Transcriptome-Wide Identification and Expression Analysis of *DIVARICATA*- and *RADIALIS*-Like Genes of the Mediterranean Orchid *Orchis italica*

Maria Carmen Valoroso, Sofia De Paolo, Giovanni Iazzetti, and Serena Aceto*

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

*Corresponding author: E-mail: serena.aceto@unina.it.

Accepted: May 24, 2017

Data deposition: This project has been deposited at GenBank under the accessions KY089088-KY089099.

Abstract

Bilateral symmetry of flowers is a relevant novelty that has occurred many times throughout the evolution of flowering plants. In *Antirrhinum majus*, establishment of flower dorso-ventral asymmetry is mainly due to interaction of TCP (CYC and DICH) and MYB (DIV, RAD, and DRIF) transcription factors. In the present study, we characterized 8 *DIV*-, 4 *RAD*-, and 2 *DRIF*-like genes from the transcriptome of *Orchis italica*, an orchid species with bilaterally symmetric and resupinate flowers. We found a similar number of *DIV*- and *RAD*-like genes within the genomes of *Phalaenopsis equestris* and *Dendrobium catenatum* orchids. Orchid DIV- and RAD-like proteins share conserved motifs whose distribution reflects their phylogeny and analysis of the genomic organization revealed a single intron containing many traces of transposable elements. Evolutionary analysis has shown that purifying selection acts on the *DIV*- and *RAD*-like coding regions in orchids, with relaxation of selective constraints in a branch of the *DIV*-like genes. Analysis of the expression patterns of *DIV*- and *RAD*-like genes in *O. italica* revealed possible redundant functions for some of them. In the perianth of *O. italica*, the ortholog of *DIV* and *DRIF* of *A. majus* are expressed in all tissues, whereas *RAD* is mainly expressed in the outer tepals and lip. These data allow for proposal of an evolutionary conserved model in which the expression of the orthologs of the *DIV*, *RAD*, and *DRIF* genes might be related to establishment of flower bilateral symmetry in the nonmodel orchid species *O. italica*.

Key words: Orchidaceae, DIV-like, RAD-like, MYB transcription factors, flower symmetry.

Introduction

Because the origin of angiosperms, bilateral flower symmetry (zygomorphy) has evolved several times in diverse lineages from the ancestral condition of radial symmetry (actinomorphy) (Citerne et al. 2010; Endress 2012). The transition from radial to bilateral symmetry is often related to the coevolution of specialized pollinator insects (Fenster et al. 2009; Ushimaru et al. 2009) and is considered a key innovation associated with angiosperm diversification (Vamosi and Vamosi 2010).

During the last 20 years, the molecular basis of flower symmetry has been studied in *Antirrhinum majus* (Coen 1996; Luo et al. 1996, 1999; Almeida et al. 1997; Galego and Almeida 2002). In this species, four genes play a key role in establishment of the asymmetry of the corolla and individual petals: *CYCLOYDEA (CYC), DICHOTOMA (DICH), RADIALIS (RAD)*, and *DIVARICATA (DIV)*. *CYC* and *DICH* encode transcription factors of the TCP family that are expressed in the dorsal domains of the snapdragon flower (Luo et al. 1996, 1999; Galego and Almeida 2002; Almeida and Galego 2005). The double mutants *cyc:dich* have radially symmetric flowers with petals that exhibit ventral identity (Almeida et al. 1997). The *DIV* and *RAD* genes encode MYB transcription factors. The DIV protein contains two MYB domains and, although expressed in both the dorsal and ventral parts of the flower, is involved in establishment of the ventral identity (Galego and Almeida 2002). RAD is a small protein containing a single MYB domain that is expressed in the dorsal part of the snapdragon flower (Corley et al. 2005). In *A. majus*, the *DIV* and *RAD* genes work antagonistically to determine the dorsoventral axis of the flower and their interaction is mediated by another MYB transcription factor, *DIV-and-RAD-interacting-factor* (*DRIF*). The RAD protein does not directly interact with DIV but binds the DRIF protein preventing DIV activation in the dorsal domain. DRIF is a co-activator of DIV that is

expressed in the dorsal and ventral parts of the flower. The DIV-DRIF complex can interact with specific DNA regions to activate other genes involved in ventralization of the flower. In the dorsal domain of the flower, RAD competes with DIV to bind DRIF, inhibiting formation of the DIV-DRIF complex. Thus, DIV cannot activate the genes involved in ventralization in the dorsal domain (Raimundo et al. 2013).

The monocot family Orchidaceae is known for its species richness (more than 20,000) and the beauty of their flowers. Although extremely diversified among species, the orchid flowers share a common basic structure: They are zygomorphic with three outer tepals, two inner lateral tepals and an inner median tepal called labellum or lip. The innermost whorl, the column, is composed of fused male and female reproductive tissues. The pollen grains are at the top of the column and the ovary, whose maturation is triggered by pollination, is at the base (Rudall and Bateman 2002). Many orchids have a resupinate flower in which the pedicel and ovary undergo a 180° rotation resulting in a reversal of the positions of the ventral and dorsal organs in the mature flower (Rudall and Bateman 2002) (fig. 1*D*).

In orchids, the MADS-box genes involved in formation of the perianth are well characterized (Mondragon-Palomino et al. 2009; Aceto and Gaudio 2011; Cantone et al. 2011; Mondragon-Palomino and Theissen 2011; Pan et al. 2011, 2014; Salemme et al. 2011, 2013a; Aceto et al. 2014; Acri-Nunes-Miranda and Mondragon-Palomino 2014), and recent studies have highlighted the role of the MADS-box gene *AGL6* in the formation of the orchid lip (Hsu et al. 2015; Huang et al. 2016). Conversely, few studies are available on orchid TCP genes (Mondragon-Palomino and Trontin 2011; De Paolo et al. 2015; Lin et al. 2016), and the MYB genes that are possibly involved in orchid flower symmetry have not been characterized. The recent release of the genomes of *Phalaenopsis equestris* and *Dendrobium catenatum* orchids (Cai et al. 2015; Zhang et al. 2016) and the availability of many orchid transcriptomes represent a valuable resource for analysis of gene families in nonmodel orchid species. The aim of this work was to identify and characterize the *DIV*- and *RAD*-like genes expressed in the inflorescence transcriptome of the Mediterranean orchid *Orchis italica* (sub-family Orchidoideae, tribe Orchidinae) that is characterized by a dense oval inflorescence with numerous pink resupinate flowers (Montieri et al. 2004) (fig. 1*A–C*). The orthologs of these genes were also identified in the *P. equestris* and *D. catenatum* (Epidendroideae) orchid genomes to perform phylogenetic and evolutionary analyses. In addition, the expression patterns of orchid *DIV*-, *RAD*-, *DRIF*-, and *AGL6*-like genes

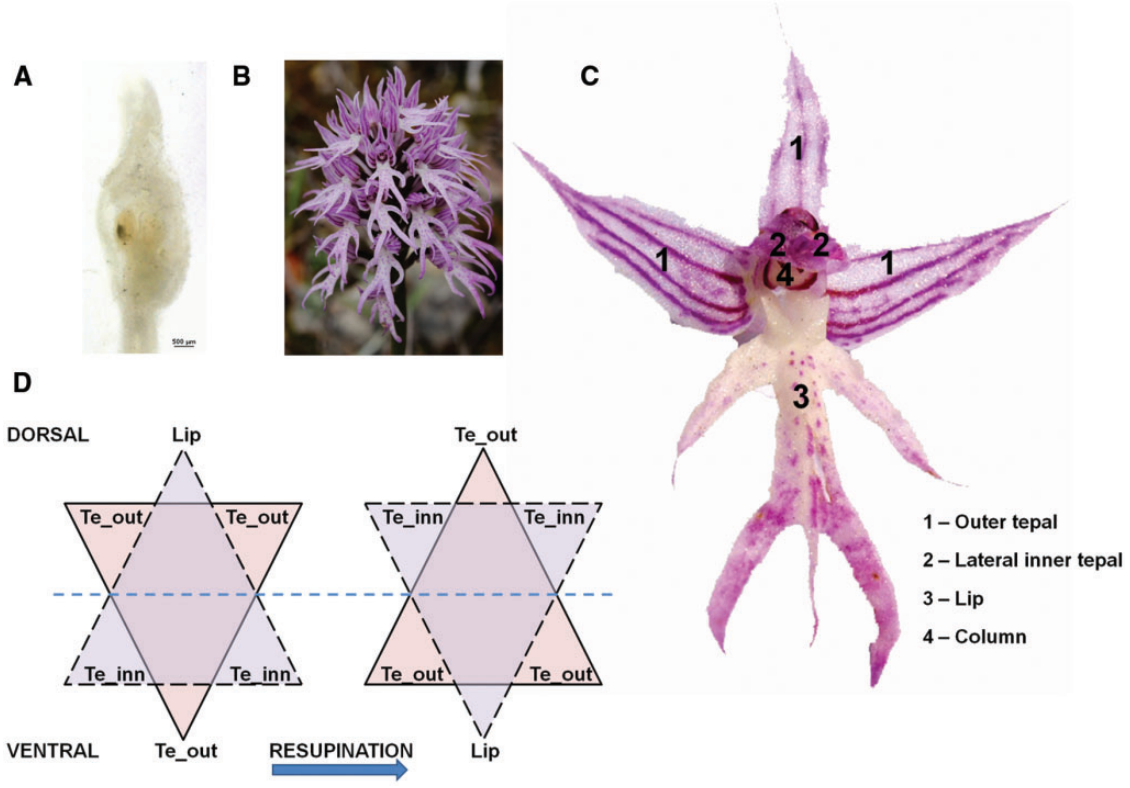

Fig. 1.—The Mediterranean orchid *Orchis italica*. (*A*) Single floret (early stage), (*B*) Inflorescence after anthesis; (*C*) Single floret after anthesis (late stage); (*D*) diagram showing the positions of the organs of the first (continuous line) and second (dotted line) whorls of the perianth before (left) and after (right) resupination;. Te_out, outer tepal; Te_inn, inner tepal.

were examined in different tissues of *O. italica* to evaluate their possible involvement in the bilateral symmetry of the orchid flower.

Materials and Methods

Plant Material

The *Orchis italica* plants used in this study were grown in the greenhouse of the Department of Biology of the University of Naples Federico II (Napoli, Italy). Two developmental stages were examined: Bud stage before anthesis (early) and mature inflorescence after anthesis (late) (fig. 1*A*–*C*). Single florets before anthesis were used for *in situ* hybridization experiments. After anthesis, single florets of the same inflorescence were collected, and some were dissected to separate the floral tissues (outer tepals, inner lateral tepals, lip, column, and ovary before pollination). Leaf tissue was also collected. These tissues were stored in RNA-later (Ambion) until nucleic acid extraction.

Identification of Orchid *DIV*- and *RAD*-Like Genes

The DIV and DVL1 (AAL78741, AAL78742) and RAD and RAD-like (AAX48042, ABI14752, ABI14753, ABI14755, AJ791699, AJ793240) proteins of *Antirrhinum majus* were used as queries to perform a TBLASTN search (*e*-value 1*e*−003) against the inflorescence transcriptomes of *Orchis italica* (Orchidoideae) (De Paolo et al. 2014) and *Ophrys* (Orchidoideae) (Sedeek et al. 2013) and against the genomes of *Phalaenopsis equestris* (Epidendroideae) (Cai et al. 2015) and *Dendrobium catenatum* (Epidendroideae) (Zhang et al. 2016). The same TBLASTN search was conducted on the coding sequences (CDSs) of *P. equestris* downloaded from ftp://ftp.genomics.org.cn/from_BGISZ/20130120/. The CDSs with significant hits were used to perform a BLASTN search on the genomic scaffolds of *P. equestris* to verify the eventual presence of introns in the *DIV*- and *RAD*-like genes and to identify the exon/intron junctions. Based on the results obtained from *P. equestris*, the scaffolds of *D. catenatum* were scanned to verify the presence and position of introns in regions encoding for DIV- and RAD-like proteins. Virtual translation of the selected nucleotide sequences of *O. italica*, *Ophrys*, *P. equestris*, and *D. catenatum* was performed to check for indels and/or stop codons in the CDSs. Finally, to identify DIV- and RAD-like orthologs in orchids, the best reciprocal hits were obtained from pairwise all-versus-all BLASTP searches conducted between the DIV- and RAD-like predicted amino acid sequences of *O. italica*, *Ophrys*, *P. equestris, D. catenatum*, and *A. majus*. In addition, the orthology of these amino acid sequences was predicted through a BLASTP search against the nonredundant protein database and an OrthoMCL analysis (Li et al. 2003). A BLASTCLUST search was conducted for paralog identification within each orchid species using the following parameters: *e*-value $\leq 10^{-10}$ and 30% minimum similarity (–S 30) over 75% of the protein (–L 0.75) (Ponting and Russell 2002; Devos et al. 2016). *DIV*- and *RAD*-like sequence information for *O. italica*, *Ophrys*, *P. equestris*, and *D. catenatum* is listed in table 1. The predicted orthology and paralogy relationships are reported in supplementary table S1, Supplementary Material online.

Analysis of the *DIV*- and *RAD*-Like Transcripts of *O. italica*

Total RNA was extracted from 10 pooled florets collected from a single inflorescence of *O. italica* after anthesis using Trizol (Ambion) followed by DNase treatment. After spectrophotometric quantification using a Nanodrop 2000 (Thermo Fisher Scientific), RNA (500 ng) was reverse transcribed using the Advantage RT-PCR kit (Clontech) and an oligo dT primer. Using the nucleotide sequences selected from the floral transcriptome of *O. italica*, primer pairs (listed in supplementary table S2, Supplementary Material online) were designed to specifically amplify the entire CDS of the *DIV*- and *RAD*-like transcripts, spanning from the 5′ to the 3′ UTR. When the *in silico* assembled *DIV*- and *RAD*-like sequence of *O. italica* did not include the 5′ and/or 3′ UTR (OITA_12910, OITA_36312, and OITA_72143), specific primers were designed to perform 5′ and 3′ RACE experiments using the FirstChoice RLM-RACE kit (Thermo Fisher Scientific). PCR amplifications were conducted using 30 ng of first strand cDNA, 0.20 mM dNTPs, 0.5 μM of each primer, 1× buffer and 2.5 U HotMaster Taq DNA polymerase (5 Prime). The thermal cycle was as follows: 94 °C for 3 min, 30 cycles of 94 °C for 30 s, 60 °C for 20 s, 65 °C for a time dependent upon the amplicon size (from 30 s to 3 min), followed by a final extension for 10 min at 65 °C. The amplicons were cloned into the pSC-A-amp/kan vector (Agilent), sequenced using the T3 and T7 plasmid primers and analyzed using an ABI 310 Automated Sequencer (Applied Biosystems).

Structural Analysis of the *DIV*- and *RAD*-Like Genes of *O. italica*

The total DNA was extracted from leaf tissue of *O. italica* using the ISOLATE II Plant DNA kit (Bioline) followed by RNase treatment and quantification. Based on the intron position of the *DIV*- and *RAD*-like genes of *P. equestris* and *D. catenatum*, specific primer pairs (listed in supplementary table S2, Supplementary Material online) were designed to amplify the introns of these genes in *O. italica*. One hundred-fifty ng of *O. italica* genomic DNA were used as a template in PCR amplifications using the specific primer pair and the Herculase II Fusion DNA polymerase (Agilent) in the amplification conditions suggested by the manufacturer. When the first reaction did not produce a sufficient amount of amplification product, a second reaction was conducted using a 1:10 (v/v) dilution of the first reaction

Table 1
The *DIV*- and *RAD*-Like Orchid Genes Examined

Subfamily	Species	Transcript_Name	cDNA_Size	ORF_Size	Intron_Size	Accession/Scaffold
DIV-like	*O. italica*	OITA_9548	1,361	864	76	KY089088
		OITA_23026[a]	737	528	111	KY089095
		OITA_3530	1,901	873	201	KY089091
		OITA_35312	1,182	894	1,615	KY089094
		OITA_13252[b]	1,694	861	3,591	KY089093
		OITA_13233	1,427	840	~5,000	KY089092
		OITA_8681	1,472	855	~6,000	KY089089
		OITA_12910	1,101	861	~10,000	KY089090
	Ophrys	OPH_5397	465	462	Unknown	
		OPH_23790	809	708	Unknown	
		OPH_29334	1,384	840	Unknown	
		OPH_22868	1,145	675	Unknown	
	P. equestris	PEQU_28155		891	246	Scaffold000684_20
		PEQU_02344		510	79	Scaffold000002_3686
		PEQU_00313		846	276	Scaffold000002_542
		PEQU_37054		891	2,090	Scaffold000306_1
		PEQU_03826		861	12,810	Scaffold000002_5659
		PEQU_21184		870	5,008	Scaffold000065_25
		PEQU_09358		885	174	Scaffold000006_57
	D. catenatum	DCAT_JSDN01S056792		888	124	JSDN01S056792
		DCAT_JSDN01S046716		525	87	JSDN01S046716
		DCAT_JSDN01S069654		876	275	JSDN01S069654
		DCAT_JSDN01S055018		882	1,327	JSDN01S055018
		DCAT_JSDN01S002760		804	8,197	JSDN01S002760
		DCAT_JSDN01S031580		867	2,705	JSDN01S031580
		DCAT_JSDN01S018459		855	3,275	JSDN01S018459
		DCAT_JSDN01S042335		891	231	JSDN01S042335
RAD-like	*O. italica*	OITA_56510	420	282	1,012	KY089097
		OITA_103296	406	285	1,056	KY089098
		OITA_32153	369	249	1,135	KY089096
		OITA_72143	543	267	No	KY089099
	Ophrys	OPH_2593	660	282	Unknown	
	P. equestris	PEQU_08237		288	1,775	Scaffold000054_59
		PEQU_31458		294	864	Scaffold000460_4
		PEQU_40552		279	1,191	Scaffold198270_12
		PEQU_25415		249	No	Scaffold000067_1
		PEQU_05110		294	No	Scaffold000840_202
	D. catenatum	DCAT_JSDN01S033743		288	1,240	JSDN01S033743
		DCAT_JSDN01S033707		285	1,178	JSDN01S033707
		DCAT_JSDN01S041675		279	1,108	JSDN01S041675
		DCAT_JSDN01S013347		324	No	JSDN01S013347
		DCAT_JSDN01S059067		303	No	JSDN01S059067

NOTE.—The size is expressed in nucleotides.
[a]Alternative splicing with intron retention resulting in a premature stop codon.
[b]Alternative splicing of an additional intron within the 5′ UTR (88 bp) that is also conserved in PEQU (86 bp) and DCAT (90 bp).

with nested primers. dATP tailing was performed on the ends of the amplicons by adding 0.5 U of DreamTaq DNA polymerase (Thermo Fisher Scientific) and 100 μM dATP to the PCR reaction and incubating at 72 °C for 15 min. The amplification products were cloned and sequenced as described above using vector- and intron-specific primers (supplementary table S2, Supplementary Material online). When the intron size exceeded 5,000 bp, direct sequencing reactions were performed on the amplification product. The nucleotide sequences of the *DIV*- and *RAD*-like genes of *O. italica* were deposited in GenBank and their accession numbers are listed in table 1.

The intron nucleotide sequences of the *DIV*- and *RAD*-like genes of *O. italica*, *P. equestris*, and *D. catenatum* were scanned for transposable/repetitive elements using CENSOR software (Kohany et al. 2006).

Phylogenetic and Evolutionary Analyses

The predicted amino acid sequences of the DIV- and RAD-like proteins of *O. italica*, *Ophrys*, *P. equestris*, and *D. catenatum* were aligned using MUSCLE (Edgar 2004) and manually adjusted. The corresponding alignments of the *DIV*- and *RAD*-like CDSs were obtained using PAL2NAL (Suyama et al. 2006). Maximum likelihood and neighbor joining (NJ) trees of the DIV- and RAD-like amino acid sequences were constructed using MEGA7 software (Tamura et al. 2013) using the JTT +G amino acid substitution model with 1,000 bootstrap replicates.

The presence of shared conserved motifs among the predicted amino acid sequences of the DIV- and RAD-like proteins of *O. italica*, *Ophrys*, *P. equestris*, and *D. catenatum* was verified using the MEME online tool (Bailey et al. 2009), with the search parameters set to any number of repetitions, an optimum width from 6 to 50 and a maximum number of motifs of 10.

The *DIV*- and *RAD*-like CDSs of *O. italica*, *Ophrys*, *P. equestris*, and *D. catenatum* were analyzed to test for variation of evolutionary rates at specific codons in the sequences and among branches with the CODEML program from PAML v.4.8 (Yang 1997). Different models (branch, sites, branch-sites, and clades) were compared with measure the ω value (ratio between nonsynonymous and synonymous substitution rates) of the *DIV*- and *RAD*-like nucleotide sequences. When ω is 1, neutral selection is acting on the examined sequences whereas purifying or positive selection drive their evolution when it is significantly lower or higher than 1, respectively (Yang and Bielawski 2000). The branch models assume one or different ω values among the branches of the DIV- and RAD-like tree. The sites models verify the presence of positively selected codons among the sequences, and the branch-sites models test for positive selection on individual codons in specific branches of the tree. The clade models check for the presence of different selective constraints between clades after gene duplication. A likelihood ratio test was applied to establish which model best fits the data.

Identification of the *DRIF*- and *AGL6*-Like Transcripts of *O. italica*

The sequence of the DRIF proteins of *A. majus* (DRIF1, AGL11918; DRIF2, AGL11919) (Raimundo et al. 2013) and of the AGL6 proteins of the *Phalaenopsis aphrodite* orchid (AGL6-1, PATC154379; AGL6-2 PATC138772) (Su et al. 2013) were used as queries to perform a TBLASTN search against the inflorescence transcriptome of *O. italica*. Validation of the expression of the identified transcripts was assessed by PCR amplification of the cDNA of inflorescence of *O. italica* using specific primer pairs (supplementary table S2, Supplementary Material online). The amplification products were cloned and sequenced as described above.

Expression Analysis

Single florets of *O. italica* collected from the same floral bud before anthesis (fig. 1*A*) were fixed in 4% (v/v) paraformaldehyde, 0.5% (v/v) glutaraldehyde, 0.1% Triton X-100, and 4% dimethylsulfoxide in phosphate–saline buffer 1× for 16 h at 4 °C (Javelle et al. 2011) and then dehydrated through an ethanol series, paraffin embedded, and sectioned at 9 μm. Digoxigenin-labeled sense and antisense RNA probes of the transcripts OITA_9548 (*DIV*), OITA_56510 (*RAD*) and OITA_10599 (*DRIF1*) 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 performed using the DIG Nucleic Acid Detection kit (Roche) following the manufacturer's instructions.

Using the protocol described above, the total RNA was extracted from floral tissues of *O. italica* after anthesis (fig. 1*B* and *C*): Outer tepals (Te_out), inner lateral tepals (Te_inn), labellum (Lip), column (Co), and ovary not pollinated (Ov). The total RNA was also extracted from leaf tissue. After reverse transcription, 30 ng of the first strand cDNA from each tissue were amplified using real-time PCR with SYBR Green PCR Master Mix (Life Technologies) in technical triplicates and biological duplicates and the 5.8S RNA as the endogenous control gene (De Paolo et al. 2015). Specific primer pairs (supplementary table S2, Supplementary Material online) were designed to evaluate the relative expression of the *DIV*-, *RAD*-, *DRIF*- and *AGL6-like* transcripts of *O. italica*, including two differentially spliced *DIV*-like transcripts. The relative expression ratio (Rn) was calculated by applying the formula:

$$Rn = \frac{(1 + E_{target})^{-CTtarget}}{(1 + E_{5.8S})^{-CT5.8S}}$$

where E is the PCR efficiency and CT the threshold cycle. The mean Rn and standard error (SE) were calculated for each tissue and the statistical significance of the mean Rn differences among and between tissues was assessed using the ANOVA test followed by the Tukey Honestly Significant Difference technique post hoc, and two-tailed *t*-tests, respectively.

Results

Identification and Structure of Orchid *DIV*- and *RAD*-Like Genes

The analysis of the inflorescence transcriptome of *O. italica* revealed 8 *DIV*-like transcripts containing two MYB domains. Their lengths range from 737 to 1901 bp and their predicted protein products vary from 175 to 297 amino acids (table 1). The PCR amplification on the inflorescence cDNA of *O. italica* followed by cloning, sequencing and BLAST analysis validated their expression in the floral tissues. The OITA_23026 and

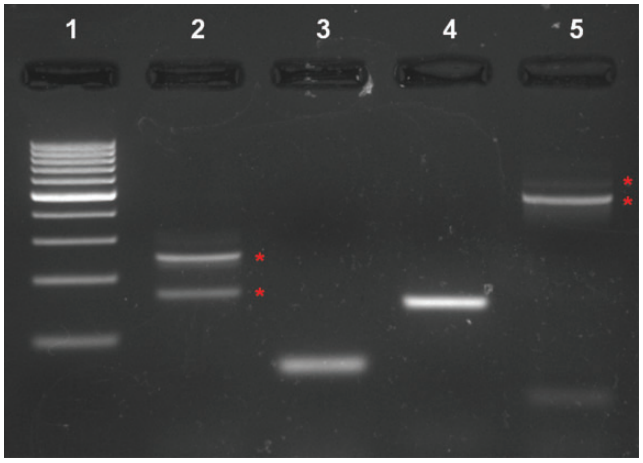

Fig. 2.—Alternative splicing of two *DIV*-like transcripts of *O. italica*. Agarose gel electrophoresis of the amplification products of fragments of the differentially spliced transcripts OITA_13252 and OITA_13252_AS (lane 2, both the isoforms); fragment of the OITA_13252 transcript (lane 3); fragment of the OITA_23026_AS transcript (lane 4); fragments of the differentially spliced transcripts OITA_23026 and OITA_23026_AS (lane 5, both the isoforms). Red asterisks indicate alternatively spliced isoforms. Amplicons were run with the GeneRuler 100 bp DNA ladder (Thermo Scientific) (lane 1). Forward primers that specifically amplify the intron-retaining isoforms were designed in the region covering the exon–intron junction (supplementary table S2, Supplementary Material online).

OITA_13252 transcripts showed two amplification products of different sizes whose sequences differ only in the presence/absence of a nucleotide fragment (111 bp in OITA_23026 and 88 bp in OITA_13252), suggesting that the different amplicons represent alternatively spliced transcripts (fig. 2). The retention of the 111 bp fragment (whose ends are 5′-GT and AG-3′) in the OITA_23026 transcript generates a premature stop codon within the CDS, resulting in a 117-residue amino acid sequence containing a single MYB domain. The two alternative transcripts of OITA_13252 differ in a region 31 nucleotides upstream of the translation starting codon in the presence/absence of an 88-nucleotide fragment (whose ends are 5′-GT and AG-3′) that does not change the predicted protein product.

The analysis of the *Ophrys* transcriptome allowed for identification of four *DIV*-like transcripts. Although they are not full length transcripts, they were included in the subsequent phylogenetic and evolutionary analyses because *Ophrys* belongs to the same sub-family of *O. italica* (Orchidoideae).

After scanning the genomes of *P. equestris* and *D. catenatum*, 7 and 8 *DIV*-like genes were identified, respectively. All contain a single intron of variable length (table 1) with canonical donor and acceptor splicing sites. PCR using genomic DNA of *O. italica* allowed for amplification of a single intron of the *DIV*-like genes in this species, in a conserved position relative to the predicted orthologs of *P. equestris* and *D. catenatum*. The intron size of the *DIV*-like genes of *O. italica* ranges from 76 to ~10,000 bp and approximately reflects the intron size of the *P. equestris* and *D. catenatum* orthologs (table 1). Due to their large size (supplementary fig. S1, Supplementary Material online), the introns of the OITA_13233 (~5,000 bp), OITA_8681 (~6,000 bp), and OITA_12910 (~10,000 bp) genes were only partially sequenced.

Nucleotide sequence analysis performed with CENSOR revealed the presence of traces of transposable/repetitive elements in the introns of OITA_13252 and OITA_13233 and their orthologs in *P. equestris* and *D. catenatum*. Table 2 reports the name and class of transposable/repetitive elements found in the introns of orchid *DIV*-like genes.

There are four *RAD*-like transcripts containing a single MYB domain in the inflorescence transcriptome of *O. italica*. Their lengths range from 543 to 420 bp and they encode for 82- to 94-amino acid proteins (table 1). The validation via PCR amplification, cloning, sequencing and BLAST analysis confirmed their expression in the inflorescence of *O. italica*.

Five *RAD*-like genes were identified in the genomes of *P. equestris* and *D. catenatum*. Three contain a single intron (table 1) with canonical donor and acceptor splicing sites and two do not have any introns. PCR amplification of the genomic DNA of *O. italica* allowed for amplification of a single intron in three of the four *RAD*-like genes (table 1). The *RAD*-like gene OITA_32153 of *O. italica* shows a premature stop codon compared with its ortholog in *P. equestris* and *D. catenatum*, shifting the intron position in the 3′ UTR.

A single *RAD*-like transcript was identified in the *Ophrys* transcriptome and was included in the phylogenetic and evolutionary analyses.

Traces of transposable/repetitive elements are present in all introns of the *RAD*-like genes of *O. italica* and their orthologs in *P. equestris* and *D. catenatum* (table 2).

Phylogenetic and Evolutionary Analyses

Figure 3*A* shows the ML tree obtained from amino acid alignment of the DIV- and RAD-like sequences identified in this study. Its topology overlaps with that of the NJ tree (data not shown). Both the DIV- and RAD-like groups have high bootstrap support values (87%). In the DIV-like group, there are three well-supported subgroups (1–3), each with *O. italica*, *P. equestris* and *D. catenatum* sequences, whereas the *Ophrys* sequences are only in subgroups 1 and 2. There are two well-supported sub-groups within the RAD-like group: The largest includes *O. italica*, *Ophrys*, *P. equestris* and *D. catenatum* sequences and the other includes only two sequences, one *P. equestris* and one *D. catenatum*.

Analysis of the conserved domains (fig. 3*B*) revealed that nine amino acid motifs are shared among groups of sequences. Motifs 1 to 3 are part of the MYB DNA-binding domain (supplementary fig. S2, Supplementary Material online) and the others have unknown function. Motifs 2 and 3 are shared by all sequences except OPH_5397; motif 1 is in all

Table 2
The Transposable Elements and Their Class in the Introns of Orchid *DIV*- and *RAD*-Like Genes

Sub-family	Species	Gene_Intron	Element_Name	Class
DIV-like	*O. italica*	OITA_13252	hAT-4_STu	DNA/hAT
			EnSpm2_SB	DNA/EnSpm/CACTA
			RTE-N1B_ATr	NonLTR/RTE
			Au	NonLTR/SINE/SINE2
		OITA_8681	RLG_scTat_3_1-I	LTR/Gypsy
	P. equestris	PEQU_21184	MuDR-6_VV	DNA/MuDR
			MtPH-E-Ia	DNA/Harbinger
			Copia-62_BD-LTR	LTR/Copia
		PEQU_03826	hAT-1N1_TC	DNA/hAT
			LTR-13_Mad	LTR
			L1-36_NN	NonLTR/L1
			Gypsy-20_RC-I	LTR/Gypsy
			Gypsy-1_RC-I	LTR/Gypsy
			Gypsy-4_TC-I	LTR/Gypsy
			Gypsy-4_TC-I	LTR/Gypsy
			Copia-4_BRa-I	LTR/Copia
			RTE-1_GM	NonLTR/RTE
			EnSpm-1B_ALy	DNA/EnSpm/CACTA
			Gypsy-133_GM-LTR	LTR/Gypsy
	D. catenatum	DCAT_JSDN01S018459	Gypsy-48_Mad-I	LTR/Gypsy
			Helitron-1_PTr	DNA/Helitron
		DCAT_JSDN01S031580	hAT-1_JC	DNA/hAT
			Gypsy-24_RC-I	LTR/Gypsy
			RTE-1_NN	NonLTR/RTE
			ENSPM1_VV	DNA/EnSpm/CACTA
		DCAT_JSDN01S002760	Helitron-1_BDi	DNA/Helitron
			Copia-45_PX-I	LTR/Copia
			Gypsy-27_VV-I	LTR/Gypsy
			RTE-1_GM	NonLTR/RTE
RAD-like	*O. italica*	OITA_32153	Copia-128_SB-I	LTR/Copia
		OITA_56510	Copia-44_ATr-I	LTR/Copia
		OITA_103296	EnSpm2_SB	DNA/EnSpm/CACTA
			EnSpm-4_HV	DNA/EnSpm/CACTA
	P. equestris	PEQU_40552	MuDR-21_VV	DNA/MuDR
			Helitron-1_GM	DNA/Helitron
		PEQU_08237	HELITRONY1D	DNA/Helitron
			WIS2_TM_LTR	LTR/Copia
			Copia-58_PX-I	LTR/Copia
			Caulimovirus-1_PTr	IntegratedVirus/Caulimovirus
	D. catenatum	DCAT_JSDN01S041675	HARB-N1_ALy	DNA/Harbinger
		DCAT_JSDN01S033707	HARB-7_FV	DNA/Harbinger

the DIV-like sequences. The distribution of these and the other motifs among the sequences strictly reflects the topology of the ML tree.

The ratio between the mean nonsynonymous and synonymous substitution rates (ω) of orchid *DIV*- and *RAD*-like coding regions indicates that purifying selection acts on these genes ($\omega \ll 1$). Different evolutionary models were compared with determine if the ω ratios differed significantly among the different branches detectable in the phylogenetic tree of the *DIV*- and *RAD*-like genes. The results obtained are shown in supplementary table S3, Supplementary Material online; supplementary table S4, Supplementary Material online, reports the statistical significance of each comparison. Within the *DIV*-like genes, the one-ratio model (that assumes an equal ω for all the branches) can be excluded in favor of the two- and three-ratio models (that consider two and three different ω values, respectively). Excluding the comparison between the three-ratio model and the two-ratio model 1 (that assumes equal ω for DIV-like subgroups 2 and 3), the three-ratio model fits the data better than the two-ratio models. The clade model 2, which assumes different selective pressure on a proportion of sites of the DIV-like subgroup 2

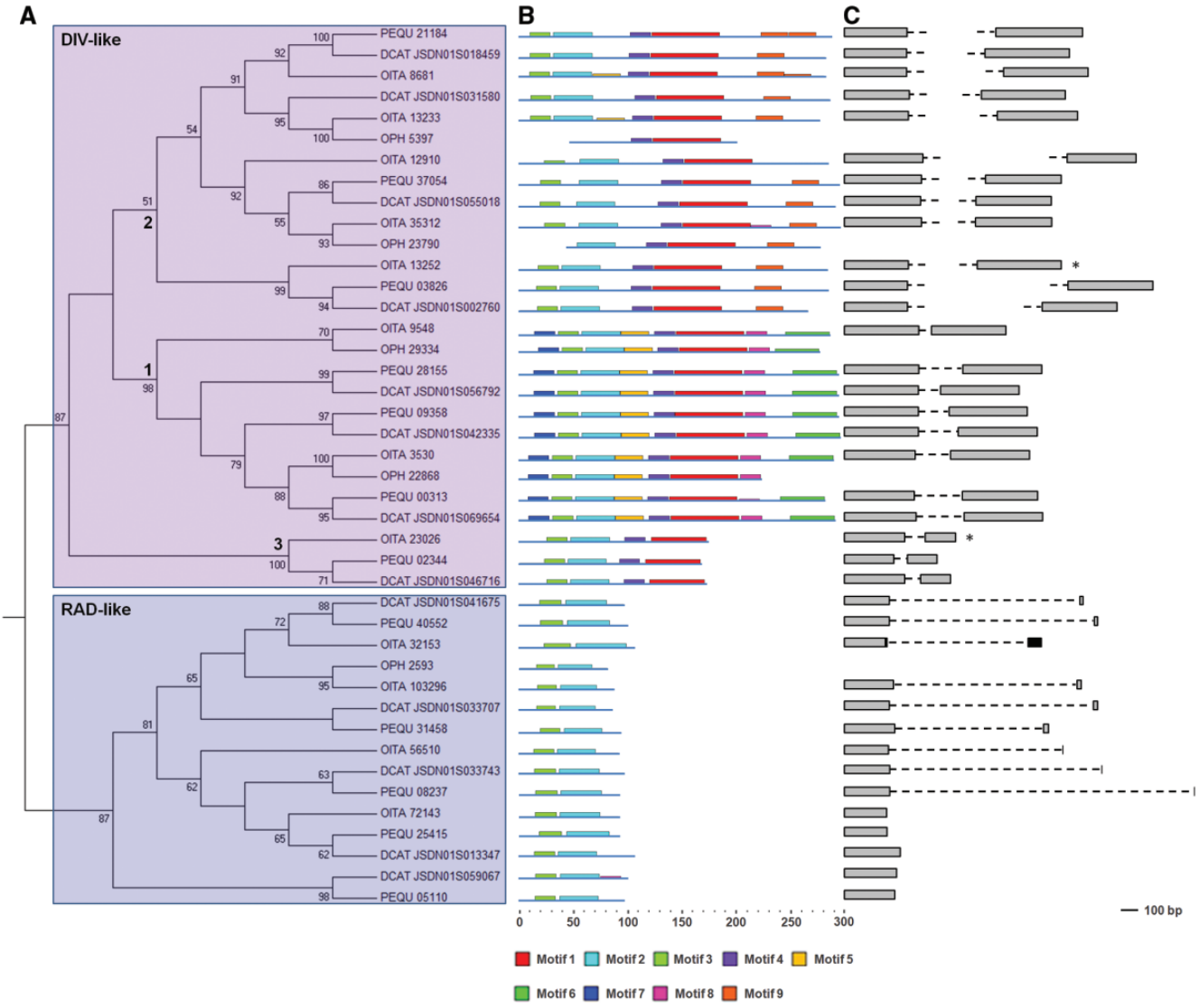

Fig. 3.—Phylogenetic analysis of orchid DIV- and RAD-like proteins. (*A*) ML tree showing three major groups in the DIV-like sub-family (1–3) and two in the RAD-like subfamily. The numbers above the nodes indicate the bootstrap percentages greater than 50%; (*B*) Diagram of the conserved motifs shared by DIV- and RAD-like orchid proteins (see supplementary fig. S2, Supplementary Material online); (*C*) Structure of orchid *DIV*- and *RAD*-like genes. Gray boxes indicate exons, dotted lines indicate introns and black boxes are UTRs. The genes marked with an asterisk are subjected to alternative splicing. The size of the boxes and lines are in scale, except for the large introns (interrupted lines).

relative to the other subgroups, is statistically more supported than its null model, showing that a significant proportion of sites (~34%) of the sequences of subgroup 2 has a different ω (0.01754) from that of the other subgroups (0.03487). The sites and branch-sites models, that allow for positive selection in specific sites and branches of the DIV-like tree, do not fit the data better than the null model of nearly neutral evolution.

Within the *RAD*-like genes, the one-ratio model can be excluded in favor of the two-ratio model. The comparison of the sites and branch-sites models does not statistically support the alternative models of positive selection. Moreover, the comparison between the clade model and its null model does not support the presence of sites with different selective pressures.

Identification of *DRIF*- and *AGL6*-Like Transcripts of *O. italica*

There are two putative *DRIF*-like transcripts in the *O. italica* inflorescence transcriptome: OITA_10599 (1250 nucleotides) and OITA_6376 (2083 nucleotides). The first encodes for a 258-amino acid protein that has 56% identity with the DRIF1 protein of *A. majus*; the second encodes for a 305 amino acid protein that has 39% identity with the DRIF2 protein of

A. majus. Higher identity scores (maximum 65% for the first and 56% for the second) after BLASTP searches against the nonredundant protein database correspond to uncharacterized proteins. Both transcripts were successfully PCR amplified from cDNA of the inflorescence of *O. italica* and their sequence corresponds to that of the transcripts in the assembled transcriptome. DRIF was only characterized in *A. majus* (Raimundo et al. 2013) and, excluding the presence of two conserved MYB domains, appears to be a poorly conserved protein. BLASTP analysis using the DRIF1 protein of *A. majus* as the query produces matches with uncharacterized proteins that have 75–78% identity in Lamiales (the same order of *Antirrhinum*) and the highest percentage of identity in monocots is approximately 60% (e.g., *Musa acuminata*). Therefore, although not well conserved, we considered OITA_10599 to be the putative ortholog of DRIF1 of *A. majus* (56% amino acid identity).

There are two *AGL6*-like transcripts in the inflorescence transcriptome of *O. italica*, like in *Phalaenopsis* (Su et al. 2013; Huang et al. 2016) and *Oncidium*, whereas in *Erycina pusilla* and *Cymbidium goeringii* there are three *AGL6*-like genes (Dirks-Mulder et al. 2017). OITA_1386 (1369 nucleotides) encodes for a 239-amino acid protein that has ~83% identity with AGL6-1 of *P. aphrodite* and OITA_4335 (1045 nucleotides) encodes for a 243-amino acid protein that has ~86% identity with AGL6-2. After a BLASTP search against the nonredundant protein database using the translated amino acid sequences of OITA_1386 and OITA_4335 as queries, the highest percentages of amino acid identity were obtained for the AGL6 proteins of the *Dendrobium hybrid cultivar* (87%) and *Cymbidium* (86%) orchids, respectively. PCR validation confirmed that these transcripts are expressed in the inflorescence tissue of *O. italica*.

Expression Patterns of *DIV*-, *RAD*-, *DRIF*-, and *AGL6*-Like Genes in *O. italica*

RNA in situ hybridization shows the expression of OITA_9548 (putative ortholog of the *A. majus DIV* gene) and OITA_10599 (putative ortholog of the *A. majus DRIF1* gene) in all the early floral tissues of *O. italica*, whereas the transcript OITA_56510 (putative ortholog of the *A. majus RAD* gene) is localized in the early lip and outer tepals (fig. 4).

After anthesis, among the *DIV*-like transcripts some have similar expression patterns and others are very different (fig. 5; supplementary fig. S3, Supplementary Material online).

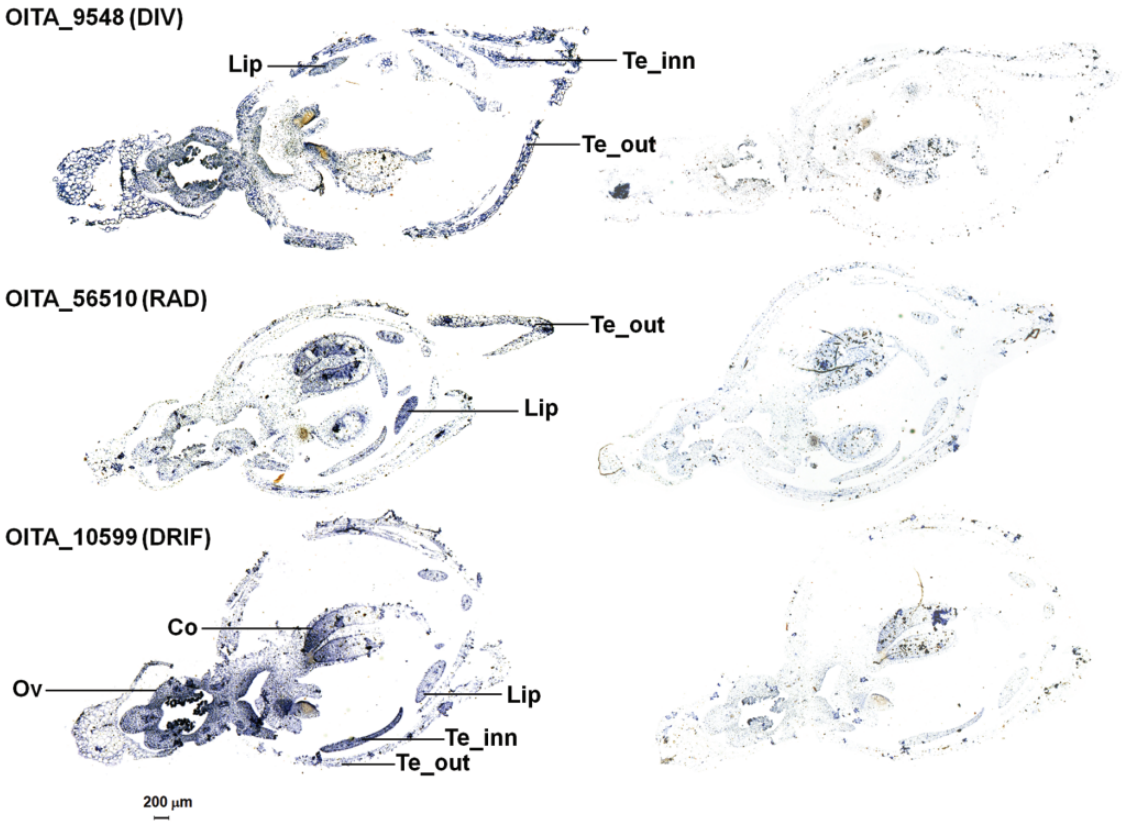

Fig. 4.—RNA *in situ* hybridization of the *DIV*, *RAD* and *DRIF* orthologs on early floral tissues of *O. italica*. Hybridizations with the antisense (left) and sense (right) probes of the transcripts OITA_9548 (*DIV*), OITA_56510 (*RAD*) and OITA_10599 (*DRIF*). Te_out, outer tepal; Te_inn, inner tepal; Co, column; Ov, ovary.

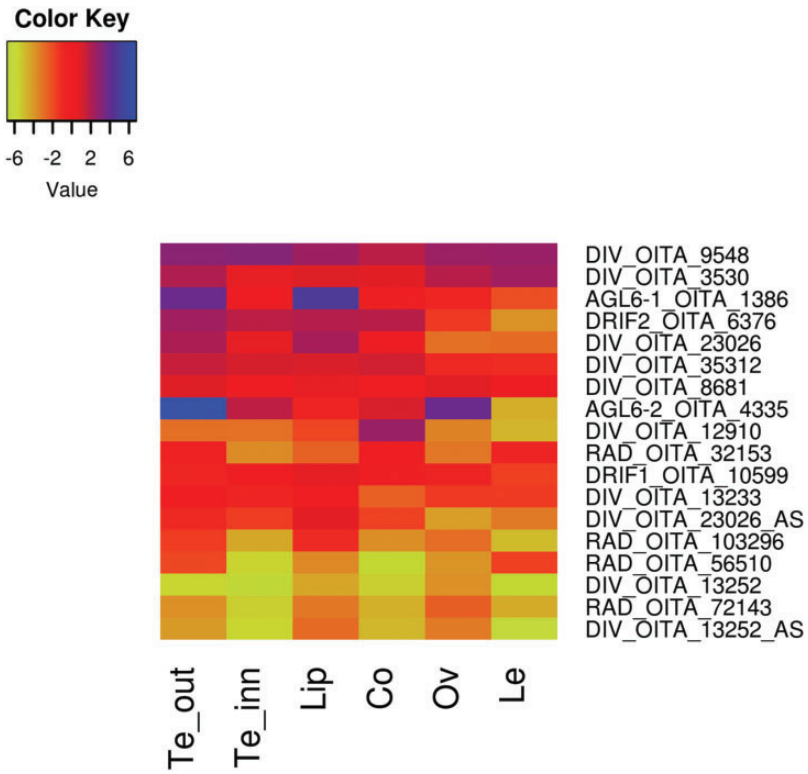

Fig. 5.—Expression profile of the *DIV*-, *RAD*-, *DRIF*-, and *AGL6*-like genes in tissues of *O. italica* after anthesis. The color-based diagram was obtained in R after conversion of the relative expression data into log2 scale. Te_out, outer tepal; Te_inn, inner tepal; Co, column; Le, leaf. The detailed relative expression data for each gene are reported in supplementary figures S3–S6, Supplementary Material online.

The OITA_9548 transcript is still expressed at high levels in all the perianth tissues (outer and inner tepals and lip). In addition, it is expressed in the column, ovary and leaf. The OITA_3530 transcript has an expression profile similar to OITA_9548 with lower levels in all tissues and with significant differences between the outer and inner tepals and lip. OITA_8681, OITA_35312, and OITA_13233 are also expressed in all tissues of the perianth; however, their expression profiles are quite different in the column, ovary and leaf, with very low OITA_35312 expression levels in the ovary and leaf and OITA_13233 levels also in the column. OITA_12910 has a very different expression profile from those of the others and is predominantly expressed in the column only. The two differentially spliced transcripts of OITA_23026 differ in their expression level pattern, which is significantly lower in all tissues (excluding leaf) for the isoform that retains the intron (OITA_23026_AS). The expression profile of the other differentially spliced transcript OITA_13252 is also different for the two isoforms. In this case, the isoform that retains the intron in the 5′ UTR (OITA_13252_AS) has significantly higher expression levels in all tissues (excluding leaf) than the isoform without the intron (fig. 6).

The *RAD*-like transcripts have similar expression profiles in the perianth tissues after anthesis, although with different expression levels. They are expressed in the outer tepals and lip with lower expression in the inner tepals. OITA_32153 is also highly expressed in the column and leaf where OITA_56510 is also expressed. All the transcripts are also expressed in the ovary (fig. 5; supplementary fig. S4, Supplementary Material online).

The OITA_10599 transcript is expressed in all floral tissues after anthesis and has a significantly higher expression in the lip whereas the OITA_6376 transcript is expressed at similar levels in all tissues except the ovary and leaf, where it is very weakly expressed (fig. 5; supplementary fig. S5, Supplementary Material online).

The *AGL6*-like transcript OITA_1386 (*AGL6-1*) is highly expressed in the lip and, at lower levels, in the outer tepals. OITA_4335 (AGL6-2) is highly expressed in the outer tepals and in the ovary at a much lower level (fig. 5; supplementary fig. S6, Supplementary Material online).

Discussion

Structure and Evolution of Orchid *DIV*- and *RAD*-Like Genes

Excluding *Ophrys*, the number of both *DIV*- and *RAD*-like genes found in orchids is very similar to that found in other plants; for example, there are up to eight *DIV*-like genes in

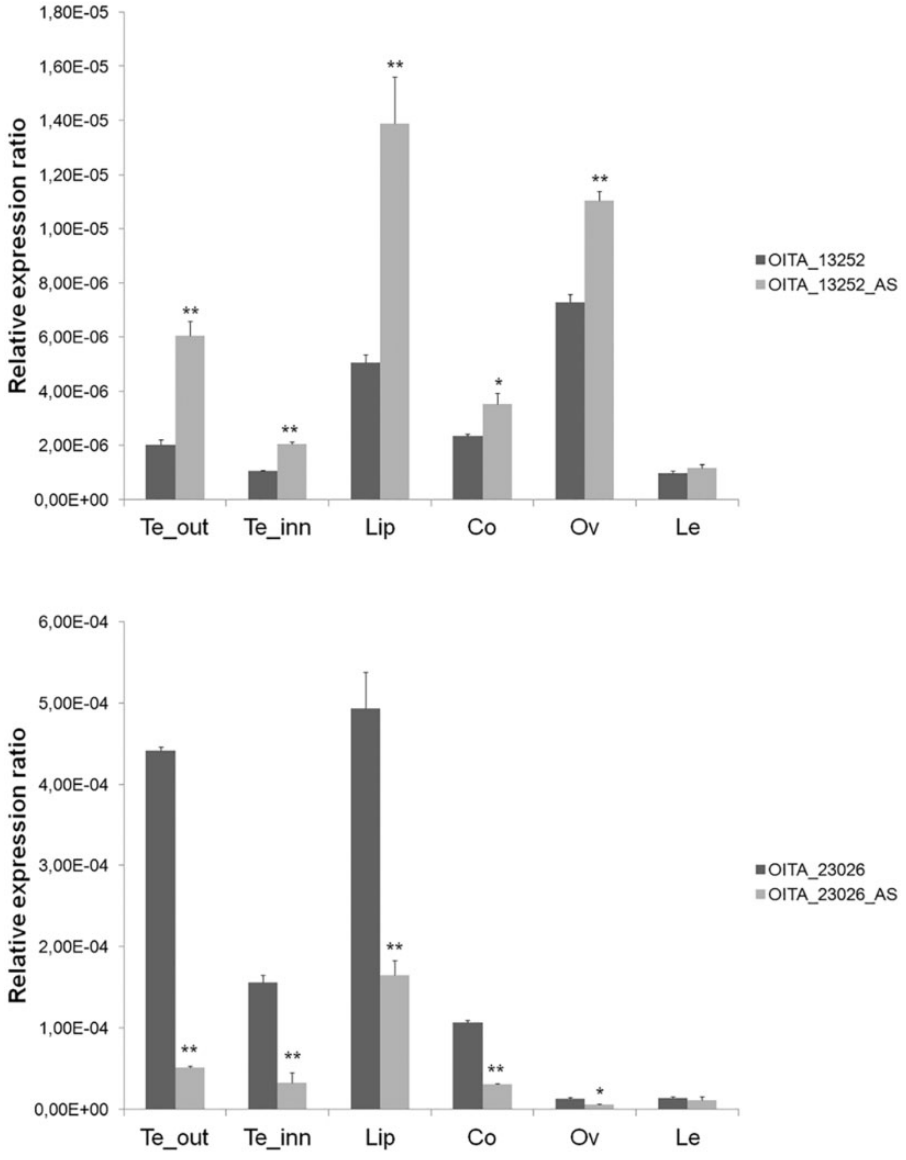

Fig. 6.—Relative expression ratio of the alternatively spliced transcripts of *DIV*-like genes OITA_13252 and OITA_23026 in tissues of *O. italica*. The asterisks indicate significant differences in the expression levels of the alternative isoforms as assessed by two-tailed *t*-test: $^{**}P < 0.01$; $^{*}P < 0.05$.

Dipsacales (Howarth and Donoghue 2009) and six *RAD*-like genes (*RAD* and five *RAD*-like) in *A. majus* and *Arabidopsis thaliana* (Baxter et al. 2007). This suggests that the number of genes identified in the present study reflects the real *DIV*- and *RAD*-like copy number of orchids. The lower number of *DIV*- and *RAD*-like transcripts found in the *Ophrys* transcriptome is probably due to the different assembly strategies used to obtain and assemble the reads for this species. The differences relative to the number of *RAD*-like genes between *O. italica* (4) and *Phalaenopsis-Dendrobium* (5) could be due to the very low expression of the homolog of the fifth *RAD*-like gene in the *O. italica* inflorescence, resulting in its absence in the assembled transcriptome. An alternative hypothesis is that there are only four *RAD*-like genes in the *O. italica* genome.

The presence of sequences of *O. italica*, *P. equestris*, and *D. catenatum* in almost all the sub-groups of the ML tree (fig. 3*A*) suggests that the duplications originating the different *DIV*- and *RAD*-like orchid genes occurred before the divergence of Orchidoideae and Epidendroideae. The split of the *DIV*-like genes into three ortholog groups (supplementary table S1, Supplementary Material online) is confirmed by the topology of the ML tree (fig. 3*A*) and is in agreement with the results obtained in Dipsacales (Howarth and Donoghue 2009), revealing that the origin of these ortholog groups predates the monocot/dicot divergence. The gene organization (intron number and size) and the distribution of the conserved motifs also reflects the division of orchid *DIV*-like genes into three groups (fig. 3*B* and *C*; supplementary fig. S2,

Supplementary Material online). Two ortholog groups are detectable in the *RAD*-like orchid genes (supplementary table S1, Supplementary Material online). This partition is in general agreement with the ML tree topology (fig. 3*A*) and with the distribution of the conserved motifs (fig. 3*B*; supplementary fig. S2, Supplementary Material online). Interestingly, some of the orchid *RAD*-like genes have a single intron whereas others are intronless (fig. 3*C*). A unique situation was found in OITA_32153 of *O. italica*, with a premature stop codon that places the intron in the 3′ UTR. Although this might appear to be an intermediate situation between intron presence and absence, its orthologs in *Phalaenopsis* and *Dendrobium* have a canonical intron in the coding sequence, suggesting that this intron shift is a derived character of *O. italica*. The intron presence/absence and the presence of an intron in the 3′ UTR in the *RAD*-like genes was also described in *A. majus* and *A. thaliana* (Baxter et al. 2007) and might reflect the evolution of this gene sub-family from an ancestral gene structure that included two exons to a more recent gene organization with a continuous open reading frame. However, it is not possible to exclude that these different gene structures originated independently in the different lineages.

The intron sizes of the orchid *DIV*- and *RAD*-like genes are conserved among the orthologs (fig. 3*C*). Orchids appear to have more large introns than other plant families, with large introns containing numerous traces of transposable elements (Cai et al. 2015; Zhang et al. 2016). Large introns of the *DIV*- and *RAD*-like orchid genes contain traces of repetitive/transposable elements (table 2), confirming the results found for other orchid genes (Salemme et al. 2013a, 2013b). In orchids, the abundance of repetitive/transposable elements within large introns might have a functional significance, affecting gene expression. The presence of repetitive/transposable elements within introns can drive antisense transcription and/or promote the formation of heterochromatin, resulting in reduction of transcriptional levels (Feschotte 2008).

The analysis of the evolutionary pressure acting on the orchid *DIV*- and *RAD*-like coding sequences shows evidence of strong purifying selection (supplementary table S3, Supplementary Material online). However, different selective constraints act on the three ortholog groups of the *DIV*-like genes, whereas the evolutionary rates of the *RAD*-like orchid genes appear more uniform.

Expression Profiles of *DIV*-, *RAD*-, *DRIF*-, and *AGL6*-Like Genes in *O. italica*

Gene expression patterns are an important tool to infer gene function, in particular in species such as *O. italica* in which it is not possible to perform functional studies based on knock-out techniques.

The transcript OITA_9548 is the putative ortholog of the *DIV* gene of *A. majus* and its expression profile in the early floral tissues of *O. italica* is similar to that of *DIV*, expressed in all the parts of the snapdragon flower (Almeida et al. 1997; Galego and Almeida 2002). The OITA_56510 transcript is the putative ortholog of the *RAD* gene of *A. majus* that is expressed in the dorsal domain of the snapdragon flower. The presence of transcript OITA_56510 in the early lip tissue of *O. italica* resembles the expression profile of the *A. majus RAD* gene. Also the expression profile of the putative ortholog of *DRIF1* of *A. majus*, OITA_10599, is similar to that described in *Antirrhinum*: It is expressed in all the early floral tissues and in the lip it could bind the RAD-like OITA_56510 and inhibit the formation of the DIV-DRIF complex, thus preventing the activation of ventralization genes. In addition, both the *RAD*-like OITA_56510 and *DRIF*-like OITA_10599 are also expressed in the outer tepals, where their interaction could have the same effect as in the lip, inhibiting ventralization. The expression profile of these three transcripts is conserved in the floral tissues after anthesis, where the presence of the transcript of the *RAD* ortholog is detectable in the ventral part of the flower of *O. italica*. This is because, after resupination, the median inner tepal (dorsal structure) twists and becomes a ventral structure. Taken together, the expression data support a model of interaction among MYB transcription factors to determine the bilateral symmetry of the flower generally conserved between *A. majus* and *O. italica*. However, some differences should be introduced to fit the uniqueness of the orchid flower. This model should be integrated with the expression pattern of the *AGL6*-like genes whose role in the development of the orchid flower might be prominent in the formation of the lip, as previously hypothesized in *Phalaenopsis* and *Oncidium* orchids (Hsu et al. 2015; Huang et al. 2016), and the outer tepals.

In general, the expression levels of the *DIV*-like genes in *O. italica* after anthesis are inversely correlated with the intron size, in agreement with the results obtained from the genome-wide analysis of *P. equestris* (Cai et al. 2015). The more efficient transcription of short mRNA molecules than that of longer paralogs is probably due to the lower energetic costs of transcribing and processing shorter transcripts (Castillo-Davis et al. 2002). In addition, the presence of traces of repetitive/transposable elements within the large introns of the *DIV*-like genes can negatively affect their expression levels (Feschotte 2008).

The *DIV*-like genes of *O. italica* are mainly expressed in the tissues of the perianth (outer and inner tepals, lip), suggesting a possible redundant role in these tissues. However, some transcripts are also expressed in the column, ovary and leaves, demonstrating the pleiotropic role of these genes, as observed in other species (Howarth and Donoghue 2009). Notably, the expression profile of the alternatively spliced *DIV*-like isoform OITA_23026_AS (that encodes a short protein with a single MYB domain) resembles that of the *RAD*-like gene OITA_103296, leading to speculation that this isoform represents an evolutionary step towards the *RAD*-like genes, possibly evolved from a *DIV*-like gene after the loss of exon 2

through alternative splicing. The expression profile of OITA_13252 is similar to that of the alternatively spliced isoform OITA_13252_AS; however, the isoform that retains the intron in the 5′ UTR (OITA_13252_AS) is expressed at higher levels. This difference in expression level suggests a transcriptional regulatory role of the small intron in the 5′ UTR, whose presence might increase the transcription efficiency.

After anthesis, the *RAD*-like genes of *O. italica* share a similar expression pattern that suggests redundant roles of these transcripts in the organs of the perianth and more specialized roles in reproductive and vegetative tissues.

Conclusions

This study is a novel analysis of MYB transcription factors that are potentially involved in the floral symmetry of orchids. The next generation sequencing approach increases the amount of data available at the transcriptomic and genomic levels to analyze gene families of interest in nonmodel species, such as orchids. The present study hypothesizes the existence of an evolutionary conserved program of gene expression between distantly related species, *O. italica* and *A. majus*, to establish bilateral symmetry. Further expression and functional analyses of *DIV*, *RAD* and *DRIF* genes, together with TCP genes, in other orchid species could verify this model.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Author Contributions

M.C.V. and S.D.P. performed the transcript and genomic DNA analyses and the expression analysis and participated in the data analysis; M.C.V. and G.I. performed the in situ hybridization experiments; S.A. designed the experiments, performed the *in silico* analyses, analyzed the data and wrote the manuscript.

Acknowledgments

The authors are grateful to Prof Luciano Gaudio for critical reading of the manuscript and to Prof Salvatore Cozzolino for plant material. This work was supported by a 2007 Regione Campania Grant L.R. N5/2002 and by funding from the University of Naples Federico II.

Literature Cited

Aceto S, Gaudio L. 2011. The MADS and the beauty: genes involved in the development of Orchid flowers. Curr Genomics 12:342–356.

Aceto S, et al. 2014. 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 9:e97839.

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

Almeida J, Galego L. 2005. Flower symmetry and shape in Antirrhinum. Int J Dev Biol. 49:527–537.

Almeida J, Rocheta M, Galego L. 1997. Genetic control of flower shape in *Antirrhinum majus*. Development 124:1387–1392.

Bailey TL, et al. 2009. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37:W202–W208.

Baxter CE, Costa MM, Coen ES. 2007. Diversification and co-option of RAD-like genes in the evolution of floral asymmetry. Plant J. 52:105–113.

Cai J, et al. 2015. The genome sequence of the orchid *Phalaenopsis equestris*. Nat Genet. 47:65–72.

Cantone C, Gaudio L, Aceto S. 2011. The PI/GLO-like locus in orchids: duplication and purifying selection at synonymous sites within Orchidinae (Orchidaceae). Gene 481:48–55.

Castillo-Davis CI, Mekhedov SL, Hartl DL, Koonin EV, Kondrashov FA. 2002. Selection for short introns in highly expressed genes. Nat Genet. 31:415–418.

Citerne H, Jabbour F, Nadot S, Damerval C. 2010. The evolution of floral symmetry. Adv Bot Res. 54:85–137.

Coen ES. 1996. Floral symmetry. EMBO J. 15:6777–6788.

Corley SB, Carpenter R, Copsey L, Coen E. 2005. Floral asymmetry involves an interplay between TCP and MYB transcription factors in Antirrhinum. Proc Natl Acad Sci USA. 102:5068–5073.

De Paolo S, Gaudio L, Aceto S. 2015. Analysis of the TCP genes expressed in the inflorescence of the orchid Orchis italica. Sci Rep. 5:16265.

De Paolo S, Salvemini M, Gaudio L, Aceto S. 2014. De novo transcriptome assembly from inflorescence of *Orchis italica*: analysis of coding and non-coding transcripts. PLoS One 9:e102155.

Devos N, et al. 2016. Analyses of transcriptome sequences reveal multiple ancient large-scale duplication events in the ancestor of Sphagnopsida (Bryophyta). New Phytol. 211:300–318.

Dirks-Mulder A, et al. 2017. Exploring the evolutionary origin of floral organs of *Erycina pusilla*, an emerging orchid model system. BMC Evol Biol. 17:89.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Endress PK. 2012. The immense diversity of floral monosymmetry and asymmetry across angiosperms. Bot Rev. 78:345–397.

Fenster CB, Armbruster WS, Dudash MR. 2009. Specialization of flowers: is floral orientation an overlooked first step?. New Phytol. 183:502–506.

Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. Nat Rev Genet. 9:397–405.

Galego L, Almeida J. 2002. Role of DIVARICATA in the control of dorsoventral asymmetry in Antirrhinum flowers. Genes Dev. 16:880–891.

Howarth DG, Donoghue MJ. 2009. Duplications and expression of DIVARICATA-like genes in dipsacales. Mol Biol Evol. 26:1245–1258.

Hsu HF, et al. 2015. Model for perianth formation in orchids. Nat Plants 1.

Huang JZ, et al. 2016. The genome and transcriptome of Phalaenopsis yield insights into floral organ development and flowering regulation. PeerJ 4:e2017.

Javelle M, Marco CF, Timmermans M. 2011. In situ hybridization for the precise localization of transcripts in plants. J Vis Exp. 57:e3328.

Kohany O, Gentles AJ, Hankus L, Jurka J. 2006. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics 7:474.

Li L, Stoeckert CJ Jr, Roos DS. 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 13:2178–2189.

Lin YF, et al. 2016. Genome-wide identification and characterization of TCP genes involved in ovule development of *Phalaenopsis equestris*. J Exp Bot. 67:5051–5066.

Luo D, Carpenter R, Vincent C, Copsey L, Coen E. 1996. Origin of floral asymmetry in Antirrhinum. Nature 383:794–799.
Luo D, et al. 1999. Control of organ asymmetry in flowers of Antirrhinum. Cell 99:367–376.
Mondragon-Palomino M, Hiese L, Harter A, Koch MA, Theissen G. 2009. Positive selection and ancient duplications in the evolution of class B floral homeotic genes of orchids and grasses. BMC Evol Biol. 9:81.
Mondragon-Palomino M, Theissen G. 2011. Conserved differential expression of paralogous DEFICIENS- and GLOBOSA-like MADS-box genes in the flowers of Orchidaceae: refining the 'orchid code'. Plant J. 66:1008–1019.
Mondragon-Palomino M, Trontin C. 2011. High time for a roll call: gene duplication and phylogenetic relationships of TCP-like genes in monocots. Ann Bot. 107:1533–1544.
Montieri S, Gaudio L, Aceto S. 2004. Isolation of the LFY/FLO homologue in *Orchis italica* and evolutionary analysis in some European orchids. Gene 333:101–109.
Pan ZJ, et al. 2011. The duplicated B-class MADS-box genes display dualistic characters in orchid floral organ identity and growth. Plant Cell Physiol. 52:1515–1531.
Pan ZJ, et al. 2014. Flower development of Phalaenopsis orchid involves functionally divergent SEPALLATA-like genes. New Phytol. 202:1024–1042.
Ponting CP, Russell RR. 2002. The natural history of protein domains. Annu Rev Biophys Biomol Struct. 31:45–71.
Raimundo J, et al. 2013. A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in Antirrhinum flower asymmetry. Plant J. 75:527–538.
Rudall PJ, Bateman RM. 2002. Roles of synorganisation, zygomorphy and heterotopy in floral evolution: the gynostemium and labellum of orchids and other lilioid monocots. Biol Rev Camb Philos Soc. 77:403–441.
Salemme M, Sica M, Gaudio L, Aceto S. 2011. Expression pattern of two paralogs of the PI/GLO-like locus during *Orchis italica* (Orchidaceae, Orchidinae) flower development. Dev Genes Evol. 221:241–246.
Salemme M, Sica M, Gaudio L, Aceto S. 2013a. The OitaAG and OitaSTK genes of the orchid Orchis italica: a comparative analysis with other C- and D-class MADS-box genes. Mol Biol Rep. 40:3523–3535.
Salemme M, Sica M, Iazzetti G, Gaudio L, Aceto S. 2013b. The AP2-like gene OitaAP2 is alternatively spliced and differentially expressed in inflorescence and vegetative tissues of the orchid Orchis italica. PLoS One 8:e77454.
Sedeek KE, et al. 2013. Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. PLoS One 8:e64621.
Su CL, et al. 2013. A modified ABCDE model of flowering in orchids based on gene expression profiling studies of the moth orchid Phalaenopsis aphrodite. PLoS One 8:e80462.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34:W609–W612.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30:2725–2729.
Ushimaru A, Dohzono I, Takami Y, Hyodo F. 2009. Flower orientation enhances pollen transfer in bilaterally symmetrical flowers. Oecologia 160:667–674
Vamosi JC, Vamosi SM. 2010. Key innovations within a geographical context in flowering plants: towards resolving Darwin's abominable mystery. Ecol Lett. 13:1270–1279.
Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 13:555–556.
Yang Z, Bielawski JP. 2000. Statistical methods for detecting molecular adaptation. Trends Ecol Evol. 15:496–503.
Zhang GQ, et al. 2016. The Dendrobium catenatum Lindl. genome sequence provides insights into polysaccharide synthase, floral development and adaptive evolution. Sci Rep. 6:19029.

Associate editor: Yves Van De Peer