, A. (2015). Floral development: Lip formation in orchids unravelled. *Nat Plants* 1**,** 15056.

Hsieh, M.H., Lu, H.C., Pan, Z.J., Yeh, H.H., Wang, S.S., Chen, W.H., and Chen, H.H. (2013). Optimizing virus-induced gene silencing efficiency with Cymbidium mosaic virus in Phalaenopsis flower. *Plant Sci* 201-202**,** 25-41.

Hsing, H.X., Lin, Y.J., Tong, C.G., Li, M.J., Chen, Y.J., and Ko, S.S. (2016). Efficient and heritable transformation of Phalaenopsis orchids. *Bot Stud* 57**,** 30.

Hsu, H.-F., Hsu, W.-H., Lee, Y.-I., Mao, W.-T., Yang, J.-Y., Li, J.-Y., and Yang, C.-H. (2015). Model for perianth formation in orchids. *Nature Plants* 1**,** 15046.

Hu, Y., Resende, M.F.R., Jr., Bombarely, A., Brym, M., Bassil, E., and Chambers, A.H. (2019). Genomics-based diversity analysis of Vanilla species using a Vanilla planifolia draft genome and Genotyping-By-Sequencing. *Sci Rep* 9**,** 3416.

Kendon, J.P., Rajaovelona, L., Sandford, H., Fang, R., Bell, J., and Sarasan, V. (2017). Collecting near mature and immature orchid seeds for ex situ conservation: 'in vitro collecting' as a case study. *Bot Stud* 58**,** 34.

Kui, L., Chen, H., Zhang, W., He, S., Xiong, Z., Zhang, Y., Yan, L., Zhong, C., He, F., Chen, J., Zeng, P., Zhang, G., Yang, S., Dong, Y., Wang, W., and Cai, J. (2017). Building a Genetic Manipulation Tool Box for Orchid Biology: Identification of Constitutive Promoters and Application of CRISPR/Cas9 in the Orchid, Dendrobium officinale. *Frontiers in Plant Science* 7.

Lau, S.E., Mazumdar, P., Hee, T.W., Song, A.L.A., Othman, R.Y., and Harikrishna, J.A. (2015a). Crude extracts of bacterially-expressed dsRNA protect orchid plants againstCymbidium mosaic virusduring transplantation fromin vitroculture. *The Journal of Horticultural Science and Biotechnology* 89**,** 569-576.

Lau, S.E., Schwarzacher, T., Othman, R.Y., and Harikrishna, J.A. (2015b). dsRNA silencing of an R2R3-MYB transcription factor affects flower cell shape in a Dendrobium hybrid. *BMC Plant Biol* 15**,** 194.

Lee, S.-H., Li, C.-W., Liau, C.-H., Chang, P.-Y., Liao, L.-J., Lin, C.-S., and Chan, M.-T. (2015). Establishment of an Agrobacterium-mediated genetic transformation procedure for the experimental model orchid Erycina pusilla. *Plant Cell Tiss Organ Cult (PCTOC)* 120**,** 211-220.

Lu, H.C., Hsieh, M.H., Chen, C.E., Chen, H.H., Wang, H.I., and Yeh, H.H. (2012). A high-throughput virus-induced gene-silencing vector for screening transcription factors in virus-induced plant defense response in orchid. *Mol Plant Microbe Interact* 25**,** 738-746.

Magrini, S., De Vitis, M., Torelli, D., Santi, L., and Zucconi, L. (2018). Seed banking of terrestrial orchids: evaluation of seed quality in Anacamptis following 4-year dry storage. *Plant Biol (Stuttg)* 0.

Mehrotra, S., and Goyal, V. (2014). Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. *Genomics Proteomics Bioinformatics* 12**,** 164-171.

Popova, E., Kim, H.H., Saxena, P.K., Engelmann, F., and Pritchard, H.W. (2016). Frozen beauty: The cryobiotechnology of orchid diversity. *Biotechnol Adv* 34**,** 380-403.

Schofield, E., Jones, E.P., and Sarasan, V. (2018). Cryopreservation without vitrification suitable for large scale cryopreservation of orchid seeds. *Bot Stud* 59**,** 13.

Su, C.L., Chao, Y.T., Yen, S.H., Chen, C.Y., Chen, W.C., Chang, Y.C., and Shih, M.C. (2013). Orchidstra: an integrated orchid functional genomics database. *Plant Cell Physiol* 54**,** e11.

Yan, L., Wang, X., Liu, H., Tian, Y., Lian, J., Yang, R., Hao, S., Wang, X., Yang, S., Li, Q., Qi, S., Kui, L., Okpekum, M., Ma, X., Zhang, J., Ding, Z., Zhang, G., Wang, W., Dong, Y., and Sheng, J. (2015). The Genome of Dendrobium officinale Illuminates the Biology of the Important Traditional Chinese Orchid Herb. *Mol Plant* 8**,** 922-934.

Yuan, Y., Jin, X., Liu, J., Zhao, X., Zhou, J., Wang, X., Wang, D., Lai, C., Xu, W., Huang, J., Zha, L., Liu, D., Ma, X., Wang, L., Zhou, M., Jiang, Z., Meng, H., Peng, H., Liang, Y., Li, R., Jiang, C., Zhao, Y.,

Nan, T., Jin, Y., Zhan, Z., Yang, J., Jiang, W., and Huang, L. (2018). The Gastrodia elata genome provides insights into plant adaptation to heterotrophy. *Nature communications* 9**,** 1615-1615.

Zhang, G.-Q., Xu, Q., Bian, C., Tsai, W.-C., Yeh, C.-M., Liu, K.-W., Yoshida, K., Zhang, L.-S., Chang, S.-B., Chen, F., Shi, Y., Su, Y.-Y., Zhang, Y.-Q., Chen, L.-J., Yin, Y., Lin, M., Huang, H., Deng, H., Wang, Z.-W., Zhu, S.-L., Zhao, X., Deng, C., Niu, S.-C., Huang, J., Wang, M., Liu, G.-H., Yang, H.-J., Xiao, X.-J., Hsiao, Y.-Y., Wu, W.-L., Chen, Y.-Y., Mitsuda, N., Ohme-Takagi, M., Luo, Y.-B., Van De Peer, Y., and Liu, Z.-J. (2016). The Dendrobium catenatum Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. *Scientific Reports* 6**,** 19029.

Zhang, G.Q., Liu, K.W., Li, Z., Lohaus, R., Hsiao, Y.Y., Niu, S.C., Wang, J.Y., Lin, Y.C., Xu, Q., Chen, L.J., Yoshida, K., Fujiwara, S., Wang, Z.W., Zhang, Y.Q., Mitsuda, N., Wang, M., Liu, G.H., Pecoraro, L., Huang, H.X., Xiao, X.J., Lin, M., Wu, X.Y., Wu, W.L., Chen, Y.Y., Chang, S.B., Sakamoto, S., Ohme-Takagi, M., Yagi, M., Zeng, S.J., Shen, C.Y., Yeh, C.M., Luo, Y.B., Tsai, W.C., Van De Peer, Y., and Liu, Z.J. (2017). The Apostasia genome and the evolution of orchids. *Nature* 549**,** 379-383.

# APPENDIX

**Nederlandse sammenvatting**

**Curriculum Vitae**

**List of publications**

**Dankwoord**

**Glossary and abbreviations**

# Nederlandse samenvatting

**Inleiding**

De orchideeënfamilie is de op één na grootste familie van bloeiende planten, met ongeveer 880 geslachten en meer dan 28.000 soorten. De familie komt bijna overal op de wereld voor, maar de meeste soorten zijn te vinden in de tropen. Orchideeën behoren tot de groep van de eenzaadlobbige, ook wel monocotylen of kortweg monocots genoemd. De andere groep bloeiende planten zijn de tweezaadlobbigen (eudicotylen). Een bekend onderzoeksmodel uit deze groep is de zandraket (*Arabidopsis thaliana*). Een orchideeënbloem is opgebouwd uit verschillende bloemkransen. Van buiten naar binnen zijn dit drie kroonbladen, gevolgd door drie kelkbladen. Samen omsluiten zij de vrouwelijke en mannelijke voortplantingsorganen. Eén kroonblad van een orchideeënbloem, de lip, is meestal anders van grootte, maar ook van kleur en vorm dan de andere twee kroonbladen. Verder bevatten de helmhokjes van veel orchideeën geen los stuifmeel, maar stuifmeelklompjes en vaak is er maar één fertiele meeldraad in plaats van drie of zes. Buiten deze gedeelde kenmerken is de diversiteit van orchideeënbloemen enorm, denk o.a. aan kleur, grootte en de vorm van de bloemblaadjes. Orchideeënbloemen zijn vooral zo divers vanwege de interactie met bestuivers.

**Bloemonderzoek aan *Erycina pusilla***

Om de enorme diversiteit aan orchideeënbloemen en -vruchten beter te begrijpen, heb ik onderzoek gedaan naar genen die betrokken zijn bij de ontwikkeling van een selectie aan bloemorganen. Door expressie van MADS-box genen te bestuderen binnen verschillende organen en door te kijken naar de uiterlijke en innerlijke kenmerken, kon de evolutionaire oorsprong opgehelderd worden. Als modelorganisme heb ik de orchideeënsoort *E. pusilla* gebruikt. *Erycina pusilla* komt in het wild in Centraal en Midden-Amerika voor. De lijn waar ik mee werkte komt uit Suriname en is al meer dan 20 jaar in kweek in Nederland. De reden voor het gebruik van *E. pusilla* is dat het een kleine en snelgroeiende orchidee is, welke steriel gekweekt kan worden in een laboratorium. De soort heeft ook een klein en diploïd genoom. Dit laatste is handig bij het volledige in kaart brengen van het genoom. Op het moment dat ik met mijn onderzoek startte was al wel een transcriptoom beschikbaar van *E. pusilla*, maar nog geen genoom.

Binnen de orchideeën zijn veel soorten bedriegers, waarmee wordt bedoeld dat zij bestuivers aantrekken, maar geen beloning aanbieden in de vorm van bijvoorbeeld nectar of olie. Door geen beloning te produceren bespaart de plant veel energie. Hoe dit zo is geëvolueerd, is niet duidelijk. *Erycina pusilla* is ook zo'n bedrieger; de bloemen hebben een vergroot middelste kelkblad, een callus bovenop de lip en twee vleugelachtige aanhangsels, ook wel stelidia genoemd, om bijen aan te trekken. Deze bijen zijn op zoek naar olie, maar vinden dit bij *E. pusilla* niet. Toch nemen ze bij hun bezoek tijdens de zoektocht naar een beloning stuifmeelklompjes mee naar een andere *E. pusilla* bloem. De bloemen zijn zo gevormd dat zij gelijkenis

vertonen met de bloemen van bomen uit de Malpighiaceae famlie, die hun bestuivers wel belonen. Orchideeën en Malpighiaceae zijn niet verwant: de eerste familie behoort tot de monocotylen, de tweede tot de dicotylen.

In **hoofdstuk 2** wordt onderzoek beschreven naar de evolutionaire oorsprong van de lip van de bloemen van *E. pusilla* en andere orchideeënsoorten. Bij de vorming van de lip van *E. pusilla* blijkt het lip-complex actief. Dit is consistent met het "Perianth Code model" voor lipidentiteit van orchideeën. Het lip-complex bestaat in *E. pusilla* uit de volgende drie MADS-box genen: *AGL6 EpMADS5*, *APETALA3 EpMADS13* en *PISTILLATA EpMADS16*.

In **hoofdstuk 3** onderzochten wij de evolutionaire oorsprong van het middelste petaloïde bloemblad van de buitenste krans, het callus en de stelidia van *E. pusilla*. We gebruikten daarvoor een combinatie van morfologische, moleculaire en fylogenetische technieken. Door de vaatbundels in een volwassen bloem te volgen, maar ook de buitenste cellaag te bestuderen en genexpressie patronen te vergelijken van alle bloemorganen, kwamen we tot een aantal belangrijke nieuwe inzichten. De vaatbundel naar het mediane bloemblad van de buitenste krans gaf aan dat het een kelkblad was met kenmerken die typisch zijn voor kroonbladen, zoals bijvoorbeeld de bolvormige epidermale cellen. De expressie van twee MADS-box genen (*AGL6 EpMADS4* en *APETALA3 EpMADS14*) was hoog in de groene zijdelingse kelkbladen van *E. pusilla*, maar laag in het mediane kelkblad en de kroonbladen. Mogelijk correleert deze lage genexpressie met het uiterlijk van een kroonblad. Het lip-complex bleek bij *E. pusilla* niet alleen in de lip maar ook in het callus actief. De vaatbundel naar de lip gaf aan dat de lip tot de kroonbladen behoort. Zes vaatbundels gaan in *E. pusilla* naar de callus, de stelidia en naar de fertiele meeldraad. Alle drie de AGAMOUS MADS-box genen komen niet tot expressie in de callus, wat consistent is met de steriliteit van dit bloemonderdeel.



MADS-box genen spelen mogelijk ook een belangrijke rol in organen in de binnenste bloemkransen die bij de voorouders van *E. pusilla* nog fertiel waren, maar in de loop van de evolutie steriel zijn geworden en uiteindelijk tot gespecialiseerde bloemorganen zijn geëvolueerd als callus en stelidia. *AGAMOUS EpMADS22* en *SEPALLATA EpMADS6* komen het sterkst tot expressie in de meeldraden en *AGAMOUS EpMADS20* en *SEEDSTICK EpMADS23* komen sterk tot expressie in de stelidia. Dit laatste suggereert dat *EpMADS22* mogelijk nodig is voor de vorming van vruchtbare meeldraden. Het mediane bloemblad van de buitenste krans van *E. pusilla* lijkt te zijn afgeleid van een kelkblad. Het callus was oorspronkelijk een meeldraad, die een bloemblad-identiteit kreeg. En de stelidia zijn afgeleid van steriel geworden meeldraden. Op deze manier bootst *E. pusilla* een niet-verwante en belonende bloem na. Duplicaties, maar ook selectie en veranderingen in expressie van verschillende MADS-box-genen speelden een belangrijke rol in de evolutie en ontwikkeling van deze bloemorganen. Dat niet alle vragen met betrekking tot bedrog van bestuivers met dit onderzoek zijn opgelost, moge duidelijk zijn. Voor *E. pusilla* zijn we echter weer een stapje verder gekomen in het begrijpen hoe deze bloemen zo zijn gevormd dat een optimale bestuiving mogelijk is, zonder een

beloning te produceren en dus energie te besparen.

**Vruchtonderzoek aan *E. pusilla*, *Cynorkis fastigiata* en *Epipactis helleborine***

Na bestuiving van een orchideeënbloem en bevruchting van de zaadknoppen verandert het onderstandig vruchtbeginsel na enige weken tot maanden in een vrucht met zaden. De vruchten beschermen de kiemende zaden tegen vraatzuchtige insecten, schimmels en Uv-straling, en zorgen er uiteindelijk ook voor dat de zaden op een efficiënte manier worden verspreid. Wanneer de vrucht rijp is, splitst deze open en worden de zaden vervolgens verspreid, meestal via de wind. Er zijn ook orchideeënvruchten die niet opensplitsen, deze worden opgegeten door dieren. De zaden worden dan op een andere locatie verspreid via de ontlasting van de verspreider. De morfologische verscheidenheid van vruchten en mechanismen van openspringen is groot binnen de orchideeënfamilie. Hoe verschillende vruchten zich ontwikkelen en welke moleculaire netwerken hierachter schuilgaan, is nog grotendeels onbekend. Het meeste onderzoek is gedaan aan de zandraket: deze plantensoort wordt al vele jaren als genetisch model gebruikt in onderzoek aan vruchten.

In **hoofdstuk 4** onderzochten wij, opnieuw met behulp van een combinatie van morfologische, moleculaire en fylogenetische technieken, de evolutie en ontwikkeling van de morfologie van orchideeënvruchten. Ook het mechanisme van opensplitsen van vruchten is door ons onderzocht om meer inzicht te krijgen in het moleculaire netwerk dat ten grondslag ligt aan de ontwikkeling van orchideeënvruchten. Wij hebben voor dit onderzoek de vruchtontwikkeling van *E. pusilla*, een epifytische orchidee met luchtwortels, vergeleken met twee soorten die wortelen in de grond, *C. fastigiata* en *E. helleborine*. Ons onderzoek leverde aanvullend bewijs voor het “Split-carpel model”. Volgens dit model bevatten de vruchten van de meeste orchideeënsoorten drie fertiele en drie steriele valven. Het opensplitsen vindt plaats aan beide zijden van de placenta’s. Tevens werden interessante verschillen waargenomen in de cellen waarmee de vruchten uiteindelijk opensplitsen. Terwijl *C. fastigiata* en *E. helleborine* daar een gelignificeerde laag ontwikkelden, deed *E. pusilla* dit niet. In de vruchten van deze laatste soort werd een lipide-achtige cellaag ontdekt, waarlangs de rijpe vrucht openscheurt.

Homologe genen van *E. pusilla* met mogelijk dezelfde functie in vruchtontwikkeling als de zandraket zijn door ons onderzocht. Het betrof vrucht-geassocieerde transcriptiefactoren, waaronder *HEC3*, *RPL* en *SPT*, en verschillende MADS-box genen. Ontdekt werd dat deze genen binnen het genetische netwerk van vruchtontwikkeling en opensplitsen goed geconserveerd lijken te zijn. Eiwit-eiwit interactiestudies lieten zien dat complexen van MADS-box eiwitten, bestaande uit FUL-, SEP- en AG-homologen, ook kunnen worden gevormd in de vruchten van *E. pusilla*. De expressie van *HEC3*, *RPL* en *SPT* in vruchten van *E. pusilla* was grotendeels hetzelfde als de expressie in de vruchten van de zandraket. Onze genexpressie-studie bracht echter ook verschillen aan het licht. Deze verschillen verklaren mogelijk het verschil in morfologie tussen de hauwtjes van de zandraket

en droge doosvruchten van de door ons onderzochte orchideeënsoorten.

**Vergelijkend transcriptoom- en kenmerkonderzoek van orchideeënvruchten**

De orchideeënfamilie staat vooral bekend om haar enorme bloemdiversiteit. Orchideeënvruchten zijn echter ook zeer divers maar veel minder goed bestudeerd. Een van de eerste doelen in het onderzoek beschreven in **hoofdstuk 5** was om genen en mogelijke netwerken van genen te detecteren die betrokken zijn bij de vruchtontwikkeling van *E. pusilla*. Dit onderzoek was gericht op een veel bredere set genen dan de transcriptiefactoren besproken in het vorige hoofdstuk. Voor het onderzoek in **hoofdstuk 5** is RNA geïsoleerd van *E. pusilla* vruchten uit verschillende ontwikkelingsstadia. Vervolgens is de genetische code van RNA-strengen afgelezen met een IlluminaHiseq sequenser. Omdat er geen referentiegenoom van *E. pusilla* beschikbaar was, is met alle verkregen data een volledig vruchten-transcriptoom gegenereerd. Hiervoor werd een bioinformatische pijplijn ontwikkeld om automatisch orchideeëntranscriptomen *de novo* te assembleren, annoteren en kwantificeren. Al snel bleek dat het *de novo* assembleren te veel ongewenste iso-vormen van genen opleverde. Tevens bleek NCBI Genbank te weinig referentiesequenties van orchideeënvruchten te bevatten voor een juiste annotatie.



We hebben vervolgens een referentiegenoom van een andere orchidee gebruikt en de *E. pusilla* transcriptoomdata daarmee vergeleken. Hiervoor is het referentiegenoom van *Phalaenopsis equestris* gekozen alsmede de annotatie van *P. equestris.* Deze aanpak leverde betere resultaten op dan de eerder uitgevoerde de novo assemblage. Voorlopige resultaten tonen aan dat genen die betrokken zijn bij verschillende processen tijdens de ontwikkeling van vruchten en zaden van *E. pusilla*, zoals pollenbuisvorming, zaadontwikkeling en lignificatie, differentieel tot expressie komen in de tijd. Momenteel worden er aanvullende transcriptomen van vruchten van orchideeënsoorten gegenereerd met een andere wijze van opensplitsen. Vergelijkende transcriptoom-analyses zullen vervolgens helpen onthullen welke ontwikkelingsgenen de morfologische diversiteit van orchideevruchten aansturen.
In **hoofdstuk 5** werd ook onderzocht of er een correlatie gevonden kon worden in de evolutie van 12 binair gescoorde eigenschappen van orchideeënvruchten. Bij dit onderzoek is gebruik gemaakt van een gecombineerde *nrITS*, *matK* en *rbcL* fylogenie. Van 41 verschillende orchideeënsoorten uit vijf subfamilies werden rijpe vruchtjes verzameld, geanalyseerd, met de hand gesneden en gekleurd met floroglucinol om lignificatie van de vrucht tijdens het rijpingsproces zichtbaar te maken. Met behulp van Reversible-jump Markov chain Monte Carlo (RJ MCMC) simulaties werd vervolgens onderzocht of er sprake was van een mogelijke gezamenlijke evolutie van vruchtkenmerken.

Er bleek binnen orchideeënvruchten veel variatie te zijn in het type (bes, droge of vlezige doosvrucht), de oriëntatie (rechtopstaand of hangend), rijpingstijd (meer of minder dan drie maanden), het aantal openingen tussen de valven (slechts één, twee, drie of meer) en de lignificatiepatronen. Deze laatste varieerden van

helemaal geen lignificatie tot verhouting van het exocarp, endocarp, en/of de cellaag waarlangs orchideeënvruchten opensplitsen.

Met behulp van statistische analyses is er bewijs van co-evolutie gevonden tussen de leefwijze van orchideeën en de rijpingstijd van de vruchten. Vruchten van epifytische orchideeën en lianen blijken een significant langere rijpingstijd te hebben, 4 maanden of meer, dan vruchten van aardorchideeën. Er bleek ook co-evolutie te hebben plaatsgevonden tussen de leefwijze en het aantal openingen van de vruchten. Vruchten van epifytische orchideeën en lianen splitsen op minder plaatsen open dan de vruchten van aardorchideeën. Verder bleek de oriëntatie van de vruchten te correleren met de omgevingstemperatuur en leefwijze. Orchideeënsoorten met hangende vruchten komen vaker voor bij gemiddelde tot hoge temperaturen dan bij lage temperaturen en meer bij epifyten en lianen dan bij aardorchideeën. Ten slotte blijkt lignificatie van de valven te zijn geëvolueerd in zowel hangende als rechtopstaande orchideeënvruchten. Het is nog niet zeker of deze correlaties blijven bestaan als meer soorten aan dit onderzoek worden toegevoegd. Voorlopig wordt daarom ingezet op een uitbreiding van de dataset. Deze uitbreiding is vooral gericht op geslachten die zowel epifytische als bodem-bewonende soorten bevatten, zoals bijvoorbeeld *Cymbidium* en *Malaxis*, en op orchideeën waarvan de vruchten niet opensplitsen, zoals bij het ondergronds bloeiende Australische geslacht *Rhizanthella*. Uiteindelijk hopen we daarmee net zoveel te weten te komen over de evolutie en ontwikkeling van orchideeënvruchten als -bloemen.

# Curriculum Vitae

Anita Mulder was born in The Hague, the Netherlands, on November 10th, 1965. In 1983 she graduated from the Veurs College in Leidschendam and started her studies on Biotechnological Research at the Van Leeuwenhoek Instituut Delft. As part of this education she performed an internship in the group of Prof. dr. M. van der Ploeg, under supervision of Prof. dr. A.K. Raap, at the Laboratory of Cytochemistry and Cytometry (Leiden University). After receiving her Bachelor of Applied Sciences degree in 1987 she started as a research technician in the same group where she worked on developing a fast and sensitive method to detect cytomegalovirus infections in blood of transplantation patients.

From 1990 she worked as a research technician at the Department of Biochemistry & Molecular Biology at the Free University of Amsterdam in the group of Prof. dr. H.A. Raué on ribosome assembly in the yeast Saccharomyces cerevisiae. In 1994 she moved to the Division of Molecular Biology and Center of Biomedical Genetics at The Netherlands Cancer Institute in Amsterdam under supervision of Prof. dr. P. Borst. Here she participated in research to find out how the parasite *Trypanosoma brucei* survives in the mammalian blood stream by regularly changing its variant surface glycoprotein coat. In 2000 she changed jobs and since then she is working as a lecturer at the Faculty of Science and Technology at the University of Applied Sciences Leiden.



From 2005 onward Anita became a member of the Innovative Molecular Diagnostics chair at the University of Applied Sciences Leiden under supervision of Dr. W.B. van Leeuwen and participated in the following projects: Monitoring of the viral load of SARS-coronaviruses in patients using multiplex real-time PCR. Medical Microbiology, LUMC, in the group of Dr. E.C.J. Claas (2005); Evaluation of the occurrence of aberrant promoter methylation in uveal melanomas. Departments of Dermatology and Ophthalmology, LUMC, in the group of Dr. P.A. van der Velden (2006); Haplotyping of ~100 DNA samples from patients from the AMC with Genotyping by Amplicon Melting Analysis. Department of Human and Clinical Genetics and The Hemoglobinopathies Laboratory, LUMC, in the group of Dr. C.L. Harteveld (2007); Diagnosing tumor cells on basis of DHX8 expression levels. Department of Molecular Cell Biology, LUMC, in the group of Dr. R.W. Dirks (2008/2009); Development of a multiplex Q-PCR to detect *Trichoderma harzianum T22* in soils samples. Commissioned by Koppert Biological Systems, Berkel en Rodenrijs. (2010-2012).

In 2012 she became a researcher in the Biodiversity chair under supervision of Dr. B. Gravendeel and worked on cloning and characterization of orchid MADS-box A and B class genes. In the same year Anita applied for a PhD grant from the Netherlands Organisation for Scientific Research (NWO) for teachers and with this grant she started her PhD in 2013 in the Endless Forms group at Naturalis Biodiversity Center under supervision of Dr. B. Gravendeel and Prof. dr. E.F. Smets, which resulted in the research described in this thesis.

# List of publications

**Dirks-Mulder A**, Ahmed I, Broek uit het M, Krol L, Menger N, Snier J, Winzum van A, Wolf de A, Wout van 't M, Zeegers JJ, Butôt R, Heijungs R, Heuven van BJ, Kruizinga J, Langelaan R, Smets EF, Star W, Bemer M, Gravendeel B (2019). Morphological and molecular characterization of orchid fruit development. *Front. Plant Sci*, 10: 137.

**Dirks-Mulder A**, Butôt R, Schaik van P, Wijnands JW, Berg van den R, Krol L, Doebar S, Kooperen van K, Boer de H, Kramer EM, Smets EF, Vos RA, Vrijdaghs A, Gravendeel B (2017). Exploring the evolutionary origin of floral organs of *Erycina pusilla*, an emerging orchid model system. *BMC Evol. Biol*, 17(1): 89.

Horn IR, Rijn van M, Zwetsloot TJ, Basmagi S, **Dirks-Mulder A**, Leeuwen van WB, Ravensberg WJ, Gravendeel B (2016). Development of a multiplex Q-PCR to detect *Trichoderma harzianum Rifai strain T22* in plant roots. *J. Microbiol. Methods*, 121: 44-49.

Gravendeel B, **Dirks-Mulder A**. (2015). Floral development: Lip formation in orchids unravelled. *Nat. Plants* 1: 15056.

Molecular diagnostic analysis of outbreak scenarios. Morsink MC, Dekter HE, **Dirks-Mulder A**, Leeuwen van WB (2012). Biochem. *Mol. Biol. Educ.*, 40 (2): 112–120.

Maat W, el Filali M, **Dirks-Mulder A**, Luyten GP, Gruis NA, Desjardins L, Boender P, Jager MJ, Velden van der PA (2009). Episodic Src activation in uveal melanoma revealed by kinase activity profiling. *Br. J. Cancer*, 101(2): 312-319.

Maat W, Velden van der PA, Out-Luiting C, Plug M, **Dirks-Mulder A**, Jager MJ, Gruis NA (2007). Epigenetic inactivation of RASSF1a in uveal melanoma. *Invest. Ophthalmol. Vis. Sci.*, 48(2): 486-490.

Toaldo CB, Kieft R, **Dirks-Mulder A**, Sabatini R, Luenen van HG, Borst P (2005). A minor fraction of base J in kinetoplastid nuclear DNA is bound by the J-binding protein 1. *Mol. Biochem. Parasitol.*, 143(1): 111-115.

Cross M, Kieft R, Sabatini R, **Dirks-Mulder A**, Chaves I, Borst P (2002). J-binding protein increases the level and retention of the unusual base J in trypanosome DNA. *Mol. Microbiol*., 46(1): 37-47.

Dooijes D, Chaves I, Kieft R, **Dirks-Mulder A**, Martin W, Borst P (2000). Base J originally found in Kinetoplastida is also a minor constituent of nuclear DNA of *Euglena gracilis*. *Nucleic Acids Res*., 28(16): 3017-3021.

Chaves I, Rudenko G, **Dirks-Mulder A**, Cross M, Borst P (1999). Control of variant surface glycoprotein gene-expression sites in *Trypanosoma brucei*. *EMBO J*., 18: 4846-4855.

Chaves I, Zomerdijk J, **Dirks-Mulder A**, Dirks RW, Raap AK, Borst P (1998). Subnucleair localization of the active variant surface glycoprotein gene expression site in *Trypanosoma brucei*. *Proc. Natl. Acad. Sci. USA*, 95: 12328-12333.

Rudenko G, Chaves I, **Dirks-Mulder A**, Borst P (1998). Selection for activation of a new variant surface glycoprotein gene expression site in *Trypanosoma brucei* can result in deletion of the old one. *Mol. Biochem. Parasitol*., 95: 97-109.

Leeuwen van F, **Dirks-Mulder A**, Dirks R.W, Borst P, Gibson W (1998). The modified DNA base ß-D-glucosyl-hydroxymethyluracil is not found in the tsetse fly stages of *Trypanosoma brucei*. *Mol. Biochem. Parsitol*., 94: 127-130.

Jeeninga RE, Delft van Y, Graaff-Vincent de M, **Dirks-Mulder A**, Venema J, Raué HA (1997). Variable regions V13 and V3 of *Saccharomyces cerevisiae* contain structural features essential for normal biogenesis and stability of 5.8S and 25S rRNA. *RNA*, 5: 476-488.

Rudenko G, McCulloch R, **Dirks-Mulder A**, Borst P (1996). Telomere exchange can be an important mechanism of variant surface glycoprotein gene switching in *Trypanosoma brucei*. *Mol. Biochem. Parasitol*., 80: 65-75.

Rudenko G, Blundell PA, **Dirks-Mulder A**, Kieft R, Borst P (1995). A ribosomal DNA promotor replacing the promotor of a telomeric VSG gene expression site can be efficiently switched on and off in *T. brucei*. *Cell*, 83: 547-553.

Nues van RW, Venema J, Rientjes JMJ, **Dirks-Mulder A**, Raué HA (1995). Processing of eukaryotic pre-rRNA: the role of transcribed spacers. *Biochem. Cell. Biol.*, 73: 789-801

Venema J, **Dirks-Mulder A**, Faber AW, Raué HA (1995). Development and application of an in vivo system to study yeast ribosomal RNA biogenesis and function. *Yeast*, 11: 145-156.



Abraham PR, **Mulder A**, Riet van 't J, Raué HA (1994). Characterization of the *Saccharomyces cerevisiae* nuclear gene CYB3 encoding a cytochrome b polypeptide of respiratory complex II. *Mol. Gen. Genet*., 242: 708-716.

Abraham PR, **Mulder A**, Riet van 't J, Planta RJ, Raué HA (1992). Molecular cloning and physical analysis of an 8.2kb segment of chromosome XI of Saccharomyces cerevisiae reveals five tightly linked genes. *Yeast*, 8: 277-238.

Kooi EA, Rutgers CA, **Mulder A**, Riet van 't J, Venema J, Raué HA (1993). The phylogenetically conserved doublet tertiary interaction in domain III of the large subunit rRNA is crucial for ribosomal protein binding. *Proc. Natl. Acad. Sci. USA*, 90: 213-216.

Jiwa NM, Rijke van de FM, **Mulder A**, Bij van der W, The TH, Rothbarth PH, Velzing J, Ploeg van der M, Raap AK (1989). An improved immunocytochemical method for the detection of human cytomegalovirus antigens in peripheral blood leucocytes. *Histochemistry*, 91: 345-349.

Jiwa NM, Gemert van GW, Raap AK, Rijke van de FM, **Mulder A**, Lens PF, Salimans MMM, Zwaan FE, Dorp van W, Ploeg van der M. (1989). Rapid detection of human cytomegalovirus DNA in peripheral blood leucocytes of viremic transplant recipients by the polymerase chain reaction. *Transplantation*, 48: 72-76.

Jiwa NM, Raap AK, Rijke van de FM, **Mulder A**, Weening JJ, Zwaan FE, The TH, Ploeg van der M. (1989). Detection of cytomegalovirus antigens and DNA in tissues fixed in formaldehyde. *J. Clin. Pathology*, 42(7): 749-754.

# Dankwoord

Waarom wil je als docent eigenlijk een promotieonderzoek doen? Deze vraag heb ik vaak moeten beantwoorden en doe het hier graag nog een keer. Voor mij is de doelstelling van mijn promotie het onderwijs. Ik vind het belangrijk dat wat ik gedaan heb, overgedragen kan worden naar het onderwijs op Hogeschool Leiden. Mede dankzij de steun vanuit onderwijs heb ik ook dit onderzoek kunnen doen en afronden. Dank aan John van der Willik, in 2013 als cluster directeur, en later aan Patrick Pijnenburg voor jullie steun en geloof in dit project. Dank ook aan Danny Dukers en Gabrielle Pinkse voor de steun support en de mogelijkheid om extra promotietijd voor mij in te roosteren. Ik had deze tijd inderdaad heel hard nodig en heb deze zo nuttig en efficiënt mogelijk besteed.

Promoveren doe je nooit alleen en in dit project heb ik dan ook, met heel veel plezier met veel studenten van de Hogeschool Leiden samengewerkt, waarbij zij stuk voor stuk allemaal hun steentje hebben bijgedragen. Stef Janson, de allereerste student die zich bij mij aanmeldde en meteen de eerste RNA master werd, met de truck van het glazen kogeltje; Kelly van Kooperen, een pionier was je, alles zelf opzetten met mooie eerste resultaten, een prima basis heb je toentertijd gelegd; Sadhana Doebar, de snelst en meest efficiënte histo-analist ooit; Peter van Schaik, respect om ‘‘even’’ 2 uur lang geconcentreerd een 384 plaat vol te pipetteren; Roel van den Berg, SEM, TEM, µCT, Aviso (mijn nachtmerrie programma) en nog veel meer, niets was te lastig voor jou, dank voor je bijdrage en fijne samenwerking ook toen je, al afgestudeerd, gewoon nog even de scan hielp verbeteren; Vinod Shankar, een enthousiaste doorzetter; Louie Krol, kwam, zag en overwon, mooie tijd ook, dank voor de hulp na afloop; Soukayna Bojaada, jij kwam als eerste vanuit het MLO het team versterken in de orchideeënkweek en hebt dit super gedaan; Jan Willem Wijnands, wat een gezellige/slimme bio-informaticus ben jij en wat heb je hard gewerkt aan die fylogenetische bomen, ook jij kwam nog even helpen een jaar later, je blijft een toffe gast; Sanne Nottelman, ook jij kwam vanuit het MLO de kweek onderhouden en deed dit prima; Martijn van ’t Wout, zo rustig en voortvarend doorwerken als jij, echt heel fijn; Jamie Zeegers, heel rustig ging je te werk en in de tussentijd rolde de resultaten maar binnen en liep alles heerlijk vlotjes; Israa Ahmed, de verbindende factor in het team van studenten van jaar 17-18, ik heb genoten; Jasmijn Snier, waar Jamie eindigde ging jij enthousiast en stug door (al die metingen van de oppervlakten), blij dat je er bij was in 17-18; Anne van Winzum, “even” een transcriptoompje analyseren, je was een echte pionier voor zowel Naturalis als de Hogeschool, wat een klus en wat ben je ver gekomen, een SucceedArmy project; Nino Menger, buiten het feit dat je de kweek als MLO student onderhield heb je deze ook nog eens geoptimaliseerd en heb je steentje bijgedragen al het onderzoek van Israa en Jasmijn, Mark uit ’t Broek ging verder waar Anne eindigde en heeft vele hobbels en valkuilen in dit proces genomen. Ook dank aan Nemi Dorst, soms lopen projecten wat door elkaar heen, en helpen we elkaar, top.

Ook heb ik mogen samenwerken met een aantal collega’s van zowel HLO, MLO als bio-informatica, onderzoek aan orchideeën heeft veel en uitdagende kanten.

Ken Kraaijeveld, dank voor de korte maar zeer productieve en voor mij leerzame periode; Jan Oliehoek, jij stond aan de basis van de fylogenetische stambomen, een hele klus waaraan Marcel Lombaerts ook aan gewerkt heeft, beide dank voor alle inspanningen en blij dat school nu ook het programma Geneious kan gebruiken en de kennis kan overdragen aan onze studenten. Anneke de Wolf, jij hebt zelf tweemaal een docent stage gedaan en was als een vis in het water (of een orchidee aan een boom) op het lab; Sigrid Beiboer, heeft met de beperkte informatie die er is naar homologe genen gezocht en dit bleek een bijna onmogelijke opdracht, dank voor al je digitale speurwerk en gepuzzel. Databases, genomen en transcriptomen analyseren, zonder de hulp vanuit de bio-informatica is dat gewoon niet te doen voor een simpele moleculair bioloog als ik, waar Ken ooit begon ging Jeroen Pijpe met een groep van enthousiaste studenten verder om deze puzzel van honderd duizenden stukjes in elkaar te zetten, dank voor al het harde werken hieraan. Ook Wouter Suring wil ik bedanken omdat ik door jouw hulp zelf lokaal in een database kon zoeken naar sequenties. Floyd Wittink, voor een bio-informaticus ben je soms digitaal lastig te contacten, maar wat ben ik blij met je bijdrage als stagebegeleider voor Anne en Mark en als inhoudsdeskundige en vraagbaak en heb je ook nog eens Stef Pieterman aan dit project toegevoegd, die ik meteen ook bedank, DE Blast2GO master.

Binnen Naturalis werkt er een team van analisten van wereldklasse, waar zouden wij als onderzoekers zijn zonder hun. Dank Bertie Joan van Heuven, Marcel Eurlings, Elza Duim, Wim Star, Rob Langelaan, Kees van den Berg, Dirk van der Marel en Frank Stokvis voor alle hulp, adviezen, gezelligheid, discussies, borrels; Arjen Speksnijder als manager van dit team; Rutger Vos voor zijn bio-informatica kennis en input; mijn mede PhD'ers Diego Bogarin, Dewi Pramanik, Adam Karremans, Richa Kusumawati en Isolde van Riemsdijk. Mijn speciale dank gaat uit naar Roland Butôt, een alleskunner, niets was te veel of te ingewikkeld, je deed vele projecten tegelijk en had altijd tijd om ook mij te helpen.

Buiten Naturalis wil ik ook een aantal collega's van het LUMC bedanken, dankzij Pieter van der Velden en Mieke Versluis mocht ik samen met een groot aantal studenten gebruik maken van hun 384 en werd de data ook nog eens opgestuurd. Karoly Szuhai heeft ons vele uren geholpen om al die mooie preparaten in te scannen en kwam met allerlei tips en ideeën voor het onderzoek. Dank ook Rogier van Vugt voor alle advies en Jaco Kruizinga voor het leveren van zoveel materiaal vanuit de Hortus botanicus Leiden voor ons onderzoek.

Ik dank ook Johan Keus, mijn steun en toeverlaat in het kweken van *Erycina,* Alexander Vrijdaghs die met grote precisie *Erycina* bloemknopjes heeft ontleed op microniveau en dit zo mooi heeft vastgelegd; Reinout Heijungs voor zijn hulp met de statistiek; Marian Bemer voor het meedenken, mee pipetteren, mee schrijven, echt super bedankt. En waar zou ik zijn geweest zonder de hulp en ondersteuning vanuit IBL? Elke week kon ik aanschuiven bij de werkbespreking in de groep van Remko Offringa en zo nu en dan de mensen verblijden met een verhaal over een orchidee; dit was buiten leerzaam ook heel gezellig, dank allemaal. Ook op het lab

werd ik aan alle kanten geholpen; Ward's kitchen met een muziekje erbij; Gerda Lamers en Marijn Luijten, dank voor alle adviezen en hulp, top analisten.
Vanuit school kreeg ik ook de nodige support, dank jullie wel Irma Lantinga, Erik Huiskens, Maria Plug, Ivo Horn, Karsten Kaspers, Bert Dekker, Hilda van Mourik, Saskia Stahlecker, Rob van Gijlswijk, Maarten Morsink, Wouter van Zon, Nicole Almering en alle andere collega's en oud-collega's van de Hogeschool Leiden voor jullie belangstelling.

Soms moet je ook even stil staan bij iemand die een groot voorbeeld is en dat is Peter ten Wolde voor mij in het onderwijs, iemand die altijd klaar staat, altijd positief is en waar je op kunt bouwen, zijn werk doet en kritisch kijkt naar alle ''snode'' onderwijsplannen, dank voor al die vele jaren dat we hebben mogen samenwerken.

Dat sport en beweging heel belangrijk voor mij is en een deel van mijn leven uitmaakt moet duidelijk zijn, maar sporten doe je samen, wel zo gezellig, dank je wel Vlietlijn hardloopgroep (met daarin ook nog de Zondag lopers, de ZZ-auto pool, Sauna dames, Winkeldames, Borrel071 lopers en de Skiërs) en de EVZV-zwemgroep voor al het sportieve plezier en gezelligheid ook na afloop. Ik dank ook al mijn Feestbeest vrienden, buurtjes, vriendinnen uit Voorschoten en natuurlijk mijn vriendin van de middelbare school Simone.

Kom ik bij het laatste en belangrijkste deel, zonder hun steun en toeverlaat had ik niet kunnen starten en was ik niet zo ver gekomen, mijn mannen Roeland, Mathijs en Reinout. Wat ook scheelt is een familie die elkaar door dik en dun steunt en blij is met de keuzes die je maakt, dus dank aan mijn ouders, schoonouders, Peter, Edith, Victor, Hennie, Liesbeth, Charlotte, Teun, Jorian en Imke.

# Glossary

| | |
|---|---|
| ABCDE model | Model describing how different floral organs are specified during development by five classes of MADS-box genes; A, B, C, D and E. |
| Abaxial | Side or surface of an organ like a petal or organ system such as a branch, facing away from the axis that bears the organ or organ system. |
| Adaxial | Side or surface of an organ like a petal or organ system such as a branch, facing towards the axis that bears the organ or organ system. |
| Anatomy | Science of the internal shape and structure of organisms and their parts. |
| Androecium | The androecium includes all the stamens of a flower and is synonymous with the male reproductive part of a flower. |
| Angiosperms | Flowering plants, producing seeds from ovules contained in ovaries that develop into fruits. |
| Anther | The pollen-bearing part of a stamen. |
| Anther cap | Cover around the pollinia in orchids or pollen in other plant families. |
| Apex | The top of a stem or root. |
| Apical | At the apex. |
| Axis | Fixed line of reference to indicate position. |
| Bud | Any stem meristem - vegetative or floral – in an embryonic stage, whether resting or not and whether protected by bud scales or not. |
| Callus | Enlarged structure at the base of the lip of an orchid flower to attract pollinators. |
| Capsule | Dry fruit derived from a compound ovary and formed by of two or more carpels. |
| CArG-box | 'C-Arich-G-box': a DNA-sequence motif bound by MADS-domain proteins, with the consensus sequence 5'-CC(A/T)6GG-3' or a similar sequence. |
| Carpel (cr) | Leaf homologue that encloses the ovules and seeds in angiosperms and develops into a fruit. |
| Clade | Group derived from a single ancestor. |
| Chloroplast | Cell organelle with two double membranes that originated by endosymbiosis and that is now involved in photosynthesis, containing circular DNA as part of the entire plant genome, a kind of plastid. |
| Column | See Gynostemium. |
| Dehiscence | Opening or splitting of the fruit at maturity. |
| Dehiscence zone | Zone between the sterile and fertile valve where a fruit splits open. |

| | |
|---|---|
| Dicot | Flowering plants that have seeds with two embryonic leaves, two cotyledons. |
| Dorsal | Of a lateral organ, relating to the side facing away from the axis in early development, i.e., the 'back; it is sometimes used to refer rather counter-intuitively to the underside of a leaf blade; synonymous with abaxial. |
| Endocarp | The inner layer of a pericarp. |
| Endosperm | Nutritive tissue in an angiosperm seed, usually triploid and formed after fertilization by the fusion of one gamete with the polar nucleus, sometimes diploid or polypoid. |
| Epiphyte | Growing on another plant, without deriving nutrients from this plant. |
| Exocarp | The outermost layer of a pericarp. |
| Fertile valve (F) | Valve with a placenta region and maturing seeds, located between two sterile valves. |
| Floral Quartet | The floral quartet model describes the interactions of MADS-domain proteins to form complexes that are called tetramers. |
| Floral whorl | A ring of floral organs, all on the same plane. |
| Funiculus | Umbilical cord that nurtures the seeds. |
| Genus | Taxonomic category, ranking above species and below family. |
| Germination | The process by which the embryo resumes growth and escapes from the confines of the seed or fruit and a young sporophyte (seedling) is established. |
| Gynoecium | The gynoecium consists of all the carpels of a flower, whether they are fused or not and is synonymous with the female reproductive part. |
| Gymnosperms | Seed-bearing plants with ovules that are not contained in ovaries and hence develop as 'naked' seeds. |
| Gynostemium | Organ formed by fusion of the androecium and gynoecium in Orchidaceae, synonymous with column. |
| Lignification | Becoming wood-like by the formation of lignin in the cell walls. |
| Lip or Labellum | The median petal of an orchid flower, often used as an attractive landing platform for pollinators. |
| Lateral | Born lateral to the axis of symmetry of the flower relative to the inflorescence axis, that is, lateral to the line joining the flower, the bract subtending it, the inflorescence axis on which it is borne. |
| Macroevolution | Evolution on a scale at or above the level of species. |

| | |
|---|---|
| MADS-box gene | Eukaryotic gene containing a MADS box, which encodes the DNA-binding and nuclear-localization domain of the respective MADS-domain transcription factors. MADS is an acronym for: MCM1, AGAMOUS, DEFICIENS and SRF. |
| Median | The plane through an axis and the axis from which it originates, e.g., in a flower, the axis of symmetry of the flower relative to the inflorescence axis, that is, on the plane joining the flower, the bract subtending it, the inflorescence axis on which it is borne. |
| MIKC | A conserved structure of the type II MADS-domain proteins. MADS-domain protein that exhibits a characteristic domain structure including a DNA-binding MADS (M) domain, an Intervening (I) domain, a keratin-like (K) domain and a C-terminal (C) domain. |
| Microevolution | Evolution on a small scale, within a single population. |
| Mesocarp | Middle layer of pericarp. |
| Monocots | Flowering plants that have seeds with only one embryonic leaf. Monocotyledons have one cotyledon in the embryo. |
| Monophyletic | A group of organisms or genes descended from a common ancestor that is a member of the group, which includes all descendants of that common ancestor. |
| Morphology | The external form and structure of organisms or its parts. |
| LinRegPCR | Program for the analysis of quantitative RT-PCR (qPCR) data resulting from monitoring the PCR reaction with fluorescent dyes. |
| Ontogeny | Development of an organism or organ, including differentiation, transformation and growth. |
| Orthologues | Two or more genes, or gene families that became distinct lineages as the result of a speciation event. |
| Outgroup | A group of organisms, or genes, that is outside the monophyletic group under consideration. |
| Ovaries | Female, ovule-bearing reproductive organs. |
| Paralogues | Homologous genes that originated from an ancestral gene through gene duplication. |
| Parenchymal cells | Fundamental tissue of plants, composed of thin-walled cells, able to divide. |
| Perianth | Sepals and petals of a flower. |
| Petal (pe) | Inner perianth whorl organ, the median is the lip in Orchidaceae. |
| Pericarp | The wall of a ripe fruit, consisting of three layers: endocarp, mesocarp, and epicarp. |

| | |
|---|---|
| Phylogenetics | Study of the evolutionary relationships between species. |
| Pistil | The female reproductive part of a flower. |
| Pollen tubes | A hollow tube that develops from a pollen grain on the stigma after pollination. |
| Pollinia | A cohesive mass of pollen. |
| Pollination | The transfer of pollen from the anther to the stigma of a flower. |
| Protocorm | Specialized structure that develops after orchid seed germination and from which a shoot develops |
| Radiation | Divergence from a central point, in particular evolution from an ancestral animal or plant group into a variety of new forms. |
| Resupination | Change of orientation of the lip during anthesis in which the lip eventually faces downwards. |
| Rostellum | Tissue that separates the anther from the fertile stigma in outcrossing orchids. |
| Sepal (se) | Outer perianth whorl organ. |
| Stamen | The male reproductive organ, it produces pollen. |
| Staminode | A non-fertile, modified stamen. |
| Stelidia | Also known as 'column wings'. Positioned next to the stigma and on both sides of the gynostemium and protecting the stigma during fruit development in *E. pusilla*. |
| Sterile valve (S) | Valve without a placenta region or seeds, located between two fertile valves. |
| Stigma | The part of the gynoecium that receives pollen. |
| Taxonomy | The science of naming, describing and classifying organisms into an hierarchical system. |
| Terrestrial | Growing in or on the ground. |
| Trichome | Any epidermal outgrowth, e.g. a hair. |
| Vascular bundle | Bundle of specialized cells transporting nutrients through the plant. |
| Viscidium | A sticky part on the stipe of a pollinium. |
| Whorl | Three or more leaves or organs (such as sepals and petals) positioned in a circle arising from a single node. |

## Abbreviations

| | |
|---|---|
| bHLH | Basic helix-loop-helix |
| BGI | Beijing Genome Institute |
| cl | callus (on the lip) |
| cr | carpel |
| DAP | Days after pollination |
| DEG | Differential Expressed Gene |
| DZ | dehiscence zone |
| Evo-devo | Evolutionary developmental biology |
| FM-index | Full-text Minute-space index |
| HOT | Homeotic Orchid Tepal |
| fs | fertile stamen |
| gm | gynostemium |
| GO | Gene Ontology |
| L-complex | Lip complex |
| LM | Light Microscopy |
| lse | lateral sepal |
| MA plot | M (log ratio) and A (mean average) |
| micro-CT | X-ray micro-computed tomography |
| ML | Maximum likelihood |
| mya | million years ago |
| mse | median sepal |
| NGS | Next Generation Sequencing |
| NCBI | National Center for Biotechnology Information |
| ORF | Open Reading Frame |
| P-code | Perianth code |
| pe | petal |
| PLB | Protocorm like body |
| PTA | Phosphotungstic acid |
| qPCR | Quantitative PCR |
| RNA-seq | RNA-sequencing |
| RPKM | Reads per Kilobase per Million mapped reads |
| RSEM | RNA-Seq Expectation Maximization |
| RT-qPCR | Real-time quantitative PCR |
| SEM | Scanning Electron Microscopy |
| SEM | Standard Error of Measurement (for qPCR) |
| SP-complex | Sepal/petal-complex |
| S(sl) | stelidium |
| TEM | Transmission Electron Microscopy |
| VB | Vascular bundle |
| WAP | Weeks after pollination |
| Y2H | Yeast two-hybrid |

## Abbreviations of genes used in this study

| | |
|---|---|
| *AG/AGL* | *AGAMOUS/ AGAMOUS-like* |
| *ALC* | *ALCATRAZ* |
| *AP1/AP2/AP3* | *APETALA1 or 2 or 3* |
| *FUL* | *FRUITFULL* |
| *HEC3* | *HECATE3* |
| *IND* | *INDEHISCENT* |
| *PI* | *PISTILLATA* |
| *PNF* | *POUNDFOOLISH* |
| *RPL* | *REPLUMLESS* |
| *SEP* | *SEPALLATA* |
| *SHP1/2* | *SHATTERPROOF1 or 2* |
| *SPT* | *SPATULA* |
| *STK* | *SEEDSTICK* |
| *SVP* | *SHORT VEGETATIVE PHASE* |
| *UBI2* | *UBIQUITIN2* |


Universiteit Leiden