# Protein expression and characterization of SEP3 from *Arabidopsis thaliana*

**Q. Shi, J. Zhou, P. Wang, X. Lin and Y. Xu**

Nurturing Station for the State Key Laboratory of Subtropical Silviculture, Zhejiang Agriculture and Forestry University, Lin'an, Zhejiang, China

Corresponding author: Y. Xu
E-mail: yxu@zafu.edu.cn

**ABSTRACT.** *SEPALLATA* (*SEP*) MADS-box genes play crucial roles in the regulation of floral growth and development. They are required for the specification of sepals, petals, stamens, and carpels as well as for floral determinacy. SEPs perform their functions through the formation of homo- or hetero-polymers, which are the molecular basis of floral quartets. *In vitro* assays indicated that SEP3 forms a tetramer after binding to DNA, but it is unclear whether DNA binding induces the tetramer, because SEP3 is often reported to form a dimer. Here, we analyzed the oligomeric status of SEP3 domains in the absence of the DNA-binding MADS-box domain. The truncated SEP3 was constructed as a fusion protein and expressed in prokaryotic cells. The purified protein fragment displayed as a tetramer in the size exclusion chromatographic column, and a glutaraldehyde cross-linking assay demonstrated that the protein contained a dimer unit. Yeast two-hybrid tests further verified that the fragments form homologous polymers *in vivo*, and that the K domain is involved in tetramer formation. Current results imply that the SEP3 protein regulates the formation of flower meristems using the tetramer as a unit, and that the DNA-binding MADS-box is dispensable for polymer formation. The C-terminal region

does not contribute to homo-tetramer formation, but it may be reserved to glue other proteins.

**Key words:** SEP3 protein; Tetramer; Yeast-two hybrid; Protein purification; *Arabidopsis thaliana*

## INTRODUCTION

The MADS-box gene family plays crucial roles in plant development and flowering formation (Agrawal et al., 2005). To specify the floral organ identity, the five MADS-box gene classes interact in a combinational way to compose the well-known "floral quartet model". In this model, the E class MADS-box *SEPALLATA3* (*SEP3*) is indispensable, and it plays central roles (Coen and Meyerowitz, 1991; Krizek and Meyerowitz, 1996). For instance, it interacts with other class E genes and *APETALA1* (*AP1*) to specify sepals, with *AP1* and *AP3-PISTILLATA* (*PI*) to specify petals, with the *AP3-PI-AGAMOUS* (*AG*) complex to specify stamens, and with *AG* to specify carpels (Pelaz et al., 2000; Pelaz et al., 2001; Kaufmann et al., 2009). Chromatin immunoprecipitation and whole-genome hybridization arrays have shown that thousands of SEP3 binding sites lie in the floral organs, and that SEP3 integrates with these target MADS-box genes to control flower development (Kaufmann et al., 2009; Hwan Lee et al., 2012; Liu et al., 2014). *In vitro* DNA hybrid tests have shown that mammalian MADS-box transcription factors bind specific DNA elements in the promoter region, which are known as CC-A rich-GG (CArG-box). For example, the serum response factor (*SRF*) and myocyte enhancer factor 2A (*MEF2A*) MADS-box genes can bind the sequence CC(A/T)6GG (West et al., 1997; West et al., 1998). Similar to *SRF* and *MEF2A*, plant MADS-box transcription factors can also bind CArG-box-like sequences, and the binding region is mainly composed of the N-terminal MADS (M) domain (West et al., 1998). As a typical MADS-box protein, SEP3 contains M, intervening (I), keratin-like (K), and C-terminus (C) structural domains. Recent *in vitro* studies indicated that SEP3 proteins could bind to two CArG-boxes (Melzer et al., 2009; Jetha et al., 2014), suggesting a tetrameric construction of SEP3. However, it is unclear whether the tetrameric SEP3 is induced by DNA binding or if it is a homo-tetramer, because the ability of SEP3 to form dimers has been previously reported (Huang et al., 1996; Mendes et al., 2013). Additionally, indirect evidence suggests that SEP3 is a tetramer or a dimer when forming a homocomplex. A recently determined crystal structure of an SEP3 fragment (75-178 aa) revealed the mechanism of SEP3 tetramerization (Puranik et al., 2014). However, the purified protein fragment contained both dimer and tetramer components. Therefore, it is necessary to clarify if SEP3 forms homo-dimers or homo-tetramers, and the role of DNA binding in polymer formation must also be determined. Here we analyzed the oligomeric status for SEP3 domains without DNA binding regions. The proteins were expressed in prokaryotic cells, and were analyzed using a size exclusion chromatographic column. The results indicated that SEP3 forms tetramers, and that DNA binding is not required for SEP3 polymer formation. Our data also suggested that the K region plays a major role in mediating SEP3 self-association, and that the C-terminus may be reserved for gluing other MADS-box proteins.

## MATERIAL AND METHODS

### Material

We used the plant Columbia *Arabidopsis thaliana* as our research subject. A yeast two-

hybrid (Y2H) kit was purchased from Clontech (Beijing), and Trans1-T1 competent cells were obtained from Trans Gen Biotech (China). The PCR product purification kit was purchased from TaKaRa (China), and the plasmid extraction kit was acquired from Sangon (China). pET protein expression vectors were purchased from Novagen (Germany), and the remaining chemical reagents were domestic.

## Construction of SEP3 domain expression plasmids

Total RNA was isolated from wild-type blossom buds using the TRIZOL reagent (China), and it was then reverse transcribed into cDNA. The open reading frame encoding SEP3 was amplified via PCR with the following upstream and downstream primer sequences: 5'-CATGTGCCATGG GAAGAGGGAGAGTAGAATTGA-3' and 5'-CGCGGATCCTTAAGTGAGAAAGATTCAAAT-3'. PCR was performed with a denaturation step at 94°C for 3 min, followed by 30 cycles of 94°C for 30 s, 56°C for 30 s, 72°C for 30 s, and an extension at 72°C for 6 min. The PCR product was T-A linked to a pMD-19 vector (TakaRa) to serve as a template. SEP3B (78-230 aa) and SEP3C (78-190 aa) were obtained by cloning the PCR products into the pET15b vector between the Nco1 and BamH1 restriction sites. The reverse primers for SEP3C and SEP3B were GGATCCTCAGTGATGATGATG ATGATGTTGGTTAGGGTTCAGCTGGA-3' and ATCCTCAGTGATGATGATGATGATGCATTCCAT CTTGTTGCCCCT-3', respectively. They both contained a non-cleavable 6XHis-tag before the stop codon. All constructs were transformed into the DH5α strain, and the plasmids were extracted and confirmed via DNA sequencing.

## Over-expression and purification of His-tagged SEP3 domain proteins

The recombinant plasmids were transferred to BL21 (DE3) Rosetta cells for over-expression analysis. To screen for optimal temperatures, 3 mL cell cultures were induced at 20° and 37°C. To scale up the expression system, 10 mL of freshly grown overnight cell cultures were used to inoculate 1 L LB medium, which was induced by adding 0.1 mM IPTG at 20°C for 15 h. The harvested cell pellets were resuspended in 10 mL lysis buffer (0.5 M NaCl, 20 mM Tris-HCl, pH 8.0) on ice, and were lysed via ultrasonication for 10 min with an 8 s pulse or 10 s interval. Each sample was centrifuged for 40 min at a speed of 36,500 $g$. The supernatant was incubated with 5 mL nickel beads on ice with slow rocking for 1 h. To remove non-specific binding proteins, the beads were washed with a gradient concentration of imidazole (4/20/80 mM) in buffer containing 200 mM NaCl, 20 mM Tris-HCl, pH 8.0. The target proteins were then eluted with 40 mL 200 mM imidazole in 200 mM NaCl and 20 mM Tris-HCl, pH 8.0. The proteins were concentrated to 500 µL using an Amicon ultracentrifuge at 4°C, and were then loaded onto the gel filtration column (Superpose 6) for final cleanup. The protein purity was checked by SDS-PAGE, and was quantified based on its optical absorbance coefficient at 280 nm.

### Glutaraldehyde cross-linking

In order to determine the homologous interactions within or between the different subunits, we performed a cross-linking reaction using glutaraldehyde, which is capable of cross-linking surface amine groups. His-tag proteins SEP3B and SEP3C were reacted with glutaraldehyde at a molar concentration ratio of 1:10 at 20°C. Resulting 10 µL samples were taken at different time points, and an equal volume of Tris-HCl buffer (100 mM, pH 8.0) was added to terminate the reaction. The samples were separated by SDS-PAGE, and were detected using the silver staining method.

## Y2H assays

Y2H assays were used to determine which SEP3 domain plays the key role in polymer formation. The experiment was designed according to the Matchmaker GAL4 Two-Hybrid System (Clontech). Five SEP3 domains were PCR amplified with primers containing two restriction enzyme loci: EcoR1 and BamH1. After cleavage and purification, the inserts were ligated to a pGADT7AD vector containing the GAL4 activation domain (AD) or a pGBKT7 vector containing the DNA binding domain (BD). The prey recombinant plasmids (AD) were transferred into AH109 cells (Clontech), and bait vectors (BD) were transferred into Y187 cells (Clontech). AD and BD cells were incubated at 30ºC on SD/-Leu selective medium and SD/-Trp plates, respectively. AD cells were mixed with BD cells and hybridized overnight. Selection of the hybrid clones was performed on SD/-Leu/-Trp plates, and protein interactions were detected on SD/-Leu/-Trp/-His/-Ade with X-α-Gal plates.

## RESULTS AND DISCUSSION

### Protein expression and purification analysis

The desired nucleotides of the full length SEP3 and the deletion construct SEP3C were verified by sequencing. The domain arrangement is shown in Figure 1A, and induction of the recombinant domain protein in *E. coli* was found to be more efficient at 20°C than at 37°C. Although the target protein expression levels remained roughly the same, the soluble fraction tended to decrease at the higher temperature (Figure 1B). Therefore, the induction was performed at 20°C. The full-length protein was not soluble under the tested conditions, and purification of the soluble recombinant protein was conducted through Ni-NTA affinity chromatography.

As shown in the SDS-PAGE results (Figure 1B), the SEP3C protein was eluted with a buffer containing 200 mM imidazole, and the eluted protein was over 90% pure. Moreover, SEP3C migrated slightly faster than the 15 kDa protein marker (Figure 1B), which was consistent with the estimated molecular weight of 13 kDa. The protein expression yield was high, and we constantly obtained over 4 mg of pure protein from per liter of cell culture. The apparent protein size in solution was determined by size exclusion chromatography, which was calibrated by ovalbumin (43 kDa) and BSA (67 and 133 kDa) markers. As shown in Figure 2, SEP3C migrated with an apparent size of ~50 kDa, demonstrating that the SEP3C domain is tetrameric in solution. Expression of the longer C-terminus construct, SEP3B, was previously reported (Zhou Jiaping, 2014), and it was also found to be tetrameric under similar conditions.

### Cross-linking analysis

To confirm the oligomeric status, a glutaraldehyde cross-linking experiment was performed. As shown in Figure 3, dimer bands appeared after the 20 min incubation, and became denser with time. This was true for both SEP3B and SEP3C proteins. Although tetrameric bands were not observed, the results do not definitively indicate that the proteins are not tetrameric. Instead, the results suggest that the dimer part of the tetrameric protein is readily cross-linked. It may be true that the distribution of the amine groups on the three dimensional structure restricts the cross-linking of the tetramers.
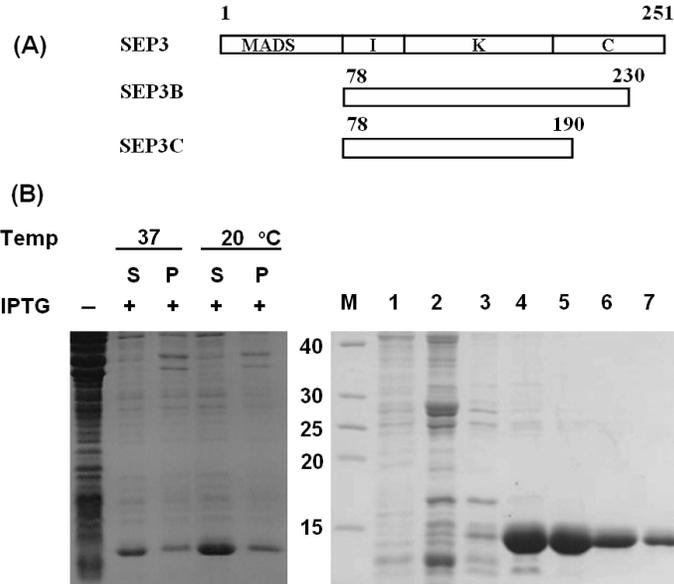
**Figure 1.** Illustrations of the SEP3 domain arrangement and protein expression profile. **A.** Full length SEP3 MADS-box protein, including MADS (M), intervening (I), keratin-like (K), and C-terminal (C) domains. SEP3C and SEP3B denote motifs with 78-190 aa and 78-230 aa, respectively. **B.** SDS-PAGE results show the induction (left) and purification (right) profiles. On the left gel, S and P denote the soluble and insoluble fractions, respectively. In the right, lanes 1-3 correspond to washes with 20, 40, and 80 mM imidazole, respectively, and lanes 4-7 correspond to elutions with 200 mM imidazole. M denotes the protein molecular-weight marker in kDa.
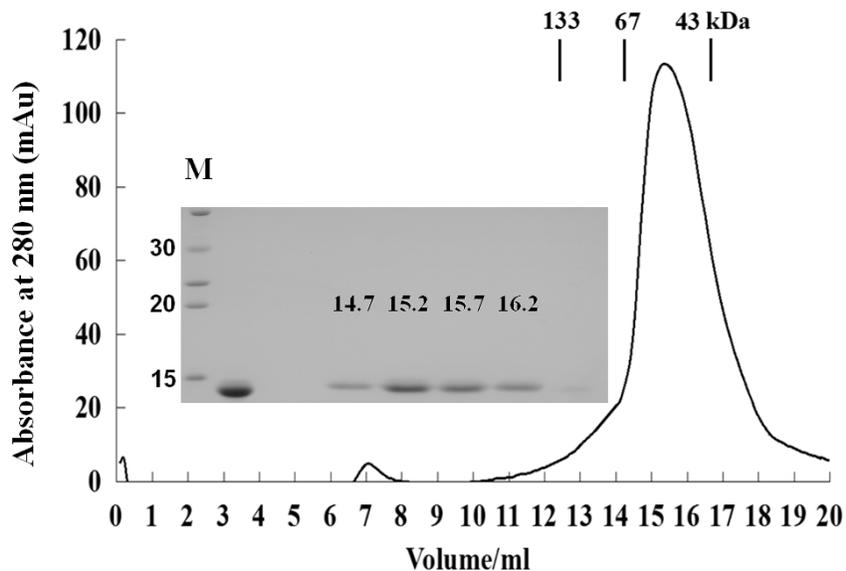


**Figure 2.** SEC profile for SEP3C protein. The SDS-PAGE insert corresponds to the SEC profile. The numbers match the elution fraction volume in SEC. M denotes the protein molecular-weight marker in kDa.
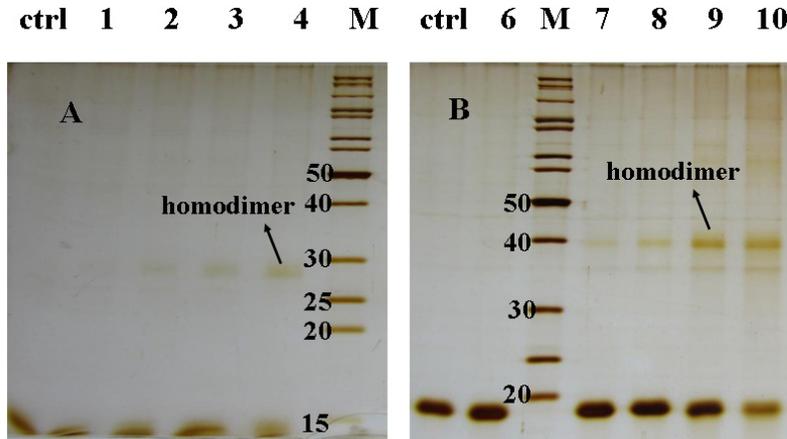
**Figure 3.** Glutaraldehyde cross-linking assays for the SEP3C and SEP3B proteins. The purified protein was incubated in either the absence or presence of glutaraldehyde at 20ºC. Silver stain was used to view SDS-PAGE results. (A) SEP3C. *Lanes 1-4* correspond to incubation times of 5 min, 20 min, 1 h, and 5 h, respectively. (B) SEP3B. *Lanes 6-10* correspond to incubation times of 5 min, 20 min, 1 h, 5 h, and 18 h, respectively. Ctrl represents the control group that was free from the cross-linking reagent.

## Y2H analysis

Y2H experiments were performed to verify if the protein domains interacted *in vivo* and to determine which domain is vital to the interaction. The domain design is shown in Figure 4A. The KC motif construct showed strong auto-activation, so it was removed from the bait list. As expected, the positive and negative controls appeared as blue and white clones, respectively (Figure 4B). The presence of a blue spot verified that the full-length SEP3 formed a homocomplex (Figure 4B), which is consistent with a previous report (de Folter et al., 2005). The KC and IK domains formed interactions with the full length SEP3. The IK domain, close to the SEP3C construct, showed a weak but positive signal for the homo-interaction, which was consistent with the *in vitro* expression assay. However, the lack of interaction with any region was indicative of the unexpected appearance of the IKC motif. Although it was unclear why IKC lacked interactions, the Y2H experiments results still suggested that the K domain outlines the core part of SEP3 tetramer formation. This is similar to many MADS-box proteins where the K domain mediates the internal complex formation (Yang and Jack, 2004; Leseberg et al., 2008). Recent crystal structure results revealed that four K domains were assembled together through their second helices, and the first helices made the dimer only (Puranik et al., 2014).

In the Y2H experimental assay, the MI motif did not associate with any region. This result, together with the *in vitro* expression assay, suggests that the MADS-box M region is dispensable for tetrameric formation. In another words, the tetramer is not induced by DNA binding. The Y2H experimental assay results also indicated that the C-terminal region was also not involved in the complex formation. This region is generally responsible for hetero-complex formation between different MADS-box proteins (Su et al., 2008; Immink et al., 2009). Therefore, the results suggest that the SEP3 C-terminus might be the glue that connects other proteins. Thus, deletion of the C-terminus amino acids will not change the oligomeric status, and the observed tetrameric property of SEP3B and SEP3C is intrinsic to the SEP3 homo-complex. However, a dimer complex was observed in the crystallized construct (75-178 aa) (Puranik et al., 2014), but not in SEP3C (78-190

aa, Figure 2). This implies that residues between 178 and 190 contribute to tetramer stabilization.

In short, our analysis demonstrated that SEP3 form homo-tetramers, mainly through K domain association. Neither the DNA-binding MADS-box nor the C-terminal region is required for the polymer formation. It is possible that the tetramer shape a platform to support the roles of N- and C-terminal domains in the development of flower meristem.
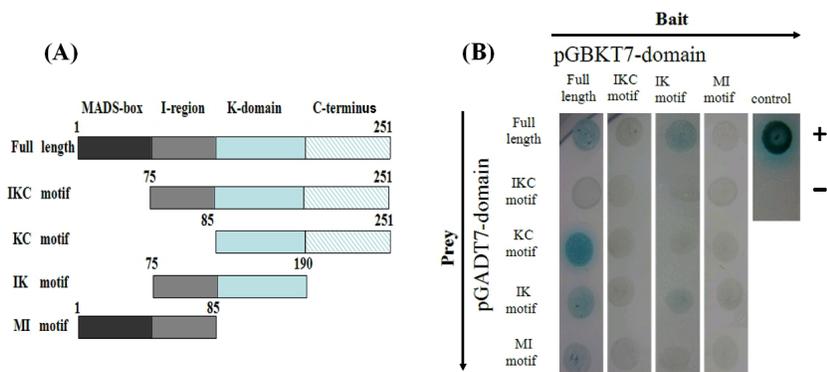


**Figure 4.** Yeast two-hybrid (Y2H) assay. **A.** Domain design. **B.** Y2H results. Schematic maps of bait and prey constructs used in the Y2H assay are shown. Hybridization results are shown on SD/-Leu/-Trp/-His/-Ade+X-a-Gal. Control "+" refers to the positive control group: pGADT7-T and pGBKT7-53 hybridization. "-" refers to the negative control group: pGADT7-T and pGBKT7-lam.

## Conflicts of interest

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

## REFERENCES

Agrawal GK, Abe K, Yamazaki M, Miyao A, et al. (2005). 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*. 59: 125-135.

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

de Folter S, Immink RG, Kieffer M, Parenicova L, et al. (2005). Comprehensive interaction map of the *Arabidopsis* MADS Box transcription factors. *Plant Cell* 17: 1424-1433.

Huang H, Tudor M, Su T, Zhang Y, et al. (1996). DNA binding properties of two *Arabidopsis* MADS domain proteins: binding consensus and dimer formation. *Plant Cell* 8: 81-94.

Hwan Lee J, Joon Kim J and Ahn JH (2012). Role of *SEPALLATA3* (SEP3) as a downstream gene of miR156-SPL3-FT circuitry in ambient temperature-responsive flowering. *Plant Signal Behav.* 7: 1151-1154.

Immink RG, Tonaco IA, de Folter S, Shchennikova A, et al. (2009). *SEPALLATA3*: the 'glue' for MADS box transcription factor complex formation. *Genome Biol.* 10: R24.

Jetha K, Theissen G and Melzer R (2014). *Arabidopsis* SEPALLATA proteins differ in cooperative DNA-binding during the formation of floral quartet-like complexes. *Nucleic Acids Res*. 42: 10927-10942.

Kaufmann K, Muino JM, Jauregui R, Airoldi CA, et al. (2009). Target genes of the MADS transcription factor *SEPALLATA3*: integration of developmental and hormonal pathways in the *Arabidopsis* flower. *PLoS Biol*. 7: e1000090.

Krizek BA and Meyerowitz EM (1996). Mapping the protein regions responsible for the functional specificities of the *Arabidopsis* MADS domain organ-identity proteins. *Proc. Natl. Acad. Sci. USA* 93: 4063-4070.

Leseberg CH, Eissler CL, Wang X, Johns MA, et al. (2008). Interaction study of MADS-domain proteins in tomato. *J. Exp. Bot.* 59: 2253-2265.

Liu D, Wang D, Qin Z, Zhang D, et al. (2014). The SEPALLATA MADS-box protein SLMBP21 forms protein complexes with JOINTLESS and MACROCALYX as a transcription activator for development of the tomato flower abscission zone. *Plant J.* 77: 284-296.

Melzer R, Verelst W and Theissen G (2009). The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in 'floral quartet'-like complexes *in vitro*. *Nucleic Acids Res*. 37: 144-157.

Mendes MA, Guerra RF, Berns MC, Manzo C, et al. (2013). MADS domain transcription factors mediate short-range DNA looping that is essential for target gene expression in *Arabidopsis*. *Plant Cell* 25: 2560-2572.

Pelaz S, Ditta GS, Baumann E, Wisman E, et al. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* 405: 200-203.

Pelaz S, Gustafson-Brown C, Kohalmi SE, Crosby WL, et al. (2001). *APETALA1* and *SEPALLATA3* interact to promote flower development. *Plant J*. 26: 385-394.

Puranik S, Acajjaoui S, Conn S, Costa L, et al. (2014). Structural basis for the oligomerization of the MADS domain transcription factor *SEPALLATA3* in *Arabidopsis*. *Plant Cell* 26: 3603-3615.

Su K, Zhao S, Shan H, Kong H, et al. (2008). The MIK region rather than the C-terminal domain of *AP3*-like class B floral homeotic proteins determines functional specificity in the development and evolution of petals. *New Phytol.* 178: 544-558.

West AG, Causier BE, Davies B and Sharrocks AD (1998). DNA binding and dimerisation determinants of *Antirrhinum majus* MADS-box transcription factors. *Nucleic Acids Res.* 26: 5277-5287.

West AG, Shore P and Sharrocks AD (1997). DNA binding by MADS-box transcription factors: a molecular mechanism for differential DNA bending. *Mol. Cell Biol.* 17: 2876-2887.

Yang Y and Jack T (2004). Defining subdomains of the K domain important for protein-protein interactions of plant MADS proteins. *Plant Mol. Biol*. 55: 45-59.

Zhou Jiaping LX and Yingwu X (2014). Expression of *Arabidopsis thaliana* SEPALLATA3 protein in a prokaryotic system. *J. Zhejiang AF Univ*. 31: 5.