UNIVERSITY OF NAPLES
FEDERICO II

DEPARTMENT OF BIOLOGY

Ph.D. in Biology
XXXV cycle

Coordinator: Prof. Sergio Esposito

“Comparative genomics and transcriptomics to identify key regulators of orchid flower symmetry and organ identity”

Anno accademico 2021-2022

Tutor

Prof. Serena Aceto

Ph.D. student

Francesca Lucibelli

INDEX

ABSTRACT

The Orchidaceae are a monocot angiosperm family with an extraordinary diversification of flower architecture and unique ecological characteristics. In addition to the scientific interest, orchids have a considerable economic value, representing the most commercialized ornamental plants. For these reasons, a complete understanding of the complex molecular network that regulates orchid flower development is an interesting challenge.

The class B MADS-box *AP3/DEF* genes lead to orchid perianth morphogenesis together with other *MADS*-box genes (*AGL6* and *SEP*-like). This work aims to identify novel candidates that may be part of the genetic program at the base of orchid flower symmetry and organ identity determination.

The *in silico* differential expression analysis on RNA-seq data of wild-type and peloric *Phalaenopsis* orchids suggested that the *YABBY DROOPING LEAF*-like gene *PeDL2* could be a novel regulator of the orchid lip specification. In wild-type *Phalaenopsis*, *PeDL2* is differentially expressed in lip and lateral inner tepals, while in the perianth organs of the peloric mutant it is expressed at similar levels. This is the first evidence of the expansion of a *DL*-like gene expression domain to the perianth, since it is generally involved in reproductive organs and leaf development. The regulatory relationship between *DL*-like and class B MADS-box genes of other angiosperms, together with the central role of the MADS-box genes in orchid perianth development, supports the hypothesis that *DL2* can be involved in the regulatory network underlying the orchid flower development. These premises prompted me to focus my Ph.D. research project on the orchid *DL*-like genes.

Genomics and transcriptomics revealed that the orchids with zygomorphic flowers have two *DL* genes, whereas the ancestral orchids with actinomorphic flowers have only one. These observations and the expression pattern of the *Phalaenopsis DL* genes suggest that the *DL* paralogs conserved their function in the development of the reproductive organs; however, after a duplication event, the neofunctionalization of the *DL2* gene could explain its acquisition of a role in orchid lip development and bilateral symmetry determination.

The *PeDL2* expression pattern is similar to that of the *AP3/DEF* genes belonging to clades 3 and 4 and opposite to that of the *AP3/DEF* genes belonging to clades 1 and 2, suggesting a possible regulatory link.

To support the hypothesis of the orchid *DL2* neofunctionalization, I evaluated the expression of the *DL*- and *DEF*-like genes outside the *Phalaenopsis* genus (*Vanilla planifolia*, *Phragmipedium logifolium*, and *Rhyncholaeliocattleya*). The obtained results led to a reformulation of the initial hypothesis, suggesting that *DL2* could have acquired a specific role in lip development only in evolutionarily more recent orchids. In the basal zygomorphic orchids, *DL2* is expressed at similar levels in all the perianth organs. Possibly, after gene duplication, *DL2* acquired a new function in the perianth specification, and later assumed a specific role in the lip.

To characterize PeDL2, I performed intracellular protein localization experiments, confirming its nuclear localization. In addition, I predicted the DLs structure and interactions by computational approach and confirmed the results by yeast two-hybrid analysis.

Finally, I evaluated the *DL2* transcriptional regulation by *in silico* screening of possible promoter interactors. I validated these results through yeast one-hybrid assay and protoplasts dsRNAi experiments. According to these analyses, *DL2* transcription could be modulated by different transcription factors; in particular, the *Phalaenopsis* clade 1 DEF-like protein MADS2 could repress *DL2* expression.

1. INTRODUCTION

1.1 Flower evolution and development

The flower is a unique structure of the angiosperms, representing the main evolutionary innovation of this group of plants that appeared ~140–250 million years ago (Mya) [1-4]. The biological role of the flower is reproduction, and the variability of their coloration and shapes are responsible for the rapid diffusion of the angiosperms, the largest and most diversified group of terrestrial plants [5]. Four organs arranged in concentric rings compose the flower structure: sepals in the external whorl, petals in the second, stamens in the third, and carpels in the fourth [6, 7].

A distinctive feature of floral diversity is symmetry. The flowers can have several planes of symmetry (radial symmetry or actinomorphy) or a single plane of symmetry (bilateral symmetry or zygomorphy). Asymmetrical flowers without planes of symmetry are less frequent (Figure 1) [8].

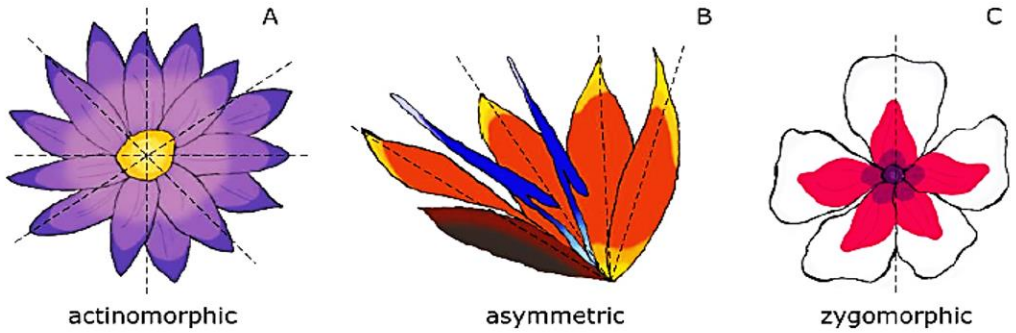


Figure 1. Different types of floral symmetry. The different symmetries are schematized in A (actinomorphy); B (asymmetry); C (zygomorphy). From *Lucibelli et al.* (2020) [9].

During evolution, numerous symmetry transitions have occurred, increasing the angiosperm diversification. Bilateral symmetry has evolved from the ancestral actinomorphic symmetry through different independent events. However, in many species of angiosperms, the reversion of symmetry from bilateral to radial is frequent [8, 10]. Since the flowers' symmetry influences pollination efficiency, an evolutionary correlation is hypothesized between pollinating insects and flower symmetry transitions. Notably, while many types of insects pollinate the radially symmetrical

flowers, specific pollinators that coevolved with the bilaterally symmetric flowers pollinate the zygomorphic flowers [11-13].

The flower morphogenesis and development are orchestrated by the action of specific transcription factors (TFs). Interestingly, during plant evolution, there was an expansion of the plant TF number in relation to the complexity of the plant (Figure 2) [14].

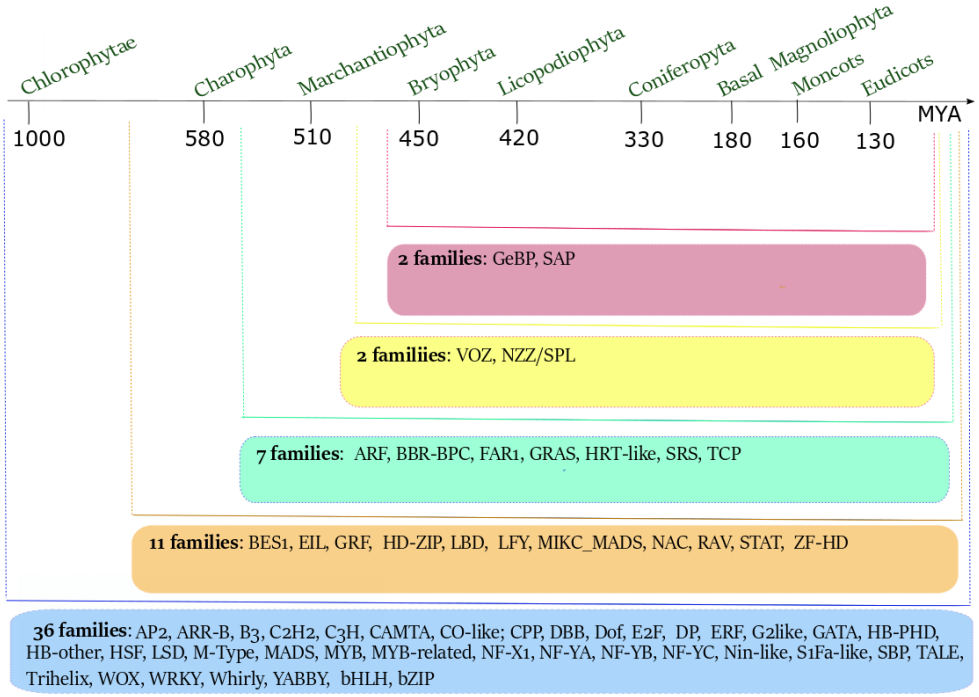


Figure 2. The expansion of transcription factor families during plant evolution. From *Lucibelli et al.* (2020) [9].

TFs regulate different mechanisms of cell activity and are fundamental to control gene expression. Members of different TF families are multifunctional proteins, and their expansion during plant evolution is associated with the acquisition of different roles [15] as result of whole genome duplication (WGD) events that can cause gene loss, sub-functionalization or neo-functionalization [16].

Floral organs differentiation is a complex molecular process based on the co-regulation of many different TFs that act in distinct tissues and stages during development. For example, the MADS-box transcription factors are one of the oldest TF families and, in angiosperms, are involved in the flowering time and in the determination of floral architecture [17-19]. On the other hand, MYBs and TCPs play a pivotal role in the molecular program at the base of flower symmetry establishment [9].

Another transcription factor family that leads flower formation and cotyledon development is the NAC family [20, 21]. Moreover, the YABBY proteins define the abaxial cell fate in flower organs regulating the seed plant vegetative and reproductive development [22].

Understanding the regulatory relationships among the different TFs families during flower development is still a fascinating challenge. Indeed, changes in the TFs biological functions led to the principal evolutionary innovations and adaptations. In particular, alterations in DNA-binding activity, protein-protein interactions, and TFs expression profile drive the evolutionary processes.

The study of non-model organisms, such as orchids, improves the knowledge about evolutionary trajectories through comparative analysis, identifying highly conserved and more recently evolved pathways. Moreover, the knowledge of the key regulators of blooming time, floral scent, coloration, and morphology is a good tool for the flowering plants' molecular breeding and the global floriculture trade [23, 24].

1.2 The Orchidaceae

Among angiosperms, Orchidaceae is one of the largest families including 763 genera and ~28,000 species [25]. The evolutionary success of the orchids is probably due to their unique characteristics in the plant kingdom such as highly specialized pollination strategies, extraordinary adaptability to different types of habitats, diversified floral morphology and zygomorphic flowers [26]. For these reasons, orchids attracted Charles Darwin's interest, as described in his work *Fertilization of Orchids*. Darwin proposed that the beauty and diversification of the orchid flowers are the results of a co-evolution with pollinating insects to attract them, obtaining cross-fertilization. Entomophily (insect pollination) increases the orchid fitness by providing greater variability fundamental for natural selection, evolution and development of new species [27].

A nice example of evolutionary relationship between orchid flowers and pollinator insects is the genus *Ophrys*. The resemblance of the *Ophrys* flower with the female wasps attracts the male insect [13].

Morphological and molecular phylogenetic analyses divide the Orchidaceae into five subfamilies that include many tribes and sub-tribes. The diversification of the subfamilies started ~90 Mya and

the earliest divergence between Orchidoideae and Epidendroideae occurred ~64 Mya (Figure 3) [17, 28-30].

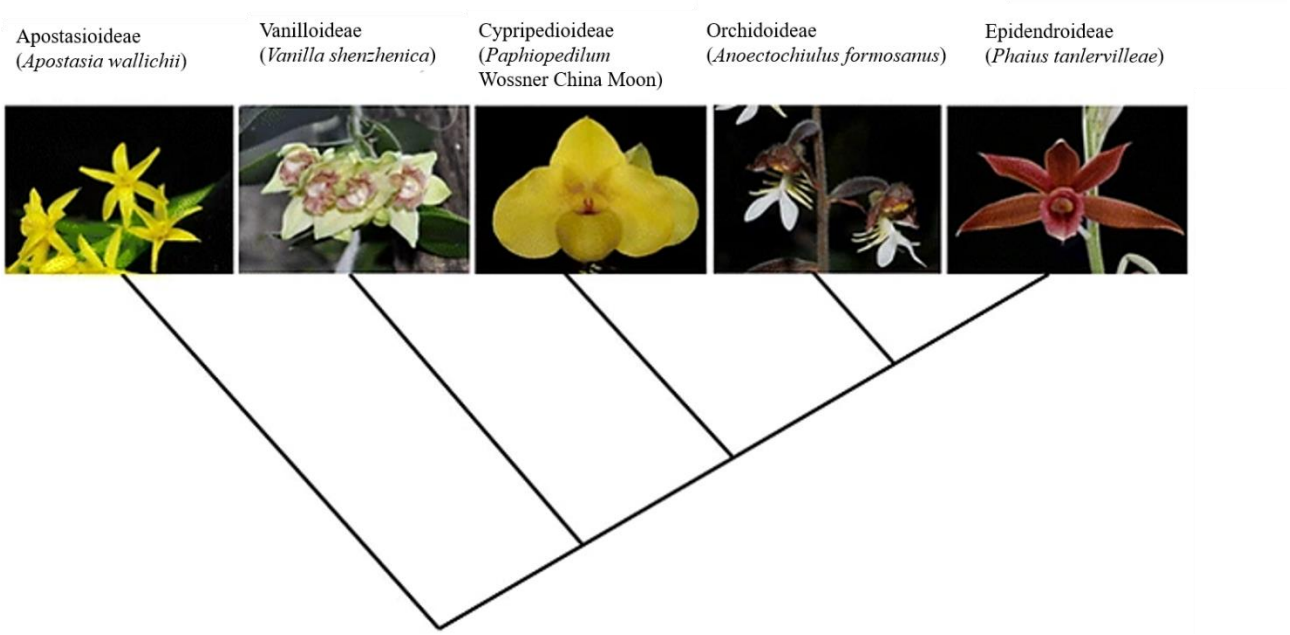


Figure 3. Orchidaceae phylogeny. From *Tsai et al* (2017)[31] .

Despite the high species number and morphological diversification, most orchid flowers share the same organization. In the outermost whorl, there are three outer tepals; the second whorl comprises three petals divided into two inner lateral tepals and a median one called lip or labellum, generally with a distinctive morphology and coloration, different from the other tepals. All orchid flowers have female (gynoecium) and male (androecium) organs fused in the gynostemium or column. In the upper part of the gynostemium there are the pollinia, whereas in the lower part there is the ovary whose maturation is activated after pollination (Figure 4) [17, 32].

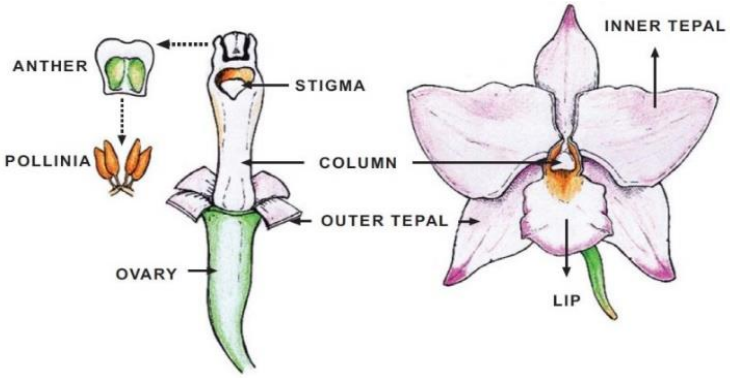


Figure 4. Orchid flower structure. From *Aceto et al.* (2011) [17].

The lip is the uppermost perianth organ. However, during flower development of many orchid species occurs resupination, a 180° rotation of the pedicel or ovary (Figure 5). This rotation shifts the lip to a ventral position in perfect opposition to the gynostemium. Consequently, lip plays a crucial role in pollination because its captivating shape and pigmentation attract insects, and represents a landing platform that directs them toward the column [33]. Lips often exhibit highly specialized structures that allow exact placement of pollens. For example, the callus is an ornamental wart-like structure on the lip of some orchid species that favours the insect's attachment to the flower. The stelidia are lateral extensions of the gynostemium, which block the insect on the lip in the position that guarantees adequate pollination. The mentum is a protuberance between the gynostemium, the lip, and the internal tepals. It pushes the insect towards the upper part of the gynostemium in a precise position that ensures the release or removal of the pollen. The presence of these structures results from adaptation to specific pollinators; therefore, they are not characteristic of all orchids and present extreme morphological diversification [34].

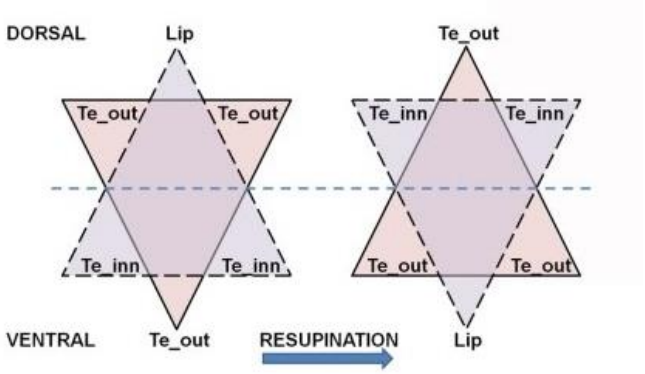


Figure 5. Orchid resupination model. From *Valoroso et al* (2017) [35].

The efficiency of orchid pollination is increased by zygomorphy, which represents the main evolutionary innovation in the orchid subfamilies Vanilloideae, Cypripedioideae, Orchidoideae and Epidendroideae, whereas the basal orchids Apostasioideae have actinomorphic flowers. Orchid bilateral symmetry is due to the lip differentiation and the developmental suppression of the adaxial stamens [32].

The peculiar features of the orchid flowers make them very attractive not only for evolutionary studies. Orchids are ornamental plants and have high economic value in global flower cultivation and represent an unique genetic resources for understanding the complex flower organogenesis processes [36-38].

1.3 Genetic basis of orchid flower development

1.3.1 Flower induction

During the flowering plant life cycle, the floral transition is a key phase of development. The shoot apical meristem (SAM) give rises leaves and branches during the vegetative phase of growth; however, when the floral induction starts, it becomes inflorescence meristem (IM) that produces floral meristem (FM) which differentiates into different type of floral organs [39].

In the different orchid species, the vegetative phase can last from one to thirteen years, and the reproductive development starts in the axillary buds with the formation of the bud primordia. As in other angiosperms, the timing of orchid floral transition is regulated by a series of signals including endogenous factors, such as hormones, and exogenous factors such as photoperiod, temperature and water availability. These signals activate a complex genetic network that regulates floral induction through the action of the floral integrator genes. Most of these genes belong to the MADS-box TF family [36].

In the model organism *Arabidopsis thaliana*, the main floral integrator genes are *FLOWERING LOCUS T* (*FT*) and the *MADS*-box *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*). They activate the expression of *LEAFY* (*LFY*) and *APETALA1* (*AP1*) that promote the FM

formation. In contrast, the *MADS*-box genes *FLOWERING LOCUS C* (*FLC*) and *SHORT VEGETATIVE PHASE* (*SVP*) repress flowering [39].

In orchids, the *FT* ortholog is highly expressed in leaves an axillary buds, promoting floral transition. Its expression pattern is affected by photoperiod and low temperature. The *FT* gene activates *SOC1* and *AP1* orthologs. Flowering repression is mediated by FLC and SVP complex that inhibits the transcription of *SOC1* (Figure 6) [36].

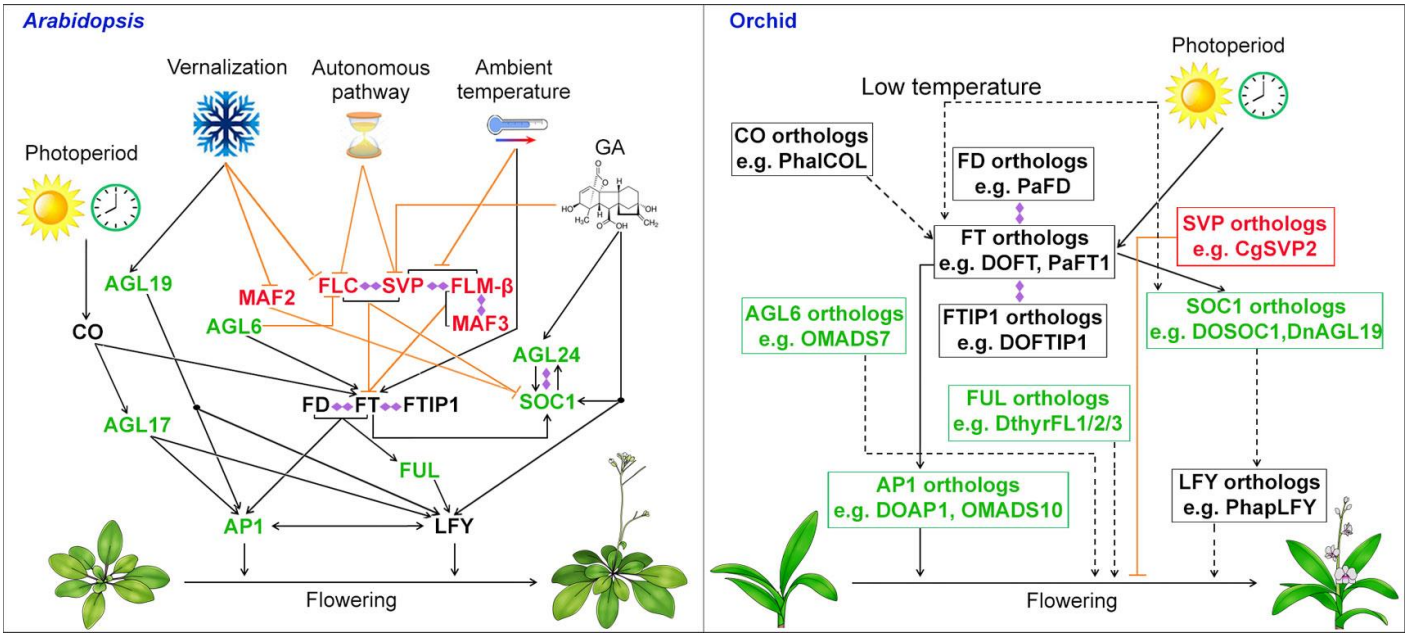


Figure 6. Floral transition in the model plant *Arabidopsis* and orchids. MADS-box factors that activate or repress flowering are indicated in green and red, respectively. Floral integrators are indicated in black. Activating and inhibiting effect are shown by black and orange lines. The dashed lines show putative positive regulation and the violet arrows show protein-protein interactions. From *Teo et al* (2019) [36].

A key role in flowering regulation is played by microRNAs (miRNAs) acting on genes involved in the transition from vegetative to reproductive stage. For example, the *Arabidopsis* miR156 and miR172 are part of the flowering time pathway. These regulatory interactions are conserved in orchids. For example, the miR172 of *Erycina pusilla* down-regulates the *APETALA2* gene, a flowering inhibitor, promoting flower induction. In contrast, miR156 inhibits the *SQUAMOSA* gene maintaining the vegetative phase and blocking the floral transition [40].

1.3.2 Floral organ identity and the ABCDE model

After the floral transition, the specification of the different floral organs is driven mainly by MADS-box TFs during flower development. The MADS-box TFs responsible for floral organ identity have a plant-specific MIKC modular protein organization (Figure 7).

Figure 7. The MADS-box TFs are present in plants, animals, and fungi [41]. Their protein structure includes the MADS-box DNA-binding domain at the N-terminus that binds the DNA CarG-box motif CC[A/T]$_6$GG [42]. There are two MADS-box groups: type I, including the *SRF* genes, and type II including the *MEF2* genes. Type I and II genes are present in plants, fungi, and animals and originated from a gene duplication event that occurred before the divergence of plants and animals [43]. However, the type II MIKC genes are plant-specific [44]. The plant type I group includes three subgroups, Malpha, Mbeta, and Mgamma, characterized by a MADS domain and a variable C-terminal domain [45]. The type II MIKC structure consists of four domains: the MADS domain (M) at the N-terminus, the I and K domains involved in protein–protein interactions and dimerization, the C-terminal (C) region that plays a role in transcriptional activation and multimeric protein complex formation [19, 46]. This lineage is divided into two clades: MIKC^C and MIKC* genes, which generally differ in the genomic structure encoding for the I domain [47]. Modified from *Gramzow et al* (2010) [48].

The first regulatory molecular network proposed to explain the flower organ formation in the model organisms *A. thaliana* and *Antirrhinum majus* (snapdragon) was the ABC model. This model identifies three different classes of floral homeotic genes (A, B, and C) that, through their combinatorial interactions and expression profiles, determine the identity of the floral organs. The addition of two other gene classes (D and E) extended the ABC to the ABCDE model [49].

Excluding *APETALA2* (*AP2*), the genes of the ABCDE model are MADS-box. The ABCDE regulatory network is quite conserved among the angiosperms, with some exceptions in specific plant groups as a consequence of the acquisition of new different floral structures during evolution [50-54].

In the basic model, the class A genes *APETALA* (*AP1* and *AP2*) are implicated in the development of the sepals in the first whorl and, together with class B genes *APETALA3* (*AP3*) and *PISTILLATA* (*PI*) in the second whorl, contribute to petal formation. In the third whorl, the class B and C genes (e.g., *AGAMOUS*, *AG*) drive the stamen development, while in the fourth whorl the class C genes

establish the carpel formation [55, 56]. The class D genes (e.g., *SEEDSTICK*, *STK* and *SHATTERPROOF*, *SHP*) lead to ovule development, and the class E genes ensure the correct specification of all floral organs [57, 58].

The MADS-box proteins form homo- and heterodimers that bind the CarG-box DNA motifs $CC[A/T]_6GG$, activating specific expression programs in the different tissues. These dimers interact with each other to form tetrameric structures, the floral quartets. In the first whorl there is the AP1/AP1/SEP/SEP complex, in the second AP1/SEP/AP3/PI, in the third AG/SEP/AP3/PI, and in the fourth AG/AG/SEP/SEP [19, 49, 59] (Figure 8).

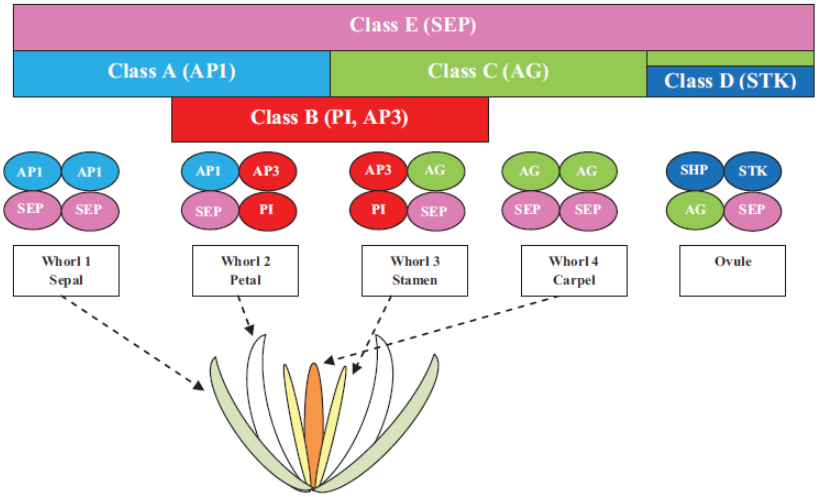


Figure 8. The basic ABCDE model. From *Aceto et al*. (2011) [17].

In agreement with the ABCDE model, the loss of function of a specific gene class results in specific changes of floral organ identity. For example, mutants of the C function have the concentric whorl organization of sepals-petals-petals-sepals. Each class of flower organ identity genes is crucial to the formation of a specific organ. For this reason, evolutionary innovations in floral structures may be traced back to modifications of the expression domains of the class A-E genes.

As described in the previous paragraph "The Orchidaceae", orchids have a petaloid perianth with similar sepals and petals. The expansion of the class B gene expression to the first whorl explains this flower feature. In addition, the orchid class A-E genes exhibit a gradient in their expression levels: high expression of a specific gene class overlaps with low expression of another gene in the adjacent

organ border (Figure 9). According to this "fading borders" model, through the progressive reduction of the expression of class B genes from the petals towards the borders of the adjacent organs, the influence for the organ identity is progressively reduced, conferring some petaloid characteristics to the adjacent floral organs (sepals) [60] [61].

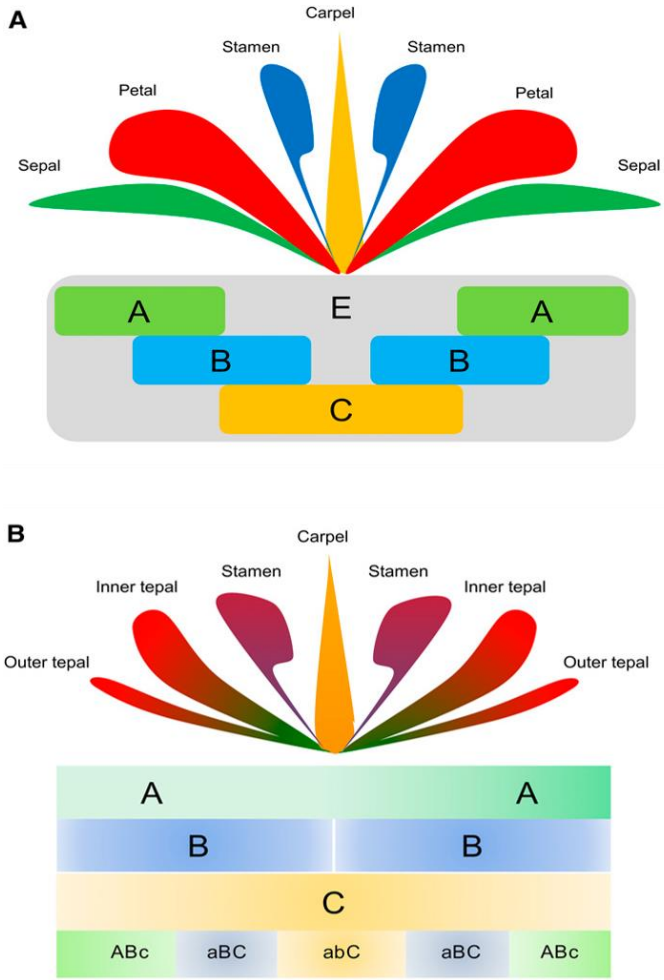


Figure 9. Basic ABCE model (A); fading borders model where "ABc," "aBC," and "abC" gene combinations establish floral organ identity. The lowercase font indicates lower functional influence (B). From *Chanderbali et al* (2016) [61].

Even though the fading borders model explains some features of the orchid perianth, it cannot explain the differentiation of the inner tepals into lateral and median (lip) and the establishment of zygomorphy. For this reason, orchid-specific models have been developed starting from the "orchid code" which has undergone changes and expansions over time, giving rise to the Homeotic Orchid Tepal (HOT) model and the Perianth code (P-code).

1.3.3 The Orchid code

The evolution of the zygomorphic orchid perianth results from changes in a developmental process that lead to the formation of some morphological key innovations, as the specialized median inner tepal called lip, crucial for pollination. As explained by the "fading borders" model, in the orchid the class B *MADS*-box genes promote the specification of the petaloid organs [7, 56], obtaining petaloid sepal (tepals) [61-64]

The class B genes are divided into two clades, *APETALA3/DEFICENS* and *PISTILLATA/GLOBOSA*, originating from a gene duplication event in an ancestor common to all angiosperms, followed by further duplications that occurred in the single clades in different families of flowering plants [44, 65-67].

The orchid flower probably originated from an actinomorphic flower with a single *AP3/DEF*-like gene equally expressed in the identical tepals. A first duplication event generated the ancestor of the clade 1 and clade 2 genes and the ancestor of the clade 3 and clade 4 genes. This initial duplication was fundamental to differentiate the outer from the inner tepals, generating an intermediate evolutionary state that still exists in Apostasiodeae, the most ancestral Orchidaceae subfamily, characterized by the absence of a differentiated lip. A second round of duplications occurred before the separation of Vanilloideae from most recent orchids and generated four different clades: clade 1 is sister to clade 2, and clade 3 is sister to clade 4. The four *DEF*-like genes have tissue-specific expression profiles acquired after duplication probably as the result of degenerative mutations in *cis*-regulatory elements and the acquisition of new *cis*-regulatory elements [68].

The orchid code model recognizes to the four *DEF*-like genes the main role in evolution and development of the orchid perianth organs. In particular, the clade 1 and clade 2 genes control the formation of the outer tepals. The differentiation of two lateral inner tepals is influenced by the combination of the high expression of the clade 1 and 2 genes and the low expression of the clade 3 and 4 genes. Lip formation is specified by the high expression of the clade 3 and 4 genes and the low expression of the genes belonging to clade 1 and 2 [17, 69] (Figure 10).

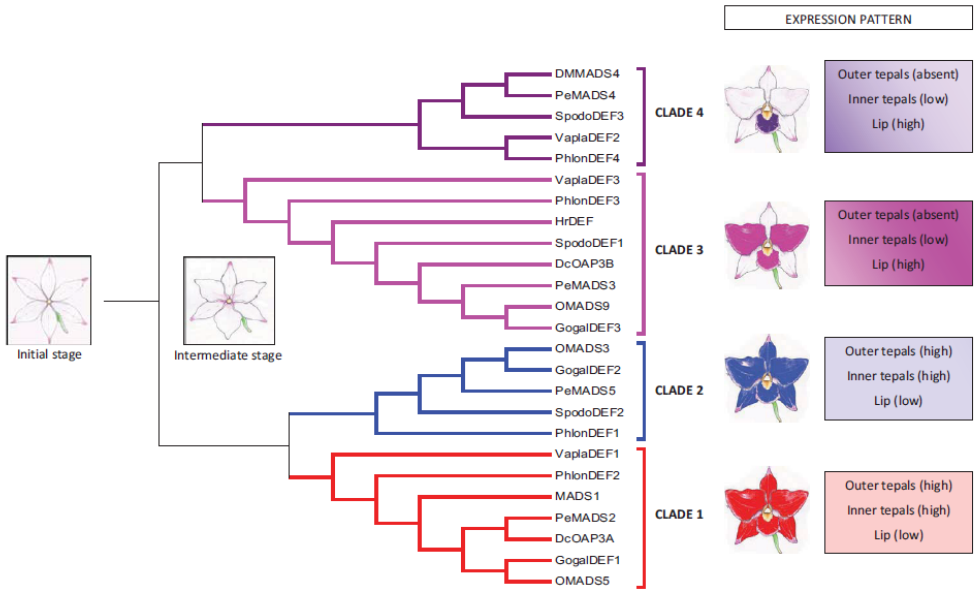


Figure 10. The orchid code molecular model starting from the possible initial and intermediate evolutionary stages of the orchid flower. From *Aceto et al* (2011) [17].

In most orchids there is only a single *GLO*-like gene expressed ubiquitously in the flower whorls. However, the identity of the orchid perianth organs depends also on the formation of heterodimers between PI/GLO-like protein and the four AP3/DEF-like proteins [17, 69].

1.3.4 Homeotic Orchid Tepal (HOT) model

The morphogenesis of the orchid perianth is a complex process that includes a series of phases during which the involved genes change their expression in space and time. The Homeotic Orchid Tepal (HOT) model describes the intricate mechanism of determination of the orchid perianth organ identities analyzing the modification of the *MADS*-box genes expression pattern under spatial and temporal conditions. According to this model, during the orchid floral primordial stage, all class B *MADS*-box genes are expressed equally. Later, during late inflorescence and floral bud stages, they became differentially expressed.

The HOT model proposes that at the early inflorescence stage, the combinatorial interaction of the MADS-box proteins results in the formation of complexes more sophisticated than the floral quartets. The gene expression profile illustrated by the HOT model at the flower bud stage agrees with the

"orchid code" model. The *PI*-like genes induce the outer tepals differentiation together with *DEF*-like clades 1 and 2. The lateral inner tepal identity is influenced by the action of the *PI*-like gene together with clades 1, 2, and 3. At the same time, lip differentiation depends on combining clades 3 and 4 with *PI*-like and other *MADS*-box genes such as *AGL6*-like and *SQUAMOSA* (*SQUA*)-like. Moreover, column and ovary development are orchestrated by B, C, and D class proteins forming different multimeric complexes [64] (Figure 11).

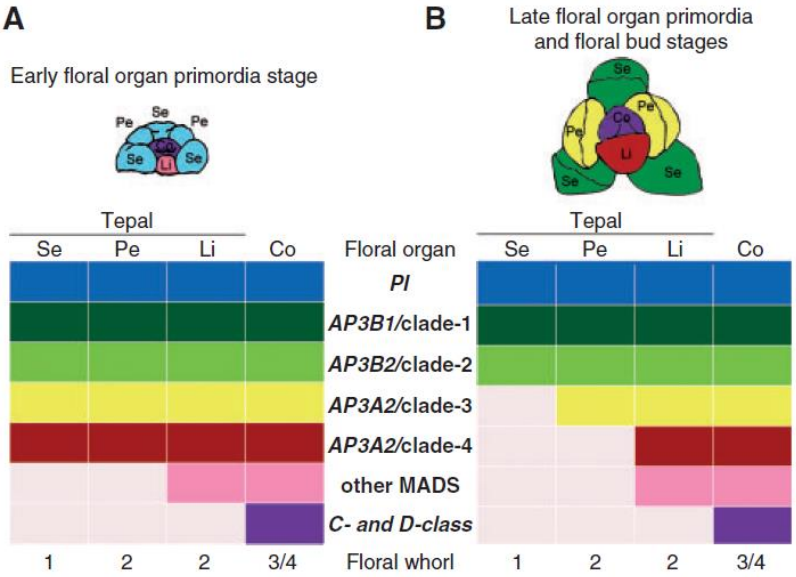


Figure 11. The Hometic Orchid Tepal (HOT) model. MADS-box genes expression profile at early (A) and late (B) orchid inflorescence stages. From *Pan et al.* (2011) [64]

1.3.5 Perianth code (P-code)

The Perianth-code (P-code) model aims to explain the detailed genetic basis of the variability of the orchid perianth structure. Based on this model, the perianth organs identity is not only regulated by the class B *AP3/DEF*-like genes, as described by the "orchid code", but also by the competition between two protein complexes of class B (PI/GLO-like and AP3/DEF-like) and E (AGL6-like) MADS-box factors. According to the P-code, in tepals (sepal/petal) and lip, there are two tetramers, both including the PI/GLO-like protein equally distributed in the perianth, and by tissue-specific proteins AP3/DEF-like and AGL6-like.

The P-code was described first in *Oncidium* orchid where the SP (sepal/petal) complex is composed of OAP3-1/OAGL6-1/OAGL6-1/OPI, and the L (lip) complex of OAP3-2/OAGL6-2/OAGL6-2/OPI. These two quartets antagonistically specify the fate of the different floral organs. The L complex promotes the lip program, which activates lip formation, while the SP complex represses it, favouring the tepal (sepal/petal) differentiation. The two complexes perform a compensatory action and their antagonism creates a balance that allows the correct development of the flower. Mutant orchids missing the SP complex have flowers with only lip-like structures, whereas the absence of the L complex promotes only tepal differentiation. The co-existence and co-absence of both complexes in the same tissue generate intermediate structures (Figure 12) [36, 70].

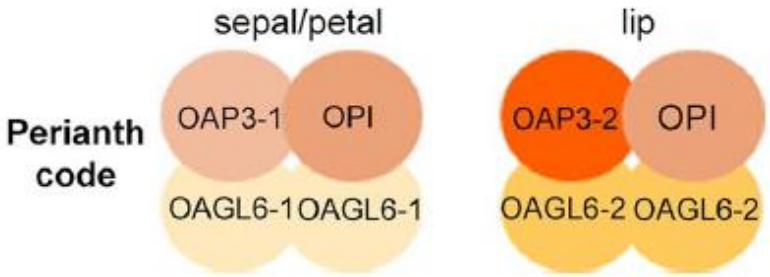


Figure 12. Perianth code model. From *Teo et. al* (2019) [36]

1.4 Orchid flower symmetry

Progress in genomic sequencing technologies have made available an increasing number of orchid genomes and transcriptomes. Multi-omics approaches permit investigating the molecular networks at the base of the orchid flower development, expanding the known molecular models [24].

1.4.1 Establishment of floral symmetry

The molecular network at the base of the floral symmetry was first analysed in the zygomorphic flowers of the snapdragon *A. majus* [71], where the key regulators of the floral symmetry are TCP (Figure 13) and MYB (Figure 14) TFs.

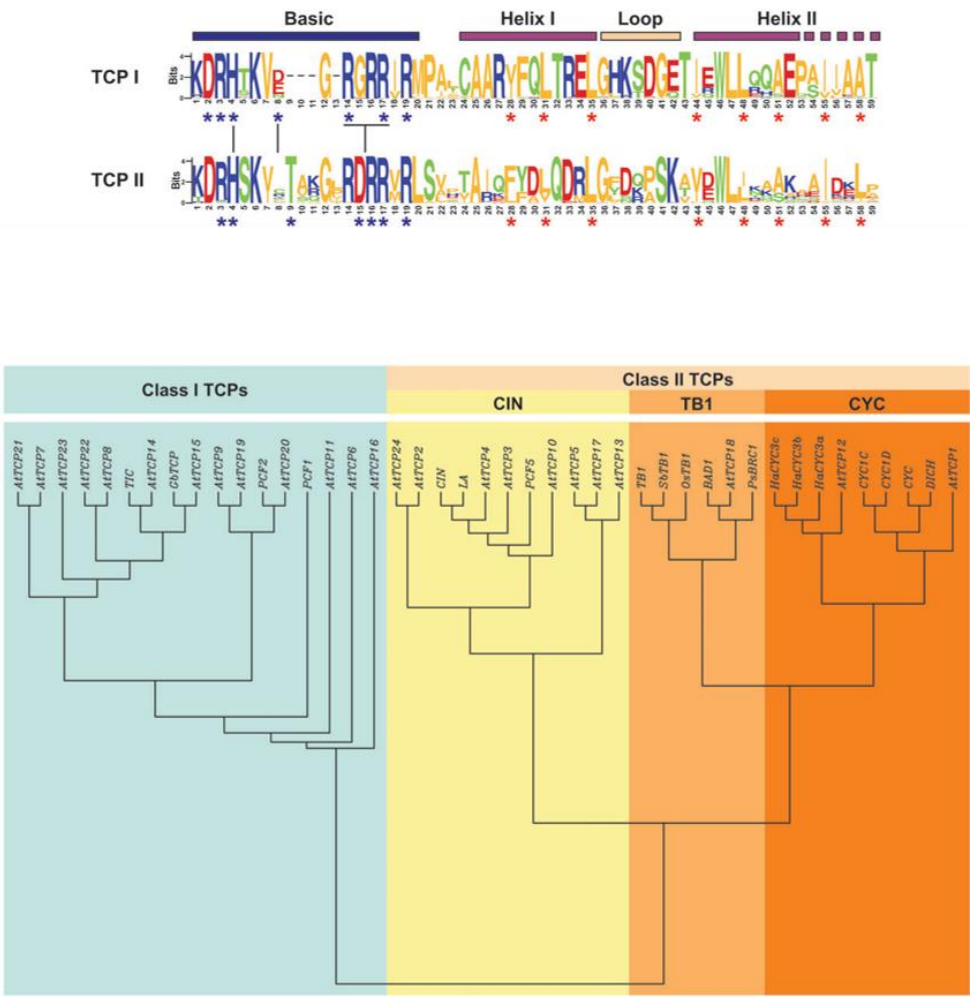


Figure 13. TCPs are a plant-specific TFs**.** The TCP proteins contain a highly conserved TCP helix-loop-helix domain of 59 amino acids at the N-terminus, essential for DNA binding, protein-protein interactions and nuclear localization [72]. Blue and red asterisks indicate the residues required for DNA binding and the hydrophobic residues, respectively. The TCP family is divided into two classes: class I (class PCF or TCP-P) and class II (class TCP-C) [73]. Class I TCPs promote cell division, growth and differentiation, while class II are generally involved in inhibiting growth and differentiation, as well as determining bilateral symmetry of the flower through interaction with MYB TFs [74, 75]. The main difference between class I and class II proteins is a 4 amino acid deletion in the class I TCP domain resulting in two different DNA binding sequences: GGNCCCAC for class I and G (T/C) GGNCCC for class II [72, 76]. In addition to the TCP domain, class II proteins have an R domain, rich in arginine, the function of which is still unknown [77]. Class II is divided into two subclasses: CIN and CYC/TB1 or ECE. In the ECE clade, a duplication event gave rise to three subgroups of *CYC1*, *CYC2* and *CYC3* genes [73, 78]. The genes belonging to the CYC2 clade are involved in the determination of the floral bilateral symmetry. In *A.majus* there are two *CYC2*-like genes, *CYC* and *DICH* [9, 79-81]. Modified from *Manassero et al.* (2013) [82].

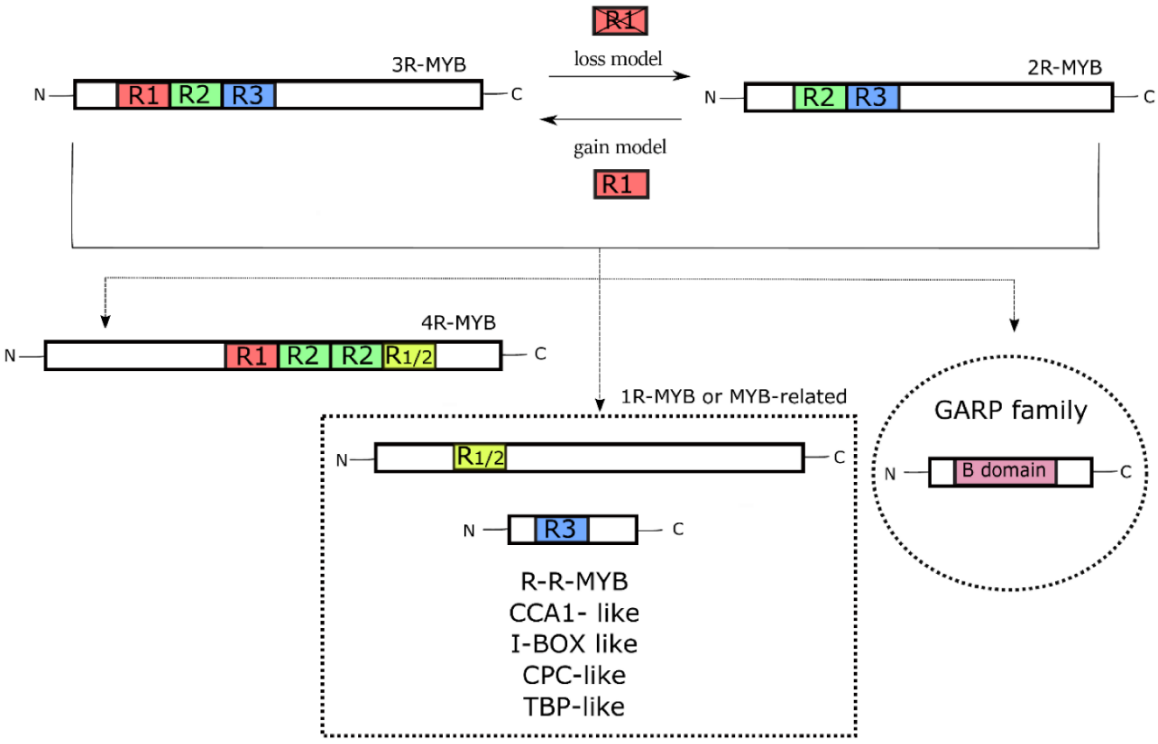


Figure 14. The MYB TF family is highly conserved in all eukaryotic organisms. MYB proteins have a variable number of MYB repeats (R). Each repeat is composed of 52 amino acids forming a helix-turn-helix motif. The R motif includes three regularly spaced tryptophan residues forming a hydrophobic region probably involved in the recognition of specific DNA sequences [83-85]. The MYB TFs are divided into 4R, 3R, 2R, and 1R-MYB types, based on the number of R domains. 3R- and 2R-MYB could have originated from the acquisition or deletion of the R1 repeat, as explained from two alternative 'gain' or 'loss' evolutionary models [86-90]. The 4R- and 1R-MYBs originated from 3R and 2R. There are five sub-groups of 1R-MYB : R-R-MYB, CCA1-like, I-box-binding-like, CPC-like, and TBP-like [83, 87, 91]. The MYBs are related to the GARP TFs for the resemblance of the B-domain to the MYB domain [92, 93]. From *Lucibelli et al* (2020) [9].

The ventralization of the snapdragon flower depends on two MYB TFs: DIV and DRIF. In the ventral part of the flower, the DRIF protein binds DIV forming a heterodimer that migrates in the nucleus activating the transcription of still unknown ventral identity genes. The DIV/DRIF complex probably regulates the *DIV* transcription itself, as suggested by the presence of the DIV target sequence (5'-GATAA-3') on the *DIV* promoter [94, 95]. In the dorsal part of the flower, the DIV/DRIF interaction is inhibited by the small MYB peptide RAD, which competes with DIV to bind DRIF. For this reason, DIV is unable to activate the ventralization program in the dorsal part of the flower [9, 96]. In the dorsal region of the snapdragon flower, the transcription of *RAD* is activated

by the TCPs CYC and DICH [79-81] that bind specific DNA sequences in the *RAD* gene promoter and intron [97] (Figure 15A).

In orchids, the regulatory module of DIV, RAD, and DRIF is conserved compared to *A. majus*. However, the *MYB* gene expression domains are rotated at 180° due to the resupination that occurs during the development of the orchid flower.

In the zygomorphic flower of *Orchis italica* and *Phalaenopsis equestris*, *DIV* and *DRIF* are ubiquitously expressed in the perianth organs and promote ventral identity in the lateral inner tepals through their protein-protein interaction. The *RAD* gene is highly expressed in the orchid lip, that is a dorsal structure, but after resupination it takes a ventral position. The RAD/DRIF interaction inhibits the formation of the DIV/DRIF complex, as in *A. majus* [9, 98] (Figure 15B). Unfortunately, the crucial CYC-like transcription factor that regulates the expression of *RAD* in *A. majus* is not yet well characterized in orchids.

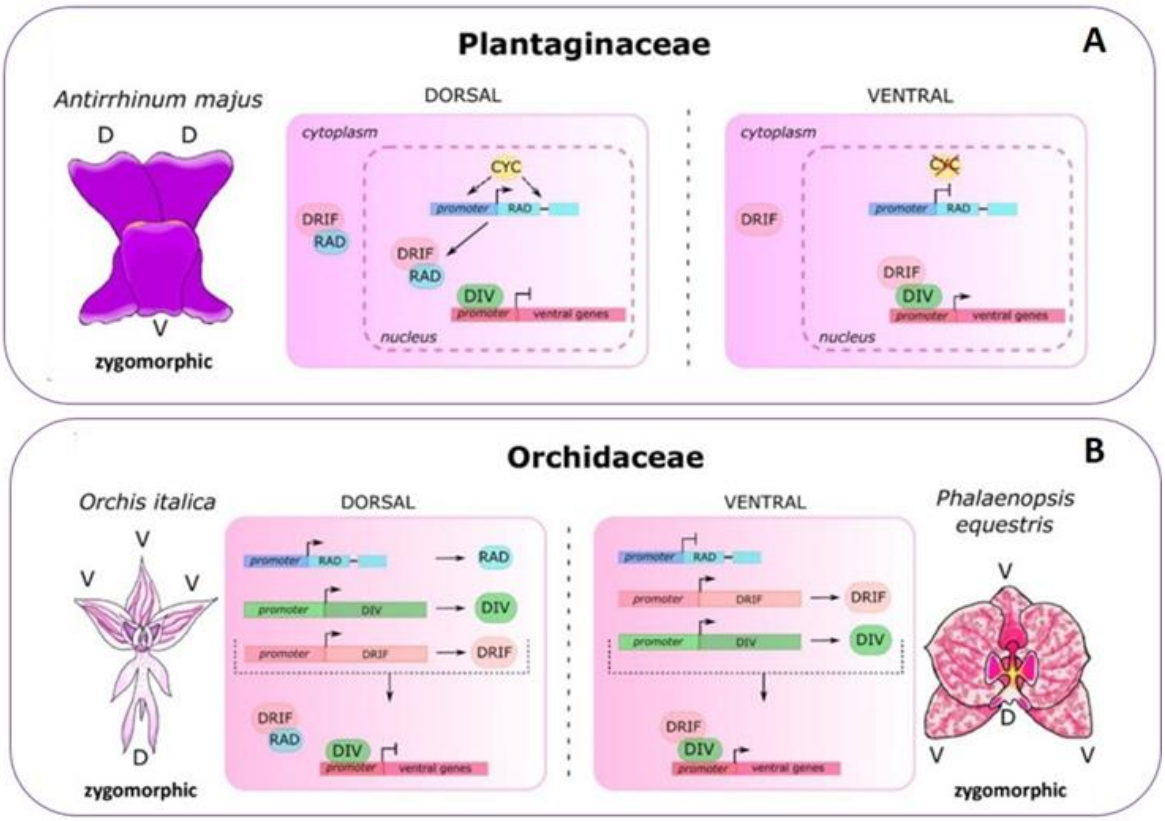


Figure 15. Molecular model at the base of floral symmetry establishment of *A. majus* (A) and orchids (B). Modified from *Lucibelli et al.* (2020) [9].

2 AIM OF THE WORK

My Ph.D. project aimed to study the complex molecular network underpinning the orchid flower development and the specification of its organ identity. I focused my analysis on the genes involved in orchid perianth morphogenesis, especially the lip, representing the crucial organ for pollination efficiency. Therefore, studying the genes involved in its development is very important in understanding the orchid family evolution and could have economic application because the modification and control of their expression profile can generate new varieties with commercial value.

According to the different molecular models that explain the organ formation of the orchid flower (Orchid code, HOT model, and P-code), duplication of the *DEF*-like class B *MADS*-box genes and their differential expression play a pivotal role in orchid perianth differentiation. In addition, our research group has recently highlighted the involvement of other TF families in orchid lip development (e.g.: TCPs and MYBs) [35, 98, 99].

Starting from this knowledge, I performed a preliminary *in silico* differential expression analysis using RNA-seq data of *Phalaenopsis* to identify new candidate genes with a role in the development of the orchid lip. These analyses indicated as best candidate the *DROOPING LEAF/CRABS CLAW* (*DL/CRC*) gene, belonging to YABBY TF family. This gene is differentially expressed between lip and lateral inner tepals in the wild type *Phalaenopsis*, whereas in the organs of the peloric mutant the expression levels are similar. This expression pattern overlaps that of the class B *AP3/DEF* MADS-box genes *PeMADS3* and *PeMADS4* and is opposite to that of *PeMADS2* and *PeMADS5*.

Based on the regulatory relationship between *DL*-like and *MADS*-box genes in many angiosperms and my *in silico* analysis results, I hypothesized that during orchid evolution the *DL*-like gene might have acquired a role in the orchid lip development together with the *AP3/DEF*- like genes.

To validate this hypothesis, I used different approaches to characterize the orchid *DL*-like gene:

- Genome- and transcriptome-wide identification of the *DL*-like genes of different orchid species and reconstruction of their gene organization. Phylogenetic and conserved motif analyses to infer the evolution of the orchid *DL*-like genes;

- Evaluation of the expression pattern of the *DL*-like and *AP3/DEF*-like genes in wild-type and peloric orchids by *in silico* differential expression analysis, real-time qPCR, *in situ* hybridization to verify if the transcriptional profile is conserved outside the *Phalaenopsis* genus;

- Determination of the subcellular compartment where the orchid DL protein localizes to confirm its nuclear position and identify the protein domain responsible of its location;

- Evaluation of possible DL protein dimer formation and search of possible protein interactors by *in silico* prediction and yeast two-hybrid analysis to shed light on the possible mechanism of action of the orchid DLs;

- *In silico* prediction of conserved regulatory motifs on the orchid *DL* promoter and validation by yeast one-hybrid analysis and protoplast double-strand RNA interference (dsRNAi) experiments to identify the TFs regulating the orchid *DL* transcription.

3 MATERIAL AND METHODS

3.1 Plant material

The orchids used in this study are shown in Figure 18 and represent different subfamilies: Epidendroideae, with the wild-type *Phalaenopsis aphrodite* (Figure 18A-B) and *Phalaenopsis* hyb. "Athens" (Figure 18D), and the peloric mutants *Phalaenopsis* hyb. "Athens" (Figure 18E) and *Phalaenopsis* hyb. "Joy Fairy Tale" (Figure 18F); Cypripedioideae, with the wild-type *Phragmipedium longifolium* (Figure 18 G); Vanilloideae, with the wild-type *Vanilla planifolia* (Figure 18C).

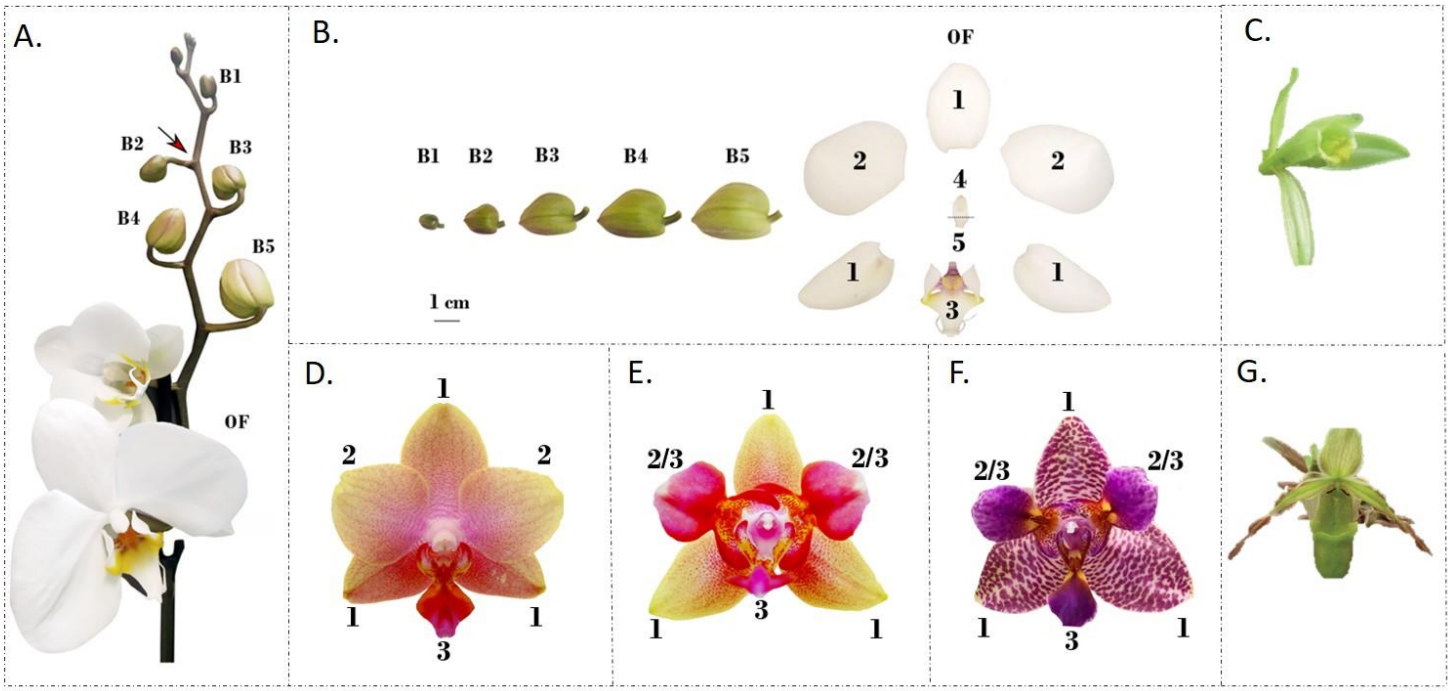


Figure 18 . Orchids used in this work. Wild-type inflorescence of *P. aphrodite* (A); floral buds at stages B1–B5 and floral organs at the OF stage of the wild-type *P. aphrodite* (B); flower of *Vanilla planifolia* (C); flower of the wild-type *Phalaenopsis* hyb. "Athens"(D); flower of the peloric mutant *Phalaenopsis* hyb. "Athens"(E); flower of the peloric mutant *Phalaenopsis* hyb. "Joy Fairy Tale" (F); flower of *Phragmipedium longifolium* (*G*). The arrow in (A) indicates the point of rotation of the pedicel during resupination. Size of the floral bud at he different developmental stages: B1 (0.5–1 cm), B2 (1–1.5 cm), B3 (1.5–2 cm), B4 (2–2.5 cm), B5 (2.5–3 cm), OF (open flower). 1, outer tepals; 2, lateral inner tepals; 3, lip; 4, column; 5, ovary; 2/3, lip-like organs. Modified from *Lucibelli et al* (2021) [100].

The orchid plants were grown in the greenhouse of the Department of Biology (University of Naples Federico II, Napoli, Italy) and of the Department of Cell Biology and Plant Biochemistry (University of Regensburg, Regensburg, Germany), under natural light and temperature.

The wild-type perianth of *Phalaenopsis* has zygomorphic symmetry with two lateral inner tepals and one median inner tepal (lip) (Figure 18A-B-D). The flowers of *Phalaenopsis* peloric mutants

have actinomorphic symmetry for the presence of two lip-like organs in substitution of the lateral inner tepals in the second whorl (Figure 18E-F). *Phalaenopsis* species used in this work and *Vanilla planifolia* exhibit a callus structure on the hypochile part of the lip and *Phragmipedium longifolium* have callus-like structures, even though placed in the middle of the lip.

Flower buds at different developmental stages were collected from three different plants of the wild-type *P. aphrodite* before anthesis: B1 (bud length 0.5-1 cm), B2 (1-1.5 cm), B3 (1.5-2 cm), B4 (2-2.5 cm) and B5 (2.5-3 cm). The tissues from open flowers (OF) were collected soon after anthesis (Figure 18A-B).

Single flowers of six wild-type *Phalaenopsis* hyb. "Athens" and of the peloric mutants were collected at the developmental stage B2. *Vanilla planifolia* and *Phragmipedium longifolium* floral buds were collected at early stage of development (bud size 1,5 cm and 1 cm, respectively).

The floral organs (outer tepals, lateral inner tepals, lip, column and ovary) were dissected from different flowers. The *P. aphrodite* lip at the B2 stage was dissected into three parts (callus, lateral lobes and central lobes) to perform tissue-specific expression analysis.

All the collected samples were immediately frozen in liquid nitrogen or immerged in RNAlater (Ambion) and stored at -80°C until RNA extraction. Total RNA was extracted from all the sampled organs using Trizol (Ambion) followed by DNase treatment. Quality and amount of the extracted RNA were checked using the 2100 Bioanalyzer system (Agilent Technologies) and Nanodrop 2000 (Thermo Scientific).

The lip calli of *P. aphrodite* (B2 stage) were collected and fixed in 4% (v/v) paraformaldehyde, 0.5% (v/v) glutaraldehyde, 0.1% Triton X-100, and 4% dimethyl sulfoxide in phosphate–saline buffer 1X for 16 h at 4° C. They were then dehydrated through ethanol series, paraffin embedded and stored at 4°C until *in situ* hybridization experiments [101].

For subcellular localization experiments, *Nicotiana benthamiana* plants were grown under natural light and temperature in the greenhouse of the Botanic Institute (Justus Liebig University Giessen, Germany).

3.2 *In silico* differential expression analysis and phylogeny

Total RNA from lateral inner tepals and lip of wild-type *Phalaenopsis* hyb. "Athens", and lip-like lateral inner tepals of peloric *Phalaenopsis* hyb. "Athens", dissected from B2 floral buds collected from three different plants was sequenced by Macrogen (Seoul, Korea). Illumina TruSeq RNA (Oligo dT) mate paired-end libraries were produced and independently sequenced in a lane with a coverage >150 million 100 bp paired-end reads, generating ~220 million paired-end reads for each sample. The

FastQC analysis revealed that 94% of reads have a quality score over 30. Trimming and mapping to the *Phalaenopsis equestris* genome v1.0 (ASM126359v1) were realized by CLC GenomicsWorkbench (v11.01).

The wild-type and peloric mutant *Phalaenopsis* hyb. “Brother Spring Dancer” KHM190 [102] Illumina raw reads were downloaded from the Sequence Read Archive (SRA). Trinity v2.3.0 software was used to assemble the paired-end reads obtained from wild-type and peloric outer tepals (accession numbers SRR1055198 and SRR1055947), inner tepals (SRR1055945 and SRR1055948), and lip (SRR1055946 and SRR1055949) [103]. The transcript functional annotation was performed using the Annocript v2.0.1 software [104], and the edgeR v3.13 software [105] was adopted to carry out differential gene expression analysis between wild-type and peloric mutant tissues.

In silico analysis of the *Phalaenopsis* hyb. "Brother Spring Dancer" transcriptome [102] revealed the presence of two *DL*-like transcripts, *PeDL1* and *PeDL2*, each with two different isoforms. To reconstruct the genomic organization of the *PeDL1* and *PeDL2* genes, BLAST analysis were conducted using as query the *DL*-like sequences identified by the transcriptome anlaysis of *Phalaenopsis* hyb. “Brother Spring Dancer” against the assembled genome of *P. equestris* (assembly ASM126359v1).

The identification of the *DL*-like genes of different orchid species was performed in the orchid-specific database Orchidbase 2.0 [106] and Orchidstra 2.0 [107] using as query a conserved YABBY domain (PFAM PF04690). The genome organization of *Dendrobium catenatum* and *Apostasia shenzhenica DL*-like genes was performend by BLAST analysis using their respective assembled genomes [108, 109]. Phylogenetic and conserved motif analyses of the orchid *DL*-like genes was performed using MEGA7 with 500 bootstrap replicates and MEME v5.5 online tool [110].

Paired-end Illumina reads from outer tepals, lateral inner tepals, and lips dissected into epichiles, and hypochiles at different developmental stages (D1, D4, D7, and D8) of *Rhyncholaeliocattleya* Beauty Girl “KOVA” were downloaded from SRA (PRJNA559603). D1 (bud length < 2 cm), D4 (bud length 4–5 cm), D7 (two days after flowering), and D8 (ten days after flowering) [111]. Reads processing and differential expression analysis were performed as previously described for *Phalaenopsis* hyb. “Brother Spring Dancer” KHM190.

3.3 Quantitative expression analysis

Reverse-transcription reactions were performed on 500 ng of total RNA using the Advantage RT-PCR kit (Clontech) and a mix of oligo dT and random hexamer primers.

To validate the nucleotide sequences of the *PeDL1* and *PeDL2* transcripts and of their alternatively spliced isoforms identified *in silico*, the wild-type *Phalaenopsis* hyb. cDNA was PCR amplified (DreamTaq, Thermo Scientific™), using gene- and isoform-specific primer pairs (Table 1-Appendix). The amplicons were cloned into pSC-A-amp/kan vector (Agilent Technologies) and sequenced using the T3 and T7 primers (Eurofins Genomics). The nucleotide sequences were deposited in GenBank with the following accession numbers: MW574592 (*PeDL1_1*), MW574593 (*PeDL1_2*), MW574594 (*PeDL2_1*), MW574595 (*PeDL2_2*).

Real-time quantitative expression analysis (qPCR) of *PeDL1* (two isoforms: *PeDL1_1* and *PeDL1_2*), *PeDL2* (two isoforms: *PeDL2_1* and *PeDL2_2)*, *PeMADS2*, *PeMADS5*, *PeMADS3*, and *PeMADS4* was performed using PowerUp™ SYBR™ Green Master Mix (Applied Biosystems™) in wild-type and peloric *Phalaenopsis* collected tissue.

Relative expression of *DL*-like and *DEF*-like genes was evaluated in all tissues of *Vanilla planifolia* and *Phragmipedium longifolium* orchid flowers, to verify if the expression profile of these genes is conserved among different orchid subfamilies/species.

18S, *Actin*, and *Elongation Factor 1α* were used as reporter genes in qPCR experiments [69]. All the qPCR experiments were conducted in technical triplicates. Normalized relative quantity (NRQ) ± SEM was calculated for each replicate to the geometric average expression of three internal control genes [112]. To evaluate the statistical significance of the differences of NRQ among the different tissues ANOVA analysis followed by Holm–Sidak post-hoc test was carried out.

3.4 RNA *in situ* hybridization

The paraffin embedded samples (see Plant material) were sectioned at 5 μm. A fragment of the *PeDL2_1* isoform was PCR amplified (DreamTaq, Thermo Scientific™) using an isoform-specific primer pair for the probe synthesis (Table 1-Appendix). The PCR amplification product was cloned into pGEM®-T Easy Vector (Promega). The digoxigenin-labeled sense and antisense RNA probes were synthesized using the T7 and SP6 RNA polymerases and the DIG RNA Labeling kit (Roche).

Hybridization and detection of the signals with alkaline phosphatase was obtained using the DIG Nucleic Acid Detection kit (Roche) following the manufacturer's instructions.

3.5 PeDL2 protein subcellular localization

The full length coding sequence (CDS) of PeDL2 and its individual domains: Zinc Finger, Intermediate, and YABBY, were PCR amplified (DreamTaq, Thermo Scientific™) (Table 1-Appendix) and sub-cloned into N-terminal GFP fusion vector pEGAD under the control of the CaMV 35S promoter [113] [114]. The borders of the PeDL2 domains (ZF aa 14–42, INT aa 43–98, YAB aa 99–148) are based on the alignment with the *Arabidopis* CRC protein domains [114] (Figure 19).

The PeDL2 complete CDS and its single domains were transiently expressed in leaves of 4 weeks old *Nicotiana benthamiana* plants by *Agrobacterium tumefaciens* GV3101 infiltration performed in the lower part of the leaf. Three days after infiltration a small disk of the infiltrated leaf was detached and stained with 1 ng/ml DAPI under vacuum. Microscopy analysis was performed with the Leica fluorescence microscope DCM5500 (Leica Microsystems GmbH, Wetzlar, Germany) using A4 filter for DAPI fluorescence and L5 filter for GFP fluorescence.

3.6 *In silico* prediction of PeDL protein structure and interactions

Prediction of the secondary structure of full length PeDLs was performed by REPPER online tool [115-117]. After this preliminary analysis, the *in silico* prediction of PeDL2 structure and interaction was carried out in collaboration with the prof. Bruno Hay Mele.

The LambdaPP service was used to get per-protein and per-residue prediction for UAB34661.1 (PeDL1) and UAB34663.1 (PeDL2) from the NCBI protein database. The sequence of PeDL1 and PeDL2 was scanned for matches against the InterPro protein signature databases, obtaining all the domains present within the two isoforms.

The AlphaFold2 AI through the ColabFold Jupyter notebook was used to perform *de-novo* structure prediction for PeDL1 and PeDL2 monomers, setting 48 recycles and no templates in the notebook.

Recently (July 2022), DeepMind expanded available prediction within the EBI Alphafold database (afdb); new entries include PeDL2 (A0A7S8F9E7) but not PeDL1 since the latter is not available in UniProt. Since the structural domains generated with ColabFold perfectly overlap the ones of the afdb entries, we will use the latter for PeDL2. For PeDL1, the ColabFold prediction will be used.

The model quality was evaluated for the predictions using the pLDDT and PAE metrics provided by AlphaFold2.

The AlphaFold2 AI through the ColabFold Jupyter notebook was used to perform *de-novo* structure prediction of the PeDL1 and PeDL1 homodimers and the PeDL1:PeDL2 heterodimer. We evaluated the model quality for the predictions using the pLDDT and PAE metrics provided by AlphaFold2. The input aa sequences are produced by sequence duplication (monomers) or joining (dimers). Sequences for UAB34661.1 (PeDL1) and UAB34663.1 (PeDL2) from the NCBI protein database was used setting 48 recycles and no templates in the notebook.

Possible PeDL2 interactors were identified based on the known co-expression profiles of the related transcripts. The putative interactors chosen were the four AP3/DEF-like factors PeMADS2 (AAR26628.1), PeMADS5 (AAR26630.1) , PeMADS3 (AAR26629.1), PeMADS4 (AAR26626.1); the AGAMOUS-like factor PeMADS1 (AAL76415.1); the YABBY factor PeYABBY7

(XP_020589671.1); and CYC-like factors PeCYC1 (KT258891.1), PeCYC2 (KT258892.1) PeCYC3 (KT258893.1).

The sequence of possible interactors was scanned for matches against the InterPro protein signature databases, obtaining all the domains present within each interactor. Then, we searched the Protein Common Interface Database for Pfam domain-domain interfaces using the information provided by the InterPro search. Finally, a search was conducted in parallel PPIDomainMiner for an estimate of interaction between PeDLx domains and possible interactor domains.

Finally, the batch ColabFold was used to perform *de-novo* structure prediction of the nine heterodimers (PeDL2:interactor), setting six recycles and no templates in the notebook. The model quality for the predictions was evaluated using the pLDDT and PAE metrics provided by AlphaFold2 and integrated our understanding with PISA metrics.

3.7 Yeast Two-Hybrid Analysis

To test the possible protein–protein interaction between PeDL2_1, PeMADS2-PeMADS6, and PeCYC1-3 and the capability to form homo and heterodimers with the different isoforms of PeDL1 and PeDL2, the GAL4-based yeast two-hybrid (Y2H) system (Matchmaker two-hybrid system; Clontech) was used. As positive control, Y2H analysis was used to check the ability of PeMADS6 to form heterodimers with PeMADS2-PeMADS5.

The full-length coding regions of PeDL1_1 (MW574592), PeDL1_2 (MW574593), PeDL2_1 (MW574594), PeDL2_2 (MW574595), PeMADS2 (AY378149), PeMADS3 (AY378150), PeMADS4 (AY378147), PeMADS5 (AY378148), PeMADS6 (AY678299), PeCYC1 (KT258891), PeCYC2 (KT258892) and PeCYC3 (KT258893) were PCR amplified (DreamTaq, Thermo Scientific™) (Table 1-Appendix). The obtained amplification products were cloned into pGBKT7 (bait vector, Clonetech) and pGADT7 (prey vector, Clonetech) in frame with the GAL4 DNA-binding domains (BD) and transcription activating domain (AD) sequences.

The *Saccharomyces cerevisiae* strain AH109 was transformed with all the prey and bait recombinant vector combinations using the LiAC/DNA/PEG transformation method [118], conducting each experiment in triplicate. The double transformed cells were plated on a synthetic defined (SD) agar medium lacking leucine and tryptophan (SD/-Leu/-Trp) and incubated at 30°C for

3-4 days, to verify the presence of the plasmids. The positive colonies were transferred on a selective SD medium lacking tryptophan, leucine, and histidine (SD/-His/-Leu/-Trp) in presence of 20 mM of 3-amino triazole (3AT).

The protein-protein interaction is demonstrated by the growth of colonies due to the reporter gene *HIS3* transcriptional activation. Moreover, we verified the self-activation of the proteins fused to the GAL4 BD by the single transformation of yeast cells and growth in SD medium without histidine and tryptophan (SD/-His/-Trp) supplemented with 20 mM 3AT. The transformation of empty pGADT7 or pGBKT7 vectors in combination with the recombinant vectors was used as negative control.

3.8 Identification of conserved transcription factor binding sites

In silico analysis of *P. equestris* and *Dendrobium catenatum* genomes [108, 119] was performed to identify *DL* transcription factor binding sites (TFBSs) conserved in orchid species. The sequences spanning ~3 kb upstream the translation start site of the *DL2* genes were selected (putative promoters). and scanned using the online tool MEME v5.3.3 [110]. The analysis was repeated on shuffled sequence as a negative control. The motifs obtained were analysed in JASPAR2020 core plants database through TOM TOM v5.3 (http://jaspar.genereg.net/, access date 18 January 2021) [120]. In addition, the search of known TFBSs within the *P. equestris, P. aphrodite* and *D. catenatum* putative promoters was conducted in PLANTPAN 3.0 [121].

The presence of conserved binding sequences of the MADS-box factors: CArG-boxes CC(A/T)6GG or variants CC(A/T)8G on *DL2* putative promoters was verified by FUZZNUC software (http:// emboss.bioinformatics.nl/cgi-bin/emboss/fuzznuc).

3.9 Yeast One-Hybrid analysis

The region of *P. aphrodite DL2* putatitive promoter 3,4 kb upstream the translation start site was divided into five fragments, which share conserved regions with the *A. thaliana CRC* promoter [122].

Genomic DNA was extracted from *P. aphrodite* leaves following the CTAB protocol [123]. Fragments of the *DL2* promoter were PCR-amplified (DreamTaq, Thermo Scientific™)(Table 1-

Appendix) and cloned into pAbAi vector (Takara Clontech, Saint-Germain-en-Laye, France) in the *Kpn*I and *Sal*I restriction sites.

The recombinant vectors were linearized by *Bst*BI or *Bbs*I and transformed into *S. cerevisiae* Y1HGold according to the Yeast Transformation System 2 manual (Takara Clontech). The transformed cells were plated on SD agar medium that lacked uracil (SD/-U) and incubated at 30°C for 3-4 days, to verify the presence of the plasmids.

The yeast autoactivation test was performed adding to SD/-U medium Aureobasidin A (Takara Clontech) (AbA) at different concentrations (100, 150, 200, 500, and 1000 ng/ml). The AbA concentration sufficient to suppress yeast growth transformed with each pAbAi construct was 150 ng/ml.

The positive colonies were transformed with the prey plasmids full-length MADS2-5/pGADT7. The transformed cells were plated on SD agar medium that lacked leucine and uracile (SD/-Leu/-U) in presence of 150 ng/ml AbA and incubated at 30°C for 3-4 days.

The possible DNA-protein interaction is demonstrated by the growth of colonies due to the transcriptional activation of the reporter gene *AUR-1C* that confers AbA resistance. The transformation of empty pGADT7 vectors was used as negative control.

3.10 Protoplast dsRNAi

Buds (~20) of three different *P. aphrodite* plants at B1 (0.5–1 cm) and B2 (1–1.5 cm) stage were collected. For each bud, the inner tepals were collected and used for tissue-specific protoplast isolation.

Flower tissue were cut into 0.5-1.0 mm strips that were immersed in the enzyme solution. The digestion occurs shaking in the dark for ~16 hours. After digestion, the protoplast and enzyme suspension was filtered to remove tissue debris, and the protoplast mixture obtained was undergone to wash steps. After ice incubation, protoplast suspension was centrifuged briefly at 200 g and the pellet was resuspended in a resuspension solution.

To test the protoplast viability, the FDA (Fluorescein diacetate) and PI (Propidium iodide) coloration were used. Living cells are stained green (FDA stained), while dead cells appear red (PI stained) [124].

To conduct a transient RNA interference experiment (RNAi) in the isolated protoplasts, dsRNA was synthetized by *in vitro* transcription. Specific region of *MADS2* genes was amplified with primers flanked by T7 promoter sequence (DreamTaq, Thermo Scientific™) (Table 1-Appendix), and the amplification product was transcribed by T7 RNA polymerase (Thermo Scientific™)[125]. The dsRNA obtained was purified by ethanol precipitation followed by DNase treatment.

Protoplast transfection was performed by adding 20 ul of dsRNA (2 µg) to 200 uL of protoplasts with a concentration of 0.5–2 x 10^5 cells/mL, together with an equal volume of PEG-calcium transfection solution. As a control, non-transfected protoplasts (NT) were used. After 10 minutes of incubation some wash steps were performed [124]. The transfected protoplasts were incubated for 16h at room temperature [125].

Protoplast RNA was extracted, reverse-transcribed, and qPCR analyses were performed as previously described (see Quantitative expression analysis).

4 RESULTS AND DISCUSSION

4.1 Transcription factors differentially expressed in the *Phalaenopsis* lip

The RNA-seq analysis conducted on the inner perianth organs of wild-type and peloric *Phalaenopsis* hyb. "Athens" (Figure 13) showed over 78% of the read pairs mapped to the *P. equestris* genome v 1.0 [119]. In particular, the 68% of the transcripts annotated, about 21,200 gene, are expressed in the floral organs considered with at least 1 TPM (transcripts per kilobase million). The lip-like structure that substitutes the lateral inner tepals in the peloric flower share the 98% of all expressed genes with the lip of the wild-type flower, suggesting that these organs have the same identity (Figure 20).

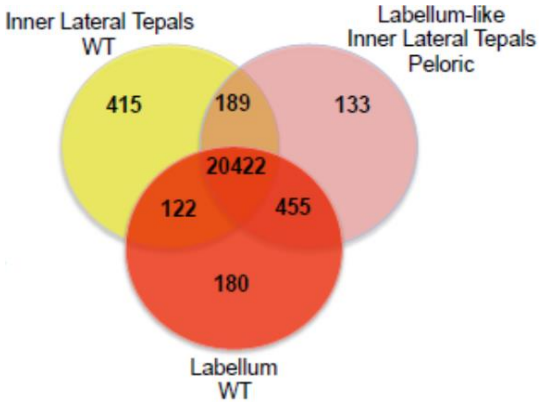


Figure 20. Transcripts expressed in perianth organs of wild-type (WT) and peloric *Phalaenopsis* hyb. "Athens" with at least 1 TPM. The number of transcripts shared and specific for each floral organ are indicated within the circles. From *Lucibelli et al*. (2021) [100].

In silico differential expression analysis shows the presence of some transcripts encoding for TFs, significantly up- or downregulated in wild-type lateral inner tepals compared to the lip. These TFs could have a role in the differentiation of these flower organs.

One of the top differentially expressed transcripts encodes for the DROOPING LEAF-like (DL-like) protein, a YABBY TF. The *DL*-like transcript is up-regulated in wild-type and peloric lip, compared to the wild-type lateral inner tepals. This result is interesting because in other angiosperms the *DL*-like genes are not expressed in the perianth organs. In different flowering plant species,

including the basal angiosperms *Amborella trichopoda* and *Calomba acquatica*, the *DL*-like genes have a conserved expression pattern in the abaxial region of the carpel. Therefore, their role in the abaxial cell fate during carpel development is an ancestral function, conserved during evolution [126].

According to the "orchid code" molecular model [69], *in silico* differential expression analysis shows up-reagultion of the *AP3/DEF*-like *PeMADS2* in wild-type lateral inner tepals. In addition, the *CYC/TB1*-like transcripts, encoding for TCP TFs involved in establishing the dorsal identity of the flower in many angiosperms[9], have low expression (under 1 TPM) in the orchid perianth, as described previously [99].

These results are in agreement with the results of the *in silico* differential expression analysis conducted using publicly available RNA-seq data of the perianth organs of wild-type and peloric mutant *Phalaenopsis* hyb. "Brother Spring Dancer" KHM190 [127]. The reads were mapped and quantified against *Phalaenopsi*s hyb. "Brother Spring Dancer" transcriptome assembled from the Illumina raw reads. The wild-type *Phalaenops*is *DL*-like transcript shows a 3 to 4 log2 fold change (FC) expression in lip compared to the lateral inner tepals. Furthermore, there is not a significative difference of *DL*-like transcript expression between lip and lip-like structures in the peloric mutant *Phalaenopsis* hyb. "Brother Spring Dancer".

In both the *in silico* analyses, the *DL*-like expression pattern in wild-type *Phalaenopsis* is similar to that of the class B *MADS*-box genes *AP3/DEF*-like (*PeMADS2-5*). In particular, it is the same as *PeMADS3* and *PeMADS4* and is opposite to that of *PeMADS2* and *PeMADS5*. This result suggests a possible correlation between these genes during lip development.

The existence of a genetic interaction between *DL*-like and class B MADS-box genes has been hypothesized in other angiosperms [128]. For example, in *Arabidospis* the *CRABS CLAW* gene (*CRC*, homolog of *DL*) is target of the class B AP3/PI proteins, which inhibit its expression during carpel development [129]. A regulatory interaction between the class B and *CRC/DL*-like genes is also demonstrated in the dicot *Physalis floridana*, where the GLO-like DOLL1 and PFGLO2 proteins regulates the *PfCRC* expression by direct binding to the CArG-box motif on the *PfCRC* promoter [130].

There are examples of DL/MADS-box interaction also in monocot plants, in particular in the Poaceae. For example, in *Oryza sativa* there is a correlation between the *DL* gene and the class B

MADS-box gene *SUPERWOMAN1* (*SPW1*). In the rice wild-type flower, the *DL* expression domain is confined to the whorl 4 regulating carpel development, while *SPW1* is expressed in whorl 2 and 3, as other class B *MADS* genes, where it induces the lodicule and stamen differentiation, respectively. In the *dl* mutants, the *SPW1* expression area expands to whorl 4, transforming carpels into stamens. In *spw1* mutants, *DL* expands its function to the whorl 3, converting stamens into carpels. In *spw1*/*dl* double mutants, carpels and stamens are missing, demonstrating that the two genes together are responsible for the identity of the organs in whorl 3 and 4 and belong to the same regulatory network [131]. In *Zea mays* there is a similar antagonistic activity between *silky1*, the ortholog of the class B *AP3/DEF* gene, and the *DL*-like co-orthologs *drl1* and *drl2*. This genetic relationship influences floral pattering and the establishment of flower zygomorphy [132]. Finally, in *Saccharum spontaneum* the CRC-like SsYABBY2 and SsMADS4 proteins directly interact, in addition to the transcriptional regulation. Probably this interaction is connected with the development of the reproductive organs [133].

These evidences suggest that the interaction between the class B *MADS-box* and *DL/CRC*-like genes evolved before the divergence between eudicots and monocots [128].

4.2 Identification of the *Phalenopsis DL*-like genes: gene structure and evolution

In silico analysis of *Phalaenopsis* hyb. "Brother Spring Dancer" transcriptome reveals the presence of two *DL*-like transcripts, *PeDL1* and *PeDL2*, each with two different isoforms. The genomic organization of *PeDL1* and *PeDL2* genes was reconstructed by BLAST analysis using as a query the longest *PeDL* transcripts (*PeDL1_1* and *PeDL2_1*) and as subject the assembled genome of *P. equestris* [119].

PeDL genes are composed by seven exons and six introns (Figure 21).

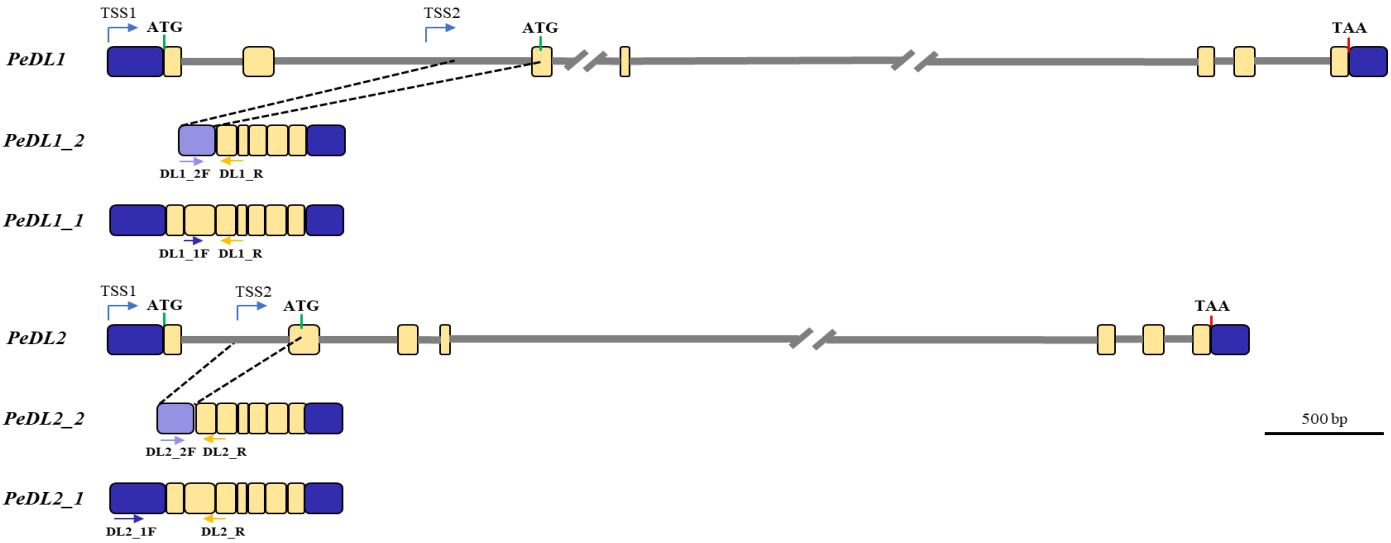


Figure 21. Reconstruction of the genomic organization of the *PeDL1* and *PeDL2* genes and schematic representation of the alternative transcripts. The blue boxes represent the 5'- and 3'-UTRs; the yellow boxes represent the coding regions, the grey lines represent the introns. Introns of unknown size are shown as interrupted lines. The green and red bars indicate the position of the translation start (ATG) and stop (TAA) codons, respectively. TSS1 and 2 are the putative alternative transcription start sites of the different isoforms. The blue and yellow arrows indicate the position of the isoform-specific primers pairs. From *Lucibelli et al.* (2021) [100].

Intron 4 is the largest and richest in repetitive sequences. This feature affected the *PeDL*1 and *PeDL2* genes correct assembly in the *P. equestris* genome. Both genes are split in two different genomic scaffolds (Scaffold000061_46 and Scaffold000061_45 for *PeDL1*; Scaffold000404_23 and Scaffold000404_21 for *PeDL2*).

Manual alignment of the short *PeDL* transcripts (*PeDL1_2* and *PeDL2_2*) against the respective scaffolds, allowed the identification of putative alternative transcription start sites for both transcripts that generate two alternative isoforms. *PeDL1_2* transcription start site is present within the intron 2 of the *PeDL1* gene and ATG start codon is within exon 3. The transcription start site of *PeDL2_2* is within intron 1 of the *PeDL2* gene, while the start codon is in exon 2 (Figure 21).

To verify the presence of these four *PeDL* transcripts and validate their sequence a PCR amplification of cDNA from perianth tissues of *Phalaenopsis* hyb. "Athens" was performed, followed by cloning and sequencing, and the sequences were deposited in GenBank with the accession numbers MW574592, MW574593 (*PeDL1_1* and *PeDL1_2*), MW574594, and MW574595 (*PeDL2_1* and *PeDL2_2*).

The YABBY TFs have a Cys2Cys2 zinc-finger domain at the N-terminus, a central portion rich in serine and proline, forming a characteristic transcription factor activation domain, and a helix-loop-helix domain at the C-terminus, the YABBY domain, resembling an HMG (high mobility group) domain, responsible for DNA binding [134].

The *PeDL1_1* and *PeDL2_1* transcripts encode for proteins composed by a Cys2Cys2 zinc-finger domain at N-terminus and a helix-loop-helix domain at C-terminus, which corresponds to the YABBY domain, separated by an intermediate domain. Both the alternative isoforms encode for proteins missing in part (PeDL2_2) or completely (PeDL1_2) the Cys2Cys2 zinc-finger domain (Figure 22).

Figure 22. Amino acid alignment of the PeDL proteins. The Cys2Cys2 zinc-finger and the YABBY domains are underlined. The asterisks indicate the variable residues within the Cys2Cys2 and YABBY domains of PeDL1_1 and PeDL2_1. From *Lucibelli et al.* (2021) [100].

The PeDL1_1 (189 aa) and PeDL2_1 (196 aa) proteins have 64.3% similarity. The YABBY domain is highly conserved in comparison to the Cys2Cys2 zinc-finger domain, which is more variable. The intermediate domain is the less conserved part of these proteins.

Unfortunately, analysis of different orchid species transcriptomes available in the orchid-specific database Orchidstra 2.0 [107] and OrchidBase 2.0 [106] have not detected the presence of *DL* alternative isoforms as observed in *Phalaenopsis*. A possible explanation could be that these transcriptomes derive from whole inflorescence, with possible under-representation of isoforms expressed in specific tissues. In contrast, in *Z. mays* two isoforms of the *DL* genes, *drl1* (https://maizegdb.org/gene_center/gene/GRMZM2G08830) and *drl2* (https://www.maizegdb.org/gene center /gene/GRMZM2G102218, access date 18 January 2021), are annotated. Interestingly, the *drl2* predicted alternative isoform encodes a small protein missing the Cys2Cys2 zinc-finger domain, as for the alternative transcripts of *Phalaenopsis*. Currently, the role of the *drl* isoforms are still unknown due to the absence of functional or expression data. Further analysis are needed to investigate their role.

Genome- and transcriptome-wide identification of the *DL*-like genes of different orchid species reveal that most orchids have two *DL*-like genes, as *Phalaenopsis*, except for the *Apostasia*, which has a single *DL*-like gene, in agreement with previous work [135]. The genomic organization of the *Phalaenopsis DL*-like genes was compared with that of *D. catenatum* and *A. shenzhenica* (Figure 23).

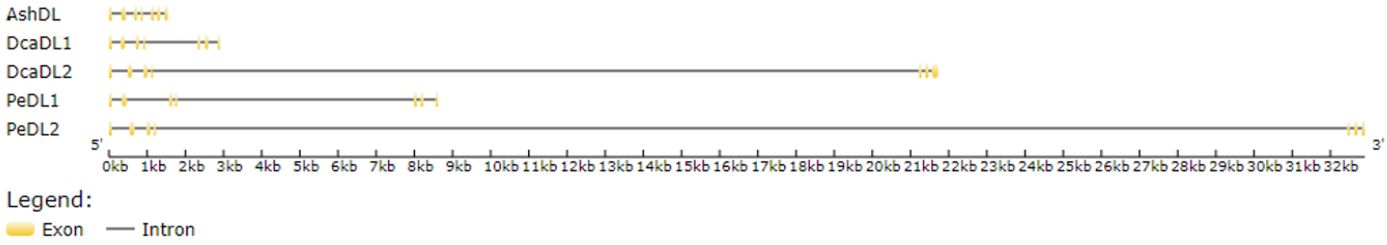
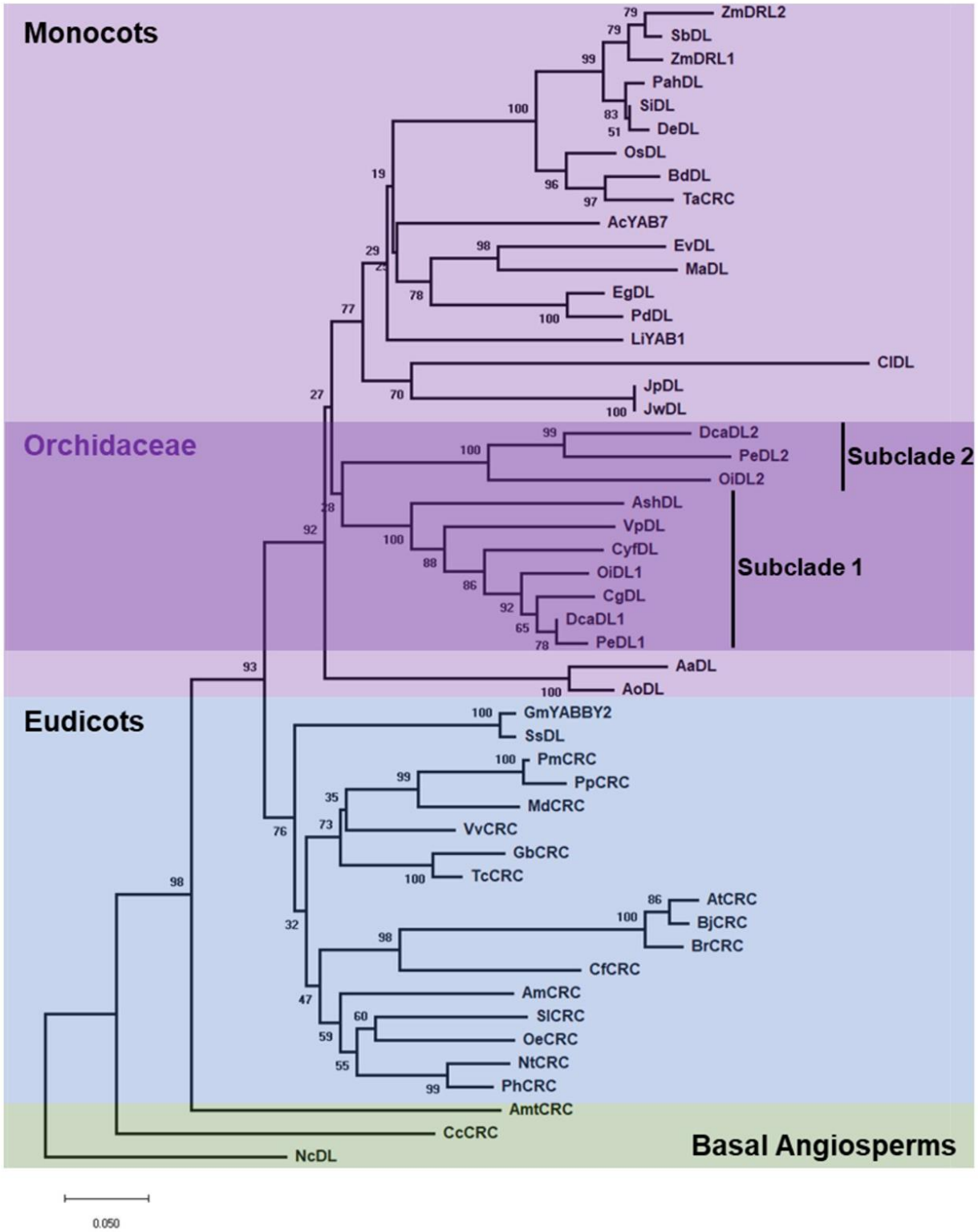


Figure23. Genomic organization of orchid *DL*-like genes.

The *DL*-like genes of *A. shenzhenica*, *D.catenatum*, and *P. equestris* exhibit a conserved genomic organization with seven exons and six introns. The exon size is conserved, while intron size is variable, due to the presence of many transposable elements characteristic of the orchid genomes.

Phylogenetic analysis of the orchid DL-like proteins reveals the presence of two subclades. The *Apostasia* DL-like protein belongs to subclade I. This result suggests that the two orchid subclades derived from an Orchidaceae-specific duplication event that occurred after the divergence of subfamilies Apostasioideae and Vanilloideae (Figure 24).

A.


Monocots
ZmDRL2
SbDL
ZmDRL1
PahDL
SiDL
DeDL
OsDL
BdDL
TaCRC
AcYAB7
EvDL
MaDL
EgDL
PdDL
LiYAB1
ClDL
JpDL
JwDL
Orchidaceae
DcaDL2
PeDL2
OiDL2
Subclade 2
AshDL
VpDL
CyfDL
OiDL1
CgDL
DcaDL1
PeDL1
Subclade 1
AaDL
AoDL
Eudicots
GmYABBY2
SsDL
PmCRC
PpCRC
MdCRC
VvCRC
GbCRC
TcCRC
AtCRC
BjCRC
BrCRC
CfCRC
AmCRC
SlCRC
OeCRC
NtCRC
PhCRC
AmtCRC
CcCRC
NcDL
Basal Angiosperms
0.050

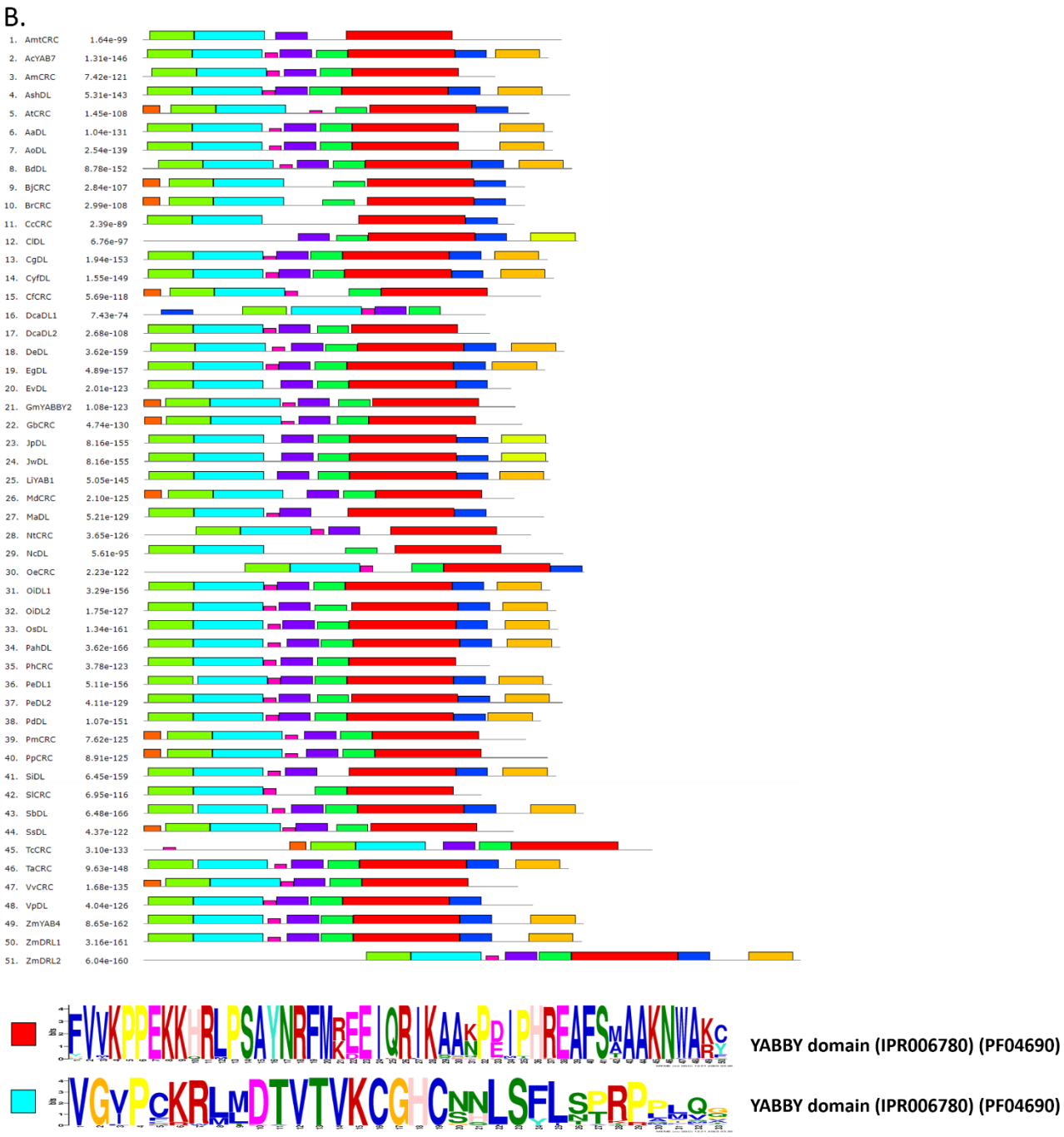


Figure 24. Neighbor-Joining tree of the DL proteins. The tree was constructed on the DL amino acid alignment of selected monocots and dicots (mainly from Chen et al., 2020 [135]) using MEGA software. The sequences were aligned using Clustal Omega. The numbers above the branches represent the bootstrap percentages (500 replicates) (A); conserved motifs of YABBY proteins (B); IPR006780, accession YABBY DOMAIN InterPro; PF04690, accession YABBY DOMAIN PFAM**. Monocots**: AcYAB7 (*Ananas comosus*, XP_020105063), AshDL (*Apostasia shenzhenica*, PKA49723), AaDL (*Asparagus asparagoides*, BAI68347), AoDL (*Asparagus officinalis*, XP_020276732), BdDL (*Brachypodium distachyon*, KQK22884), ClDL (*Carex littledalei*, KAF3326740), CgDL (*Cymbidium goeringii*, ADI58463), CyfDL (*Cypripedium formosanum*, CFTC009810), DcDL1 and DcDL2 (*Dendrobium catenatum*, PKU69929 and PKU86051, respectively), DeDL (*Digitaria exilis*, CAB3496665), EgDL (*Elaeis guineensis*,

XP_010939213), EvDL (*Ensete ventricosum*, RRT32785), JpDL (*Juncus prismatocarpus*, BAJ14106), JwDL (*Juncus wallichianus*, BAJ14110), LiYAB1 (*Lilium longiflorum*, ABP35569), MaDL (*Musa acuminata*, XP_009395243), OiDL1 and OiDL2 (*Orchis italica*, OITC006016 and comp7559, respectively), OsDL (*Oryza sativa*, AY494713), PahDL (*Panicum hallii*, XP_025797474), PeDL1 and PeDL2 (*Phalaenopsis equestris*, MW574592 and MW574594, respectively), PdDL (*Phoenix dactylifera*, XP_008785629), SiDL (*Setaria italica*, XP_012704688), SbDL (*Sorghum bicolor*, XP_021307124), TaCRC (*Triticum aestivum*, AAQ11881), VpDL (*Vanilla planifolia*, VPTC001074), ZmDRL1 and ZmDRL2 (*Zea mays*, GRMZM2G088309 and GRMZM2G102218, respectively). **Eudicots**: AmCRC (*Antirrhinum majus*, AAS10180), AtCRC (*Arabidopsis thaliana*, NP_177078.1), BjCRC (*Brassica juncea*, AAZ23116.1), BrCRC (*Brassica rapa*, XP_009105464), CfCRC (*Cynophalla flexuosa*, AAW83045), GmYABBY2 (*Glycine max*, XP_003517857), GbCRC (*Gossypium barbadense*, KAB2041458), MdCRC (*Malus domestica*, XP_008339784), NtCRC (*Nicotiana tobacum*, AAW83046), OeCRC (*Olea europaea*, XP_022846429), PhCRC (*Petunia x hybrid*, AAW83048), PmCRC (*Prunus mume*, XP_008243820), PpCRC (*Prunus persica*, XP_007223999.2), SlCRC (*Solanum lycopersicum*, XP_004239032.1), SsDL (*Spatholobus suberectus*, TKY69532), TcCRC (*Theobroma cacao*, EOY01637), VvCRC (*Vitis vinifera*, XP_010650015). **Basal Angiosperms**: AmtCRC (*Amborella trichopoda*, CAI47004.1), CcCRC (*Cabomba caroliniana*, BAJ83622), NcDL (*Nymphaea colorata*, XP_031490637). Modified from *Lucibelli et al.* (2021) [100].

Generally, duplicated genes can be lost or undergo functional diversification that increases the organism biological complexity [16]. The differential expression of a *DL*-like transcripts in *Phalaenopsis* lip, revealed by *in silico* analysis, suggests a possible recruitment of this duplicated gene in the development of this flower organ.

As in other angiosperms the expression of *DL*-like genes has never been detected in perianth organs, it is possible to hypothesize that, after an orchid-specific duplication event, there was a sub- or neo-functionalization of the duplicated *DL*-like gene. The presence of only one *DL*-like member in the ancestral orchid *A. shenzhenica* (Apostasioideae), which has an actinomorphic flower, supports a possible involvement of the duplicated *DL*-like gene in lip differentiation and the consequent determination of bilateral symmetry.

In addition to the conserved role in carpel development *DL*/*CRC*-like genes have acquired novel functions also in other angiosperms [128, 136]. For example, the *CRC*-like genes play a role in the formation of the nectary in rosids and asterids, such as *A. thaliana* and *Nicotiana tabacum*. However, no expression in nectaries is observed in basal eudicots, suggesting that this function is originated after the divergence from the latter [137, 138]. Studies on an early-diverging eudicot, *Eschscholzia californica*, show a conserved expression in the gynoecium and a novel role in placenta development

and ovule initiation [139]. Furthermore, the grasses *DL*-like genes promote the determination of the carpel identity and the midrib formation[128]. The *DL*-like genes leaf expression could have been acquired in the monocot lineage, as demonstrated by the presence of *DL* transcripts in the leaves of *Asparagus asparagoides* and *Lilium longiflorum* [136, 140, 141]. The innovative functions acquired by the *DL*/*CRC* homologs during angiosperms evolution could derive from changes in *cis*-regulatory sequence that influences their expression profile, alteration of the coding region and establishment of new regulatory networks [122, 142].

4.3 Expression pattern of *DL*- and *DEF*-like genes of wild-type and peloric orchids

4.3.1 Differential expression of the *PeDL1* and *PeDL2* genes

Quantitative real-time PCR experiments were performed to verify the expression profile of *PeDL1* and *PeDL2*. The reactions were conducted on cDNA obtained from different floral tissue of wild-type *Phalaenopsis* hyb. "Athens" collected at B2 stage (1-1.5 cm) (Figure 18D).

The expression pattern of the two *PeDL* genes in the perianth organs does not completely overlap. The different isoforms of *PeDL* (*PeDL1_1*, *PeDL1_2*, *PeDL2_1*, *PeDL2_2*) are expressed in the column and ovary, confirming their conserved role in the reproductive organs. Nevertheless, the *PeDL2* transcripts are more expressed in the lip than in the lateral inner tepals of the wild-type flower, confirming the results of the *in silico* differential expression analysis (Figure 25).

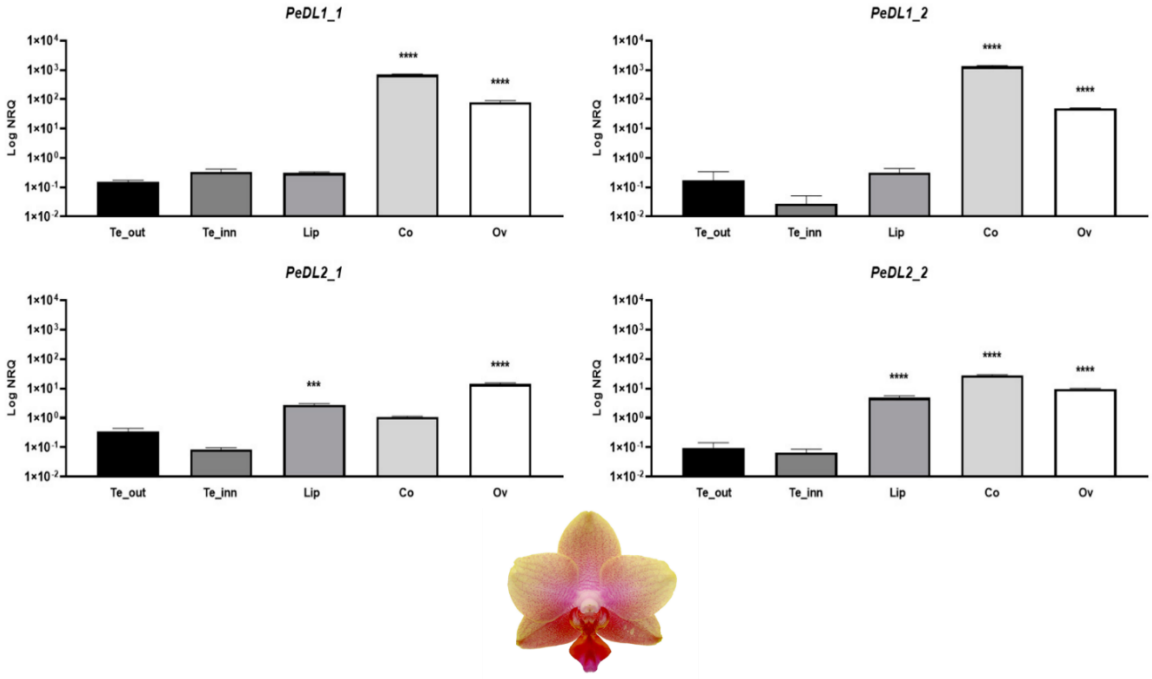


Figure 25. Relative expression of the different isoforms of the *PeDL1* and *PeDL2* genes of wild-type *Phalaenopsis* hyb. "Athens" floral organs at the B2 developmental stage (1–1.5 cm). The expression is reported as logarithm of the normalized relative quantity (Log NRQ). The bars represent the SEM of the biological and technical replicates. The asterisks indicate the statistically significant difference of the expression compared to outer tepals. p-Values *** <0.001, **** <0.0001. Te_out, outer tepals; Te_inn, lateral inner tepals; Co, column; Ov, ovary. Modified from *Lucibelli et al.* (2021) [100].

A time course expression analysis was conducted to verify the expression profile of the *Phalaenopsis DL*-like genes through the different stages of flower development. Real-time RT-PCR was performed on the perianth tissues of *P. aphrodite* collected at different floral bud size and after the anthesis (Figure 18A-B).

The differential expression of the *DL2_1* isoform between the lip and the internal tepals is particularly evident in the early stages of flower development. During the subsequent stages it decreases, with a statistically significant negative correlation between expression level and stage (Spearman correlation r = -1, p = 0.0028). This result suggests a potential role of this transcript in the initial phases of the lip morphogenesis (Figure 26).

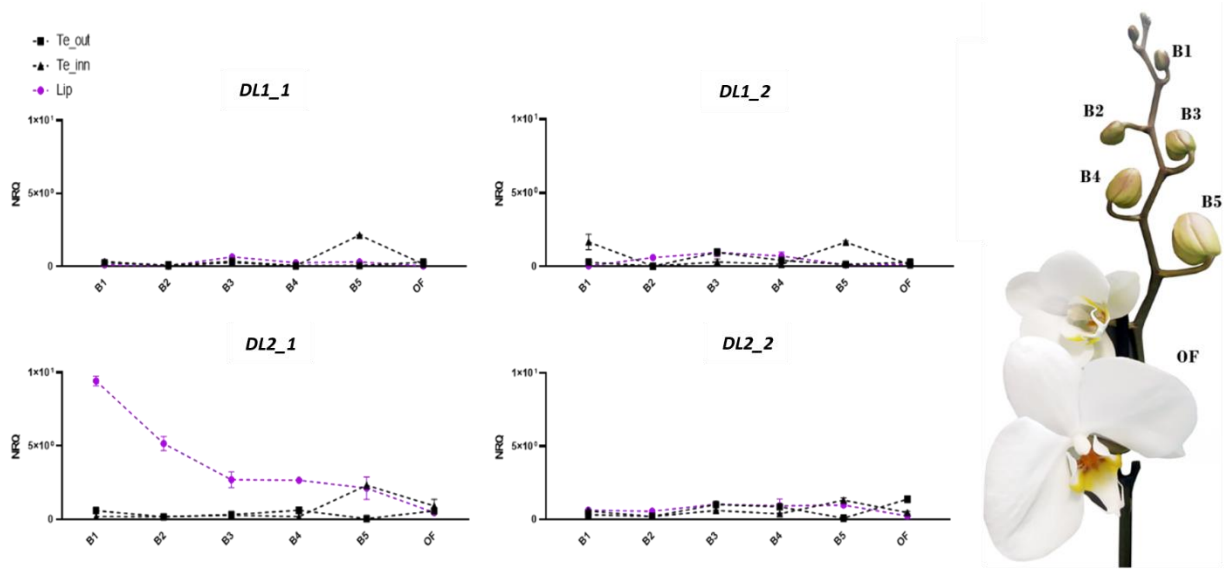


Figure 26. Relative expression of the different *DL* isoforms in the wild-type *P. aphrodite* perianth at different developmental stages. The expression is reported as normalized relative quantity (NRQ). The bars represent the SEM of the biological and technical replicates. Bud size of the developmental stages: B1 (0.5–1 cm), B2 (1–1.5 cm), B3 (1.5–2 cm), B4 (2–2.5 cm), B5 (2.5–3 cm), OF (open flower). Te_out, outer tepals; Te_inn, lateral inner tepals. Modified from *Lucibelli et al.* (2021) [100].

A further suggestion of the recruitment of the *DL2* gene in the specification of the lip is given by the expression analysis of both the *PeDL2* isoforms (*PeDL2_1* and *PeDL2_2*) conducted on the perianth tissues of the peloric mutant orchids *Phalaenopsis* hyb. "Athens" and *Phalaenopsis* hyb. "Joy Fairy Tale", which have lost bilateral symmetry due to the presence of three lips in the second floral whorl.

At the floral bud stage B2 (1-1.5cm) of the the peloric mutant *Phalaenopsis* hyb. "Athens", *PeDL2* gene increases its expression level in the lip-like structure compared to the wild-type lateral inner tepals. In particular, the mean difference of the *PeDL2_1* expression between wild-type lateral inner tepals and lip is −2.71 and decrease to −1.93 between lip-like structures and lip in the peloric mutant. At the same time, for *PeDL2_2* the mean difference of expression between wild-type lateral inner tepals and lip is −4.82 and is reduced to −0.83 between lip-like structures and lip (Figure 27).

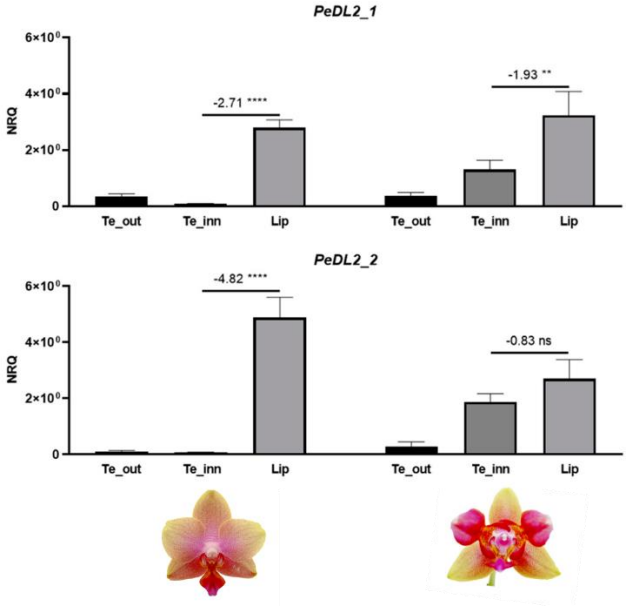


Figure 27. *PeDL2_1* and *PeDL2_2* relative expression in the perianth tissue of the wild-type and peloric *Phalaenopsis* hyb. "Athens" at the B2 developmental stage (1–1.5 cm). The expression is reported as normalized relative quantity (NRQ). The vertical bars represent the SEM of the biological and technical replicates. The numbers above the horizontal lines are the mean differences of the expression between lateral inner tepals and lip (Te_inn - Lip). p-Values ** <0.01, **** <0.0001; ns, not significant. Te_out, outer tepals; Te_inn, lateral inner tepals or lip-like structures that substitute the lateral inner tepals in the peloric mutant. Modified from *Lucibelli et al.* (2021) [100].

In the peloric mutant *Phalaenopsis* "Joy Fairy Tale", the isoforms *PeDL2_1* and *PeDL2_2* do not have a significant difference in expression between the lip-like structure and lip (Figure 28).

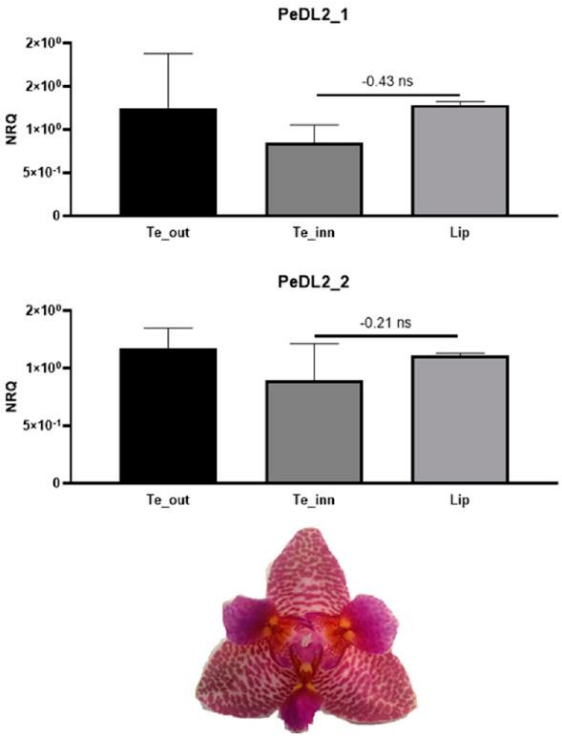


Figure 28. *PeDL2_1* and *PeDL2_2* relative expression in the perianth tissue of the peloric *Phalaenopsis* hyb. "Joy Fairy Tale" at the B2 developmental stage (1–1.5 cm). The expression is reported as normalized relative quantity (NRQ). The vertical bars represent the SEMs of the biological and technical replicates. The numbers above the horizontal lines are the mean differences of the expression between lip-like structures and lip (Te_inn-Lip). ns, not significant. Te_out, outer tepals; Te_inn, lip-like structures that substitute the lateral inner tepals in the peloric mutant. From *Lucibelli et al.* (2021) [100].

The isoforms *PeDL1_ 1* and *PeDL1_1* do not have a significant difference in expression among the perianth tissue of wild-type and peloric *Phalaenopsis* orchids (Figure 29). This result suggests that, unlike *PeDL2*, the paralog gene *PeDL1* does not play a role in the differentiation of the orchid perianth.

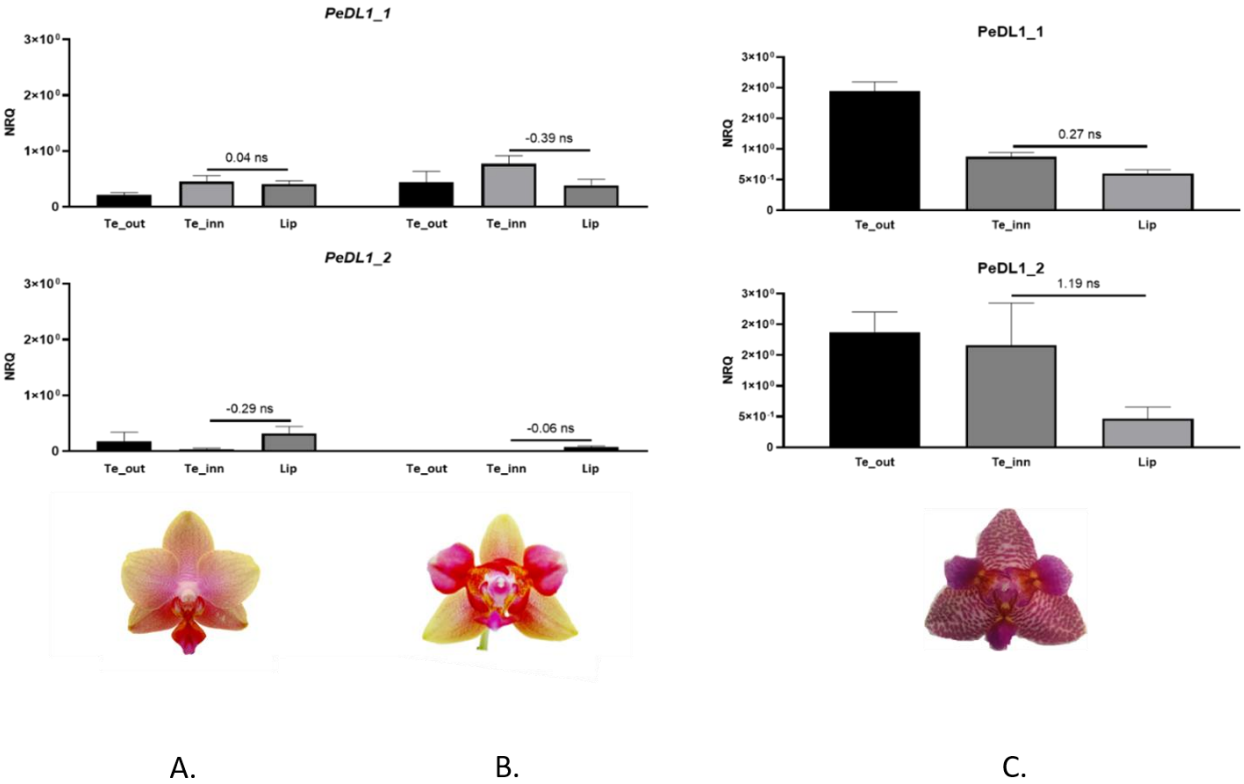


Figure 29. *PeDL1_1* and *PeDL1_2* relative expression in the perianth tissue of the wild-type (A) and peloric (B) *Phalaenopsis* hyb. "Athens" and of the peloric *Phalaenopsis* hyb. "Joy Fairy Tale" (C) at the B2 developmental stage (1-1.5 cm). The expression is reported as normalized relative quantity (NRQ). The vertical bars represent the SEM of the biological and technical replicates. The numbers above the horizontal lines are the mean differences of the expression between lateral inner tepals and lip (Te_inn - Lip). ns, not significant. Te_out, outer tepals; Te_inn, lateral inner tepals or lip-like structures that substitute the lateral inner tepals in the peloric mutant. From *Lucibelli et al.* (2021) [100].

4.3.2 Differential expression of the *PeMADS2-PeMADS5* genes in *Phalaenopsis* perianth

As highlighted in before, the regulatory relationship between the *DL*-like and the class B *MADS*-box genes has been demonstrated or hypothesized in other angiosperms [128]. The *PeDL2* expression profile supports the hypothesis that this gene could enrich the molecular model of the "orchid code" contributing in the specification of the lip through a cooperation with the *DEF*-like genes.

Real-time PCR experiments were performed to verify the expression pattern of the *Phalaenopsis DEF*-like *MADS*-box genes *PeMADS2-PeMADS5* and compared it with that of *PeDL2*. In particular, the *PeMADS2-PeMADS5* expression was examined in perianth tissues of wild-type and peloric *Phalaenopsis* hyb. "Athens" (Figure 18D-E) and of the peloric *Phalaenopsis* hyb. "Joy Fairy Tale" (Figure 18F). According to the "orchid code", in wild-type *Phalaenopsis* the clade 1 and 2 genes (*PeMADS2* and *PeMADS5*) show an high expression in the lateral inner tepal compared to the lip. In contrast, the clade 3 and 4 genes (*PeMADS3* and *PeMADS4*) are more expressed in the lip than in the lateral inner tepals.

qPCR experiments on the peloric mutant *Phalaenopsis* hyb "Athens" and *Phalaenopsis* hyb. "Joy Fairy Tale" confirmed the role of this gene in the determination of the perianth organ identity. In the peloric *Phalaenopsis* "Athens" there is a reduction of the mean difference between the expression levels of the four *DEF*-like genes (*PeMADS2-PeMADS5*) in the inner perianth tissues. In particular, in the lip-like structures the *PeMADS2* and *PeMADS5* expression decreases compared to the wild-type lateral inner tepals, while the *PeMADS3* and *PeMADS4* expression increases (Figure 30).

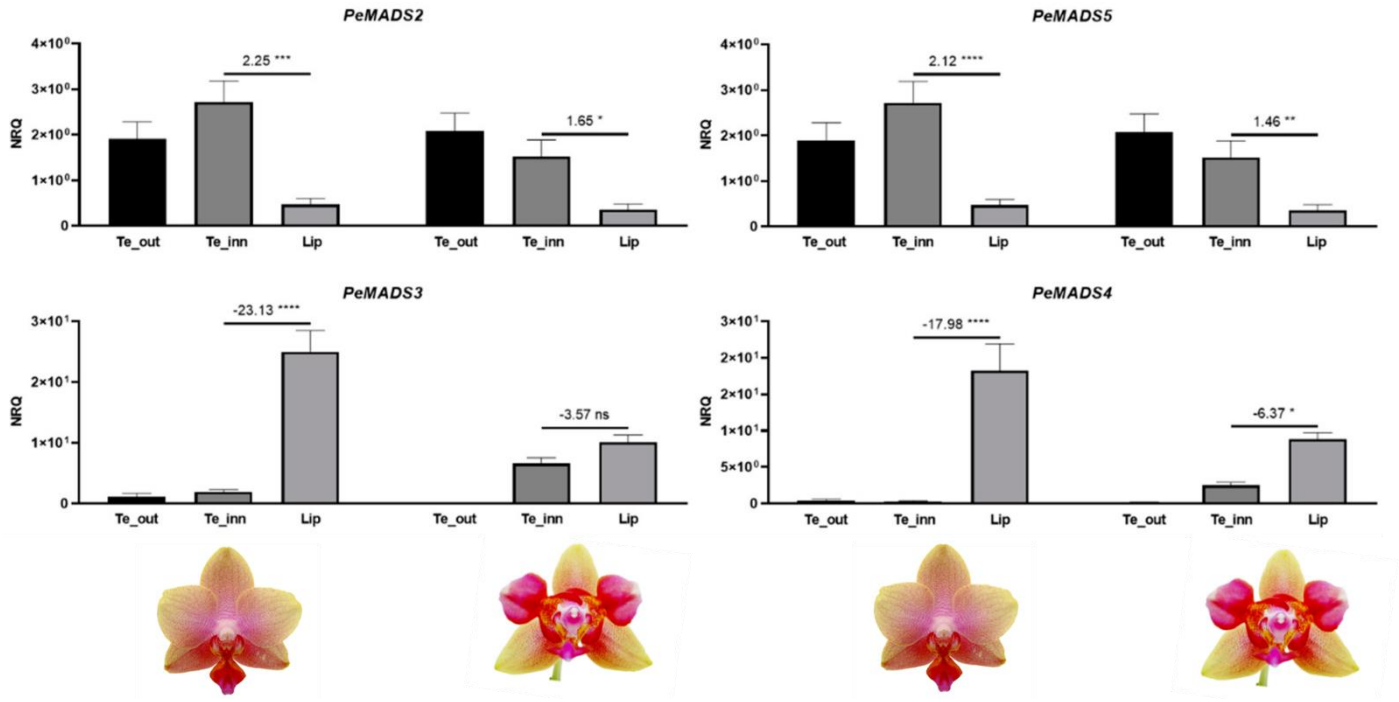


Figure 30. *DEF*-like genes *PeMADS2-5* relative expression in the perianth tissue of *Phalaenopsis* hyb. "Athens" wild-type and peloric mutant at the B2 developmental stage (1–1.5 cm). The expression is reported as normalized relative quantity (NRQ). The vertical bars represent the SEM of the biological and technical replicates. The numbers above the horizontal lines are the mean differences of the expression between lateral inner tepals and lip (Te_inn-Lip). p-Values * <0.05, ** <0.01, *** <0.001, **** <0.0001; ns, not significant. Te_out, outer tepals; Te_inn, lateral inner tepals or lip-like structures that substitute the lateral inner tepals in the peloric mutant. Modified from *Lucibelli et al.* (2021) [100].

In the peloric *Phalaenopsis* hyb. "Joy Fairy Tale", there is not a significant difference in the expression level between the lip-like structures and lip, except for *PeMADS4*, which is more expressed in the lip-like structures (Figure 31).

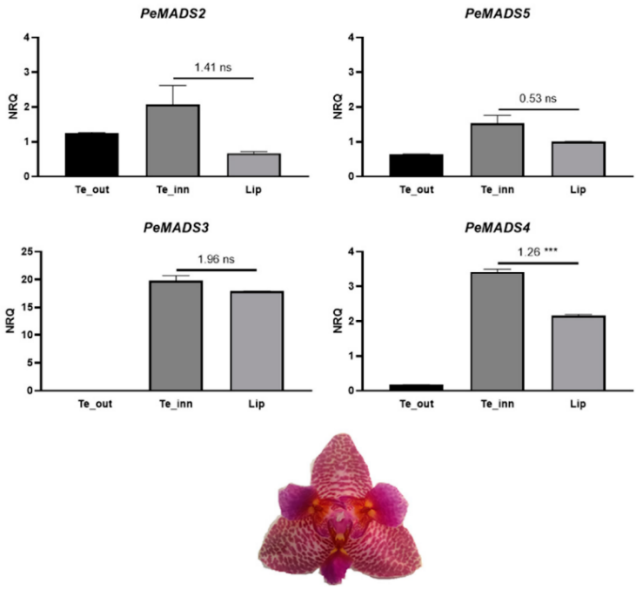
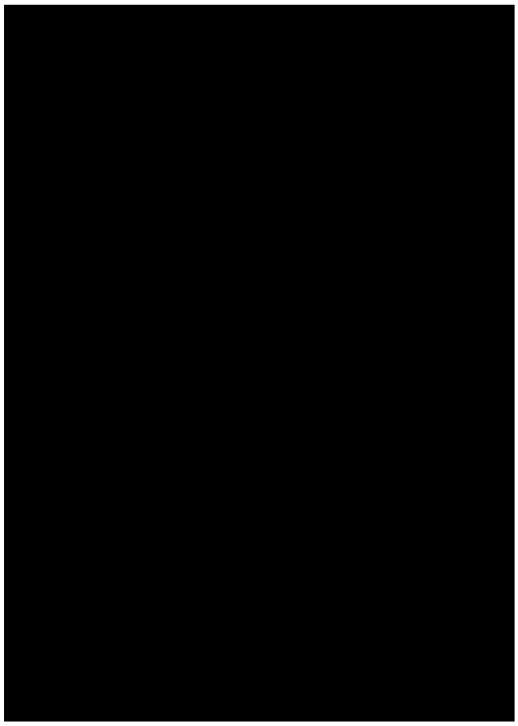
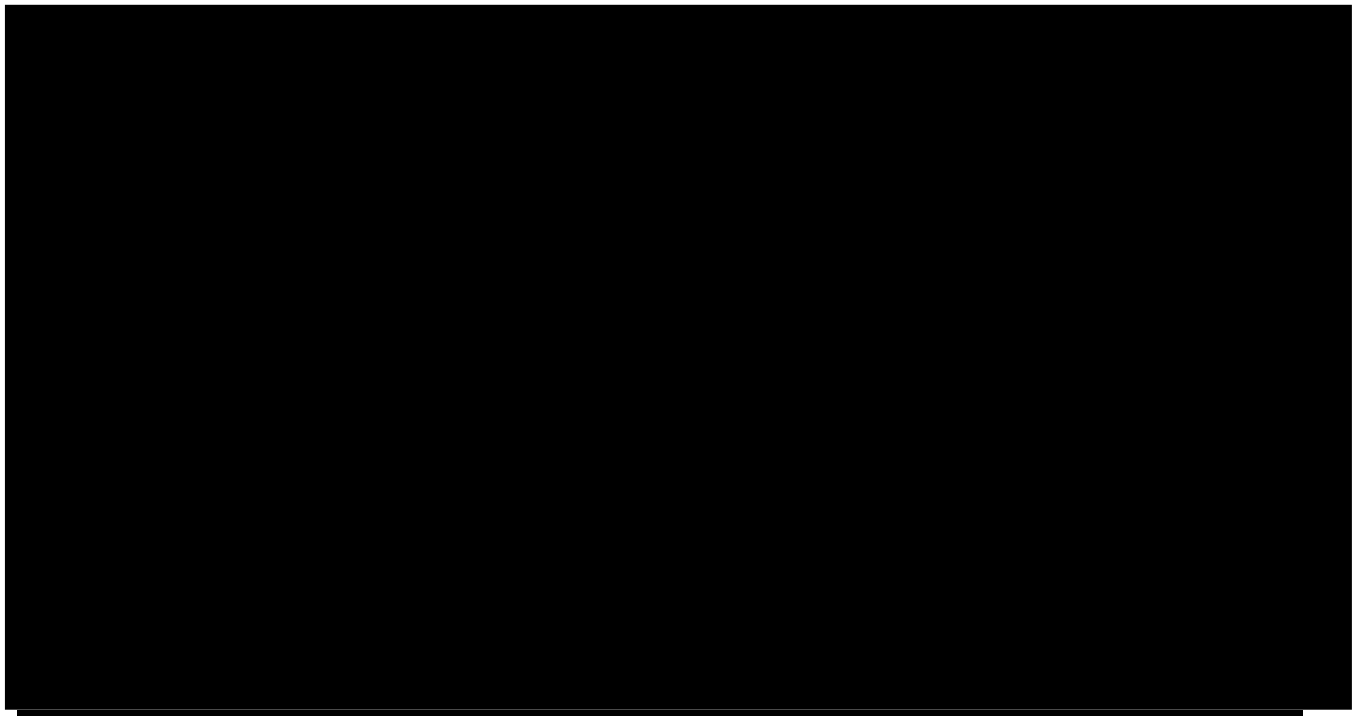


Figure 31. *DEF*-like genes *PeMADS2–5* relative expression in the perianth tissue of *Phalaenopsis* hyb. "Joy Fairy Tale" at the B2 developmental stage (1–1.5 cm). The expression is reported as normalized relative quantity (NRQ). The vertical bars represent the SEM of the biological and technical replicates. The numbers above the horizontal lines are the mean differences of the expression between lateral inner tepals and lip-like structures (Te_inn-Lip). p-Values *** <0.001; ns, not significant. Te_out, outer tepals; Te_inn, lip-like structures that substitute the lateral inner tepals in the peloric mutant. Note the different scale for *PeMADS3.* From *Lucibelli et al.* (2021) [100].

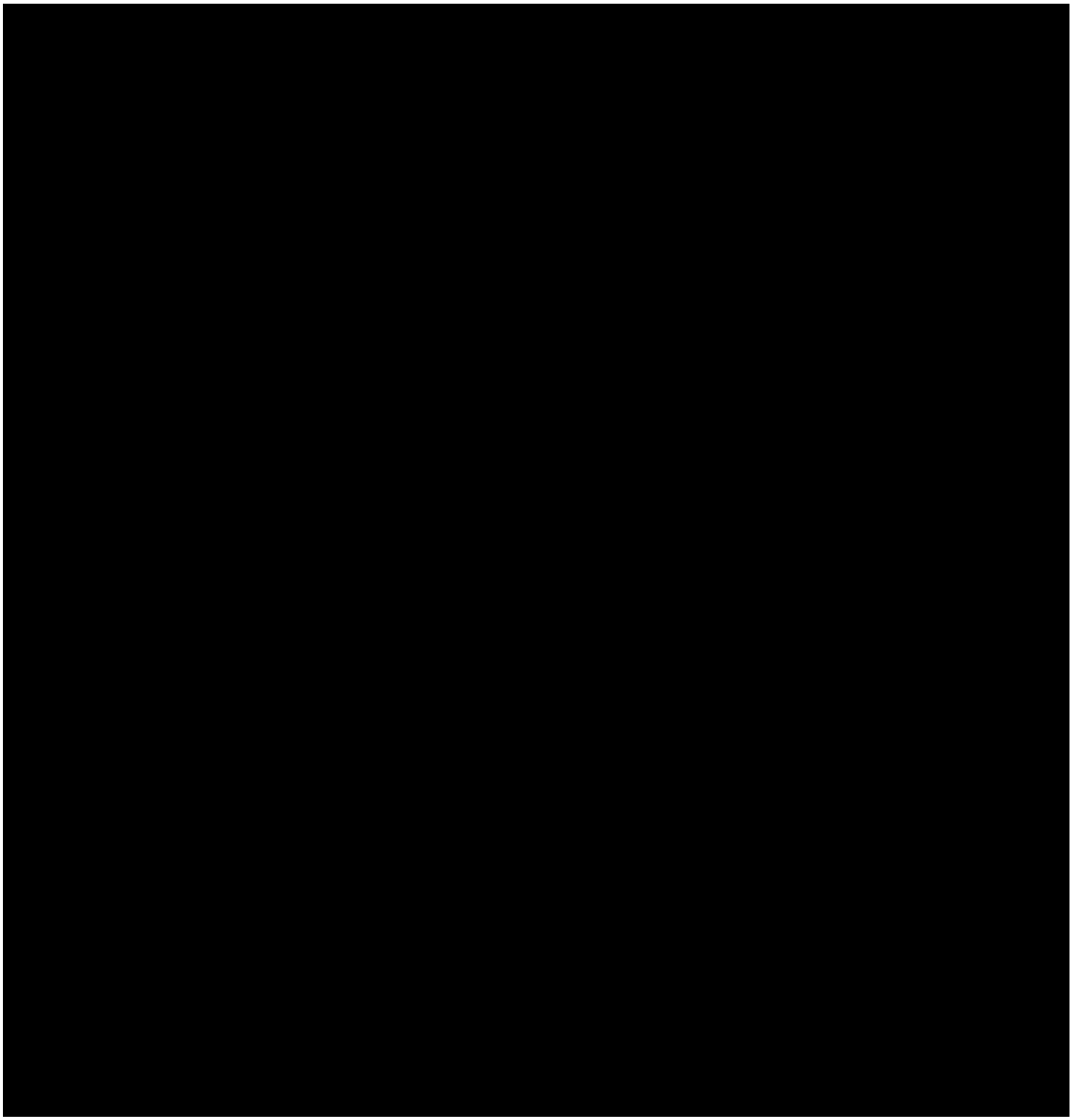
The expression analysis shows that that in wild-type *Phalaenopsis* the genes *PeMADS2* and *PeMADS5* have an expression profile opposite to *PeDL2*, while *PeMADS3* and *PeMADS4* genes are mainly expressed in the lip together with *PeDL2*, confirming the *in silico* differential expression results.





.

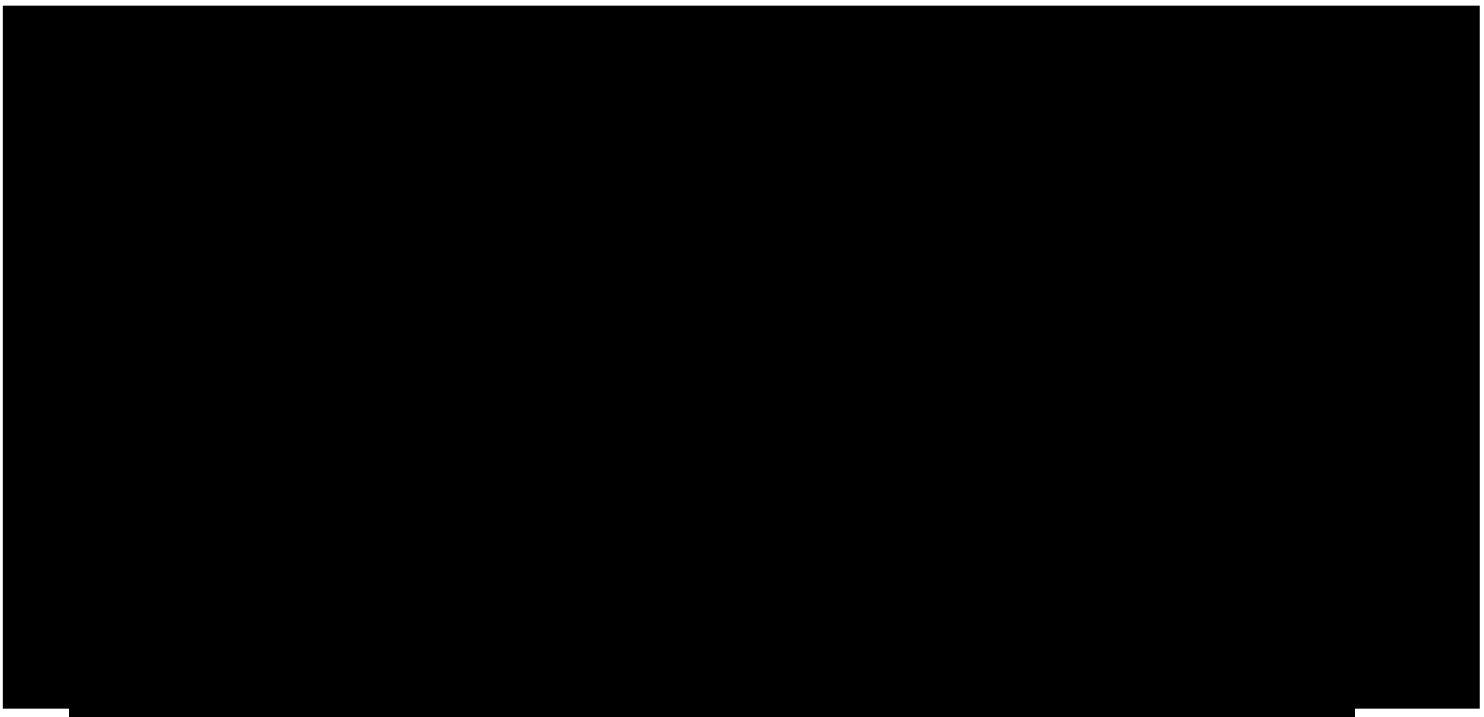
[redacted]



[redacted]

.

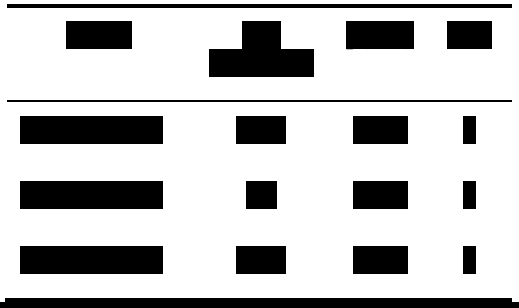
.



[redacted]

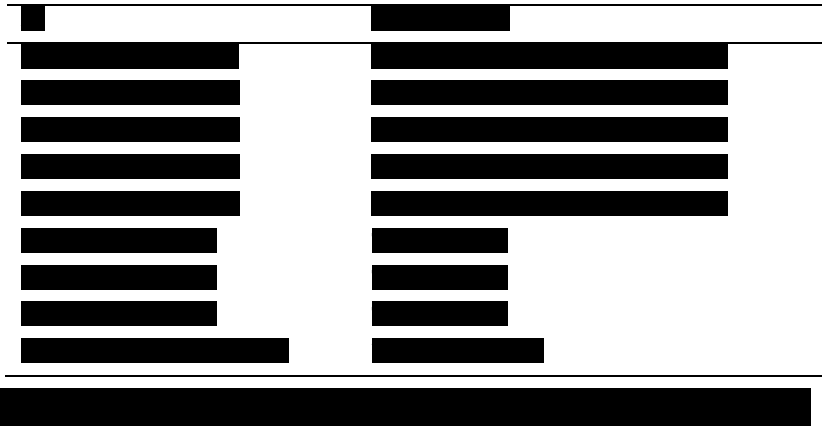
[redacted]

[redacted]

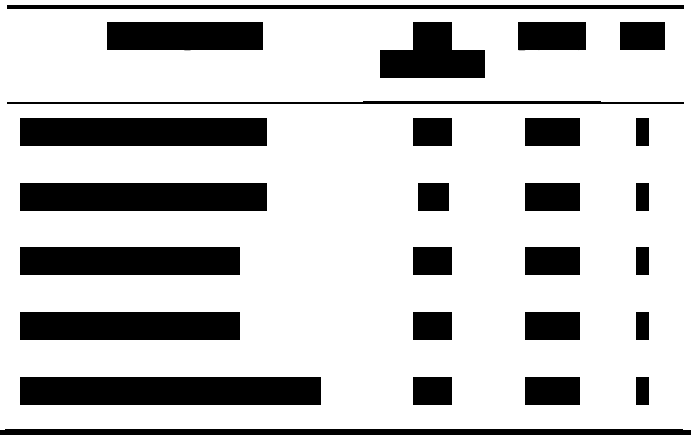


[redacted]

[redacted]



[redacted]



4.5.4 *In vivo* PeDL protein-protein interactions

The Yeast Two-Hybrid (Y2H) analysis was performed to test the possible protein-protein interactions of PeDL2 with other proteins to verify the *in silico* analysis results.

No protein-protein direct interaction was observed between PeDL2 and PeMADS2-5 (Figure 47A). The ability of PeDL2 to bind the class B GLO protein PeMADS6, equally expressed in all the perianth organs, was also checked, revealing the absence of direct interaction (Figure 47A). These results are coherent with the *in silico* results, although the ability of PeDL2 to directly bind any of the PeMADS2-5 proteins cannot be excluded, as it could require the formation of a multimeric protein complex.

In contrast to the *Arabidopsis* CRC, both the isoforms of PeDL1 and PeDL2 are not able to form homo- nor heterodimers in all the possible combinations (Figure 47 C), in agreement with a recent study on the *DL*-like genes of *Phalaenopsis* [148]. This result indicates that the ability of DL proteins to form homo- and heterodimer is not conserved among plants, probably for the sequence divergence after gene duplication.

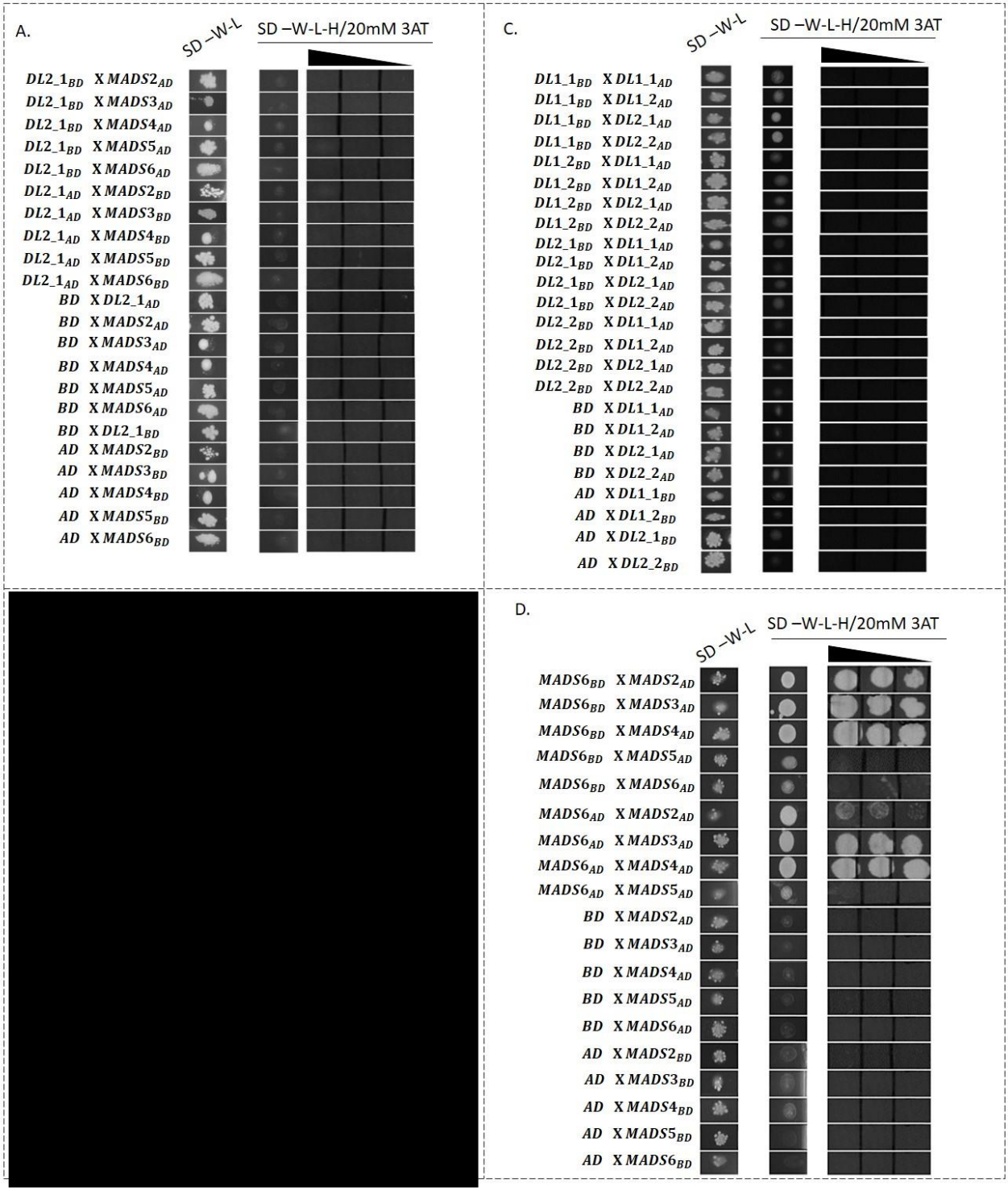


Figure 47. Y2H analysis. Interactions of PeDL2 and (A) PeMADS2-6, (B) PeCYC1-3; (C) interaction of the different isoforms of PeDL1/2; (D) interaction of PeMADS2-5 and PeMAD6 as positive control. CYC1 factor autoactivation was observed (B). After double transformations, yeast growth in absence of tryptophan and leucine (SD -W-L) indicates the plasmid presence; yeast growth in medium lacking tryptophan, leucine and histidine (SD -W-L-H) indicates interaction between the two tested proteins. Double transformations conducted using one of the vectors empty are negative controls. 1:10, 1:100 and 1:1000 are the dilution factor applied to the yeast inoculate. BD, GAL4 DNA-binding domain (pGBKT7 vector); AD, GAL4 activation domain (pGADT7 vector). Modified from *Lucibelli et al.* (2021) [100].

Protein interaction may not occur in yeast, for example, due to the lack of a chaperone for proper protein folding, which is only present in the specific proteins host, or if a protein does not undergo all the appropriate post-translational modifications in yeast. Currently, we are evaluating the PeDL interactions by Bimolecular Fluorescence Complementation (BiFC) to validate the *in silico* prediction.

4.6 *DL2* transcriptional regulation

4.6.1 *In silico* identification of conserved regulatory motifs within the *DL2* putative promoter

To understand which TF could regulate the *DL2* expression, *in silico* analysis of the *DL2* putative promoters of *P. equestris* (*PeDL2*) and *D. catenatum* (*DcDL2*) was performed to identify TFBSs conserved between the two orchid species, both belonging to the Epidendroideae subfamily.

Nucleotide sequences 3000 bp upstream of the *DL2* translation start site were downloaded and scanned using the MEME suite online tool [110]. Two motifs (Motifs 1 and 3) show a conserved position within the ~300 bp upstream of the translation start site. The reliability of this result is supported by the absence of these motifs in the shuffled sequences of the putative promoters.

The conserved motifs were analysed through TOMTOM in the JASPAR Core Plants database [120]. The Motifs 1 contains a putative binding site for a TCP protein, while Motif 3 presents a putative binding site for an SBP-type zinc-finger protein (Figure 48).

Promoter sequence analysis conducted with PLANTPAN 3.0 revealed the presence of further TFBSs belonging to AP2/ERF, MYB/SANT, and MADS-box transcription factor families (CArG-boxes). In particular, FUZZNUC software identified two variants of CArG-boxes: $CC(A/T)_7G$ and $C(A/T)_8G$. One $CC(A/T)_7G$ site is located in both the *Phalaenopsis* and *Dendrobium DL2* putative promoters, while four $C(A/T)_8G$ are detected in the *PeDL2* promoter and six in the *DcDL2* promoter (Figure 48). These results are in agreement with the presence of CArG-boxes in all promoters of the Solanaceae *CRC*-like genes [130], suggesting that the regulation between *CRC/DL*-like genes and class B MADS box factors could occur at transcriptional level in orchids and this regulatory pathway could be conserved among the angiosperms.

Moreover, the existence of the TCP and SBP TFBSs, conserved in sequence and position on the *PeDL2* and *DcDL2* putative promoters, hints that the regulation of *DL2* transcription could be modulate by different transcription factors.

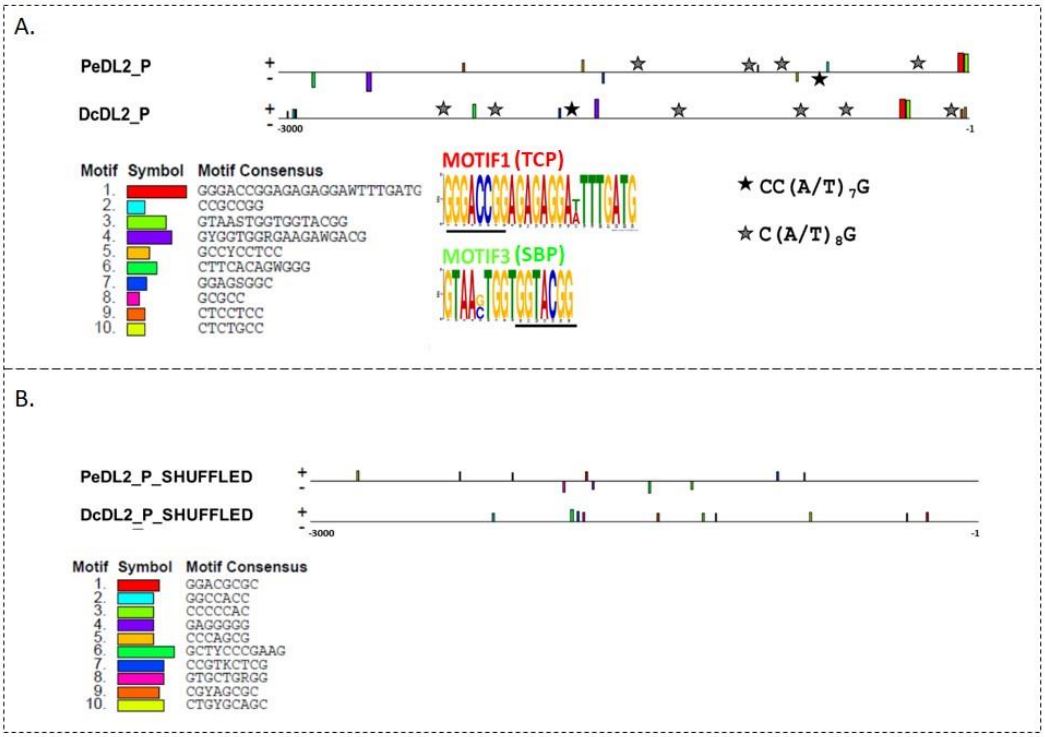


Figure 48. Conserved motifs within the putative promoters of the *DL2* genes of *P. equestris* and *D. catenatum*. (A) *PeDL2_P* and *DcDL2_P* are the nucleotide sequences spanning 3,000 bp upstream the ATG translation start site, numbered from -1 to -3000. In the sequence logo of the Motif 1 and 3, the predicted binding site of the TCP factor (JASPAR IDs MA1096.1 and MA1035.1) and of the SBP-type zinc-finger (JASPAR ID MA0955.1) are underlined. The black and grey stars indicate the CArG-box variants $CC(A/T)_7G$ and $C(A/T)_8G$, respectively. (B) Conserved motifs within the shuffled sequences of the putative promoters of the *DL2* genes of *P. equestris* and *D. catenatum.* Note that these motifs do not correspond to those found within the primary (not-shuffled) sequences. PeDL2_P_SHUFFLED and DcDL2_P_SHUFFLED are the shuffled nucleotide sequences spanning 3000 bp upstream the ATG translation start site of the *PeDL2* and *DcDL2* genes, numbered from -1 to -3000. Modified from *Lucibelli et al.* (2021) [100].

4.6.2 [redacted]

[redacted]

[redacted]

[redacted]

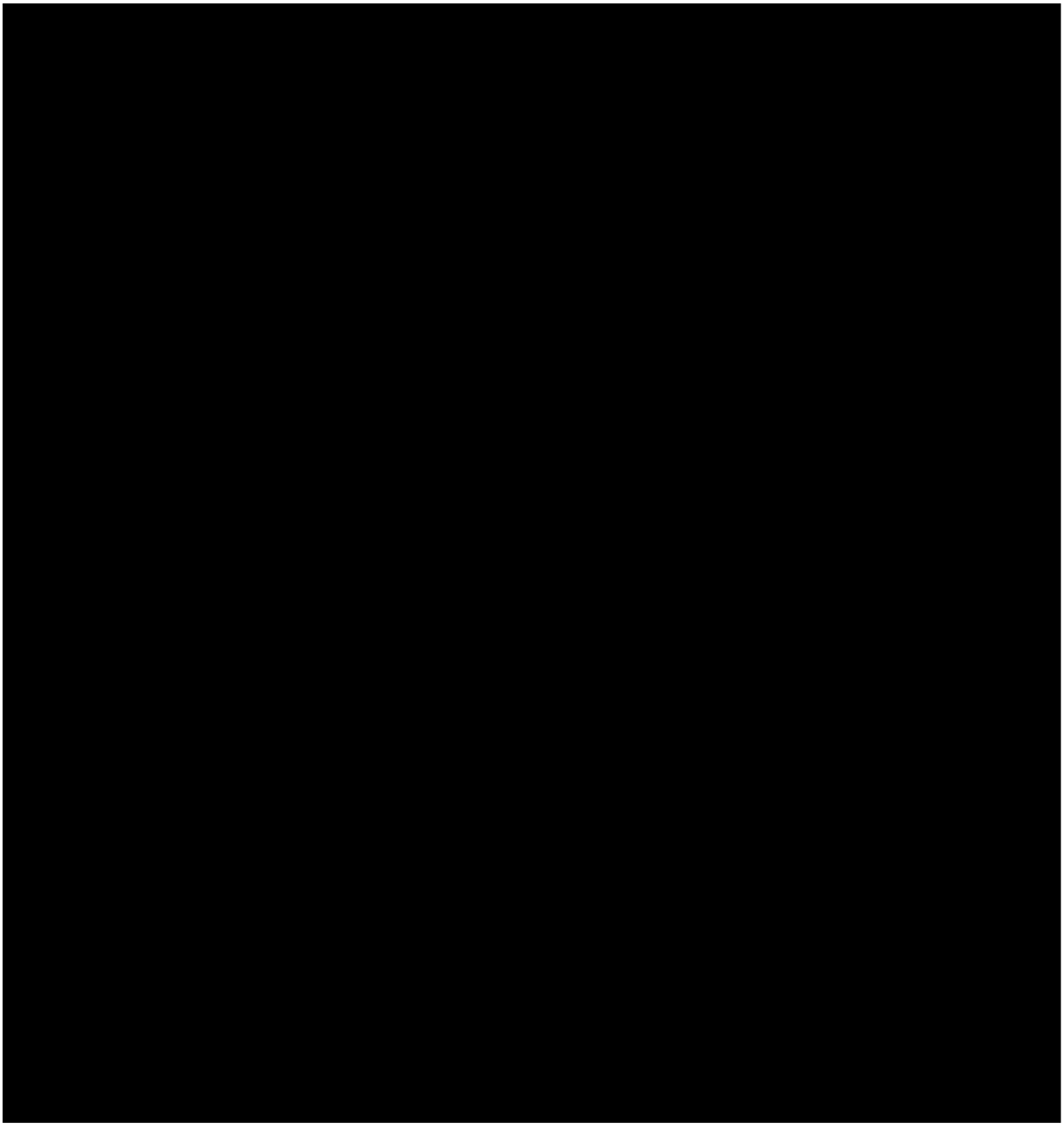
[redacted]

[redacted].

[redacted]

[redacted]

[redacted]



.

[redacted]

[redacted]

5. CONCLUSION

Orchidaceae is one of the most numerous families of angiosperms representing a unique ecological and genetic source. The diversity of the orchid flower morphology, the specialized developmental programs, the pollination syndromes, and the colonization of every type of habitat make orchid research a keystone in the study of biodiversity conservation and evolution. Moreover, orchids have a significant economic value because, as ornamental plants, their commercialization and cultivation is a substantial part of the worldwide floriculture trade. Therefore understanding the molecular pathways that underlie the development of the orchid flower has a relevant scientific and economic role.

In this scenario, my Ph.D. project aimed to enrich the information about the molecular basis of orchid flower development. Current knowledge recognizes that MADS-box transcription factors play a pivotal role in flower organogenesis, together with TCPs and MYBs. In particular, according to the molecular model "orchid code", the differential expression of the class B MADS-box *DEF*-like genes in the orchid perianth is fundamental for the floral organ specification.

The results of my research project suggest that in orchids, the YABBY *DROOPING-LEAF*-like gene *DL2*, after gene duplication, have acquired a role in the complex flower developmental program, contributing to orchid perianth differentiation. In the most ancestral subfamily Apostasioideae, which does not have a differentiated lip, *DL* is a single-copy gene. After the divergence between Apostsioideae and Vanilloideae, a gene duplication event occurred, and in each subfamily with differentiated lip there are two *DL* genes. Probably, the *DL2* duplicated gene acquired new functions in the development of the zigomorphic perianth. [redacted].

Considering the similar expression profile of the orchid *DL2* and *DEF* genes and the regulatory connection between class B MADS and *DL*-like genes identified in many flowering plants, my study proposes *DL2* as a good candidate to expand the original “orchid code”(Figure 54).

[redacted]

[redacted]

APPENDIX

Use	Gene	Forward (5'-3')	Reverse(5'-3')
PCR	PeDL1	AAGTCCTACTCCTTCATTGCTTACA	AGAAAAACAACATAAGCTGCTCGTT
	PeDL2	ACTCTCCACCATCTTCCTCC	CCCATGCATGCTATATCAAATGT
qPCR	PeDL1_1	GTGTGGGCACTGCAATCATCT	TCTTCTCTGGAGGTTTGACAAC
	PeDL1_2	GGAAACAGGACTGGTTGTGTGA	TCTTCTCTGGAGGTTTGACAAC
	PeDL2_1	ACTCTCCACCATCTTCCTCC	ATGGCCTCTGAAGCAGCGTTTGT
	PeDL2_2	AAGTTGGCTGTGAAATGCATGAT	ATGGCCTCTGAAGCAGCGTTTGT
	PeMADS2	GGGAAACTTACCGCGCTCTA	GATTGGGCTGATTCGGATGA
	PeMADS3	ATTTCGTACCCAACCAAGCC	TGCAGTGCTAGACCCTACTT
	PeMADS4	GATCTTCGCCTCGCTTGATA	ACCACAGAATCACACATAGCA
	PeMADS5	GGAAGGGTTGGGCGTTAAAG	GTTCATGCGTTAGGGCTCTG
	VpDL1	GTGGGCACTGCAATCATCTC	GCTTTTGGTGCTTGCTCCA
	VpDL2	CTGTTCTCGCGGTTGGAGT	CTCTGAAGCAGCGTTTGTTGA
	VpDEF1	ACTTAGGATTGTTCGCCAGA	TTGTAAACGCTGCTTGGAT
	VpDEF2	GCAGACGGATACCTACAAGAAGAAG	AGACGAAGATCATGGGAGGAATAT
	VpDEF3	CAGCACTCAAACAGCACTATAA	CTCATGCAAGCCGAAGGT
	PlDL1	ACAGTGAAGTGTGGGCACTG	GATGCTGCCGAGGTATGTCC
	PlDL2	CTGTTCTCGCGGTTGGAGT	CTCTGAAGCAGCGTTTGTTGA
	PlDEF1	GGGGGAAGAGAATGCAAATA	GCTTGGCTGGACTTCACTGT
	PlDEF2	TTGGATGGAGTCGCTTAGAATTGCTTCG	CGCAAACTCTGTTGTGTTCTTCTG
	PlDEF3	AGAAGATAGAAAACCCCACGAA	AGTTGATTCCAGATACTTTGCTCG
	PlDEF4	TCCGAGCAACTACGAGGGTA	ATCCAAATCCCATTCGATGA
	18S	TTAGGCCACGGAAGTTTGAG	ACACTTCACCGGACCATTCAA
	Actin	GGTATTGTTCTTGATTCTGGTGATGGTGTCA	GTCTTGGCAGTTTCCAACTCTTGCTCATAATC
	EF 1α	TGTGAAGAAGAAATGAAGTG	AACAACAGACTCAAAGACCT
In situ	PeDL2_1	ACTCTCCACCATCTTCCTCC	ATGGCCTCTGAAGCAGCGTTTGT
Subcellular localization	PeDL2	AAAAGAATTCATGGACTATGGCTCTCCTTCAGAC	AAAAGGATCCTTATTCCATGCGACGCTGTTGT
	PeDL2_ZF	ACATGAATTCTGTGCCTACTGCACCACTGTT	TTTGGATCCGCAGTGGCCACACTTAACTG
	PeDL2_INT	ACATGAATTCAATTCCCTCTCTTTTCTCAATCCC	AAAGGATCCAAAGGGAGCTGCTTTTGGGC
	PeDL2_YAB	ACATGAATTCGTCATTAAACCTCCGGAGAAGAAG	TATGGATCCATCGCATTTAGCCCAATTCTTTGC
Y2H	PeMADS2_AD	CCGGAATTCCGGATGGGGAGGGGGAAGATAGAGA	AGAGCTCAATTATGCAAGGCTAAGATCATGTG
Y1H	PeMADS2_BD	CCGGAATTCCGGATGGGGAGGGGGAAGATAGAGA	TCCGTCGACTTATGCAAGGCTAAGATCATGTG
	PeMADS3_AD	CCGGAATTCCGGATGGGGAGGGGGAAGATCGAGA	AGAGCTCAATCAGGCGAGACGTAGATCATG
	PeMADS3_BD	CCGGAATTCCGGATGGGGAGGGGGAAGATCGAGA	TCCGTCGACTCAGGCGAGACGTAGATCATG
	PeMADS4_AD	CCGGAATTCCGGATGGGGAGGGGGAAGATAGAGA	AGAGCTCAATCACGATCTTCGCCTCGCTTGA
	PeMADS4_BD	CCGGAATTCCGGATGGGGAGGGGGAAGATAGAGA	TCCGTCGACTCACGATCTTCGCCTCGCTTGA
	PeMADS5_AD	CCGGAATTCCGGATGGGGAGAGGGAAGATAGAGA	AGAGCTCAATCAATCAAAGCCAAACTCATGAC
	PeMADS5_BD	CCGGAATTCCGGATGGGGAGAGGGAAGATAGAGA	TCCGTCGACTCAATCAAAGCCAAACTCATGAC

	PeMADS6_AD	CCGGAATTCCGGATGGGTCGGGGAAAGATAGAGA	AGAGCTCAATTACTTATTTCCCTGCA
	PeMADS6_BD	CCGGAATTCCGGATGGGTCGGGGAAAGATAGAGA	TCCGTCGACTTACTTATTTCCCTGCA
	PeDL1_1_AD/BD	CCGGAATTCCGGATGGATCTGGTTTCTTCAAGGGA	AAGTCGACTTAGTTCTGGCGATCCATTTGTTT
	PeDL1_2_AD/BD	CCGGAATTCCGGATGCAGGATCCATGTAATGATTGT	AAGTCGACTTAGTTCTGGCGATCCATTTGTTT
	PeDL2_1_AD/BD	CCGGAATTCCGGATGGACTATGGCTCTCCTTCAGAC	AGGATCCAATTATTCCATGCGACGCTGTTGTC
	PeDL2_2_AD/BD	AGGATCCAAATGACGGACACAGTGACAGTTAAG	AAGTCGACTTATTCCATGCGACGCTGTTGTC
	CYC1_AD/BD	CCGGAATTCCGG ATGCTACGTACTATGTTTCCTCAA	AGGATCCAATTAACGTAGAAGCTCACCAAGAAG
	CYC2_AD/BD	CCGGAATTCCGGATGTTTTCTCAGGCTAATGAGCTC	AGGATCCAATTAGACATGATCAAACATTGATGGCG
	CYC3_AD/BD	CCGGAATTCCGGATGTTTTCTCAGCCTTATTATACATCC	AGGATCCAATTAGACACGAGTCACATGATCA
	Pro_PaDL2_1	TCTAGGTACCGTAAGGGCTCACACTAGCCGCTA	AAAGTCGACAAGCTAAGAGGGTTGAAGTTCC
	Pro_PaDL2_2	TCTAGGTACCCGAAGGTGTGATGGTTGGTGGG	AAAGTCGACCAATCCATTCGATGATCAAAGGT
	Pro_PaDL2_3	TCTAGGTACCACCTTTGATCATCGAATGGATTG	AAAGTCGACAGTTCTAGGGAGGGGCTATGT
	Pro_PaDL2_4	TCTAGGTACCACATAGCCCCTCCCTAGAACT	AAAGTCGACATTACCAAGCAGAGAGGCCT
	Pro_PaDL2_5	TCTAGGTACCAGGCCTCTCTGCTTGGTAA	AAAGTCGACTCGACACATTGTTCCCCACA
Protoplasts dsRNAi	dsRNA MADS2	TAATACGACTCACTATAGGGGGGAAACTTACCG CGCTCTA	TAATACGACTCACTATAGGGGATTGGGCTGATTCG GATGA
	MADS2	GGGAAACTTACCGCGCTCTA	TTATGCAAGGCTAAGATCATG
	PeDL2_1	ACTCTCCACCATCTTCCTCC	ATGGCCTCTGAAGCAGCGTTTGT
	PeDL2_2	AAGTTGGCTGTGAAATGCATGAT	ATGGCCTCTGAAGCAGCGTTTGT

Table 1. List of the primer sequences used. qPCR, quantitative real time PCR; Y2H, Yeast Two-Hybrid analysis; Y1H, Yeast One-Hybrid analysis; BD, GAL4 DNA-binding domain (pGBKT7 vector); AD, GAL4 activation domain (pGADT7 vector).

TFBS	TF FAMILY	POSITION	STRAND	SEQUENCE	*PaDL2* PROMOTER FRAGMENT
TFmatrixID_1139	MADS box	299	+	TCACCA	1
TFmatrixID_0508	MADS box ; MIKC	345	+	CCTTTTTTTGA	1
TFmatrixID_1134	MADS box ; MIKC	952	+	ATTTTCCACTTCCTTCAC	1
TFmatrixID_0508	MADS box ; MIKC	991	+	TTCAGTTTTGG	1
TFmatrixID_0501	MADS box ; MIKC	1414	+	TTGCCCTTTTTGAGA	3
TFmatrixID_1139	MADS box	1538	+	TCACCA	3
CWWWWWWWWG	MADS box	2012	+	CAATTTTTAG	4
TFmatrixID_1136	MADS box ; MIKC	2189	+	ATTCCAAAGTCAGGAA	4
CWWWWWWWG	MADS box	2214	+	CATTTTATG	4
TFmatrixID_0503	MADS box ; MIKC	2422	+	ATTTTTCTAATTTTCTTTCTC	4
CWWWWWWWWG	MADS box	2534	+	CAAATTTATG	4
CWWWWWWWWG	MADS box	2638	+	CATATTTTAG	4
CWWWWWWWWG	MADS box	2796	-	CATTTTTTTGG	5
CWWWWWWWG	MADS box	2998	+	CATAATTTG	5
CWWWWWWWWG	MADS box	3189	+	CATTAAATAG	5

Table 2. List of the TFBSs on the putative promoter of the *PaDL2* gene identified using PLANTPAN 3.0.

REFERENCES

1. Beaulieu JM, O'Meara BC, Crane P, Donoghue MJ: **Heterogeneous Rates of Molecular Evolution and Diversification Could Explain the Triassic Age Estimate for Angiosperms.** *Syst Biol* 2015, **64:**869-878.
2. Foster CSP, Sauquet H, van der Merwe M, McPherson H, Rossetto M, Ho SYW: **Evaluating the Impact of Genomic Data and Priors on Bayesian Estimates of the Angiosperm Evolutionary Timescale.** *Syst Biol* 2017, **66:**338-351.
3. Magallon S, Gomez-Acevedo S, Sanchez-Reyes LL, Hernandez-Hernandez T: **A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity.** *New Phytol* 2015, **207:**437-453.
4. Sauquet H, von Balthazar M, Magallon S, Doyle JA, Endress PK, Bailes EJ, Barroso de Morais E, Bull-Herenu K, Carrive L, Chartier M, et al: **The ancestral flower of angiosperms and its early diversification.** *Nat Commun* 2017, **8:**16047.
5. Soltis PS, Soltis DE: **Flower diversity and angiosperm diversification.** *Methods Mol Biol* 2014, **1110:**85-102.
6. Lohmann JU, Weigel D: **Building beauty: the genetic control of floral patterning.** *Dev Cell* 2002, **2:**135-142.
7. Coen ES, Meyerowitz EM: **The war of the whorls: genetic interactions controlling flower development.** *Nature* 1991, **353:**31-37.
8. Citerne HJ, F. ; Nadot, S. ; Damerval, C.: **The evolution of floral symmetry.** *Advances in Botanical Research* 2010, **54:**85-137.
9. Lucibelli F, Valoroso MC, Aceto S: **Radial or Bilateral? The Molecular Basis of Floral Symmetry.** *Genes (Basel)* 2020, **11**.
10. Endress PK: **Symmetry in Flowers: Diversity and Evolution.** *Int J Plant Sci* 1999, **160:**S3-S23.
11. Fenster CB, Armbruster WS, Dudash MR: **Specialization of flowers: is floral orientation an overlooked first step?** *New Phytol* 2009, **183:**502-506.
12. Cronk Q, Moller M: **Genetics of floral symmetry revealed.** *Trends Ecol Evol* 1997, **12:**85-86.
13. Neal PRD, A. ; Giurfa, M.: **Floral Symmetry and Its Role in Plant-Pollinator Systems: Terminology, Distribution, and Hypotheses.** *Annual Review of Ecology and Systematics* 1998, **29:**345-373.
14. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G: **PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants.** *Nucleic Acids Res* 2017, **45:**D1040-D1045.
15. Xiong Y, Liu T, Tian C, Sun S, Li J, Chen M: **Transcription factors in rice: a genome-wide comparative analysis between monocots and eudicots.** *Plant Mol Biol* 2005, **59:**191-203.
16. Airoldi CA, Davies B: **Gene duplication and the evolution of plant MADS-box transcription factors.** *J Genet Genomics* 2012, **39:**157-165.
17. Aceto S, Gaudio L: **The MADS and the Beauty: Genes Involved in the Development of Orchid Flowers.** *Curr Genomics* 2011, **12:**342-356.
18. Bowman JLS, D.R.; Meyerowitz, E.M.: **Genetic interactions among floral homeotic genes of Arabidopsis.** *Development* 1991, **112:**1–20.
19. Riechmann JL, Meyerowitz EM: **MADS domain proteins in plant development.** *Biol Chem* 1997, **378:**1079-1101.

20. Sablowski RW, Meyerowitz, E.M.,: **A homolog of NO APICAL MERISTEM is an immediate target of the floral homeotic genes APETALA3/PISTILLATA.** *Cell* 1998, **92:**93-103.
21. Aida M, Ishida T, Fukaki H, Fujisawa H, Tasaka M: **Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant.** *Plant Cell* 1997, **9:**841-857.
22. Golz JF, Hudson A: **Plant development: YABBYs claw to the fore.** *Curr Biol* 1999, **9:**R861-863.
23. Romani F, Moreno JE: **Molecular mechanisms involved in functional macroevolution of plant transcription factors.** *New Phytol* 2021, **230:**1345-1353.
24. Zhang D, Zhao XW, Li YY, Ke SJ, Yin WL, Lan S, Liu ZJ: **Advances and prospects of orchid research and industrialization.** *Hortic Res* 2022, **9:**uhac220.
25. Liu C, Leng J, Li Y, Ge T, Li J, Chen Y, Guo C, Qi J: **A spatiotemporal atlas of organogenesis in the development of orchid flowers.** *Nucleic Acids Res* 2022.
26. Cozzolino S, Widmer A: **Orchid diversity: an evolutionary consequence of deception?** *Trends in Ecology and Evolution* 2005, **20:**487–494.
27. Darwin C: **On the Various Contrivances by Which British and Foreign Orchids are Fertilised by Insects** *Cambridge University Press* 2011.
28. Cameron KM, Chase MW, Whitten WM, Kores PJ, Jarrell DC, Albert VA, Yukawa T, Hills HG, Goldman DH: **A phylogenetic analysis of the Orchidaceae: evidence from rbcL nucleotide.** *Am J Bot* 1999, **86:**208-224.
29. Gorniak M PO, Chase MW: **Phylogenetic relationships within Orchidaceae based on a low-copy nuclear coding gene, Xdh: Congruence with organellar and nuclear ribosomal DNA results.** *Molecular phylogenetics and evolution* 2010, **56:**784-795.
30. Givnish TJ SD, Ames M, Lyon SP, Hunter SJ, Zuluaga A, Iles WJ, Clements MA, Arroyo MT, Leebens-Mack J, Endara L, Kriebel R, Neubig KM, Whitten WM, Williams NH, Cameron KM. : **Orchid phylogenomics and multiple drivers of their extraordinary diversification.** *Proceedings Biological sciences* 2015, **282**.
31. Tsai WC, Dievart A, Hsu CC, Hsiao YY, Chiou SY, Huang H, Chen HH: **Post genomics era for orchid research.** *Bot Stud* 2017, **58:**61.
32. Rudall PJ, Bateman RM: **Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots.** *Biol Rev Camb Philos Soc* 2002, **77:**403-441.
33. Bateman RM, Rudall PJ: **The Good, The Bad, The Ugly: Using Naturally Occurring Terata to Distinguish the Possible from the Impossible in Orchid Floral Evolution.** *Aliso* 2006, **22:**481-496.
34. Pramanik D, Dorst N, Meesters N, Spaans M, Smets E, Welten M, Gravendeel B: **Evolution and development of three highly specialized floral structures of bee-pollinated Phalaenopsis species.** *Evodevo* 2020, **11:**16.
35. Valoroso MC, De Paolo S, Iazzetti G, Aceto S: **Transcriptome-Wide Identification and Expression Analysis of DIVARICATA- and RADIALIS-Like Genes of the Mediterranean Orchid Orchis italica.** *Genome Biol Evol* 2017, **9**.
36. Teo ZWN, Zhou W, Shen L: **Dissecting the Function of MADS-Box Transcription Factors in Orchid Reproductive Development.** *Front Plant Sci* 2019, **10:**1474.
37. Yu H, Goh CJ: **Molecular genetics of reproductive biology in orchids.** *Plant Physiol* 2001, **127:**1390-1393.

38. Da Silva JAT, Aceto, S., Liu, W., Yu, H., and Kanno, A.: **Genetic control of flower development, color and senescence of Dendrobium orchids.** *Scientia Horticulturae* 2014, **175:**74–86.
39. Liu C TZ, Yu H: **Coming into bloom: the specification of floral meristems.** *Development* 2009, **136:**3379–3391.
40. Lin CS, Chen JJ, Huang YT, Hsu CT, Lu HC, Chou ML, Chen LC, Ou CI, Liao DC, Yeh YY, et al: **Catalog of Erycina pusilla miRNA and categorization of reproductive phase-related miRNAs and their target gene families.** *Plant Mol Biol* 2013, **82:**193-204.
41. Theissen G, Kim JT, Saedler H: **Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes.** *J Mol Evol* 1996, **43:**484-516.
42. Schwarz-Sommer Z, Huijser P, Nacken W, Saedler H, Sommer H: **Genetic Control of Flower Development by Homeotic Genes in Antirrhinum majus.** *Science* 1990, **250:**931-936.
43. Alvarez-Buylla ER, Pelaz S, Liljegren SJ, Gold SE, Burgeff C, Ditta GS, Ribas de Pouplana L, Martinez-Castilla L, Yanofsky MF: **An ancestral MADS-box gene duplication occurred before the divergence of plants and animals.** *Proc Natl Acad Sci U S A* 2000, **97:**5328-5333.
44. Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter KU, Saedler H: **A short history of MADS-box genes in plants.** *Plant Mol Biol* 2000, **42:**115-149.
45. Parenicova L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B, et al: **Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world.** *Plant Cell* 2003, **15:**1538-1551.
46. Kaufmann K, Melzer R, Theissen G: **MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants.** *Gene* 2005, **347:**183-198.
47. Henschel K, Kofuji R, Hasebe M, Saedler H, Munster T, Theissen G: **Two ancient classes of MIKC-type MADS-box genes are present in the moss Physcomitrella patens.** *Mol Biol Evol* 2002, **19:**801-814.
48. Gramzow L, Theissen G: **A hitchhiker's guide to the MADS world of plants.** *Genome Biol* 2010, **11:**214.
49. Theissen G, Saedler H: **Plant biology. Floral quartets.** *Nature* 2001, **409:**469-471.
50. Ambrose BA, Lerner DR, Ciceri P, Padilla CM, Yanofsky MF, Schmidt RJ: **Molecular and genetic analyses of the silky1 gene reveal conservation in floral organ specification between eudicots and monocots.** *Mol Cell* 2000, **5:**569-579.
51. Whipple CJ, Ciceri P, Padilla CM, Ambrose BA, Bandong SL, Schmidt RJ: **Conservation of B-class floral homeotic gene function between maize and Arabidopsis.** *Development* 2004, **131:**6083-6091.
52. Whipple CJ, Zanis MJ, Kellogg EA, Schmidt RJ: **Conservation of B class gene expression in the second whorl of a basal grass and outgroups links the origin of lodicules and petals.** *Proc Natl Acad Sci U S A* 2007, **104:**1081-1086.
53. Ferrario S, Immink RG, Angenent GC: **Conservation and diversity in flower land.** *Curr Opin Plant Biol* 2004, **7:**84-91.
54. Irish VF, Litt A: **Flower development and evolution: gene duplication, diversification and redeployment.** *Curr Opin Genet Dev* 2005, **15:**454-460.
55. Jack T: **Plant development going MADS.** *Plant Mol Biol* 2001, **46:**515-520.

56. Theissen G: **Development of floral organ identity: stories from the MADS house.** *Curr Opin Plant Biol* 2001, **4:**75-85.
57. Theissen G, Melzer R: **Molecular mechanisms underlying origin and diversification of the angiosperm flower.** *Ann Bot* 2007, **100:**603-619.
58. Heijmans K, Ament K, Rijpkema AS, Zethof J, Wolters-Arts M, Gerats T, Vandenbussche M: **Redefining C and D in the petunia ABC.** *Plant Cell* 2012, **24:**2305-2317.
59. Liu Z, Mara C: **Regulatory mechanisms for floral homeotic gene expression.** *Semin Cell Dev Biol* 2010, **21:**80-86.
60. Valoroso MC, Censullo MC, Aceto S: **The MADS-box genes expressed in the inflorescence of Orchis italica (Orchidaceae).** *PLoS One* 2019, **14:**e0213185.
61. Chanderbali AS, Berger BA, Howarth DG, Soltis PS, Soltis DE: **Evolving Ideas on the Origin and Evolution of Flowers: New Perspectives in the Genomic Era.** *Genetics* 2016, **202:**1255-1265.
62. Nakamura T, Fukuda T, Nakano M, Hasebe M, Kameya T, Kanno A: **The modified ABC model explains the development of the petaloid perianth of Agapanthus praecox ssp. orientalis (Agapanthaceae) flowers.** *Plant Mol Biol* 2005, **58:**435-445.
63. Tsai WC, Kuoh CS, Chuang MH, Chen WH, Chen HH: **Four DEF-like MADS box genes displayed distinct floral morphogenetic roles in Phalaenopsis orchid.** *Plant Cell Physiol* 2004, **45:**831-844.
64. Pan Z, Cheng C, Tsai W, Chung M, Chen W, Hu J, HH. C: **The Duplicated B-class MADS-Box Genes Display Dualistic Characters in Orchid Floral Organ Identity and Growth.** *Plant and Cell biology* 2011, **52:**1515–1531.
65. Kramer EM, Dorit RL, Irish VF: **Molecular evolution of genes controlling petal and stamen development: duplication and divergence within the APETALA3 and PISTILLATA MADS-box gene lineages.** *Genetics* 1998, **149:**765-783.
66. Kim S, Yoo MJ, Albert VA, Farris JS, Soltis PS, Soltis DE: **Phylogeny and diversification of B-function MADS-box genes in angiosperms: evolutionary and functional implications of a 260-million-year-old duplication.** *Am J Bot* 2004, **91:**2102-2118.
67. Zahn LM L-MJ, DePamphilis CW, Ma H, Theissen G. : **To B or Not to B a flower: the role of DEFICIENS and GLOBOSA orthologs in the evolution of the angiosperms.** *J Hered* 2005, **96:**225–240.
68. Mondragon-Palomino M, Theissen G: **MADS about the evolution of orchid flowers.** *Trends Plant Sci* 2008, **13:**51-59.
69. Mondragon-Palomino M, Theissen G: **Conserved differential expression of paralogous DEFICIENS- and GLOBOSA-like MADS-box genes in the flowers of Orchidaceae: refining the 'orchid code'.** *Plant J* 2011, **66:**1008-1019.
70. Hsu H, Hsu W, Lee Y, Mao W, Yang J, Li J, Yang C: **Model for perianth formation in orchids.** *Nature Plants* 2015, **1**.
71. Carpenter R, Coen ES: **Floral homeotic mutations produced by transposon-mutagenesis in Antirrhinum majus.** *Genes Dev* 1990, **4:**1483-1493.
72. Kosugi S, Ohashi Y: **DNA binding and dimerization specificity and potential targets for the TCP protein family.** *Plant J* 2002, **30:**337-348.
73. Martin-Trillo M, Cubas P: **TCP genes: a family snapshot ten years later.** *Trends Plant Sci* 2010, **15:**31-39.

74. Li C, Potuschak T, Colon-Carmona A, Gutierrez RA, Doerner P: **Arabidopsis TCP20 links regulation of growth and cell division control pathways.** *Proc Natl Acad Sci U S A* 2005, **102:**12978-12983.
75. Preston JC, Hileman LC: **Developmental genetics of floral symmetry evolution.** *Trends Plant Sci* 2009, **14:**147-154.
76. Viola IL, Uberti Manassero NG, Ripoll R, Gonzalez DH: **The Arabidopsis class I TCP transcription factor AtTCP11 is a developmental regulator with distinct DNA-binding properties due to the presence of a threonine residue at position 15 of the TCP domain.** *Biochem J* 2011, **435:**143-155.
77. Cubas P, Lauter N, Doebley J, Coen E: **The TCP domain: a motif found in proteins regulating plant growth and development.** *Plant J* 1999, **18:**215-222.
78. Howarth DG, Donoghue MJ: **Phylogenetic analysis of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots.** *Proc Natl Acad Sci U S A* 2006, **103:**9101-9106.
79. Luo D, Carpenter R, Vincent C, Copsey L, Coen E: **Origin of floral asymmetry in Antirrhinum.** *Nature* 1996, **383:**794-799.
80. Luo D, Carpenter R, Copsey L, Vincent C, Clark J, Coen E: **Control of organ asymmetry in flowers of Antirrhinum.** *Cell* 1999, **99:**367-376.
81. Busch A, Zachgo S: **Flower symmetry evolution: towards understanding the abominable mystery of angiosperm radiation.** *Bioessays* 2009, **31:**1181-1190.
82. Manassero NG, Viola IL, Welchen E, Gonzalez DH: **TCP transcription factors: architectures of plant form.** *Biomol Concepts* 2013, **4:**111-127.
83. Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, Zhaofeng G, Zhiqiang L, Yunfei Z, Xiaoxiao W, Xiaoming Q, et al: **The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family.** *Plant Mol Biol* 2006, **60:**107-124.
84. Ogata K, Morikawa S, Nakamura H, Hojo H, Yoshimura S, Zhang R, Aimoto S, Ametani Y, Hirata Z, Sarai A, et al.: **Comparison of the free and DNA-complexed forms of the DNA-binding domain from c-Myb.** *Nat Struct Biol* 1995, **2:**309-320.
85. Kanei-Ishii C, Sarai A, Sawazaki T, Nakagoshi H, He DN, Ogata K, Nishimura Y, Ishii S: **The tryptophan cluster: a hypothetical structure of the DNA-binding domain of the myb protooncogene product.** *J Biol Chem* 1990, **265:**19990-19995.
86. Kranz H, Scholz K, Weisshaar B: **c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages.** *Plant J* 2000, **21:**231-235.
87. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L: **MYB transcription factors in Arabidopsis.** *Trends Plant Sci* 2010, **15:**573-581.
88. Jiang C, Gu J, Chopra S, Gu X, Peterson T: **Ordered origin of the typical two- and three-repeat Myb genes.** *Gene* 2004, **326:**13-22.
89. Lipsick JS: **One billion years of Myb.** *Oncogene* 1996, **13:**223-235.
90. Rosinski JA, Atchley WR: **Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin.** *J Mol Evol* 1998, **46:**74-83.
91. Zhang C, Ma R, Xu J, Yan J, Guo L, Song J, Feng R, Yu M: **Genome-wide identification and classification of MYB superfamily genes in peach.** *PLoS One* 2018, **13:**e0199192.
92. Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, et al: **Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes.** *Science* 2000, **290:**2105-2110.

93. Hosoda K, Imamura A, Katoh E, Hatta T, Tachiki M, Yamada H, Mizuno T, Yamazaki T: **Molecular structure of the GARP family of plant Myb-related DNA binding motifs of the Arabidopsis response regulators.** *Plant Cell* 2002, **14:**2015-2029.
94. Raimundo J, Sobral R, Laranjeira S, Costa MMR: **Successive Domain Rearrangements Underlie the Evolution of a Regulatory Module Controlled by a Small Interfering Peptide.** *Mol Biol Evol* 2018, **35:**2873-2885.
95. Sengupta A, Hileman LC: **Novel Traits, Flower Symmetry, and Transcriptional Autoregulation: New Hypotheses From Bioinformatic and Experimental Data.** *Front Plant Sci* 2018, **9:**1561.
96. Raimundo J, Sobral R, Bailey P, Azevedo H, Galego L, Almeida J, Coen E, Costa MM: **A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in Antirrhinum flower asymmetry.** *Plant J* 2013, **75:**527-538.
97. Costa MM, Fox S, Hanna AI, Baxter C, Coen E: **Evolution of regulatory interactions controlling floral asymmetry.** *Development* 2005, **132:**5093-5101.
98. Valoroso MC, Sobral R, Saccone G, Salvemini M, Costa MMR, Aceto S: **Evolutionary Conservation of the Orchid MYB Transcription Factors DIV, RAD, and DRIF.** *Front Plant Sci* 2019, **10:**1359.
99. De Paolo S, Gaudio L, Aceto S: **Analysis of the TCP genes expressed in the inflorescence of the orchid Orchis italica.** *Sci Rep* 2015, **5:**16265.
100. Lucibelli F, Valoroso MC, Theissen G, Nolden S, Mondragon-Palomino M, Aceto S: **Extending the Toolkit for Beauty: Differential Co-Expression of DROOPING LEAF-Like and Class B MADS-Box Genes during Phalaenopsis Flower Development.** *Int J Mol Sci* 2021, **22**.
101. Chen TK, Yang HT, Fang SC, Lien YC, Yang TT, Ko SS: **Hybrid-Cut: An Improved Sectioning Method for Recalcitrant Plant Tissue Samples.** *J Vis Exp* 2016.
102. Lee CYV, K.K.; Huang, J.Z.; Lee, C.P.; Lin, C.P.; Cheng, T.C.; Chang, B.C.; Chin, S.W.; Chen, F.C.: **PhalDB: A comprehensive database for molecular mining of the Phalaenopsis genome, transcriptome and miRNome.** *Genet Mol Res* 2018, **17**.
103. Grabherr MG HB, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A: **Full-length transcriptome assembly from RNA-Seq data without a reference genome.** *Nat Biotechnol* 2011, **29:**644–652.
104. Musacchia F, Basu S, Petrosino G, Salvemini M, Sanges R: **Annocript: a flexible pipeline for the annotation of transcriptomes able to identify putative long noncoding RNAs.** *Bioinformatics* 2015, **31:**2199-2201.
105. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics* 2010, **26:**139-140.
106. Tsai WC, Fu CH, Hsiao YY, Huang YM, Chen LJ, Wang M, Liu ZJ, Chen HH: **OrchidBase 2.0: comprehensive collection of Orchidaceae floral transcriptomes.** *Plant Cell Physiol* 2013, **54:**e7.
107. Chao YT, Yen SH, Yeh JH, Chen WC, Shih MC: **Orchidstra 2.0-A Transcriptomics Resource for the Orchid Family.** *Plant Cell Physiol* 2017, **58:**e9.
108. Zhang GQ, Xu Q, Bian C, Tsai WC, Yeh CM, Liu KW, Yoshida K, Zhang LS, Chang SB, Chen F, et al: **The Dendrobium catenatum Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution.** *Sci Rep* 2016, **6:**19029.
109. Zhang GQ, Liu KW, Li Z, Lohaus R, Hsiao YY, Niu SC, Wang JY, Lin YC, Xu Q, Chen LJ, et al: **The Apostasia genome and the evolution of orchids.** *Nature* 2017, **549:**379-383.

110. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS: **MEME SUITE: tools for motif discovery and searching.** *Nucleic Acids Res* 2009, **37:**W202-208.
111. Li BJ ZB, Wang JY, Tsai WC, Lu HC, Zou LH, Wan X, Zhang DY, Qiao HJ, Liu ZJ, Wang Y: **New insight into the molecular mechanism of colour differentiation among floral segments in orchids.** *Commun Biol* 2020, **3**.
112. Acri-Nunes-Miranda R, Mondragon-Palomino M: **Expression of paralogous SEP-, FUL-, AG- and STK-like MADS-box genes in wild-type and peloric Phalaenopsis flowers.** *Front Plant Sci* 2014, **5:**76.
113. Cutler SR, Ehrhardt DW, Griffitts JS, Somerville CR: **Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency.** *Proc Natl Acad Sci U S A* 2000, **97:**3718-3723.
114. Gross T, Broholm S, Becker A: **CRABS CLAW Acts as a Bifunctional Transcription Factor in Flower Development.** *Front Plant Sci* 2018, **9:**835.
115. Jones DT: **Protein secondary structure prediction based on position-specific scoring matrices.** *J Mol Biol* 1999, **292:**195-202.
116. Gruber M, Söding, J., and Lupas, A. N.: **REPPER–repeats and their periodicities in fibrous proteins.** *Nucleic Acids Res* 2005, **33**.
117. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, Lozajic M, Gabler F, Soding J, Lupas AN, Alva V: **A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core.** *J Mol Biol* 2018, **430:**2237-2243.
118. Gietz RD, Schiestl RH, Willems AR, Woods RA: **Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure.** *Yeast* 1995, **11:**355-360.
119. Cai J, Liu X, Vanneste K, Proost S, Tsai WC, Liu KW, Chen LJ, He Y, Xu Q, Bian C, et al: **The genome sequence of the orchid Phalaenopsis equestris.** *Nat Genet* 2015, **47:**65-72.
120. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS: **Quantifying similarity between motifs.** *Genome Biol* 2007, **8**.
121. Chow CN, Lee TY, Hung YC, Li GZ, Tseng KC, Liu YH, Kuo PL, Zheng HQ, Chang WC: **PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants.** *Nucleic Acids Res* 2019, **47:**D1155-D1163.
122. Lee JY, Baum SF, Alvarez J, Patel A, Chitwood DH, Bowman JL: **Activation of CRABS CLAW in the Nectaries and Carpels of Arabidopsis.** *Plant Cell* 2005, **17:**25-36.
123. Doyle JJDJL: **Isolation of Plant DNA from Fresh Tissue.** *Focus* 1999, **12:**13-15.
124. Lin HY, Chen JC, Fang SC: **A Protoplast Transient Expression System to Enable Molecular, Cellular, and Functional Studies in Phalaenopsis orchids.** *Front Plant Sci* 2018, **9:**843.
125. An CI, Sawada A, Fukusaki E, Kobayashi A: **A transient RNA interference assay system using Arabidopsis protoplasts.** *Biosci Biotechnol Biochem* 2003, **67:**2674-2677.
126. Fourquin C, Vinauger-Douard M, Fogliani B, Dumas C, Scutt CP: **Evidence that CRABS CLAW and TOUSLED have conserved their roles in carpel development since the ancestor of the extant angiosperms.** *Proc Natl Acad Sci U S A* 2005, **102:**4649-4654.
127. Huang JZ, Lin CP, Cheng TC, Chang BC, Cheng SY, Chen YW, Lee CY, Chin SW, Chen FC: **A de novo floral transcriptome reveals clues into Phalaenopsis orchid flower development.** *PLoS One* 2015, **10:**e0123474.

128. Ishikawa M, Ohmori Y, Tanaka W, Hirabayashi C, Murai K, Ogihara Y, Yamaguchi T, Hirano HY: **The spatial expression patterns of DROOPING LEAF orthologs suggest a conserved function in grasses.** *Genes Genet Syst* 2009, **84:**137-146.
129. Wuest SE, O'Maoileidigh DS, Rae L, Kwasniewska K, Raganelli A, Hanczaryk K, Lohan AJ, Loftus B, Graciet E, Wellmer F: **Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA.** *Proc Natl Acad Sci U S A* 2012, **109:**13452-13457.
130. Gong P, Song C, Liu H, Li P, Zhang M, Zhang J, Zhang S, He C: **Physalis floridana CRABS CLAW mediates neofunctionalization of GLOBOSA genes in carpel development.** *J Exp Bot* 2021, **72:**6882-6903.
131. Nagasawa N, Miyoshi M, Sano Y, Satoh H, Hirano H, Sakai H, Nagato Y: **SUPERWOMAN1 and DROOPING LEAF genes control floral organ identity in rice.** *Development* 2003, **130:**705-718.
132. Strable J, Vollbrecht E: **Maize YABBY genes drooping leaf1 and drooping leaf2 regulate floret development and floral meristem determinacy.** *Development* 2019, **146**.
133. She Z, Huang X, Aslam M, Wang L, Yan M, Qin R, Chen Y, Qin Y, Niu X: **Expression characterization and cross-species complementation uncover the functional conservation of YABBY genes for leaf abaxial polarity and carpel polarity establishment in Saccharum spontaneum.** *BMC Plant Biol* 2022, **22:**124.
134. Romanova MA, Maksimova AI, Pawlowski K, Voitsekhovskaja OV: **YABBY Genes in the Development and Evolution of Land Plants.** *Int J Mol Sci* 2021, **22**.
135. Chen YY, Hsiao YY, Chang SB, Zhang D, Lan SR, Liu ZJ, Tsai WC: **Genome-Wide Identification of YABBY Genes in Orchidaceae and Their Expression Patterns in Phalaenopsis Orchid.** *Genes (Basel)* 2020, **11**.
136. Nakayama H, Yamaguchi T, Tsukaya H: **Expression patterns of AaDL, a CRABS CLAW ortholog in Asparagus asparagoides (Asparagaceae), demonstrate a stepwise evolution of CRC/DL subfamily of YABBY genes.** *Am J Bot* 2010, **97:**591-600.
137. Bowman JL, Smyth DR: **CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains.** *Development* 1999, **126:**2387-2396.
138. Lee JY, Baum SF, Oh SH, Jiang CZ, Chen JC, Bowman JL: **Recruitment of CRABS CLAW to promote nectary development within the eudicot clade.** *Development* 2005, **132:**5021-5032.
139. Orashakova S, Lange M, Lange S, Wege S, Becker A: **The CRABS CLAW ortholog from California poppy (Eschscholzia californica, Papaveraceae), EcCRC, is involved in floral meristem termination, gynoecium differentiation and ovule initiation.** *Plant J* 2009, **58:**682-693.
140. Yamaguchi T, Nagasawa N, Kawasaki S, Matsuoka M, Nagato Y, Hirano HY: **The YABBY gene DROOPING LEAF regulates carpel specification and midrib development in Oryza sativa.** *Plant Cell* 2004, **16:**500-509.
141. Wang A, Tang J, Li D, Chen C, Zhao X, Zhu L: **Isolation and functional analysis of LiYAB1, a YABBY family gene, from lily (Lilium longiflorum).** *J Plant Physiol* 2009, **166:**988-995.
142. Fourquin C, Vinauger-Douard M, Chambrier P, Berne-Dedieu A, Scutt CP: **Functional conservation between CRABS CLAW orthologues from widely diverged angiosperms.** *Ann Bot* 2007, **100:**651-657.

143. Dirks-Mulder A, Butot R, van Schaik P, Wijnands JW, van den Berg R, Krol L, Doebar S, van Kooperen K, de Boer H, Kramer EM, et al: **Exploring the evolutionary origin of floral organs of Erycina pusilla, an emerging orchid model system.** *BMC Evol Biol* 2017, **17:**89.
144. Givnish TJ, Spalink D, Ames M, Lyon SP, Hunter SJ, Zuluaga A, Iles WJ, Clements MA, Arroyo MT, Leebens-Mack J, et al: **Orchid phylogenomics and multiple drivers of their extraordinary diversification.** *Proc Biol Sci* 2015, **282**.
145. Burdach J, O'Connell MR, Mackay JP, Crossley M: **Two-timing zinc finger transcription factors liaising with RNA.** *Trends Biochem Sci* 2012, **37:**199-205.
146. Alvarez J, Smyth DR: **CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS.** *Development* 1999, **126:**2377-2386.
147. Lin YF, Chen YY, Hsiao YY, Shen CY, Hsu JL, Yeh CM, Mitsuda N, Ohme-Takagi M, Liu ZJ, Tsai WC: **Genome-wide identification and characterization of TCP genes involved in ovule development of Phalaenopsis equestris.** *J Exp Bot* 2016, **67:**5051-5066.
148. Chen YY, Hsiao YY, Li CI, Yeh CM, Mitsuda N, Yang HX, Chiu CC, Chang SB, Liu ZJ, Tsai WC: **The ancestral duplicated DL/CRC orthologs, PeDL1 and PeDL2, function in orchid reproductive organ innovation.** *J Exp Bot* 2021, **72:**5442-5461.
149. Tsai WC, Pan ZJ, Hsiao YY, Jeng MF, Wu TF, Chen WH, Chen HH: **Interactions of B-class complex proteins involved in tepal development in Phalaenopsis orchid.** *Plant Cell Physiol* 2008, **49:**814-824.
150. Smaczniak C, Immink RG, Angenent GC, Kaufmann K: **Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies.** *Development* 2012, **139:**3081-3098.
151. M; F, GP; DS, K; S, R; EK, FM; Q: **One Protoplast Is Not the Other!** *Plant Physiol* 2011, **156**.