The MADS-box genes expressed in the inflorescence of *Orchis italica* (Orchidaceae)

Maria Carmen Valoroso, Maria Concetta Censullo, Serena Aceto

Department of Biology, University of Naples Federico II, Napoli, Italy

* serena.aceto@unina.it

Abstract

The Orchidaceae family, which is one of the most species-rich flowering plant families, includes species with highly diversified and specialized flower shapes. The aim of this study was to analyze the MADS-box genes expressed in the inflorescence of *Orchis italica*, a wild Mediterranean orchid species. MADS-box proteins are transcription factors involved in various plant biological processes, including flower development. In the floral tissues of *O. italica*, 29 MADS-box genes are expressed that are classified as both class I and II. Class I MADS-box genes include one Mβ-type gene, thereby confirming the presence of this type of MADS-box genes in orchids. The class II MIKC* gene is highly expressed in the column, which is consistent with the conserved function of the MIKC* genes in gametophyte development. In addition, homologs of the SOC, SVP, ANR1, AGL12 and OsMADS32 genes are expressed. Compared with previous knowledge on class II MIKC*C* genes of *O. italica* involved in the ABCDE model of flower development, the number of class B and D genes has been confirmed. In addition, 4 class A (*AP1/FUL*) transcripts, 2 class E (*SEP*) transcripts, 2 new class C (*AG*) transcripts and 1 new *AGL6* transcript have been identified. Within the *AP1/FUL* genes, the sequence divergence, relaxation of purifying selection and expression profiles suggest a possible functional diversification within these orchid genes. The detection of only two *SEP* transcripts in *O. italica*, in contrast with the 4 genes found in other orchids, suggests that only two *SEP* genes could be present in the subfamily Orchidoideae. The expression pattern of the MIKC*C* genes of *O. italica* indicates that low levels at the boundary of the domain of a given MADS-box gene can overlap with the expression of genes belonging to a different functional A-E class in the adjacent domain, thereby following a “fading borders” model.

Introduction

Among the flowering plants, the monocot family Orchidaceae is one of the most species-rich and widespread; this family has adapted to different habitats and exhibits highly specialized reproductive strategies [1]. The Orchidaceae family includes five subfamilies (Apostasioideae, Cypripedioideae, Vanilloideae, Epidendroideae and Orchidoideae) and numerous tribes and subtribes [2]. One of the most attractive orchid structures is the flower that assumes an
enormous variety of shapes and colors among the species although it has a generally conserved structural organization. Three outer tepals, two lateral inner tepals and an inner median tepal (labellum or lip) define a zygomorphic perianth. Male and female reproductive tissues are fused to form the gynostemium or column, and pollen grains are located at the top of this structure. The ovary is located at the base of the column, and its maturation is triggered by pollination [3].

MADS-box genes play a crucial role in the evolution of flower architecture. This family of transcription factors is divided into two lineages, types I and II, that differ in genomic organization, developmental roles, evolutionary rate and level of functional redundancy [4].

Type I MADS-box proteins contain the MADS domain and are divided in $\alpha$, $\beta$ and $\gamma$ based on sequence divergence at the C-terminus [4]. These proteins are involved in seed, embryo and female gametophyte development [5]. Type II MADS-box proteins are the most studied lineage of MADS-box genes given their involvement in different plant developmental processes, including flower formation. These transcription factors are characterized by four domains, including three variably conserved domains and one variable domain, that form the MIKC structure. The highly conserved MADS domain has DNA-binding activity and recognizes the CArG-box motifs in the target genes [6]. The less conserved I and K domains are involved in protein–protein interactions and the formation of protein complexes [7, 8]. The variable C domain has a role in the formation of protein complexes and confers specific functional activity [8, 9].

A duplication event involving the 5’ region of the exon encoding for the K domain followed by neofunctionalization gave rise to two classes of MIKC-type genes: MIKC$^C$ and MIKC$^C$ [10]. The MIKC$^C$ genes are involved in male gametophyte development [11], whereas the MIKC$^C$, the most studied MIKC genes, play different roles in various processes of plant growth and the establishment and maintenance of floral organs [12].

The MIKC$^C$ genes involved in flower organ formation are divided into five functional classes (from A to E), and their activity is described by the ABCDE model of flower development. The MADS-box genes involved in this model form homo- and heterodimers (floral quartets) that regulate specific expression programs in different floral whorls [13].

The ABCDE floral quartet model is applicable to different plant species, where it is well conserved [14–17]; however, analyses of non-model species have revealed differences related to the different structures of the flower [18]. For example, in orchids, the class B MADS-box genes exhibit an expression profile expanded to the first floral whorl, explaining the presence of petaloid sepals [19–21]. In addition, the orchid AP3/DEF lineage of the class B genes underwent two subsequent duplication events that have played an important role in the evolutionary origin of the current structure of the orchid flower described by the “orchid code” and the “homeotic orchid tepal” (HOT) models [22–24]. The more recent “P-code” model integrates the function of class B and AGL6 MADS-box genes to explain the formation of the orchid perianth, particularly the lip [25].

Orchis italica is an orchid species belonging to the subfamily Orchidoideae. It is one of the most widespread Mediterranean orchids with a white-purple cluster inflorescence and a lip with flaps that assume an anthropomorphic form (Fig 1). Previous studies have analyzed the structure, expression and evolution of some genes of O. italica involved in flower development. In particular, the floral meristem identity gene LFY [26], the AP2/ERF gene AP2 (class A) [27] and the MADS-box genes PI/GLO, AP3/DEF (class B), BG (class C), STK (class D) and AGL6 [20, 21, 27–33] have been assessed. However, the other MADS-box genes of O. italica have not yet been studied in contrast to other orchid species belonging to different subfamilies, mainly Epidendroideae [34–39]. Thus, the aim of this work was to identify the MADS-box genes.
expressed in the inflorescence of *O. italica*, analyze their expression in floral tissues, evaluate their evolutionary rate and compare the results with those reported in other orchids.

**Materials and methods**

**Plant material**

The *Orchis italica* plants used in this study (Fig 1A) are grown in the greenhouse at the Department of Biology, University of Naples Federico II. After anthesis, single florets were collected.
and outer tepals, lateral inner tepals, lips (Fig 1B), column and ovary (before pollination) were dissected and stored in RNALater™ (Ambion) at −80˚C for subsequent RNA extraction. Tissues from 10 flowers of two different plants were separately pooled and used for expression analysis.

Isolation of the MADS-box transcripts of *O. italica*

To identify the transcripts encoding for the MADS-box proteins expressed in the floral tissues, a TBLASTN search was conducted against the inflorescence transcriptome of *O. italica* [40] using the MADS domain of the proteins PI/GLO, AP3/DEF, AG and STK of *O. italica* as queries [20, 21, 30–32]. The transcripts with significant hits were virtually translated to exclude those missing the domains downstream the MADS and/or containing premature stop codons. The MADS-box coding sequences (CDSs) identified in the genome of the orchids *Phalaenopsis equestris* (Epidendroideae) [36], *P. aphrodite* [37] and *Apostasia shenzenica* (Apostasioideae) [35] were downloaded from GenBank and Orchidstra2 (http://orchidstra2.abrc.sinica.edu.tw/orchidstra2/pagenome.php) along with MADS-box protein and nucleotide sequences of other plant species (S1 Table).

To annotate the selected transcripts of *O. italica*, BLASTN and BLASTX searches were conducted using the transcripts as queries against the Viridiplantae nucleotide and protein database, respectively.

The nucleotide sequence of all the transcripts of *O. italica* analyzed in this study was verified by PCR amplification and sequencing using the cDNA of inflorescence of *O. italica* as template [33]. The sequences of the primers used are reported in S2 Table.

Phylogenetic and evolutionary analysis

The amino acid alignment of the MADS-box proteins of *O. italica, Phalaenopsis* and *A. shenzenica* was performed using MAFFT [41] and manually adjusted. The maximum likelihood (ML) tree was constructed using PhyML [42], computing the branch support values with approximate likelihood ratio tests.

The amino acid sequences of the MADS-box proteins of *O. italica* belonging to the AP1/FUL, SEP and AGL6 subfamilies were separately aligned to the homolog proteins of different orchids and other plant species (S1 Table), and phylogenetic trees were generated as described above. Protein-based nucleotide alignment of the orchid AP1/FUL, SEP and AGL6 subfamilies was produced using PAL2NAL [43].

The evolutionary rates of the AP1/FUL, SEP and AGL6 CDSs of *O. italica* and other orchids were analyzed using CODEML program from PAML v.4.9 [44]. Branch and branch-sites evolutionary models were tested, and the ratio between nonsynonymous and synonymous substitution rates (ω) was measured. The branch models assume one (one-ratio) or different (two- and three-ratios) ω values among the tree branches. The branch-site models test for positive selection on individual codons in specific branches of the tree and are compared with the respective null models that assume purifying or nearly neutral selection in the same branches. To establish which model best fits the data, a likelihood ratio test was applied comparing the alternative branch and branch-site models (allowing different ω or positive selection, respectively) with the corresponding null models.

miRNA putative target analysis

The psRNATarget web server [45] was used to predict the presence of miRNA putative target sites on the MADS-box transcripts of *O. italica*. The analysis (maximum expectation value 0.0)
Expression analysis
Total RNA was extracted from outer tepals (Te_out), inner lateral tepals (Te_inn), labellum (Lip), column (Co), and ovary (Ov) of *O. italica* after anthesis using the PureLink RNA Mini Kit (Invitrogen). Total RNA was also extracted from leaf tissue (Le). To evaluate the expression profile of the MADS-box transcripts identified in *O. italica*, reverse transcription and real time PCR reactions were conducted as previously reported [33, 46] using the specific primer pairs listed in S2 Table. For each transcript and tissue, the mean relative expression and standard error (SE) were calculated among three technical and two biological replicates. ANOVA followed by Tukey’s post hoc test was performed to assess the statistical significance of the mean relative expression differences among the tissues.

Results and discussion
The inflorescence transcriptome of *O. italica* [40] includes twenty-nine transcripts encoding for MADS-box proteins. This number is lower than that of the MADS-box genes present in the genome of *P. equestris* (51), *D. catenatum* (63) and *A. shenzenica* (36) [35, 36, 38] because it does not include the transcripts specifically expressed in not-floral tissues (e.g., leaf, root, stem, etc.). BLAST analysis and phylogenetic reconstruction (Fig 2) demonstrated that the MADS-box genes expressed in the inflorescence of *O. italica* belong to both class I and II. Excluding three class I and three class II genes (SOC, ANR and one class A), they show floral specific expression or are expressed in floral tissues at levels higher than in leaves (Fig 3).

**Phylogenetic and expression analysis of Class I MADS-box in O. italica**
The four class I proteins of *O. italica* contain a conserved MADS domain (S1A Fig). Among them, two belong to the Mα type, one to Mγ and one to Mβ (Fig 2), confirming the presence of Mβ genes in orchids recently reported in *P. aphrodite* [37] and previously proposed to be absent from the orchid genomes [35]. The class I MADS-box transcripts of *O. italica* are poorly expressed in all the floral tissues (S2 Fig), as expected based on their role in embryo and endosperm maturation [5] and consistent with the low expression in the floral organs of most of the class I MADS-box genes in the orchid *A. shenzenica* [35]. Despite at low levels, all the class I genes identified in *O. italica* (especially the two Mα) are expressed in the column (at whose top are located the pollen grains) (Fig 3, S3 Fig). This pattern supports their involvement in the development of the orchid pollinia, which is consistent with the results obtained in *P. aphrodite*, where two Mα and three Mγ genes are specifically expressed in the pollen [37]. Their different expression in tepals, lip, ovary and leaf suggests a possible role also in other orchid floral and vegetative organs.

**Phylogenetic and expression analysis of Class II MADS-box in O. italica**
**MIKC**
The only MIKC* transcript found in the transcriptome of *O. italica* belongs to the P-subclade (Fig 2). This transcript encodes for a MIKC protein exhibiting a K domain longer than that of the MIKC* transcription factors (S1B Fig) as reported for the MIKC* proteins [10]. This MIKC* gene is expressed at high levels in the column (Fig 3, S3 Fig), which is consistent with the conserved function of the MIKC* genes in gametophyte development [10, 47], as hypothesized for the orchid species *Erycina pusilla* [34] and *P. aphrodite* [37].
SOC, SVP, ANR1, AGL12 and OsMADS32

Among the twenty-four MIKC<sup>C</sup> transcripts identified in the inflorescence transcriptome of O. italica, two are homologs of MADS-box genes involved in the regulation of flowering (SOC1 and SVP). Their expression in the floral tissues of O. italica is weak (S2 and S4 Figs), which is consistent with the expression profile detected in the flowers of the orchids Dendrobium [48].

https://doi.org/10.1371/journal.pone.0213185.g002

Fig. 2. ML tree of the orchid MADS-box proteins. The different classes and groups of MADS-box proteins of O. italica, P. aphrodite, P. equestris and A. shenzenica are highlighted with different colors. The accessions of O. italica are noted in pink.
and *Erycina pusilla* [34] and other plant species [50–52]. Two other transcripts only weakly expressed in the inflorescence of *O. italica* are homologs of MADS-box genes involved in root development (*AGL12* and *ANR1*) [53, 54]. The *AGL12* gene is present in the genome of terrestrial orchids (e.g., *Apostasia*) and seems to be absent in epiphytic orchids [35]. Its presence in *O. italica*, a terrestrial orchid species, supports its role in the formation of true roots typical of terrestrial plant species. Another MIKC<sup>C</sup> transcript expressed at low levels in the floral tissues of *O. italica* is a homolog of *OsMADS32* (also named *CFO1*), a monocot-specific MADS-

![Figure 3. Heatmap of the relative gene expression of the MADS-box genes in the floral tissues of *O. italica*. The scale for each gene is independent of each other and set to 1 as the highest value. Te_out, outer tepals; Te_inn, lateral inner tepals; Co, column; Ov, ovary; Le, leaf.](https://doi.org/10.1371/journal.pone.0213185.g003)
box gene that regulates flower organ identity [55, 56]. In rice and wheat, this gene is expressed in the early stages of the inflorescence and late seed development [50, 57], which could explain the low expression detected in the mature floral tissues of *O. italica*.

The MIKC<sup>C</sup> transcripts most expressed in the floral tissues of *O. italica* belong to the A-E classes of MADS-box genes involved in the ABCDE model of flower development (S2 Fig).

**PI/GLO and AP3/DEF**

Six transcript homologs of the class B genes, including two PI/GLO and four AP3/DEF, have been previously analyzed in *O. italica* [20, 21, 28–31]. Their expression (in particular of the AP3/DEF genes) is not restricted to the second and third floral whorls. Indeed, their expression is extended to the outer tepals (first whorl) (Fig 3), as described for orchids and other species exhibiting petaloid tepals in the first floral whorl [58]. In addition, the orchid AP3/DEF genes have been extensively studied from the evolutionary perspective and are considered the main genes responsible of the diversification of the orchid perianth, as described in the “orchid code” and HOT models [22–24].

**AG and STK**

Also class C (AG) and D (STK) MADS box genes of *O. italica* have been previously analyzed [32]. However, in the present work, two additional transcripts belonging to class C have been identified (Fig 2). Both these transcripts encode for proteins that contain AG-motifs I and II at the C-terminus (S1C Fig) and share 69% of amino acid residues. The protein encoded by the transcript OLcomp1784 exhibits 85% identity with the class C OitaAG (and 62% with the class D OitaSTK), whereas the protein encoded by the transcript OLcomp16674 exhibits 71% identity with OitaAG (and 52% with OitaSTK). A transcript similar to OLcomp16674 (84% identity) is also present in the transcriptome of *Ophrys sphegodes*, an orchid belonging to the Orchidoideae subfamily, as *O. italica*. The expression of class C and D transcripts in *O. italica* is high in the column and ovary (Fig 3, S5 Fig), and this pattern is consistent with the expression of the C and D genes in other orchid species, as they are involved in development of the female reproductive structure [59–61]. The newly identified transcript OLcomp16674 is also detectable in the other floral organs albeit at low levels (S5 Fig). The presence of three AG genes and one STK gene in *O. italica* is consistent with the copy number of class C and D MADS-box genes of *Erycina pusilla* [34, 39], as well as their expression pattern [39].

Excluding two AGL6 transcripts [33], class A and E genes of *O. italica* have never been studied. In the inflorescence of *O. italica*, four class A, two class E, and three AGL6 transcripts are expressed, increasing the number of AGL6 transcripts previously reported in *O. italica*.

**AP1/FUL**

The four class A transcripts of *O. italica* belong to the two monocot FUL-like lineages previously identified [62] (Fig 4). Within each FUL-like lineage, orchid-specific duplications expanded the copy number of the AP1/FUL genes, defining four well supported orchid FUL clades (1a, 1b, 2 and 3). Three AP1/FUL transcripts of *O. italica* encode for proteins exhibiting the complete FUL-like motif LPPWML at the C-terminus (S1D Fig). The protein encoded by the transcript OLcomp3679 lacks the FUL-like motif as observed in other AP1/FUL orchid proteins of species belonging to Epidendroideae [59] and Apostasioideae (S1D Fig). This finding suggests that the divergence of the C-terminus of some orchid AP1/FUL proteins is ancient given that Apostasioideae are the most basal orchid subfamily [2]. The expression of the AP1/FUL transcripts OLcomp3679, 9283 and 2508 is high in column and ovary, and the two latter transcripts are also detectable in the outer and inner tepals (Fig 3, S5 Fig). This expression
profile is very similar to that of EpMADS10, 11 and 12 of E. pusilla [34, 39]. In contrast, expression of the OIcomp11046 transcript is increased in outer and inner tepals compared with column and ovary (Fig 3, S5 Fig). Although evolutionary analysis did not reveal positive selection signals within the API/FUL coding sequences of orchids, relaxation of purifying selection is strongly supported (Table 1, S3 Table). In fact, the $\omega$ value of the orchid API/FUL genes (0.23) is significantly increased compared with orchid SEP (0.11) and AGL6 (0.14) genes. This result together with a previous report of diversifying selection detected in the orchid API/FUL genes [39], the divergence of the C-terminus sequence of some members of this clade in orchids and the peculiar expression profile of OIcomp11046 suggest a possible functional diversification after duplication within these orchid genes.

Fig 4. ML tree of the Class A MADS-box proteins. The orchid branches are highlighted in pink, and orchid FUL clades 1–3 are indicated. The accessions of O. italica are noted in pink. The numbers indicate the statistical support of the branches.
SEP

Two transcripts expressed in the inflorescence of O. italica encode for SEP proteins that harbor both the SEP I and II motifs (S1E Fig) and share 60% identity. In orchids, four clades of SEP genes have been reported that were generated from orchid-specific duplications within monocot SEP clades [34, 39, 59]. One SEP sequence of O. italica belongs to the clade that includes EpMADS9 and PeSEP3 (SEP clade 3), whereas the other SEP sequence is grouped with EpMADS7 and PeSEP4 (SEP clade 4) (Fig 5). BLAST search for SEP transcripts in the transcriptome of Ophrys sphegodes (Orchidoideae) [63] produced two significant hits, both exhibiting high sequence identity with the SEP transcripts identified in O. italica. In addition, a recent study identified two SEP transcripts expressed in the floral buds of Habenaria radiata (Orchidoideae) [64]. This evidence suggests that only two SEP genes are present or that only two SEP genes are expressed in the floral tissues in the subfamily Orchidoideae. Unfortunately, genome data are not currently available for orchid species belonging to Orchidoideae, rendering it difficult to discriminate between the two hypotheses. In both cases, the expression of SEP genes in O. italica is detectable in all floral organs (Fig 3, S6 Fig), and OIcomp1006 exhibits significantly increased expression in lip. This pattern is similar to that reported in wild type H. radiata [64] and other orchid species, such as the Epidendroideae P. equestris [39] and E. pusilla [34, 39], and this finding is consistent with the expression pattern in all the floral whorls of class E genes involved in the formation of all the organs of the flower.
The three AGL6 transcripts expressed in the inflorescence of *O. italica* encode for proteins harboring both conserved AGL6-motifs I and II at the C-terminus (S1F Fig). In orchids, two main AGL6 clades have been reported: one includes other monocot species and one includes only orchid species [25, 37]. The three AGL6 transcripts found in *O. italica* belong to these two clades, and the tree topology clearly show a duplication within the orchid-specific AGL6 clade that gave rise to two different, well supported paralog groups: AGL6 clades 2 and 3 (Fig 6).

Both the clades include basal orchid species (Apostasioideae, Cypripedioideae or Vanilloideae) in addition to Epidendroideae and Orchidoideae, demonstrating the ancient origin of these paralog groups. The newly identified AGL6 transcript OIcomp8204 is the paralog of OIcomp1386 [33]. These genes have an overlapping expression profile (Fig 3, S6 Fig), suggesting a redundant functional role. They are mostly expressed in lip and outer tepals, whereas OIcomp4335 exhibits high expression in the ovary and outer tepals and lower expression in lip. However, comparing the expression of these three AGL6 genes, similar levels are detected in lips, whereas OIcomp4335 expression is significantly increased in tepals (Fig 7). Orchid AGL6 genes together with AP3/DEF genes drive the formation of the orchid perianth. In particular,
the “P-code” model proposes the formation of SP and L protein complexes that include combinations of different AGL6 and AP3/DEF proteins. The SP complex determines the formation of the tepals, and the L complex determines the formation of lip [25]. Given that the AGL6 transcripts of *O. italica* are expressed at similar levels in lip, whereas the four AP3/DEF genes exhibit different expression levels (Fig 7) [30], it is possible to hypothesize that the modulation of the expression level of the genes encoding for the different AP3/DEF proteins is responsible for the formation of the L or SP complex within the specific floral whorl.

The presence of putative miRNA target sites on the twenty-nine MADS-box transcripts of *O. italica* identified in this study was predicted using the psRNATarget web server with a stringent parameter setting (maximum expectation 0.0). The analysis gave positive results only for the transcripts DEF2 and DEF4, known targets of the *O. italica* miRNAs homologs of miR5179 [30]. The other MADS-box transcripts do not have any putative target site, showing that they are not regulated by miRNAs.

**Conclusions**

Current advances in transcriptome and genome sequencing are highlighting the molecular programs that underlie floral evolution and development of non-model species. For many species belonging to basal angiosperms, magnoliids and basal eudicots, the “fading borders model” proposes a gradient of the expression levels of floral homeotic genes to explain flower development [65]. In particular, low expression levels at the boundary of the domain of a given
MADS-box gene overlap with the expression of a different homeotic gene in the adjacent domain. The analysis of all the MADS-box genes expressed in the inflorescence of *O. italica* demonstrate that they follow a “fading borders” scheme and that their expression is generally conserved among orchids of different subfamilies. Given that class I and MIKC\textsuperscript{C} genes are understudied compared with MIKC\textsuperscript{C} genes, it will be interesting to focus forthcoming studies on these classes of MADS-box genes in orchids to understand their developmental role and evolution.

**Supporting information**

S1 Fig. Amino acid motifs of the orchid MADS-box proteins. The sequences of selected orchid species are grouped according to the MADS-box type. A, class I MADS domain; B, I domain and part of the K domain of the MIKC\textsuperscript{C} proteins; C-F, C-terminus of AG/STK, AP1/FUL, SEP and AGL6 proteins, respectively. The sequences of A and B are also aligned with the
corresponding region of one DEF protein of *O. italica*. The conserved motifs of each clade of MADS-box proteins are indicated with a black line. The accessions of *O. italica* are noted in bold.

**(TIF)**

**S2 Fig. Relative expression of the MADS-box genes expressed in the inflorescence of *O. italica*.** Each column of the MADS-box genes indicates the relative expression of 10 floral organs in two cDNA pools (10 floral organs from two different plants), both of which are amplified in triplicate. The error bars represent the standard error of the mean. Te_out, outer tepals; Te_inn, inner tepals; Co, column; Ov, ovary.

**(TIF)**

**S3 Fig. Relative expression of the class I and MIKC + MADS-box genes in floral tissues of *O. italica*.** Each column of the class I and MIKC + genes represents the relative expression of 10 floral organs in two cDNA pools (10 floral organs from two different plants), both of which are amplified in triplicate. The error bars represent the standard error of the mean. Te_out, outer tepals; Te_inn, inner tepals; Co, column; Ov, ovary.

**(TIF)**

**S4 Fig. Relative expression of the SOC, SVP, ANR1, AGL12 and OsMADS32 genes in floral tissues of *O. italica*.** Each column of the *SOC, SVP, ANR1, AGL12* and *OsMADS32* genes represents the relative expression of 10 floral organs in two cDNA pools (10 floral organs from two different plants), both of which are amplified in triplicate. The error bars represent the standard error of the mean. Te_out, outer tepals; Te_inn, inner tepals; Co, column; Ov, ovary.

**(TIF)**

**S5 Fig. Relative expression of the AG and class A MADS-box genes in floral tissues of *O. italica*.** Each column of the class C *AG* and class A *API/FUL* genes indicates the relative expression of 10 floral organs in two cDNA pools (10 floral organs from two different plants), both of which are amplified in triplicate. The error bars represent the standard error of the mean. Te_out, outer tepals; Te_inn, inner tepals; Co, column; Ov, ovary.

**(TIF)**

**S6 Fig. Relative expression of the AGL6 and class E MADS-box genes in floral tissues of *O. italica*.** Each column of the *AGL6* and the class E *SEP* genes shows the relative expression of 10 floral organs in two cDNA pools (10 floral organs from two different plants), both of which are amplified in triplicate. The error bars represent the standard error of the mean. Te_out, outer tepals; Te_inn, inner tepals; Co, column; Ov, ovary.

**(TIF)**

**S1 Table. List of sequences used in the alignments, phylogenetic and evolutionary analyses.** (XLSX)

**S2 Table. List of the primers used in the present work.** (XLSX)

**S3 Table. Likelihood ratio statistics for the comparison of the evolutionary models of the orchid class A, *AGL6* and class E coding regions.** df, degrees of freedom; ns, not significant. (XLSX)
Acknowledgments
The authors are grateful to Prof. Giovanni Scopece for plant material and help in flower dissec-
tion and to Prof. Luciano Gaudio for critical reading of the manuscript.

Author Contributions
Conceptualization: Serena Aceto.
Data curation: Maria Carmen Valoroso, Maria Concetta Censullo.
Formal analysis: Serena Aceto.
Investigation: Maria Carmen Valoroso, Maria Concetta Censullo.
Methodology: Maria Carmen Valoroso, Maria Concetta Censullo.
Project administration: Serena Aceto.
Supervision: Serena Aceto.
Validation: Maria Carmen Valoroso.
Writing – original draft: Serena Aceto.
Writing – review & editing: Maria Carmen Valoroso, Serena Aceto.

References
1. Cozzolino S, Widmer A. Orchid diversity: an evolutionary consequence of deception? Trends Ecol Evol. 2005; 20(9):487–94. https://doi.org/10.1016/j.tree.2005.06.004 WOS:000232128300008. PMID: 16701425
2. Gustafsson ALS, Verola CF, Antonelli A. Reassessing the temporal evolution of orchids with new fossils and a Bayesian relaxed clock, with implications for the diversification of the rare South American genus Hoffmannseggellia (Orchidaceae: Epidendroideae). Bmc Evol Biol. 2010; 10. Art 177 10.1186/1471-2148-10-177. WOS:000280368000001.
3. 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. 2002; 77(3):403–41. https://doi.org/10.1101/S1467931020059396 WOS:000178058500004. PMID: 12227521
4. Parenicova L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, et al. Molecular and phyloge-netic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell. 2003; 15(7):1538–51. https://doi.org/10.1105/tpc.011544 PMID: 12837945; PubMed Central PMCID: PMC165399.
5. Masiero S, Colombo L, Grini PE, Schnittger A, Kater MM. The emerging importance of type I MADS box transcription factors for plant reproduction. The Plant cell. 2011; 23(3):865–72. https://doi.org/10.1105/tpc.110.081737 PMID: 21378131; PubMed Central PMCID: PMC3082269.
6. Tilly JJ, Allen DW, Jack T. The CArG boxes in the promoter of the Arabidopsis floral organ identity gene APETALA3 mediate diverse regulatory effects. Development. 1998; 125(9):1647–57. PMID: 9521903.
7. Reichmann JL, Krizek BA, Meyerowitz EM. Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. Proceedings of the National Academy of Sciences of the United States of America. 1996; 93(10):4793–8. PMID: 8643462; PubMed Central PMCID: PMC39358.
8. Kaufmann K, Melzer R, Theissen G. MiKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. Gene. 2005; 347(2):183–98. https://doi.org/10.1016/j. gene.2004.12.014 PMID: 15777618.
9. Reichmann JL, Meyerowitz EM. MADS domain proteins in plant development. Biological chemistry. 1997; 378(10):1079–101. PMID: 9372178.
10. Kwantes M, Liesbch D, Verelst W. How MiKC* MADS-box genes originated and evidence for their conserved function throughout the evolution of vascular plant gametophytes. Molecular biology and evolution. 2012; 29(1):293–302. https://doi.org/10.1093/molbev/msr200 PMID: 21813465.
11. Zobell O, Faigl W, Saedler H, Munster T. MIKC* MADS-box proteins: conserved regulators of the gametophytic generation of land plants. Molecular biology and evolution. 2010; 27(5):1201–11. https://doi.org/10.1093/molbev/msq005 PMID: 20080864.

12. Aceto S, Gaudio L. The MADS and the Beauty: Genes Involved in the Development of Orchid Flowers. Curr Genomics. 2011; 12(5):342–56. https://doi.org/10.2174/138920211794629754 PMID: 22294877; PubMed Central PMCID: PMC3145264.

13. Theissen G, Melzer R, Rumpler F. MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. Development. 2016; 143(18):3259–71. https://doi.org/10.1242/dev.134080 WOS:000393450700007 PMID: 27624831

14. 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. Molecular cell. 2000; 5(3):569–79. PMID: 10882141.

15. 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(24):6083–91. https://doi.org/10.1242/dev.01523 PMID: 15537689.

16. 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. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104(3):1081–6. https://doi.org/10.1073/pnas.0606431104 PMID: 17210918; PubMed Central PMCID: PMC1783367.

17. Ferrario S, Immink RG, Angenent GC. Conservation and diversity in flower land. Current opinion in plant biology. 2004; 7(1):84–91. PMID: 14732446.

18. Irish VF, Litt A. Flower development and evolution: gene duplication, diversification and redeployment. Current opinion in genetics & development. 2005; 15(4):454–60. https://doi.org/10.1016/j.gde.2005.06.001 PMID: 15964755.

19. Kanno A, Nakada M, Akita Y, Hirai M. Class B gene expression and the modified ABC model in non-gametophytic generation of land plants. Molecular biology and evolution. 2010; 27(5):1201–11. https://doi.org/10.1093/molbev/msq005 PMID: 19162144

20. Cantone C, Sica M, Gaudio L, Aceto S. The OrcPI locus: Genomic organization, expression pattern, and noncoding regions variability in Orchis italica (Orchidaceae) and related species. Gene. 2009; 434(1–2):9–15. https://doi.org/10.1016/j.gene.2008.12.015 WOS:000264589800002 PMID: 19162144

21. Salemme M, Sica M, Gaudio L, Aceto S. Expression pattern of two paralogs of the PI/GLO-like locus during Orchis italica (Orchidaceae, Orchidinae) flower development. Development genes and evolution. 2011; 221(4):241–6. https://doi.org/10.1007/s00427-011-0372-6 PMID: 21892737.

22. Mondragon-Palomino M, Theissen G. Why are orchid flowers so diverse? Reduction of evolutionary constraints by paralogues of class B floral homeotic genes. Ann Bot-London. 2009; 104(3):583–94. https://doi.org/10.1093/aob/mcn258 WOS:000268807600020 PMID: 19141602

23. 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(6):1008–19. https://doi.org/10.1111/j.1365-313X.2011.04560.x WOS:000292103900008 PMID: 21435045

24. Pan PZ, Cheng CC, Tsai WC, Chung MC, Chen WH, Hu JM, et al. The duplicated B-class MADS-box genes display dualistic characters in orchid floral organ identity and growth. Plant Cell Physiol. 2011; 52(9):1515–31. https://doi.org/10.1093/pcp/pcr092 PMID: 21757456.

25. Hsu HF, Hsu WH, Lee YI, Mao WT, Yang JY, Li JY, et al. Model for perianth formation in orchids. Nat Plants. 2015; 1(5). Artn 15046 https://doi.org/10.1038/Nplants.2015.46 WOS:000364398500001

26. Montieri S, Gaudio L, Aceto S. Isolation of the LFY/FLO homologue in Orchis italica and evolutionary analysis in some European orchids. Gene. 2004; 333:101–9. https://doi.org/10.1016/j.gene.2004.02.015 WOS:000222074800012 PMID: 15177685

27. Salemme M, Sica M, Iazzetti G, Gaudio L, Aceto S. The AP2-like gene OltaAP2 is alternatively spliced and differentially expressed in inflorescence and vegetative tissues of the Orchid Orchis italica. PLoS One. 2013; 8(10):e77454. https://doi.org/10.1371/journal.pone.0077454 PMID: 24204832; PubMed Central PMCID: PMC3804621.

28. Aceto S, Cantone C, Chiaiese P, Ruotolo G, Sica M, Gaudio L. Isolation and Phylogenetic Footprinting Analysis of the S'-Regulatory Region of the Floral Homeotic Gene OrcPI from Orchis italica (Orchidaceae). J Hered. 2010; 101(1):124–31. https://doi.org/10.1093/jhered/esp062 WOS:000272687700016 PMID: 19861638

29. Aceto S, Montieri S, Sica M, Gaudio L. Molecular evolution of the OrcPI locus in natural populations of Mediterranean orchids. Gene. 2007; 392(1–2):299–305. https://doi.org/10.1016/j.gene.2007.01.005 WOS:000246078300034 PMID: 17321699
30. Aceto S, Sica M, De Paolo S, D’Argenio V, Cantiello P, Salvatore F, et al. The analysis of the inflorescence miRNome of the orchid Orchis italica reveals a DEF-like MADS-box gene as a new miRNA target. PloS one. 2014; 9(5):e97839. https://doi.org/10.1371/journal.pone.0097839 PMID: 24832004; PubMed Central PMCID: PMC4022656.

31. Cantone C, Gaudio L, Aceto S. The PI/GLO-like locus in orchids: duplication and purifying selection at synonymous sites within Orchidinae (Orchidaceae). Gene. 2011; 481(1):48–55. https://doi.org/10.1016/j.gene.2011.04.004 PMID: 21549812.

32. Salemme M, Sica M, Gaudio L, Aceto S. The OitaAG and OitaSTK genes of the orchid Orchis italica: a comparative analysis with other C- and D-class MADS-box genes. Molecular biology reports. 2013; 40(5):3523–35. https://doi.org/10.1007/s11033-012-2426-x PMID: 23277396.

33. 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 biology and evolution. 2017; 9(6). https://doi.org/10.1093/gbe/exv101 PMID: 28541415; PubMed Central PMCID: PMC5499889.

34. Lin CS, Hsu CT, Liao DC, Chang WJ, Chou ML, Huang YT, et al. Transcriptome-wide analysis of the MADS-box gene family in the orchid Erycina pusilla. Plant biotechnology journal. 2016; 14(1):284–98. https://doi.org/10.1111/pbi.12383 PMID: 25917508.

35. Zhang GQ, Liu KW, Li Z, Lohaus R, Hsiao YY, Niu SC, et al. The Apostasia genome and the evolutionary origin of floral organs of Erycina pusilla, an emerging orchid model system. BMC Evol Biol. 2017; 17:ARTN 89 https://doi.org/10.1186/s12862-016-0844-4 WOS:000397333900005.

36. De Paolo S, Salvemini M, Gaudio L, Aceto S. De novo transcriptome assembly from inflorescence of Orchis italica: analysis of coding and non-coding transcripts. PloS one. 2014; 9(7):e102155. https://doi.org/10.1371/journal.pone.0102155 PMID: 25025767; PubMed Central PMCID: PMC4090910.

37. Chao YT, Chen WC, Chen CY, Ho HY, Yeh CH, Kuo YT, et al. Chromosome-level assembly, genetic and physical mapping of Phalaenopsis aphrodite genome provides new insights into species adaptation and resources for orchid breeding. Plant biotechnology journal. 2018. https://doi.org/10.1111/pbi.12936 PMID: 29704444.

38. Zhang GQ, Xu Q, Bian C, Tsai WC, Liu KW, et al. The Dendrobium catenatum Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. Scientific reports. 2016; 6:19023. https://doi.org/10.1038/srep19023 PMID: 26754549; PubMed Central PMCID: PMC4709516.

39. Dirks-Mulder A, Butot R, van Schaik P, Wijnands JWPM, van den Berg R, Krol L, et al. Exploring the evolutionary origin of floral organs of Erycina pusilla, an emerging orchid model system. Bmcevol Biol. 2017;17:17. ARTN 89 https://doi.org/10.1186/s12862-016-0844-4 WOS:000397333900005.

40. De Paolo S, Salvemini M, Gaudio L, Aceto S. De novo transcriptome assembly from inflorescence of Orchis italica: analysis of coding and non-coding transcripts. PloS one. 2014; 9(7):e102155. https://doi.org/10.1371/journal.pone.0102155 PMID: 25025767; PubMed Central PMCID: PMC4090910.

41. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution. 2013; 30(4):772–80. https://doi.org/10.1093/molbev/mst010 PMID: 23329690; PubMed Central PMCID: PMC3603318.

42. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic biology. 2010; 59(3):307–21. https://doi.org/10.1093/sysbio/syq010 PMID: 20525638.

43. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic acids research. 2006; 34(Web Server issue):W609–12. https://doi.org/10.1093/nar/gkl315 PMID: 16845082; PubMed Central PMCID: PMC1538804.

44. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Molecular biology and evolution. 2007; 24(8):1586–91. https://doi.org/10.1093/molbev/msm088 PMID: 17431113.

45. Dai XB, Zhuang ZH, Zhao PX. psRNATarget: a plant small RNA target analysis server (2017 release). Nucleic acids research. 2018; 46(W1):W49–W54. https://doi.org/10.1093/nar/gky316 WOS:000438574100009. PMID: 29718424

46. De Paolo S, Gaudio L, Aceto S. Analysis of the TCP genes expressed in the inflorescence of the orchid Orchis italica. Scientific reports. 2015; 5:16265. https://doi.org/10.1038/srep16265 PMID: 26531864; PubMed Central PMCID: PMC4632031.

47. Liu Y, Cui S, Wu F, Yan S, Lin X, Du X, et al. Functional conservation of MIKC+ Type MADS box genes in Arabidopsis and rice pollen maturation. The Plant cell. 2013; 25(4):1289–303. https://doi.org/10.1105/tpc.112.110949 PMID: 23613195; PubMed Central PMCID: PMC3663266.

48. Ding L, Wang Y, Yu H. Overexpression of DOSOC1, an ortholog of Arabidopsis SOC1, promotes flowering in the orchid Dendrobium Chao Parya Smile. Plant & cell physiology. 2013; 54(4):595–608. https://doi.org/10.1093/pcc/pcp026 PMID: 23396600.
49. Liu XR, Pan T, Liang WQ, Gao L, Wang XJ, Li HQ, et al. Overexpression of an Orchid (Dendrobium nobile) SOC1/TM3-Like Ortholog, DnAGL19, in Arabidopsis Regulates HOS1-FT Expression. Frontiers in plant science. 2016; 7:99. https://doi.org/10.3389/fpls.2016.00099 PMID: 26904066; PubMed Central PMCID: PMC4746357.

50. Arora R, Agarwal P, Ray S, Singh AK, Singh VP, Tyagi AK, et al. MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. Bmc Genomics. 2007; 8. Art 242 https://doi.org/10.1186/1471-2164-8-242 WOS:000248702600001. PMID: 17640358

51. Diaz-Riquelme J, Lijavetzky D, Martinez-Zapater JM, Carmona MA. Genome-Wide Analysis of MIKCC-Type MADS Box Genes in Grapevine. Plant Physiol. 2009; 149(1):354–69. https://doi.org/10.1104/pp.108.131052 WOS:000262261500040. PMID: 18997115

52. Hu LF, Liu SQ. Genome-wide analysis of the MADS-box gene family in cucumber. Genome. 2012; 55(3):245–56. https://doi.org/10.1139/g2012-009 WOS:000301043300007. PMID: 22376137

53. Zhao T, Ni ZF, Dai Y, Yao YY, Nie XL, Sun QX. Characterization and expression of 42 MADS-box genes in wheat (Triticum aestivum L.). Mol Genet Genomics. 2006; 276(4):334–50. https://doi.org/10.1007/s00438-006-0147-3 WOS:000241553200003. PMID: 16858583

54. Dodsworth S., Petal Sepal, or Tepal? B-Genes and Monocot Flowers. Trends Plant Sci. 2017; 22(1):8–10. https://doi.org/10.1016/j.tplants.2016.11.006 WOS:000392892400005. PMID: 27894712

55. Sedeek KEM, Qi WH, Schauer MA, Gupta AK, Poveda L, Xu SQ, et al. Transcriptome and Proteome Data Reveal Candidate Genes for Pollinator Attraction in Sexually Deceptive Orchids. PloS one. 2013; 8(5). ARTN e64621 https://doi.org/10.1371/journal.pone.0064621 WOS:0003279534200016. PMID: 23095287

56. Skinner M, Johansen LB, Pedersen KB, Frederiksen S, Johansen BB. Cloning and transcription analysis of an AGAMOUS- and SEEDSTICK ortholog in the orchid Dendrobium thrysiflorum (Reichb. f.). Gene. 2006; 366(2):266–74. https://doi.org/10.1016/j.gene.2005.08.014 WOS:000236312600008. PMID: 16236468

57. 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(4):1255–65. https://doi.org/10.1534/genetics.115.182964 WOS:000373959100004. PMID: 27053123