Journal of Plant Physiology xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Journal of Plant Physiology

journal homepage: www.elsevier.de/jplph

The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors

Athanasios Tsaftaris[a,b,*], Konstantinos Pasentsis[a], Antonios Makris[a], Nikos Darzentas[a], Alexios Polidoros[a,1], Apostolos Kalivas[a,2], Anagnostis Argiriou[a]

[a] *Institute of Agrobiotechnology, Center for Research and Technology Hellas, 6th Km Charilaou Thermi Road, Thermi GR-570 01, Greece*
[b] *Department of Genetics and Plant Breeding, Aristotle University of Thessaloniki, Thessaloniki GR-541 24, Greece*

ARTICLE INFO

Article history:
Received 11 August 2010
Received in revised form 22 March 2011
Accepted 26 March 2011

Keywords:
Crocus sativus L.
MADS-box genes
Monocots
RCA-RACE
SEPALLATA

ABSTRACT

To further understand flowering and flower organ formation in the monocot crop saffron crocus (*Crocus sativus* L.), we cloned four MIKC^c type II MADS-box cDNA sequences of the E-class *SEPALLATA3* (*SEP3*) subfamily designated *CsatSEP3a/b/c/c_as* as well as the three respective genomic sequences. Sequence analysis showed that cDNA sequences of *CsatSEP3* c and c_as are the products of alternative splicing of the *CsatSEP3c* gene. Bioinformatics analysis with putative orthologous sequences from various plant species suggested that all four cDNA sequences encode for *SEP3*-like proteins with characteristic motifs and amino acids, and highlighted intriguing sequence features. Phylogenetically, the isolated sequences were closest to the *SEP3*-like genes from monocots such as *Asparagus virgatus*, *Oryza sativa*, *Zea mays*, and the dicot *Arabidopsis* SEP3 gene. All four isolated *C. sativus* sequences were strongly expressed in flowers and in all flower organs: whorl1 tepals, whorl2 tepals, stamens and carpels, but not in leaves. Expression of *CsatSEP3a/b/c/c_as* cDNAs was compared in wild-type and mutant flowers. Expression of the isolated*CsatSEP3*-like genes in whorl1 tepals together with E-class *CsatAP1/FUL* subfamily and B-class *CsatAP3* and *CsatPI* subfamilies of genes, fits the ABCE "quartet model," an extended form of the original ABC model proposed to explain the homeotic transformation of whorl1 sepals into whorl1 tepals in *Liliales* and *Asparagales* plants such as *C. sativus*. This conclusion was also supported by the interaction of the CsatSEP3b protein with CsatAP1/FUL and CsatAP3 proteins. In contrast, expression of both B-class *CsatAP3* and *CsatPI* genes and the C-class *CsatAGAMOUS* genes together with E-class *CsatSEP3*-like genes in carpels, without any phenotypic effects on carpels, raises questions about the role of these gene classes in carpel formation in this non-grass monocot and requires further experimentation. Finally, taking advantage of the size and sequence differences in amplified genomic sequences of the triploid *C. sativus* and comparing them with the respective sequences from *C. tomasii*, *C. hadriaticus* and *C. cartwrightianus*, three putative wild-type diploid progenitor species, we examined the origin of CsatSEP3a sequence.

Introduction

Most angiosperm flowers contain four distinct organ types, from the outside to the inside, namely: sepals in whorl1, petals in whorl2, stamens in whorl3, and carpels in whorl4. Depending upon the taxa, this model can be adapted, placing species-specific homologous organs in each whorl. For example, plants such as orchids, roses, asparagus, maize, crocus (among many others) have varied and distinctive flowers. A common departure from the typical flower, observed in several monocots including saffron crocus (*Crocus sativus* L.), is that there is no clear distinction between sepals and petals, and the petaloid organs in whorls 1 and 2 are therefore referred to as tepals (Bowman, 1997).

C. sativus is a monocot triploid sterile species belonging to the *Iridaceae* family of *Asparagales*, whose whorl4 red stigmatic styles constitute saffron, a commercially very important popular food additive with delicate aroma and attractive color, also used for coloring and medical purposes. The flower of saffron is bisexual. The perianth consists of six petaloid tepals in whorls 1 and 2. The

* Corresponding author at: Institute of Agrobiotechnology, Center for Research and Technology Hellas, 6th Km Charilaou Thermi Road, Thermi GR-570 01, Greece. Tel.: +30 2310 498271; fax: +30 2310 498270.
E-mail address: tsaft@certh.gr (A. Tsaftaris).
[1] Present address: Department of Genetics and Plant Breeding, Aristotle University of Thessaloniki, Thessaloniki GR-541 24, Greece.
[2] Present address: Cotton and Industrial Plants Institute, National Agricultural Research Foundation, Sindos GR-574 00, Greece.

0176-1617/$ – see front matter © 2011 Elsevier GmbH. All rights reserved.
doi:10.1016/j.jplph.2011.03.015

androecium consists of three distinct stamens in whorl3 and the gynoecium consists of a single compound pistil with three stigmas in whorl4. Several phenotypic flower mutants have been described, such as flowers with varying numbers of styles and stamens, as well as plants with double or even triple flowers, indicative perhaps of indeterminate inflorescences in this mutant phenotype (Grilli Caiola et al., 2004).

C. sativus blooms only once a year during mid-November and is harvested by hand. After wind separation of tepals, the red stigmas are separated from the yellow stamens by hand. Consequently, the cultivation of this crop for its flowers, and specifically for its stigmas, is very labor-intensive, leading to high costs (Tsaftaris et al., 2004, 2007). Thus, understanding flower development in crocus could help us not only to decode tepal formation in this monocot, but also reveal ways to increase yield and lower production costs.

The characteristics of the different floral organs during flower development are determined by actions of floral organ identity genes. The proposed eudicot Arabidopsis- and snapdragon-based model envisioned the action of flower-specific genes, belonging to three distinct functions termed A, B and C (hence the name of the original "ABC" model), where A-function genes alone could determine sepals, A+B function genes determine petals, B+C function genes determine stamens, and C-function genes alone determine carpels (Coen and Meyerowitz, 1991). Parallel studies found that, with the exception of the APETALA2 of A-function, all genes responsible for the A, B, and C functions involved in flower development belong to a large family of MADS-box transcription factors, which specify flowering time, floral meristem identity and space-time regulation of flower organ formation (Jack, 2001, 2004; Theissen, 2001; Immink and Angenent, 2002).

Expression of ABC genes throughout a plant does not transform leaves into floral organs and thus, the ABC functions, though necessary, are not sufficient to superimpose floral organ identity on a leaf development program (Theissen and Saedler, 2001).

An additional class of MADS-box genes identified later, the E-function *SEPALLATA* (*SEP*) genes of *Arabidopsis*, provide redundant control over the ABC system. *SEP* genes constitute one of the three phylogenetically determined clades of the E-superclade genes (the other two are the *AP1*/*SQUA*/*FUL*-like and the *AGL6*-like clades) (Becker and Theissen, 2003). Of these three subfamilies of E-class genes, only the *AGL6*-like subfamily of genes has been found in gymnosperms. *SEP* MADS-box genes are present in all angiosperms, but not in gymnosperms, and are required for the regulation of floral meristem determinacy and the specification of flower organs (Rijpkema et al., 2009).

The three different *SEP* genes in *Arabidopsis* (*SEP1*, *SEP2* and *SEP3*) that were identified through their sequence similarity to *AGAMOUS* (*AG*) (Ma et al., 1991; Mandel and Yanofsky, 1998), when expressed together with the ABC genes, are sufficient for specification of petals, stamens and carpels. The ectopic expression of *SEP3*, together with BC genes, may turn leaves into floral organs (Honma and Goto, 2001). *SEP1*/*2*/*3* are still expressed in B- and C-function loss-of-function mutants, and the initial expression patterns of B- and C-function genes are not altered in the *sep1*/*2*/*3* triple mutant, indicating that *SEP1*/*2*/*3* do not act downstream of such genes and are not required for their initial activation (Pelaz et al., 2000). Moreover, *SEP4*, a fourth *SEP*-like gene discovered in *Arabidopsis*, is probably still expressed in these triple mutants and seems sufficient to confer the sepal-like structure; in quadruple *sepl*/*2*/*3*/*4* mutants, leaf-like structures replace flowers (Ditta et al., 2004).

Thus, *SEP* genes comprise a separate E-class subfamily of floral homeotic genes that provide an additional layer of function by encoding proteins forming higher-order complexes together with A-, B-, or C-class transcription factors, several of which are sufficient to transform leaves into floral organs (Honma and Goto, 2001). These findings led to the suggestion of a protein-based combinatorial "floral quartet model" that could explain how the different floral organ identity genes interact at the molecular level (Theissen, 2001; Theissen and Saedler, 2001; Melzer et al., 2009). More specifically, the ABCE model postulates that sepals are specified by A-function protein (AP1/SQUA) activity together with E-function, petals by A+B+E, stamens by B+C+E, and carpels by C+E (Goto et al., 2001; Theissen, 2001; Theissen and Saedler, 2001). Thus, E-function activity, and consequently, *SEP* gene expression, is required for the formation of whorl1, 2, 3 and 4 organs. It should be noted that, for the ABC model, the A-function genes AP1 or SQUA in *Arabidopsis* and snapdragon, respectively, together with FRU, constitute a clade of the E-superclade or class of genes; the other two clades of the E-family are the *SEP*-like and the AGL6-like subfamilies (Shan et al., 2009; Melzer et al., 2010).

Recently, Immink et al. (2009), using yeast three-hybrid and fluorescent techniques to study MADS-box protein interactions, showed strong indications that higher-order complex formation is a general and essential molecular mechanism for plant MADS-box protein functionality. The authors describe SEP proteins and SEP3 in particular as the "glue" protein for higher-order complexes and flower organ formation. SEP3, in addition to flower organ formation and flower organ identity, was also recently found to regulate flowering (Dornelas et al., 2010) and flower organ growth (Kaufmann et al., 2009). Genes involved in organ size are targets of SEP3 complexes. Two recent, parallel studies from the group of Theissen using *in vitro* assays demonstrated that homotetramers of *Arabidopsis* SEP3 are sufficient to loop DNA by binding to two neighboring CArG elements and that SEP3 forms stronger floral-quartet-like structure with the B-class proteins AP3 and PI than the structure formed by AP3 and PI themselves (Melzer and Theissen, 2009; Melzer et al., 2009).

SEP-like genes in grasses showing relatively heterogeneous expression patterns strongly suggest that they are also heterogeneous in function (Becker and Theissen, 2003). In maize, for example, numerous *SEP3*-like genes with distinguishable expression patterns that have been suggested to be involved in determining the alternative identity of spikelet primordial, the upper versus the lower floret within each spikelet primordium, or conferring determinacy to the spikelet or upper floret meristem have been described (Cacharron et al., 1999; Theissen et al., 2000; Theissen, 2001). Similar observations have been reported for rice plants in which *OSMADS1*, the putative ortholog of the maize *ZMM8*, plays an important role in floral meristem determination during the early development of rice florets (Agrawal et al., 2005; Cui et al., 2010; Gao et al., 2010; Kobayashi et al., 2010). Two *SEP1*-like (or *AGL2*-like) MADS-box genes, named *LMADS3* and *LMADS4* from lily (*Lilium longiflorum*), with extensive homology of *LMADS3* to the *Arabidopsis* SEPALLATA3, were expressed in the inflorescence meristem and in floral buds of different developmental stages. *LMADS4* mRNA is also expressed in vegetative leaf and in the inflorescence stem, where *LMADS3* expression is absent (Tzeng et al., 2003).

Previous results from our group with saffron crocus have indicated that expression of B-class paleo *AP3*-like genes (*CsatAP3*-like) is extended in whorl1 and may be important for the homeotic transformation of whorl1 sepals into tepals in this species (Tsaftaris et al., 2006) similarly to other monocot species (Kanno et al., 2003). We have also demonstrated the presence of *PI*-like genes in *C. sativus* (*CsatPI*/*GLO*-like), also in all four whorls of the flower, including whorl1 tepals (Kalivas et al., 2007). Among MADS-box genes, we also studied the *CsatAP1*/*FUL*-like genes, one of the three subfamilies of the E-superclade of MADS box genes of *C. sativus*. Three such *CsatAP1*/*FUL*-like genes were cloned and characterized. All three genes are expressed not only in all flower organs, as found for other *AP1*/*FUL*/*SQUA* members in other plant species, but they were

Please cite this article in press as: Tsaftaris A, et al. The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors. J Plant Physiol (2011), doi:10.1016/j.jplph.2011.03.015

expressed in crocus leaves, too (Tsaftaris et al., 2004). *CsatAP1/FUL* genes were the only MADS-box genes to be expressed in the leaves from all the MADS-box genes examined till now in this species.

Here, extending our studies with crocus MADS-box genes, we report the isolation and characterization of *SEP3*-like genes, the second subfamily of the E-superclade MADS-box genes designated *CsatSEP3*-like genes. Four SEP3-like cDNAs corresponding to three isolated genomic sequences were isolated, characterized and their expression was studied. We also detected interaction between CsatSEP3-like protein described here and two other MADS-box proteins, CsatAP1/FRU (Tsaftaris et al., 2004) and CsatAP3 (Tsaftaris et al., 2006) that could explain the formation of tetramers between FUL-SEP3-AP3/PI in whorl1 and whorl2 tepals. Finally, the expression of all four *CsatSEP3* sequences was compared in wild-type and a double flower, as well as in a stamenless flower mutant. Taking advantage of size and sequence differences in isolated genomic sequences of *CsatSEP3a*, we examined the presence of each genomic region in the three putative diploid progenitors of *C. sativus*, namely: *C. cartwrightianus*, *C. tomasii* and *C. hadriaticus*, to obtain more evidence for the origin of the three genomes in the cultivated triploid crocus, *C. sativus* (Grilli Caiola et al., 2004).

Materials and methods

Plant material

Crocus sativus field-growing plants were collected from Kozani, Greece. Tissues from wild-type saffron and two mutant plants found in the field, one lacking stamens and one with double the number of flower organs (tepals, stamens, carpels) were collected. Sampling took place during the late flowering season in October. Tissues were separated and immediately frozen in liquid nitrogen and stored at −80 °C until use. Tissues from *C. cartwrightianus* Herb., and *C. hadriaticus* Herb., subsp. *hadriaticus* were collected from the CROCUSBANK project database (http://www.crocusbank.org/). Genomic DNA from *C. tomasii* Ten. was kindly provided by Dr. Ole Seberg.

RNA isolation and cDNA synthesis

Total RNA from flowers and leaves were extracted using Trizol reagent (Invitrogen) following the manufacturer's instructions. First-strand cDNA synthesis was performed with 200 U SuperScript II-RT (Invitrogen) in a 20 μl reaction volume containing a pool of 5 μg total RNA from flowers and leaves (2.5 μg each), 0.5 μg NdT-Adaptor primer (5′-TGACTAATGAATGTGGTAATGA$(T)_{18}$-3′) 5′ phosphate (synthesized at VBC-genomics, Austria), 0.5 mM dNTPs, 1× first strand buffer (Invitrogen) and 10 mM DTT (Invitrogen). The reactions were incubated for 50 min at 42 °C followed by 15 min at 70 °C to inactivate the RT. After the addition of 2 units of RNAseH (Invitrogen) the reactions were incubated for 20 min at 37 °C. The reactions were purified using the QIAquick PCR purification kit (Qiagen).

Rolling-cycle amplification-RACE (RCA-RACE) of cDNAs

Circularization followed with rolling circle amplification-RACE (RCA-RACE) of *C. sativus* cDNA was performed as has been previously described (Polidoros et al., 2006). Briefly, after circularization, the reactions were purified using the QIAquick PCR purification kit (Qiagen). Rolling circle amplification reactions were performed in 50 μl volume containing 1/6 (5 μl) of the circularized cDNA, 1 mM dNTPs, 200 μg/ml BSA (NEB), 1× phi29 DNA Polymerase reaction buffer (NEB), 10 u phi29 DNA Polymerase (NEB) and 1 μM NSFInF primer (5′-TTCATTAGTCAG-3′), with two PTO linkages on the 3′-end (custom synthesized at VBC-genomics). The reactions were incubated for 1 h at 30 °C followed by the addition of 1 μM NSFInvR (5′-TGTGGTAATGAT-3′) and further incubated for 1 h at 30 °C. Finally heat inactivation was performed for 10 min at 60 °C.

For the direct cloning approach with RCA-RACE using degenerate inward primers, where the forward primer extends versus the 3′ end and the reverse primer extends versus the 5′ end, a 1 μl aliquot of a 1/10 dilution of the rolling circle amplification reaction was used for the PCR reactions. The reactions were performed in 50 μl volume containing 0.2 mM dNTPs, 1 U DyNAzyme II DNA polymerase (Finnzymes), 0.4 μM MADS-2F degenerate primer (5′-GTKCTYTGYGAYGCYGAGGT-3′) designed from the conserved amino acid sequence VLCDAEV in the sense orientation, and either 0.4 μM MADS5R degenerate primer (5′-AAYGTNACYTGNCGRTTDAT-3′) designed from the conserved amino acid sequence INRQVTF in antisense orientation or 0.4 μM MADS6R degenerate primer (5′-CGRTTDATYTTRTTYTCDAT-3′) designed from the conserved amino acid sequence IENKINR in antisense orientation. The cycling conditions were: 3 min at 94 °C, 40 cycles of 45 s at 94 °C, 45 s at 50 °C, 1.5 min at 72 °C and a final extension step of 10 min at 72 °C.

For the second approach, degenerate outward primers were used. In this case, the Forward primer extends versus the 5′ end and Reverse primer extends versus the 3′ end. In order to isolate the family of SEPALLATA like genes, one (1) μl of the RCA-RACE library from crocus, prepared with the InVUP primer as described in a previous paper (Polidoros et al., 2006), was used as template for inverse PCR, under the same conditions as above. Specifically, we used primers MADS-2F and MADSPromR degenerate primer (5′-CTTVSAGWAGGTSACYTG-3′) corresponding to the conserved amino acid sequence QVTFSK in antisense orientation. Several PCR products ranging from 900 bp to 1300 bp were excised from an agarose gel, ligated into pCR 2.1-TOPO vector using the PCR II-TOPO TA cloning kit (Invitrogen) and transformed in DH5a competent cells. Several individual clones were screened for the presence of an insert and sequenced. Finally based on the sequence information obtained from the isolated sequences, gene-specific primers were designed and used for PCR reactions to complete the small missing parts (39–57 nucleotides between the two outward-oriented primers) and obtain the full length coding sequence of the genes. Having as template 1/30 of the single stranded cDNA used in the RCA-RACE experiment, the reactions also contained 0.2 mM dNTPs, 1 U DyNAzyme II DNA polymerase (Finnzymes) and 0.4 μM of the corresponding primers. The cycling conditions were: 2 min at 94 °C, 35 cycles of 45 s at 94 °C, 30 s at 54 °C, 1.5 min at 72 °C and a final extension step of 10 min at 72 °C. The primers used were: Sep3-F (5′-AATAAATTGGGCTCTCAGAA-3′) and Sep3-2R (5′-ATCGAAGGGCTGATAATTAACC-3′) for amplifying *CsatSEP3a*, Contig3F (5′-TTCTAGAGAGAGAAATTAGGTAGT-3′) and Contig3R (5′-ATGTTCTTTTGATCGATTTGGGAC-3′), for amplifying *CsatSEP3b* and Contig5/6F (5′-GAAGGAGAGAAATCGTTGGTAATT-3′) and Contig5/6R (5′-GCAAGTAACCAAGAGCAAATCACT-3′) for amplifying *CsatSEP3c* and *CsatSEP3c_as*. A single PCR product from each primer pair was cloned as above. Several individual clones were screened for the presence of an insert and sequenced (Macrogen, Korea).

Isolation of genomic sequences

Genomic DNA was isolated from leaves using a modified CTAB-based protocol (Li et al., 2001). For amplification of the *C. sativus SEP3* genes, specific primers were designed based on the information gained from the isolated cDNAs. PCR amplification was performed using 0.1 μg genomic DNA as template, 200 nM primers each, 1 mM dNTPs and 1 unit DyNAzymeII polymerase (Finnzymes, Espoo, Finland) in 50 μl total volume reaction. PCR conditions were 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 58 °C for 30 s, 72 °C

Please cite this article in press as: Tsaftaris A, et al. The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors. J Plant Physiol (2011), doi:10.1016/j.jplph.2011.03.015

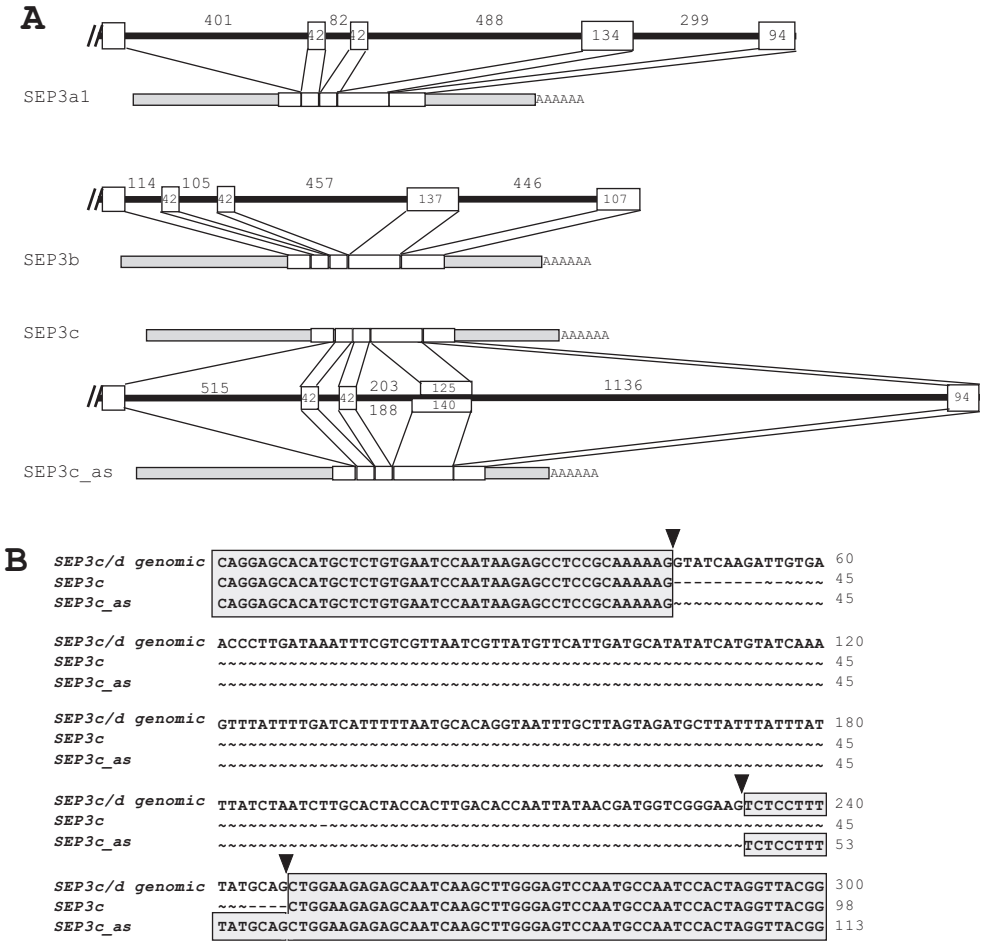

Fig. 1. (A) Genomic organization of the SEP3 isolated genes. Exons are presented with white boxes. Numbers indicate number of base pairs. (B) Alignment of *SEP3c*, *SEP3d* cDNA sequences and the corresponding codifying genomic sequence (*SEP3c*/*d* genomic). Grey boxes indicate exonic sequences and triangles indicate the splicing sites.

for 2 min and a final extension period at 72 °C for 10 min. The PCR products were separated on a 1% agarose gel and cloned into the pCR 2.1-TOPO vector using the TOPO TA cloning Kit (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. Five individual clones for each *CsatSEP3a*, *CsatSEP3b*, *CsatSEP3c* and *CsatSEP3c_as* were sequenced. The same set of primers and PCR conditions, used to isolate the three genomic sequences corresponding to the last four exons and introns from *C. sativus*, were used to amplify genomic sequences of the *SEP3* genes from genomic DNA isolated from *C. cartwrightianus*, *C. tomasii*, and *C. hadriaticus* subsp. *hadriaticus*. Genomic sequences for *SEP3A* like genes were cloned as described and three individual clones were sequenced from each crocus species. Sequencing was performed by MACROGEN (Korea) on a ABI3730 DNA analyzer.

Bioinformatics analysis

We collected all *Viridiplantae* proteins from UniProt version 15.4 of 2009 with a statistically significant hit for the MADS-box domain (PF00319: SRF-type transcription factor (DNA-binding and dimerisation domain)) from the Pfam database (Finn et al., 2008). After sequence fragments were filtered out, the resulting set was the subject of an all-against-all similarity detection step using the BLAST algorithm (Altschul et al., 1997). We used the bidirectional best hit approach (Overbeek et al., 1999) to discover putative orthologous proteins. We also asked for any orthologous sequence to be one of at least ten sequences with the MADS-box domain from the corresponding monocot or dicot species, so that we reduced the chance of best hits due to lack of sequence context. We thus concatenated our *C. sativus* sequences and their putative orthologs, and the four SEP proteins of *Arabidopsis* and their putative orthologs.

The multiple alignment of the sequence set described above was created with version v6.713b (2009/09/24) of the MAFFT algorithm [http://www.ncbi.nlm.nih.gov/pubmed/15661851] and the G-INS-I global alignment strategy with a maximum of 1000 iterations, edited and visualized with JalView version 2.6 (Waterhouse et al., 2009). The phylogenetic tree was constructed with the Geneious (trial/free version 5.3.3 – www.geneious.com) Tree Builder with the Neighbor Joining method, the default Jukes-Cantor genetic distance model, and bootstrap resampling with 1000 replicates and 50% support threshold. The resulting tree was edited and visualized with FigTree version 1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/).

The multi-RELIEF algorithm (Ye et al., 2008) was used on the resulting alignment, with 100 iterations and sample size of 10, and useful alignment positions with scores equal or greater than 0.8 (arbitrary threshold based on manual observations) were selected.

Expression analysis

The expression analysis of the isolated genes was performed with RT-PCR. Total RNA from leaves, flowers, whorl1 tepals, whorl2 tepals, stamens, and carpels of a mutant flower lacking stamens and a double flower, were extracted using the RNeasy plant mini kit

Please cite this article in press as: Tsaftaris A, et al. The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors. J Plant Physiol (2011), doi:10.1016/j.jplph.2011.03.015

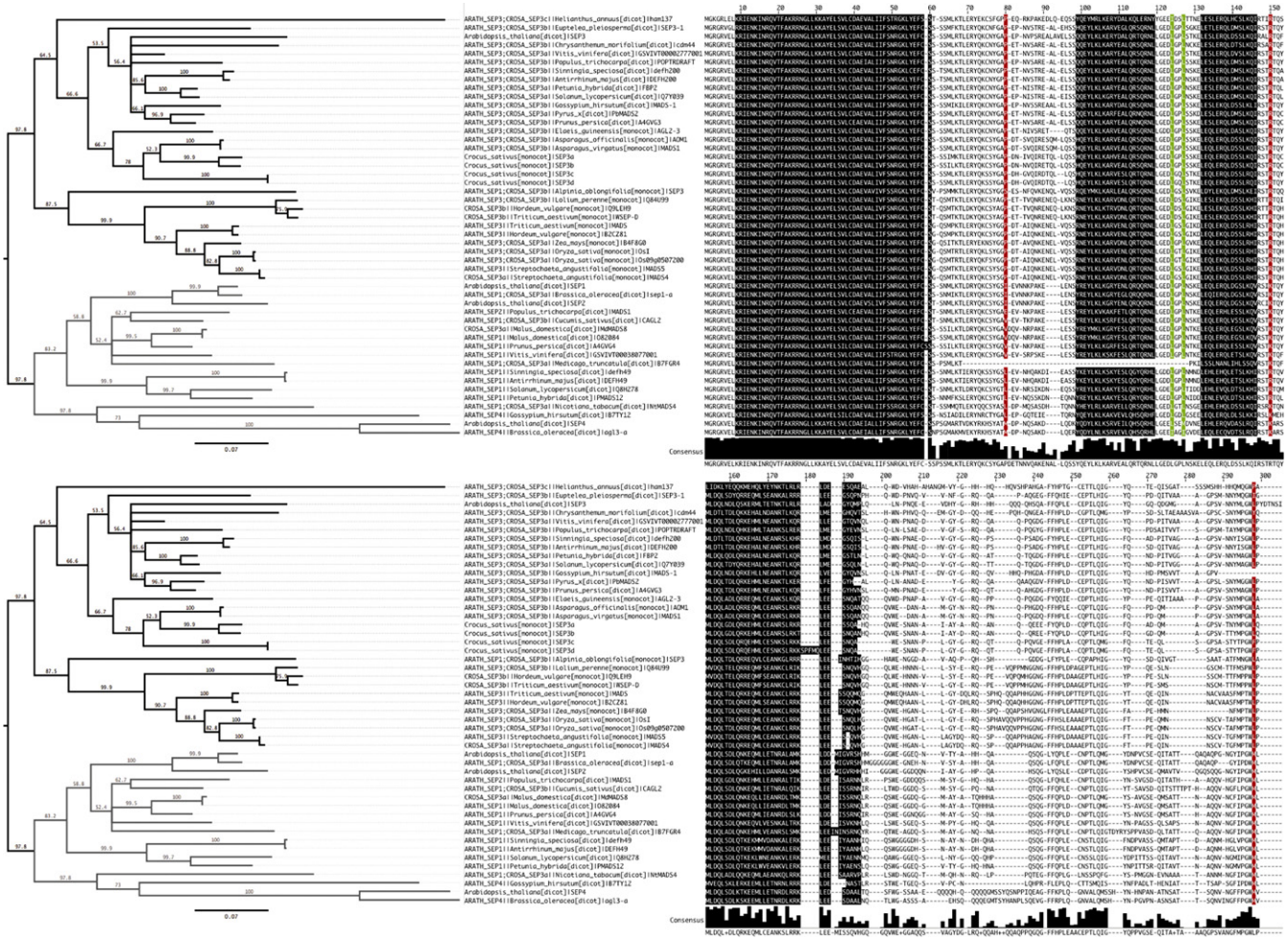

Fig. 2. Multiple alignment and phylogenetic tree of the four crocus SEP3-like proteins, *Arabidopsis* SEP1, SEP2, SEP3, and SEP4 and selected orthologs of monocot and dicot plant species from the UniProt database. Highlighted on the alignment with dark grey are the regions of the MADS-box domain (9–59), and the three subdomains of the K-box (K1: 99–119; K2: 132–146; K3: 154–194). Informative residues identified with the multi-RELIEF algorithm are highlighted in red, while two highly conserved leucines of the K1–K2 interhelical region are highlighted in green. The bottom feature line plots the consensus percentage support, with the consensus itself underneath. The phylogenetic tree shows the SEP3-like branch in black and the SEP1/SEP2/SEP4-like branch in grey. Percentage support from the bootstrap resampling is shown on the branches.

(Qiagen). On-column digestion of DNA during RNA purification was performed using the RNase-Free DNase Set (Qiagen). One (1) μg of total RNA from each sample was used in a reverse transcription reactions containing 200 U SuperScript II-RT, 0.5 μg NdT-Adaptor primer, 0.5 mM dNTPs, 1× first strand buffer and 10 mM DTT in a 20 μl reaction volume. The reactions were incubated for 50 min at 42 °C followed by 15 min at 70 °C to inactivate the RT. PCR was performed with a template of 1/20 of the sample cDNAs supplemented with 0.2 mM dNTPs,1 u DyNAzyme II DNA polymerase and 0.4 μM of the corresponding primers. A control PCR having as template 50 ng genomic DNA from crocus was included for each primer pair. Successful cDNA synthesis was monitored by amplifying a fragment of the actin-beta gene as described in Tsaftaris et al. (2006). The primers used were: Sep3-F/Sep3-2R for *CsatSEP3a* (annealing at 50 °C), CsSEP3B-F (5′-CTTGAAAGACAACTTGATTCGTCG-3′)/Contig3R for *CsatSEP3b* (annealing at 56 °C), CsSEP3C-F (5′-GAGCCTCCGCAAAAAGCTGGAAG-3′)/CsSE3C/D-R (5′-GCAAGTAA-CCAAGAGCAAATCAC-3′) for *CsatSEP3c* (annealing at 58 °C) and CsCEP3D-F (5′-GAGCCTCCGCAAAAAGTCTCCTT-3′)/CsSE3C/D-R for *CsatSEP3c_as* (annealing at 58 °C). The cycling conditions were: 2 min at 94 °C, 35 cycles of 30 s at 94 °C, 45 s at annealing temperature, 45 s at 72 °C and a final extension step of 10 min at 72 °C. The PCR products were separated on 2% agarose gels, where amplification products of the expected size could be observed.

Assessment of protein–protein interactions

To clone *CsatAP3b* (GenBank: AY948340), *CsatSEP3b* (GenBank: EU424138) and *CsatAP1*/*FULc* (GenBank: AY337930, also indicated in database as KCAP1c) open reading frames the 5′CsatAP3b-EcoRI (5′-GAATTCATGGGGAGAGGAAGGATTGAG-3′) and 3′CsatAP3b-XhoI (5′-CTGGAGTCATGCAAGACTCAGATCATG-3′) primers were used in PCR amplifications so as to incorporate restriction sites (5′EcoRI and 3′XhoI) that would enable the in-frame expression as fusion proteins. For *CsatAP1*/*FULc* the primers used were 5′CsatAP1/FULc-EcoRI (5′-GAA TTC ATG GGA AGG GGT AGG GTA CAG-3′) and 3′CsatAP1/FULc-XhoI (5′-CTCGAG CTATCCATTAACAGAGCGAAT-3′), and for *CsatSEP3b* 5′-CsatSEP3b-EcoRI (5′-GAATTCATGGGGAGAGGGAGAGTCGAG-3′) and 3′-CsatCEP3b-XhoI (5′-CTCGAGTCATTGCAACCATCCAGGCAT-3′). The PCR products were resolved in 1% agarose gel and the DNA fragments were gel extracted, purified, cloned into the pCRII-TOPO vector and sequenced. To subclone *CsatAP3b*, *CsatSEP3b* and *CsatAP1*/*FULc* fragments into suitable plasmids, the pCRII plasmids containing the inserts were restriction digested with EcoRI and XhoI. The digestion products were resolved in an agarose gel and the inserts were extracted, purified, and ligated to the pGilda (Gal1p-LexA-MCS, cen ori, HIS3), pJG4-5 (Gal1p-B42AD-MCS, 2 μ ori, TRP1) two-hybrid vectors. The *CsatAP1*/*FULc* insert was

Please cite this article in press as: Tsaftaris A, et al. The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors. J Plant Physiol (2011), doi:10.1016/j.jplph.2011.03.015

additionally subcloned into the pYES-mycGST vector expressing a myc tag-GST fusion protein in yeast.

The "bait" and "library" two hybrid constructs were co-transformed into the yeast reporter strain EGY48 (*ura3 trp1 his3* 6LexA-operator-LEU2) by lithium acetate transformation. Transformants were selected in glucose complete media lacking histidine and tryptophan (CM-his, tryp). Assays for the suitability of "bait" to test interactions were performed on galactose-raffinose complete media lacking leucine (CM-leu) (Kilili et al., 2004). The library (pJG4-5) constructs were co-transformed with the pYESmycGST or the pYESmycGST-CsatAP1/FULc construct and plated on glucose complete media lacking uracil and tryptophan (CM-ura, trp). Fusion protein expression was induced by transferring cells on galactose-based media which specifically induced the expression of the LexA, B42AD-HA, and GST fusion proteins. Protein expression was assessed by resolving the yeast cell protein extracts by SDS-PAGE, western blotting and detection using antibodies recognizing the LexA, Hemaglutinin (HA) and myc tag moieties.

For protein binding assays, yeast cells stably co-expressing the GST fusion of CsatAP1/FULc or the GST together with the HA-fusions of CsatSEP3b, CsatAP3b and CsatAP1/FULc were used to evaluate the capacity of these proteins to form complexes between them. The yeast protein OYE2 was used as a negative control. The well-known interaction between OYE2 and actin was used as positive control for the assay. Protein extracts from 50 ml cell cultures grown to late log phase in galactose based medium were brought to 1 ml volume in TBS-1% Triton X-100. In each tube 50 μl GSH-agarose were added and incubated at RT by constant mixing for 2 h. The GSH-agarose beads were washed extensively in TBS-Triton X-100 (>6×) by repeated centrifugations. At the end of washes the pellets were mixed with 20 μl of 2× SDS-PAGE loading buffer, the samples were denatured by boiling and the supernatants were resolved in a 12% gel and transferred on membrane by western blotting. Co-precipitating proteins were identified by probing the membranes with antibody recognizing the HA epitope tag (Simons and Vander Jagt, 1981).

Results

Isolation of crocus SEP3-like cDNAs

Thirty individual clones obtained from cloning of MADS-box genes using RCA-RACE methods with inserts ranging from 900 bp to about 1300 bp were sequenced. After initial analysis, four sequences were discarded because of PCR and circularization artifacts. BLAST similarity searches revealed that 26 remaining clones showed homology to MADS-box genes from different plant species; these sequences were grouped to seven distinct contigs using the combination of Phred/Phrap/Consed software (http://www.phrap.org/). While differences were found among them, two were nearly identical except for a 15 bp insertion/deletion (Fig. 1). Preliminary sequence analysis using BLAST revealed that four of the seven contigs were similar to other E-class *SEP* genes, and more specifically, to *SEP3*-like MADS-box genes from several plant species (a finding later confirmed by detailed bioinformatics described herein). Therefore, the four *SEP3*-like sequences from *C. sativus* were selected for further study.

To verify the four isolated cDNA sequences and complete small missing parts between inverse primers (see Materials and Methods), gene-specific primers were designed and used in RT-PCR, as described in Materials and Methods section. All three primer pairs produced single and clear RT-PCR products. Five to six individual clones for each one of the four contigs were sequenced and analyzed. The sequences obtained together with the sequences of the original clones, generated with RCA-RACE, were reassembled to four different sequences that were termed *CsatSEP3a*, *CsatSEP3b*, *CsatSEP3c* and *CsatSEP3c_as* (GenBank accession numbers EU424137, EU424138, EU424139, EU424140, respectively; UniProt accession numbers B3FTV4_CROSA, B3FTV5_CROSA, B3FTV6_CROSA, B3FTV7_CROSA, respectively). It is worth noting that part of the 3′ end of *CsatSEP3c* was identified in the recently published partial collection of saffron stigma-specific 6,803 ESTs (D'Agostino et al., 2007) under the cluster ID cr.saCl001179. The other three sequences were not represented in that partial collection.

To assess whether the four isolated *SEP3*-like cDNAs corresponds to four different genomic sequences, primers were used to isolate part of their genomic sequences covering the regions where differences among the cDNA sequences were found. Three different genomic sequences were isolated, corresponding to *CsatSEP3a*, *CsatSEP3b* and *CsatSEP3c*/*c_as* containing the last four exons and introns of the genes (Fig. 1A). Alignment between the four isolated cDNAs and the three genomic sequences revealed that the same genomic sequence encodes for the two *CsatSEP3c* and *CsatSEPc_as* sequences differing in 15 bp as a result of an alternative splicing, thus, "as" indicates alternative splicing (Fig. 1B). As also shown in Fig. 1A, similar genomic organization was found for all the isolated genes with two (*CsatSEP3a* and *CsatSEP3b*) presenting higher similarity, whereas *CsatSEP3c*/*c_as* was more distant.

Amino acid sequence and phylogenetic analysis

We collected 1895 sequences from the UniProt protein database containing the MADS-box domain, as described in the Materials and Methods. The *Arabidopsis* SEP3 protein (UniProt accession number: O22456) was the top *Arabidopsis* hit for all four crocus proteins with 69.26% identity to CsatSEP3a, 67.74% to CsatSEP3b, 67.35% to CsatSEP3c, and 65.31% to CsatSEP3c_as; the other three *Arabidopsis* SEPALLATA proteins SEP1 (UniProt: Q5XXN7), SEP2 (UniProt: P29384), and SEP4 (UniProt: P29383) were consistently in the hit list, with SEP4 furthest away.

Using BLAST-derived orthology information, we aligned our four sequences with *Arabidopsis* SEP1–4, along with putative orthologs from different monocot and dicot plant species (Sup. 1). As shown in Fig. 2, the resulting alignment highlighted interesting features, including the near-perfect conservation of the MADS-box domain (alignment positions 9–60), along with high conservation of the three subdomains of the K-box as described in (Yang and Jack, 2004) (K1: 99–119; K2: 132–146; K3: 154–194). Also very highly conserved are two hydrophobic leucines (alignment positions 124 and 126) in the K1–K2 interhelical region (Yang and Jack, 2004).

The SEP3-specific interacting arginine (R) at position 150 (Yang and Jack, 2004) was found in all *C. sativus* SEP3-like sequences as well as their putative orthologs, but was replaced by a lysine in the majority of the *Arabidopsis* SEP1, SEP2, and SEP4 and their putative orthologs. Other potential specificity-determining residues, as recognized by the multi-RELIEF algorithm, existed at position 80 with a proline in all SEP3 orthologs compared to isoleucines, leucines, methionines, valines and a glutamic acid in the rest, and at position 298 where the SEP3 orthologs had a leucine in the great majority of cases, whereas the rest exclusively featured a (biochemically similar) methionine. It should be noted that the significance of these residues in these positions cannot be evaluated, but are mentioned to support the placement of the *C. sativus* sequences in the SEP3 group (since they share the same residues in these positions with the rest of the SEP3 orthologs – an interesting exception is the phenylalanine at position 298 in CsatSEP3c and CsatSEP3c_as), and as potentially interesting features for further research.

As also shown in Fig. 2, the phylogenetic tree of putative orthologs clearly places two *Asparagus* proteins (AOM1 in *Asparagus*

Please cite this article in press as: Tsaftaris A, et al. The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors. J Plant Physiol (2011), doi:10.1016/j.jplph.2011.03.015

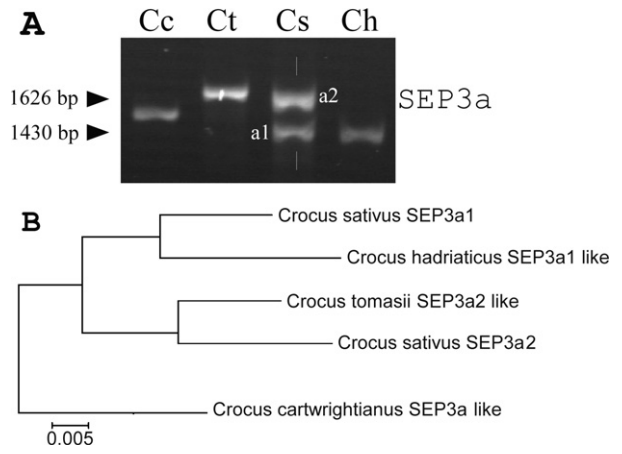

Fig. 3. (A) PCR products of SEP3a, SEP3b and SEP3c from *C. cartwrithgianus* (Cc), *C. tomasii* (Ct), *C. sativus* (Cs) and *C. hadriaticus* (Ch) Arrows indicate the bands amplified from *C. sativus*. (B) Phylogenetic tree of the isolated *SEP3A* sequences from the putative progenitors of *C. sativus*.

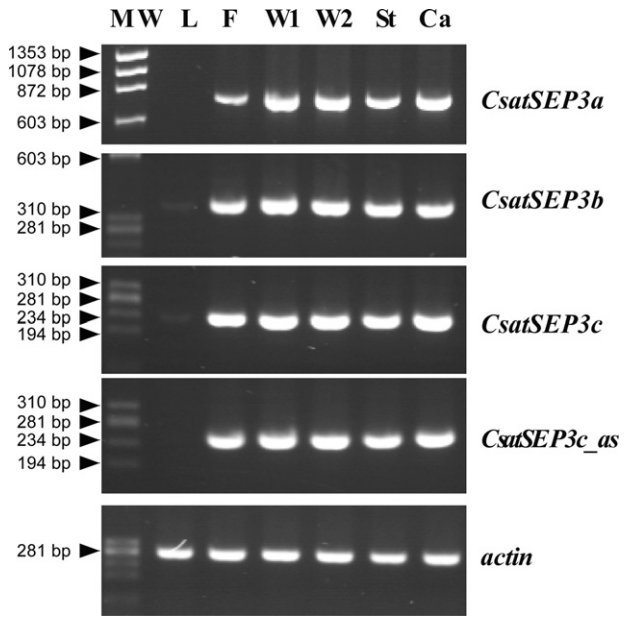

Fig. 4. PCR amplification products on cDNA from Leaves (L), Flowers (F), Whorl1 tepals (W1), whorl2 Tepals (W2), Stamens (St) and Carpels (Ca) using the *CsatSEP3a*-, *b*-, *c*-, *c_as* and β-actin-gene specific primers. MW indicates the λHindIII/ΦX174HaeIII molecular weight marker.

officinalis and MADS1 in *Asparagus virgatus*) as the closest putative ortholog to all four predicted *C. sativus* SEP3-like protein sequences.

Characterization of the genomic sequences of C. sativus CsatSEP3a and its putative progenitors

Amplification of the *CsatSEP3a* genomic region from *C. sativus* using gene specific primers showed two bands of different size designated *CsatSEP3a1* and *CsatSEP3a2* (Fig. 3A), while only a single band was observed for *CsatSEP3b* and *CsatSEP3c*/*c_as* (data not shown). To identify the origin of the *CsatSEP3a1* and *CsatSEP3a2* isolated sequences, we attempted to amplify SEP3a like genomic sequences from three putative progenitors of *C. sativus* namely *C. cartwrightianus*, *C. tomasii* and *C. hadriaticus*) (Seberg and Petersen, 2009; Grilli Caiola et al., 2010). As shown in Fig. 3, based on size differences, it is likely that the *CsatSEP3a1* gene derived from *C. hadriaticus* and *CsatSEP3a2* from *C. tomasii*. To obtain evidence for the origin of *CsatSEP3a1* and *CsatSEP3a2*, the two amplified bands from *C. sativus* and those from *C. cartwrightianus*, *C. tomasii* and *C. hadriaticus* were cloned and sequenced. Alignment of the two *CsatSEP3a* genomic sequences from *C. sativus* and the three putative progenitors revealed higher similarity with *C. tomasii* and *C. hadriaticus* as expected. Interestingly, only one *SEP3a* cDNA was isolated from *C. sativus* (named *CsatSEP3a1* Acc. EU424137) identical in the coding region of the *CsatSEP3a1* genomic sequence and more similar to the genomic sequence isolated from *C. hadriaticus*. This observation suggests that only the *CsatSEP3a1* gene, originated from *C. hadriaticus*, is expressed in *C. sativus*. Although these are preliminary data, they indicate that it is likely that one of the *C. sativus* triploid progenitors is *C. hadriaticus*, while *C. tomasii* could be another, as shown in the phylogenetic tree in Fig. 3B.

Expression analysis

The expression patterns of the isolated *SEP3*-like genes were compared between different crocus organs by RT-PCR. As shown in Fig. 4, the transcripts were mainly present in flowers and not in leaves (although a weak amplification product for *CsatSEP3b* and *CsatSEP3c* was observed in leaves). We also investigated expression patterns in different flower organs. As also shown in Fig. 4, RT-PCR experiments performed with cDNA synthesized from tepals, stamens and carpels resulted in the identification of all four transcripts in all mature flower organs.

Due to its triploidy, cultivated *C. sativus* is a sterile perennial species asexually propagated for thousands of years following its domestication, and numerous mutants preserved in the fields have been identified and described (Grilli Caiola et al., 2004). Among them, indeterminate flower phenotypes with double (Fig. 5) or even triple flowers have been found. Since the *SEP* gene has been found to play a role in flowering and flower determinacy (Liu et al., 2009) in addition to its role in flower organ formation according to the ABCE model (Malcomber and Kellogg, 2004), the expression of all four *CsatSEP3*-like transcripts was examined in one such mutant with double flowers. As shown in Fig. 5, there were no expression differences among the four *SEP*-like transcripts in WT and the double flower (DF) mutant. Also, no differences were found when another mutant (M), lacking the stamen part of the *C. sativus* flower, was examined (Fig. 5), where according to the ABCE model, SEP like proteins are expected to form tetrameric complexes together with B-class CsatAP3/PI-like and C-class CsatAG-like proteins.

Evaluation of protein–protein interactions

To assess the capacity of the isolated CsatSEP3-like proteins to interact with other MADS-box proteins isolated from crocus, and therefore add further evidence to their characterization, we selected *CsatSEP3b* along with previously published *CsatAP3b* and *CsatAP1*/*FULc* to test on the "yeast two-hybrid" system and on a cell-based GST-fusion co-precipitation approach. The three proteins were expressed as fusions to the LexA protein and the B42AD-HA (hemaglutinin tag) and were transformed singly and in pairs in yeast reporter cells (Golemis et al., 2001). Unfortunately, the LexA fusions of the MADS-box proteins exhibited a capacity to autoactivate the LEU2 reporter, making the two-hybrid assay unsuitable (data not shown). Thus, we subsequently employed a cell-based direct co-precipitation assay based on GST fusions. The cDNA of *CsatAP1*/*FULc* was expressed as the C-terminal fusion to a myc epitope tag and to the GST protein under the control of the inducible galactose promoter (Fig. 6A). The B42AD-HA fusions of CsatSEP3b, CsatAP3b and CsatAP1/FULc and *OYE2* (yeast Old Yellow Enzyme 2, used as control (Odat et al., 2007) were stably co-expressed with

Please cite this article in press as: Tsaftaris A, et al. The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors. J Plant Physiol (2011), doi:10.1016/j.jplph.2011.03.015

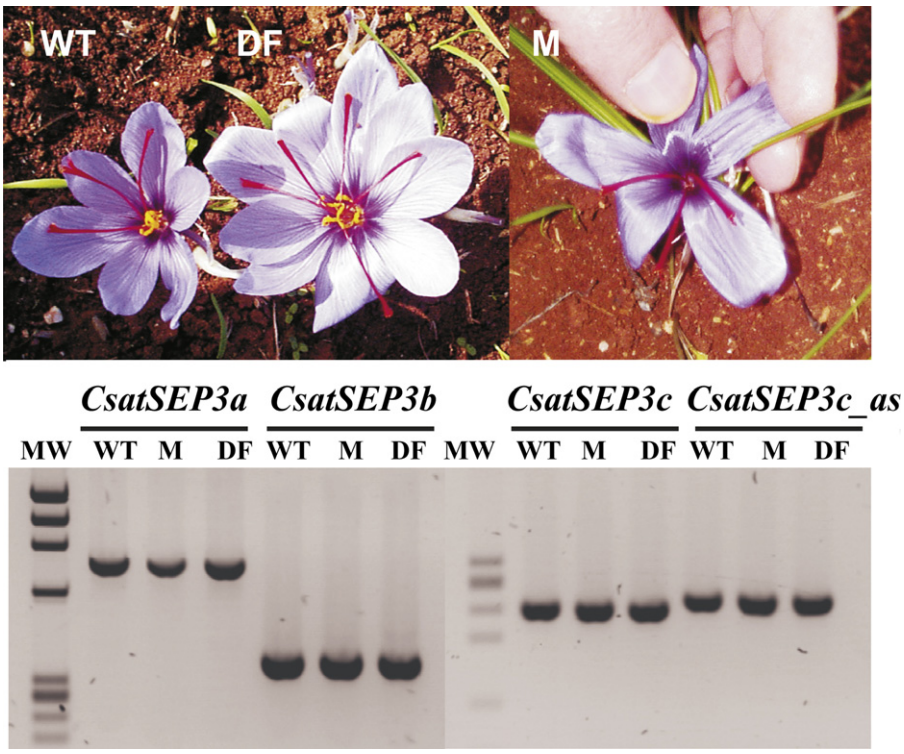

Fig. 5. Gene expression analysis with RT-PCR in carpels of wild-type (WT) flower, a mutant (M) *C. sativus* flower with no stamens and a double flower (DF) using *CsatSEP3a*-, *b*-, *c*-, *c_as* gene specific primers. In the upper panel a WT, a DF and a M flower are presented.

either the empty GST or the CsatAP1/FULc fusion. Yeast offers the advantage in co-precipitation assays that plasmids are maintained stably and proteins are consistently expressed in modest quantities compared to bacterially expressed GST fusions. Interactions take place within the intracellular environment of a eukaryotic cell.

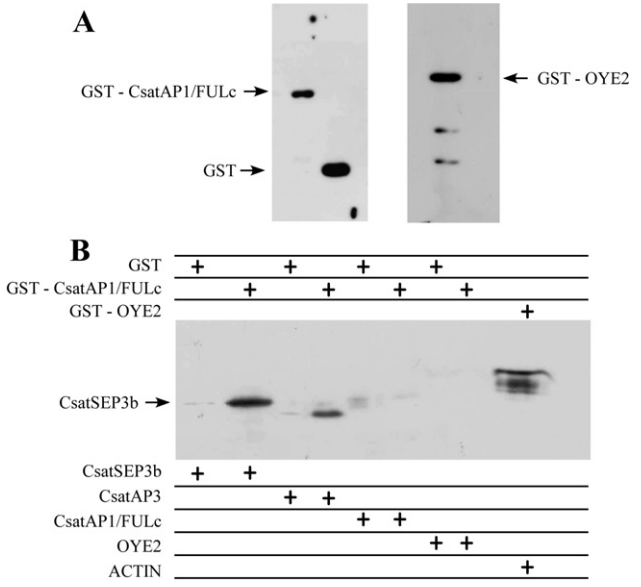

Fig. 6. *CsatAP1/FULc* interacts specifically with *CsatSEP3b* and *CsatAP3b*. (A) Protein extracts from EGY48 yeast cells expressing the GST-CsatAP1/FULc fusion, the GST and the GST-OYE2 fusion were resolved in a 12% SDS-PAGE and transferred on a PVDF membrane. The proteins were detected using an antibody recognizing the myc epitope tag. (B) Protein extracts from cells expressing GST, GST-CsatAP1/FULc or GST-OYE2 in pairs with the B42AD-HA fusions of *CsatSEP3b*, *CsatAP3b*, *CsatAP1/FULc*, *OYE2* and actin were used in co-precipitation assays. GSH agarose beads pulled down the GST fusions. Co-precipitating protein complexes were resolved in 12% SDS-PAGE and transferred to membrane. Detection of bound proteins was done using an anti-hemaglutinin (HA) tag antibody.

Protein extracts from cells expressing the various combinations were incubated with glutathione (GSH) agarose beads, which bind strongly to the GST component. Proteins that specifically bind to the CsatAP1/FULc part of the GST fusion would co-precipitate and could be detected using an HA antibody recognizing their N-terminal fusion. The B42AD-HA fusion of the yeast protein OYE2 was used as negative control, whereas the well-known interaction between a GST-OYE2 and B42AD-HA-actin was used as positive control for the assay.

As shown in Fig. 6B, GST-CsatAP1/FULc specifically co-precipitated CsatSEP3b and CsatAP3b but did not appear to interact significantly with itself to form a homodimer. As expected, GST alone could not precipitate the two proteins and GST-CsatAP1/FULc did not co-precipitate the unrelated OYE2 protein, confirming overall the specificity of the interactions observed.

Discussion

SEP-like genes play an important role in flower development, encoding one member of the transcriptional regulatory complex controlling flower organ formation (Immink et al., 2009).

In this work, we attempted to isolate *SEP*-like genes from *C. sativus*. Despite our extensive efforts, the only *SEP* genes isolated from crocus flowers were *SEP3*-like, as bioinformatics analysis with putatively orthologous sequences strongly suggests. Indeed, as shown Fig. 2, outside the promiscuous MADS-box domain, there are conserved amino acids (e.g. the SEP3-specific arginine (R) at position 147 of the alignment), but also intriguing differentiation. Such sequence deviations are primarily found in CsatSEP3c and CsatSEP3c_as (see Results), which could indicate some form of sub-functionalization, albeit not evident in our expression analysis in different flower organs and mutant flowers. These differences in the sequences can also be explained by the triploid genome of saf-

Please cite this article in press as: Tsaftaris A, et al. The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors. J Plant Physiol (2011), doi:10.1016/j.jplph.2011.03.015

fron crocus. *C. sativus* is triploid ($2n = 3x = 24$), generally assumed to be of autotriploid or hybrid origin (Grilli Caiola et al., 2004, 2010; Frizzi et al., 2007). Most researchers assume that one of the *C. sativus* parents must have been *C. cartwrightianus* donating the $2n$ $2 \times 8 = 16$ chromosomes part of the genome, while another unknown diploid parent contributed the other $1n$ $1 \times 8 = 8$ chromosomes. But recently, the hypothesis, based on sequence similarities, that *C. sativus* is an allotriploid consisting of three different genomes, two perhaps of close similarity and a third one more distant, is becoming credible (Grilli Caiola, 2005; Grilli Caiola et al., 2010). For instance, in previous work from our group (Tsaftaris et al., 2004, 2006; Kalivas et al., 2007) with *CsatAP1*/*FUL*, *CsatAP3*, *CsatPI* and *CsatAG*-like genes, three forms were frequently isolated. Two show very close similarity, presumably originating from *C. cartwrightianus* or *C. tomasii*, progenitors contributing maybe the two more similar genomes and one more distantly related, presumably originated from the other progenitor species, reinforcing the idea that *C. sativus* is an alloploid species and explaining the presence of differences in the sequences of the isolated *SEP3*-like genes. Experiments are underway in our and other laboratories to clarify the progenitor species of the triploid cultivated *C. sativus*, and complement the preliminary results presented in this paper, to better understand the origin of the different classes of MADS-box genes in this crop and the origin of this crop itself.

Unlike the sequence and probably the functional conservation, the expression pattern of *SEP*-like transcripts differs between monocots and eudicots. Within monocots, *SEP*-like genes have been most intensively studied in grasses, including the important cereals maize and rice (Cui et al., 2010; Gao et al., 2010). Similarly to other monocots, expression of the isolated *SEP3*-like genes of *C. sativus* was detected in all four whorls of flower organs. This pattern of extended expression of E-class genes together with B-class genes reported previously (Tsaftaris et al., 2004, 2006) is compatible with tepal formation in whorl1. *Muscari armeniacum* is another member of *Asparagales* that has petaloid organs in the outer two whorls and the expression of B-class genes extended to whorl1, similar to saffron crocus (Nakada et al., 2006). Thus, the extended expression of B- and E-class genes in whorl1 fits the extended ABCE model proposed to explain tepal formation in tulip and other nongrass monocots such as *C. sativus* (van Tunen et al., 1993; Kanno et al., 2003, 2007). This is further supported by the co-precipitation of CsatSEP3b and CsatAP3b proteins to form heterodimers with CsatAP1/FULc. This result is supported indirectly by our earlier findings concerning the expression of a *C. sativus* NAP-like gene, a member of the plant specific NAC family transcription factor proteins (Kalivas et al., 2010). It is known, for example, that heterodimers of AP3/PI MADS-box transcriptional factors induce the expression of NAP-like genes by recognizing and binding on the CArG motifs present on the NAP promoter sequences (Sablowski and Meyerowitz, 1998). In the promoter sequence of *CsatNAP*-like, isolated and characterized previously from our group, two nearby putative CArG boxes were found, presumably targets of AP3/PI heterodimers that can induce the expression on *CsatNAP*-like gene in whorl1, where *CsatNAP* was found to be expressed, indicating indirectly that a functional AP3/PI protein could be present in whorl1 tepals in *C. sativus*, too. Combining this observation with the presence of *CsatSEP3* transcripts in the whorl1 and the protein–protein interaction experiments of CsatSEP3b with CsatAP3 and CsatAP1/FULc presented in this paper (Fig. 6), we suggest that, in whorl1 of saffron crocus, the formation of the quartet protein complex between SEP3, AP3/PI and AP1/FUL is feasible. However, expression of genes like *CsatAP3*/*PI*, *CsatAP1*/*FUL* and *CsatSEP3* in whorl4 (as observed in *C. sativus* wild-type and mutant flowers and other non-grass monocots), together with *CsatAG*-like MADS-box gene expression, is a result difficult to explain (Kanno et al., 2007). To further our understanding of saffron crocus flower formation and particularly whorl4 carpel formation, it would therefore be interesting to examine MADS-box protein accumulation and interaction of crocus B-, C- and E-class function genes in whorl4.

In summary, four *C. sativus SEP3*-like cDNAs, transcribed from three genes, were isolated, their sequences and phylogenetic relationships described, and their expression patterns and potential for protein interactions with other *C. sativus* MADS-box proteins characterized. All isolated cDNAs encode for proteins belonging to the mixed eudicot–monocot *SEP3* subclade of MADS-box proteins and although there seems to be no reason for the absence of other crocus *SEP*-like genes, these were the only such genes isolated after extensive search with different methods. Overall, these data suggest that the inferred role of the isolated E-class genes in *C. sativus* is compatible with that of E-function MADS-box transcription factors.

Acknowledgments

We thank Mr. Yiannis Patsios for his help in collecting plant material in the field. Financial support for the Institute of Agrobiotechnology/CERTH from the General Secretariat of Research and Technology of Greece and the EU project CROCUSBANK are also acknowledged. Dr. Ole Seberg and Dr. Eleni Maloupa are also acknowledged for providing biological material.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jplph.2011.03.015.

References

Agrawal GK, Abe K, Yamazaki M, Miyao A, Hirochika H. Conservation of the E-function for floral organ identity in rice revealed by the analysis of tissue culture-induced loss-of-function mutants of the OsMADS1 gene. Plant Mol Biol 2005;59:125–35.
Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–402.
Becker A, Theissen G. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol 2003;29:464–89.
Bowman JL. Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. J Biosci 1997;22:515–27.
Cacharron J, Saedler H, Theissen G. Expression of MADS box genes ZMM8 and ZMM14 during inflorescence development of *Zea mays* discriminates between the upper and the lower floret of each spikelet. Dev Genes Evol 1999;209:411–20.
Coen ES, Meyerowitz EM. The war of the whorls: genetic interactions controlling flower development. Nature 1991;353:31–7.
Cui R, Han J, Zhao S, Su K, Wu F, Du X, et al. Functional conservation and diversification of class E floral homeotic genes in rice (*Oryza sativa*). Plant J 2010;61:767–81.
D'Agostino N, Pizzichini D, Chiusano ML, Giuliano G. An EST database from saffron stigmas. BMC Plant Biol 2007;7:53.
Ditta G, Pinyopich A, Robles P, Pelaz S, Yanofsky MF. The SEP4 gene of *Arabidopsis thaliana* functions in floral organ and meristem identity. Curr Biol 2004;14:1935–40.
Dornelas MC, Patreze CM, Angenent GC, Immink RG. MADS: the missing link between identity and growth? Trends Plant Sci 2010;16:89–97.
Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz HR, et al. The Pfam protein families database. Nucleic Acids Res 2008;36:D281–288.
Frizzi G, Miranda M, Pantani C, Tammaro F. Allozyme differentiation in four species of the *Crocus cartwrightianus* group and in cultivated saffron (*Crocus sativus*). Biochem Syst Ecol 2007;35:859–68.
Gao X, Liang W, Yin C, Ji S, Wang H, Su X, et al. The SEPALLATA-like gene OsMADS34 is required for rice inflorescence and spikelet development. Plant Physiol 2010;153:728–40.
Golemis A, Serebriiskii I, Finley RJ, Kolonin MG, Gyuris J, Brent R. Interaction trap/two-hybrid system to identify interacting proteins. Curr Protoc Mol Biol 2001, 2009; Chapt. 19: Unit 19.2.
Goto K, Kyozuka J, Bowman JL. Turning floral organs into leaves, leaves into floral organs. Curr Opin Genet Dev 2001;11:449–56.
Grilli Caiola M. Embryo origin and development in *Crocus sativus* L. (Iridaceae). Plant Biosyst 2005;139:335–43.
Grilli Caiola M, Caputo P, Zanier R. RAPD analysis in *Crocus sativus* L. accessions and related Crocus species. Biol Plant 2004;48:375–80.

Grilli Caiola M, Leonardi D, Canini A. Seed structure in *Crocus sativus* L. × *C. cartwrightianus* Herb., *C. thomasii* Ten., and *C. hadriaticus* Herb. at SEM. Plant Syst Evol 2010;285:111–20.
Honma T, Goto K. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature 2001;409:525–9.
Immink RG, Angenent GC. Transcription factors do it together: the hows and whys of studying protein–protein interactions. Trends Plant Sci 2002;7:531–4.
Immink RG, Tonaco IA, de Folter S, Shchennikova A, van Dijk AD, Busscher-Lange J, et al. SEPALLATA3: the 'glue' for MADS box transcription factor complex formation. Genome Biol 2009;10:R24.
Jack T. Relearning our ABCs: new twists on an old model. Trends Plant Sci 2001;6:310–6.
Jack T. Molecular and genetic mechanisms of floral control. Plant Cell 2004;16(Suppl.):S1–17.
Kalivas A, Pasentsis K, Argiriou A, Darzentas N, Tsaftaris AS. famRCA-RACE: a rolling circle amplification race for isolating a family of homologous cDNAs in one reaction and its application to obtain nac genes transcription factors from crocus (*Crocus sativus*) flower. Prep Biochem Biotechnol 2010;40:177–87.
Kalivas A, Pasentsis K, Polidoros AN, Tsaftaris AS. Heterotopic expression of B-class floral homeotic genes PISTILLATA/GLOBOSA supports a modified model for crocus (*Crocus sativus* L.) flower formation. DNA Seq 2007;18:120–30.
Kanno A, Nakada M, Akita Y, Hirai M. Class B gene expression and the modified ABC model in nongrass monocots. Scientific World J 2007;7:268–79.
Kanno A, Saeki H, Kameya T, Saedler H, Theissen G. Heterotopic expression of class B floral homeotic genes supports a modified ABC model for tulip (*Tulipa gesneriana*). Plant Mol Biol 2003;52:831–41.
Kaufmann K, Muino JM, Jauregui R, Airoldi CA, Smaczniak C, Krajewski P, et al. Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the *Arabidopsis* flower. PLoS Biol 2009;7:e1000090.
Kilili KG, Atanassova N, Vardanyan A, Clatot N, Al-Sabarna K, Kanellopoulos PN, et al. Differential roles of tau class glutathione S-transferases in oxidative stress. J Biol Chem 2004;279:24540–51.
Kobayashi K, Maekawa M, Miyao A, Hirochika H, Kyozuka J. PANICLE PHYTOMER2 (PAP2), encoding a SEPALLATA subfamily MADS-box protein, positively controls spikelet meristem identity in rice. Plant Cell Physiol 2010;51:47–57.
Li H, Luo J, Hemphill JK, Wang J, Gould J. A rapid and high yielding DNA miniprep for cotton (*Gossypium* spp.). Plant Mol Biol Rep 2001;19. pp. 183a–183e.
Liu C, Xi W, Shen L, Tan C, Yu H. Regulation of floral patterning by flowering time genes. Dev Cell 2009;16:711–22.
Ma H, Yanofsky MF, Meyerowitz EM. AGL1-AGL6, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. Genes Dev 1991;5:484–95.
Malcomber ST, Kellogg EA. Heterogeneous expression patterns and separate roles of the *SEPALLATA* gene *LEAFY HULL STERILE1* in grasses. Plant Cell 2004;16:1692–706.
Mandel MA, Yanofsky MF. The Arabidopsis *AGL9* MADS box gene is expressed in young flower primordia. Sex Plant Reprod 1998;11:22–8.
Melzer R, Theissen G. Reconstitution of 'floral quartets' in vitro involving class B and class E floral homeotic proteins. Nucleic Acids Res 2009;37:2723–36.
Melzer R, Verelst W, Theissen G. The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in 'floral quartet'-like complexes in vitro. Nucleic Acids Res 2009;37:144–57.
Melzer R, Wang YQ, Theissen G. The naked and the dead: the ABCs of gymnosperm reproduction and the origin of the angiosperm flower. Semin Cell Dev Biol 2010;21:118–28.
Nakada M, Komatsu M, Ochiai T, Ohtsu K, Nakazono M, Nishizawa NK, et al. Isolation of *MaDEF* from *Muscari armeniacum* and analysis of its expression using laser microdissection. Plant Sci 2006;170:143–50.
Odat O, Matta S, Khalil H, Kampranis SC, Pfau R, Tsichlis PN, et al. Old yellow enzymes, highly homologous FMN oxidoreductases with modulating roles in oxidative stress and programmed cell death in yeast. J Biol Chem 2007;282:36010–23.
Overbeek R, Fonstein M, D'Souza M, Pusch GD, Maltsev N. The use of gene clusters to infer functional coupling. Proc Natl Acad Sci USA 1999;96:2896–901.
Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF. B and C floral organ identity functions require *SEPALLATA* MADS-box genes. Nature 2000;405:200–3.
Polidoros AN, Pasentsis K, Tsaftaris AS. Rolling circle amplification-RACE: a method for simultaneous isolation of 5′ and 3′ cDNA ends from amplified cDNA templates. Biotechniques 2006;41:35–40.
Rijpkema AS, Zethof J, Gerats T, Vandenbussche M. The petunia AGL6 gene has a SEPALLATA-like function in floral patterning. Plant J 2009;60:1–9.
Sablowski RW, Meyerowitz EM. A homolog of no apical meristem is an immediate target of the floral homeotic genes APETALA3/PISTILLATA. Cell 1998;92:93–103.
Seberg O, Petersen G. How many loci does it take to DNA Barcode a *Crocus*? PLoS One 2009;4.
Shan H, Zahn L, Guindon S, Wall PK, Kong H, Ma H, et al. Evolution of plant MADS box transcription factors: evidence for shifts in selection associated with early angiosperm diversification and concerted gene duplications. Mol Biol Evol 2009;26:2229–44.
Simons PC, Vander Jagt DL. Purification of glutathione S-transferases by glutathione-affinity chromatography. Methods Enzymol 1981;77:235–7.
Theissen G. Development of floral organ identity: stories from the MADS house. Curr Opin Plant Biol 2001;4:75–85.
Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, et al. A short history of MADS-box genes in plants. Plant Mol Biol 2000;42:115–49.
Theissen G, Saedler H. Plant biology. Floral quartets. Nature 2001;409:469–71.
Tsaftaris AS, Pasentsis K, Iliopoulos I, Polidoros AN. Isolation of three homologous AP1-like MADS-box genes in crocus (*Crocus sativus* L.) and characterization of their expression. Plant Sci 2004;166:1235–43.
Tsaftaris AS, Polidoros AN, Pasentsis K, Kalivas A. Tepal formation and expression pattern of B-class paleoAP3-like MADS-box genes in crocus (*Crocus sativus* L.). Plant Sci 2006;170:238–46.
Tsaftaris AS, Polidoros AN, Pasentsis K, Kalivas A. Cloning, structural characterization, and phylogenetic analysis of flower MADS-box genes from crocus (*Crocus sativus* L.). Scientific World J 2007;7:1047–62.
Tzeng TY, Hsiao CC, Chi PJ, Yang CH. Two lily SEPALLATA-like genes cause different effects on floral formation and floral transition in *Arabidopsis*. Plant Physiol 2003;133:1091–101.
van Tunen AJ, Eikelboom W, Angenent GC. Floral organogenesis in *Tulipa*. Flower Newslett 1993;16:33–7.
Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 2009;25:1189–91.
Yang Y, Jack T. Defining subdomains of the K domain important for protein–protein interactions of plant MADS proteins. Plant Mol Biol 2004;55:45–59.
Ye K, Anton Feenstra K, Heringa J, Ijzerman AP, Marchiori E. Multi-RELIEF: a method to recognize specificity determining residues from multiple sequence alignments using a Machine-Learning approach for feature weighting. Bioinformatics 2008;24:18–25.

Please cite this article in press as: Tsaftaris A, et al. The study of the E-class *SEPALLATA3*-like MADS-box genes in wild-type and mutant flowers of cultivated saffron crocus (*Crocus sativus* L.) and its putative progenitors. J Plant Physiol (2011), doi:10.1016/j.jplph.2011.03.015