國立台灣大學生命科學院生命科學系

碩士論文

Department of Life Science

College of Life Science

National Taiwan University

Master Thesis

Transcriptional Factors Responsive to CYCLOIDEA in zygomorphic flower of *Sinningia speciosa*

大岩桐兩側對稱花中受 CYCLOIDEA 調控之轉錄因子

王佩琦

Jocelin Muliawan

指導教授：王俊能 博士

Advisor：Chun-Neng Wang, Ph.D

中華民國 108 年 06 月

June 2019

doi:10.6342/NTU201901184

國立臺灣大學碩士學位論文
口試委員會審定書

Transcriptional Factors Responsive to CYCLOIDEA in zygomorphic flower of *Sinningia speciosa*

大岩桐兩側對稱花中受CYCLOIDEA調控之轉錄因子

本論文係王佩琦君（R06B21035）在國立臺灣大學生命科學系所完成之碩士學位論文，於民國一百零八年六月十四日承下列考試委員審查通過及口試及格，特此證明

口試委員：

國立台灣大學 生命科學系	王俊能 博士	王俊能
國立台灣大學 生物科技研究所	陳仁治 博士	陳仁治
國立中興大學 生命科學系	王隆祺 博士	王隆祺
中央研究院 農生中心	林崇熙 博士	林崇熙
國立成功大學 熱帶植物科學研究所	蔡文杰 博士	蔡文杰

生命科學系 系主任 黃偉邦 （簽名）

doi:10.6342/NTU201901184

中文摘要

兩側對稱性花被認定是被子植物演化的主要趨勢，其花從正面可畫出單一個對稱軸，將花分成兩個鏡像半部，背側，兩側和腹側花瓣沿著此對稱軸排列。兩側對稱花使傳粉者從固定的角度進入花中，以促進精確的花粉傳播和柱頭接收，從而大大提高繁殖成功率。在金魚草中，TCP 轉錄因子 *CYCLOIDEA*（*CYC*）在侷限在背部花瓣上表現，*CYC* 透過調節細胞增殖和細胞延長的作用，促使背側花瓣發育，使其在外型上與兩側及腹側花瓣相異。然而，CYC 啟動了那些下游基因，以及它們如何合作以產生背部辨識的花瓣形狀和大小是未知的。野生型大岩桐(*Sinningia speciosa*)為兩側對稱花朵，然而在人為栽培的大岩桐中，兩側對稱卻可輕易地轉換成輻射對稱，這說明了花對稱的發育模組可能是很容易改變。

為了找出 *CYC* 可能的下游基因，我們從大岩桐 'Espirito Santo'（SsES）的轉錄組(RNA-seq)中篩選出背腹側瓣之間的差異性表達的轉錄因子（DE-TFs）。其中，篩出 9 個背側高表達的轉錄因子(包括 *SsCYC*)，其 5 端調節區(regulatory region)都有鑑定出 TCP 結合位點，同時也透過 qRT-PCR 再次驗證這 9 個轉錄因子確實侷限在背側花瓣表現，因此，這 9 個轉錄因子很有可能就是 *SsCYC* 的下游基因。為了證明 SsCYC 對這九個轉錄因子的調節能力，在煙草(*Nicotiana benthamiana*)原生質體的暫時性表達系統中，以雙熒光素酶測定檢測 SsCYC 和報告子（候選 TF 的 5 端調節區）之間的相互作用。結果發現，SsCYC 能夠自我調節，並且活化 *RADIALIS-like*（*SsRL2*）基因，該基因是金魚草中 *RADIALIS* 的直系同源基因，但其功能尚不清楚。有趣的是，SsCYC 還活化乙烯反應轉錄激活因子 *SsERF1* 並抑制乙烯反應轉錄抑制因子 *SsERF3* 和 ovate 家族轉錄抑制因子 SsOFP6，其功能目前也尚未知。

SsERF1 和 *SsERF3* 的可以調控乙烯信號傳導途徑的下游基因。它們可能透過調控 *EXPANXIN*（*EXPA*）基因、木葡聚醣內轉葡糖基酶/水解酶（xyloglucan endotransglucosylase/hydrolase）基因和內切-1,4-β-D-葡聚醣酶（EGase）基因來使細

doi:10.6342/NTU201901184

胞壁變的鬆散，進而改變背側花瓣細胞的延長。同時，這三個基因也在大岩桐轉錄組中被鑑定為背側表達基因，這也符合我們在大岩桐中觀察到背側花瓣的細胞有較大的細胞面積，因此背側花瓣相較於腹側花瓣長度較長，這也被認為是大岩桐花發育成兩側對稱的原因之一。

關鍵詞：大岩桐; 兩側對稱性; SsCYC; 5 端調節區; TCP 結合位點;下游轉錄因子; 細胞延長

doi:10.6342/NTU201901184

Abstract

Floral zygomorphy (bilateral symmetry), in which the dorsal, lateral and ventral petals are arranged along a single plane, dividing flower into two mirror-image halves, has been selected as the major trend in angiosperm evolution. Zygomorphic flowers allow the pollinators to enter the flower in fixed angle to facilitate exact pollen deposition and stigma reception, thus greatly enhance reproductive success. In *Antirrhinum*, TCP transcription factor, *CYCLOIDEA* (*CYC*) is strictly expressed at the dorsal petals and it can function to regulate cell proliferation and expansion for generating dorsal identity. However, what the downstream of CYC are and how they cooperate to generate the petal shape and size for the dorsal identity are largely unknown. The wild type *Sinningia speciosa* exhibits zygomorphic symmetry, yet reversal to actinomorphic (radial symmetry) is common, indicating that the developmental module for floral zygomorphy might be easily altered.

In order to discover CYC downstream, differentially expressed transcription factors (DE-TFs) between dorsi-ventral petals were screened from the RNA-seq data of *S. speciosa* 'Espirito Santo' (SsES). Among them, nine TFs, including *SsCYC* itself, have their 5' regulatory regions been identified with TCP binding sites and their dorsal restricted expression was confirmed by qRT-PCR. To demonstrate the possible regulation

doi:10.6342/NTU201901184

of SsCYC on these TFs, dual-luciferase assay transiently expressed in protoplasts of *Nicotiana benthamiana* leaves was used to examine the interaction between the effector (SsCYC) and the reporter (5' regulatory region of the candidate TFs). It was found that SsCYC was able to auto-regulate itself and also upregulate a *RADIALIS-like* (*SsRL2*) gene which is the orthologue of *RADIALIS* in *Antirrhinum*, but its function is unknown. Interestingly, SsCYC also up-regulated the ethylene response transcriptional activator, *SsERF1* and down-regulated the ethylene response transcriptional repressor, *SsERF3* and an ovate family transcriptional repressor, *SsOFP6* whose function is unknown.

The finding of *SsERF1* and *SsERF3* as SsCYC responsive TFs could be linked to their function as downstream regulators of ethylene signaling pathway. They might alter dorsal cell expansion via regulation of *EXPANXIN* (*EXPA*) genes, xyloglucan endotransglucosylase/hydrolase (XTH) encoding gene and endo-1,4-β-D-glucanase (EGase) encoding gene to loosen the cell wall, since these three genes were identified as the dorsal expressed genes in the RNA-seq data of SsES. This suggestion is also reflected by the observation that the dorsal petals of SsES have larger cell area, thus are longer in length compared to the ventral petals, which is considered as one of the factors that generates floral zygomophy in this flower.

KEYWORDS: *Sinningia speciosa*; floral zygomorphy; SsCYC; 5' regulatory region; TCP binding sites; downstream TFs; cell expansion

doi:10.6342/NTU201901184

Table of Contents

doi:10.6342/NTU201901184

List of Tables

doi:10.6342/NTU201901184

List of Figures

doi:10.6342/NTU201901184

List of Supplementary Data

doi:10.6342/NTU201901184

doi:10.6342/NTU201901184

Abbreviation

AD primer	Arbitrary degenerate primer
bHLH	basic helix-loop-helix
C:I	Choloroform : Isoamyl alcohol
CIB	*Cryptochrome 2-interacting bHLH*
CTAB	Hexadecyl trimethyl-ammonium bromide
CYC	*CYCLOIDEA*
DEGs	Differentially expressed genes
DE-TF	Differentially expressed TF
DICH	*DICHOTOMA*
DIV	*DIVARICATA*
DRIFs	DIV-and-RAD-interacting-factors
EGase	Endo-1,4-β-D-glucanase
ERF	*Ethylene response factor*
EXPA	*EXPANSIN*
GFP	*Green Fluorescence Protein*
GS Primer	Gene specific primer
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LB	Luria Bertani
NaOAc	Acetic acid sodium salt
NGAL	*NGATHA-Like*
OFP	*OVATE FAMILY PROTEINS-like*
P:C:I	Phenol : Choloroform : Isoamyl alcohol
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PVPP	Polyvinylpolypyrrolidone
RAD	*RADIALIS*
RL	*RADIALIS-like 2*
SEFA-PCR	Self-Formed Adaptor PCR
SP primer	Specific primer
SsA	*S. speciosa* ‘Avanti’
SsAN	*S. speciosa* ‘Avenida Niemeyer’
SsES	*S. speciosa* ‘Espirito Santo’
SsPF	*S. speciosa* ‘Pink Flower’
TAIL-PCR	Thermal Asymmetric Interlaced PCR

doi:10.6342/NTU201901184

TCP	*TEOSINTE BRANCHED1*, *CYCLOIDEA* and *PROLIFERATING CELL FACTORS*
TF	Transcriptional Factor
TR	Transcriptional regulator
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside
XTH	Xyloglucan endotransglucosylase/hydrolase

doi:10.6342/NTU201901184

Introduction

Floral symmetry has been considered as the important feature that influences the interaction between plant and pollinator. Generally, there are two main types of floral symmetry, which are zygomorphic (bilateral/mono-symmetry) and actinomorphic (radial/poly-symmetry) that usually could be determined by face-on view of flower perianth. Zygomorphic flowers are characterized by having the dorsal, lateral and ventral petals arranged along a single plane, dividing flower into two mirror-image halves (one dividing plane) whereas the actinomorphic flowers have their perianth arranged into more than one dividing planes. The emergence of zygomorphic symmetry from its actinomorphic ancestral has been correlated with plant-pollinator specific interaction (Spencer and Kim, 2018; Hileman 2014).

The complexity of the floral image in zygomorphic flowers improves the pollinators recognition and discrimination, by limiting the pollination to particular species, preventing the inefficient pollinating species. This restriction then results in reproductive barriers that lead to speciation in both plants and pollinators, often suggested as plant-pollinator co-evolution. While the pollinators of actinomorphic flowers may approach the flowers from any direction, zygomorphic flowers provide these visitors additional horizontal/vertical information which increases the precision of pollen placement on, and

doi:10.6342/NTU201901184

stigma contact with, the pollinator's body. This precision thus results in a higher proportion of pollen reaching the stigma. Therefore, floral zygomorphy provides more efficient pollination, which is then suggested as a reproductive advantage during the angiosperm evolution. Although the zygomorphic flowers have evolved many times from the actinomorphic ancestors, the reversals to actinomorphic have also been observed. This suggests that the developmental module for floral zygomorphy might be easily altered (Neal et al., 1998; Spencer and Kim, 2018).

The molecular mechanism underlying the floral zygomorphy is centered on *CYCLOIDEA* (*CYC*) dorsi-ventral asymmetric expression. *CYC* is belong to TCP transcription factor family, in ECE-CYC2 clade. This TF family is characterized by the amino acid basic helix–loop–helix (bHLH) motif in its encoded proteins. The TCP is named after *TEOSINTE BRANCHED1* (*TB1*) from maize (*Zea mays*), *CYCLOIDEA* (*CYC*) from snapdragon (*A. majus*), and *PROLIFERATING CELL FACTORS 1* and *2* (*PCF1* and *PCF2*) from rice (*Oryza sativa*). Based on the differences within the TCP domain, TCP transcription factors are classified into TCP class I that consists of rice PCF proteins and TCP class II that consists of TB1 and CYC proteins. Outside of the TCP domain, there is 18–20 residue arginine-rich motif (the R domain) which is found in some of class II TCPs, but absent in almost all of the class I TCPs. The TCP class II is then further divided into CYC/TB1 (ECE) and CIN clades. The ECE clade

doi:10.6342/NTU201901184

is characterized with glutamic acid-cysteine-glutamic acid motif found between TCP and R domains. Upon duplication, this clade is divided into CYC1, CYC2 and CYC3. The *CYC2* gene group is considered as the major regulator of floral symmetry (Martín-Trillo and Cubas, 2009).

In *Antirrhinum majus*, the dorsal specific expression of *AmCYC* and its close related protein *DICHOTOMA* (*AmDICH*) generates specific dorsal shape and size by regulating the cell proliferation and expansion, and additionally inhibiting the stamens growth. Both of the genes inhibit the expression of ventral determinant gene, *DIVARIVATA* (*AmDIV*). This inhibition is mediated through *RADIALIS* (*AmRAD*), as the AmRAD protein competes with AmDIV for the interaction with DIV-and-RAD-interacting-factors (DRIFs). Interaction of AmDIV with DRIF is important for the activation of genes that are important for ventral identity. In dorsal petal, AmDIV interaction with DRIF is distracted by the presence of AmRAD, thus making the AmDIV become restricted to be only in ventral petals (**Supplementary Fig. S1A**; Spencer and Kim, 2018). The absence of *AmCYC* and *AmDICH* in *A. majus cyc;dich* double mutant causes no restriction of AmDIV to the dorsal area, thus the mutant flowers become ventralized and have actinomorphic appearance. Moreover, *cyc* mutant alone produces semipeloric flowers and the *dich* single mutant only alters dorsal petal shape. As the mutation analysis shows that *AmCYC* has stronger phenotypic effect campared to *AmDICH*, then *CYC* is considered as

doi:10.6342/NTU201901184

the key regulator of floral zygomorphy (Corley et al. 2005; Luo et al. 1999). Besides in *A. majus*, other *CYC2*-like genes in the core eudicots also play the major function in controlling floral zygomorphy due to the strong dorsoventrally asymmetric expression. Species showing dorsal or along with lateral expression of *CYC* generally have zygomorphic flowers, whereas the absence of *CYC* or ubiquitous *CYC* expression in all petals results in actinomorphic flowers in some species. Therefore, the progression of CYC expression (absent- ubiquitous − dorsal/lateral − dorsal) plays an important role during the transition of actinomorphic to zygomorphic, and also its reversal (**Supplementary Fig. S1B-E**; Spencer and Kim, 2018).

The wild type of *Sinningia speciosa* flower exhibits zygomorphic symmetry, while the commercial type has actinomorphic symmetry. The floral zygomorphy in the wild type is regulated by the dorsal specific expression of a single copy of *CYC2*-like gene (**Supplementary Fig. S2**; Ye, 2018, unpublished work), *SsCYC*. The actinomorphic mutant of this flower is caused by 10 bases deletion in this gene, causing it to be inactive (Dong et al., 2018). Since *CYC* acts as the key regulator of floral zygomorphy in *S. speciosa* as if in *A. majus*, then *S. speciosa* could be a comparable model to study the floral zygomorphy regulation. However, the floral zygomorphy regulation of CYC through RAD, as described in *A. majus* is not conserved in all zygomorphic lineages (Baxter et al., 2007; Costa et al., 2005, Hsu et al., 2018) which means that CYC might

doi:10.6342/NTU201901184

regulate other genes for the generation of zygomorphic symmetry. Yet, what these target genes of CYC are and how they cooperate to generate petal shape and size for floral zygomorphy are still largely unknown.

The bHLH domain of TCP TF has the capability for binding to GC rich DNA sequences and also for protein-protein interaction. The basic region of this domain mediates the interaction between the protein and targeted DNA sequences, whereas the HLH region provides protein-protein interaction by forming homo- or hetero-dimer (Atchley and Fitch 1997). Both two classes of TCP TF have distinct but overlapping consensus of DNA binding sequences (TCP binding sites); GGNCCCAC for class I and GTGGNCCC for class II, with GGNCCC serves as the core sequence (Koshugi and Ohashi, 2002). TCP binding sites have been reported to role as the cis-elements that mediate TCP TFs regulation of their targets, such as *AmRAD* (Costa et al., 2005), *CYCLIN* (Li *et al*. 2005), *PCNA* (Kosugi and Ohashi 1997), *LIPOXIGENASE2* (Schommer *et al*. 2008), *CIRCADIAN ASSOCIATED1* (Pruneda-Paz *et al*. 2009), etc. Therefore, the presence of TCP binding sites could be the indicator for determining TCP TFs' targets, including CYC's targets (Koshugi and Ohashi 2002).

The recent study in *S. speciosa* 'Espirito Santo' (SsES) has revealed that there were 630 dorsi-ventral differentially expressed genes (DEGs) (Pan, Z.J., unpublished data). Among these DEGs, there might be some genes that have certain influences in patterning

doi:10.6342/NTU201901184

the floral zygomorphy of *S. speciosa*, including SsCYC downstream. In order to minimize the scope for the screening of SsCYC downstream, this study focused mainly on the TFs activated by SsCYC. TF is known for its effect on a single developmental module which influences only the morphology of a single organ. As the consequence, TF is naturally selected as the source of phenotypic variation. Therefore, mapping SsCYC target TFs will provide a better rationale of how the floral zygomorphy in *S. speciosa* is established. As TCP binding sites serve as the important elements that might mediate CYC regulation, the identification of SsCYC targets from the DE-TFs of SsES relied on the presence of the TCP binding sites at their 5'regulatory regions. In order to narrow down to SsCYC activation target TFs, the identification was focused on the TFs that had similar expression pattern with *SsCYC*, which were the dorsal-expressed TFs. The regulation of these TFs by SsCYC was then demonstrated by dual-luciferase assay, transiently expressed in the protoplasts of *Nicotiana benthamiana* leaves, with SsCYC as the effector and the 5' regulatory region fused with firefly luciferase as the reporter.

Flowers with zygomorphic symmetry often have their petals could be distinguished into dorsal, lateral and ventral parts due to the different shape and size within these regions, which leads to the hypothesis that the petal identity of each region should have some effects to the establishment of floral zygomorphy. Petal identity itself is determined by two factors, which are cell elongation as well as rate and direction of cell division. The

doi:10.6342/NTU201901184

cell elongation in the basal part is important for determining the final size and shape of the petal, while the rate and direction of cell division determine the shape and size of the distal region. This mechanism requires a quite complex hormonal regulation (Irish, 2008; van Es, 2018). Jasmonic acid influences the petal size of *Arabidopsis* through post-transcriptional regulation of BIGPETAL (BPE), TF that regulates cell expansion (Brioudes et al., 2009). Auxin, ethylene and gibberellin also affect cell proliferation and elongation during petal development by integrating in certain TF regulations (Chandler 2011).

In this study, several TFs were found to be responsive to SsCYC. Instead of *SsRAD*, orthologue of *Anthirrhinum RAD*; another *RAD-like* gene (*SsRL2*) whose function was unknown, was identified as SsCYC downstream. Interestingly, two *ethylene responsive factors* (*SsERF*s) were also found to be regulated by SsCYC. ERFs are known as the integral components of signaling cascade that regulate different kinds of downstream genes of various developmental and stress responsive pathways. As the downstream component of ethylene signaling pathway, ERFs also interact with other hormone pathways, such as jasmonic acid, ABA, auxin, salicylic acid, gibberellins, and brassinosteroids (Müller and Munné-Bosch, 2015). Moreover, an ovate family protein (*SsOFP6*) whose overexpression in *Arabidopsis* results in flat, thick and cyan leaves (Wang et al., 2011), enhanced apical dormancy of the plant, was also responsive to

doi:10.6342/NTU201901184

SsCYC. Taken together, these results led to the suggestion that SsCYC might work through these TFs to affect the dorsal petals cell growth of SsES, developing the floral zygomorphy of this flower.

doi:10.6342/NTU201901184

Materials and Methods

Plant material and growth condition

Sinningia speciosa 'Espirito Santo' was obtained from Dr. Cecilia Koo Botanic Conservation Center, Pingtung, Taiwan. The seeds were cultivated under 16/8 hours (day/night) cycle at 24℃ with 70% relative humidity. Floral bud developmental stage was determined based on dorsal corolla tube length. Floral bud stage 5 which has 8-10 mm length of dorsal tube was used for transcription factor (TF) isolation. Dissected dorsal and ventral petals from floral bud stage 5 were used for expression pattern validation of the dorsal expressed TFs. Finally, the leaves were used for 5' regulatory region isolation. All samples were frozen in liquid nitrogen and stored at -80℃.

Prediction of transcription factor

RNA-seq data of *S. speciosa* 'Espirito Santo' floral bud stage 5 was provided and analyzed by Dr. Zhao-Jun, Pan. Based on RNA-seq analysis, 630 genes were found to have dorsi-ventral differential expression (DEGs) (p-value<0.05; log2FC ≥1). In order to find the TFs among these DEGs, TF prediction was performed using iTAK online (v1.6) (http://itak.feilab.net/cgi-bin/itak/online_itak.cgi) for nucleotide sequences. The identified TFs were analyzed using NCBI BLASTX

doi:10.6342/NTU201901184

(https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastx&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome) for annotation.

Prediction of TCP binding site

Since *SsCYC* downstream regulation might be facilitated by the presence of TCP binding sites at the 5'regulatory region of its target genes, screening for the binding sites was done for each of the predicted TF. The 5' regulatory region of each TF was retrieved from *S. speciosa* 'Avenida Niemeyer' draft genome using RStudio software (Version 1.1.463; RStudio Inc., 2009) and Linux Interface (done by Ya-Chi, Nien). TCP binding consensus was summarized from the paper 'TCP Transcription Factors: Evolution, Structure, and Biochemical Function' (González-Grandío and Cubas, 2016) that has compiled most of TCP binding sites found in the *in vitro* and *in vivo* experiments in numerous studies. Screening for the presence of each of the summarized TCP binding consensus (**Supplementary Table S1**) was done for all the predicted regulatory regions using fuzznuc (http://emboss.bioinformatics.nl/cgi-bin/emboss/fuzznuc) for both strands of complementary sequence.

doi:10.6342/NTU201901184

Total RNA extraction and reverse transcription

The total RNA from whole floral bud stage 5 and dissected dorsal and ventral petals of floral bud stage 5 were extracted using Trizol Reagent (Invitrogen, Waltham, MA, USA) according to manufacturer's protocol. The RNA quality was measure using NanoDrop Spectrophotometer. Synthesis of complementary DNA (cDNA) was done using Superscript IV (Invitrogen, Waltham, MA, USA) according to manufacturer's protocol.

Isolation of the transcription factor of *S. speciosa* 'Espirito Santo'

TFs that have been predicted to contain TCP binding sites at their 5' regulatory regions were isolated in order to get their full length coding sequences. The sequence of each TF was amplified with PCR using Phusion® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA) (**Supplementary Table S2**) and the products were purified by gel extraction (Viogene, GP1002), following the manufacturer's protocol. The purified products were proceed to A-tailing in order to increase ligation efficiency. A-tailing was done by adding 0.3 μL of TaKaRa Ex Taq DNA Polymerase (Takara Bio, USA), 3 μL of Ex Tag buffer and 0.6 μL of 2mM dATP into 10 μL of purified product and the mixture was then incubated at 72℃ for 1 hour. The A-tailed products were purified by PCR Clean Up system (Viogene, GP1002) and were ligated to T&A™ cloning vector (Yeastern Biotech Co, Taipei, Taiwan) with following recipe:

doi:10.6342/NTU201901184

Ligation mixture component	
vector: insert molar ratio	1:3
Vector fragments end conc.	3-30 fmol
Insert fragments end conc.	9-90 fmol
10x Ligation Buffer A	2.0 μL
10x Ligation Buffer B	2.0 μL
yT4 DNA ligase	1.0 μL
ddH_2O to final volume of	20 μL

The ligation mixture was incubated overnight. The next day, transformation was done using the heat shock method. About 2 μL of vector containing DNA of interest was mixed with 20 μL of competent cell, *Escherichia coli* HIT-DH5α (Real Biotech Corporation, Taipei, Taiwan) and was chilled on ice for 20 minutes. Then, the mixture was thawed at 42℃ for 1 minute for heat shock and quickly chilled on ice. After heat shock procedure, 50 μL of LB broth was added to the mixture, followed by incubation at 37℃ for 1 hour. The bacterial solution was added with 100 μL of 0.1 M IPTG and 20 μL of 80 mg/mL X-Gal, and spread on LB agar plate contained Ampicillin (100 μg/mL). The plate was incubated for 16-18 hours. Colonies containing the insert were selected by using colony PCR. After confirmation, the colonies containing the correct insertion size of DNA was cultured in 3 mL of LB broth contained Ampicillin (100 μg/mL) by shaking at 37℃ for 16-18 hours. The plasmids were extracted using Mini Plus Plasmid DNA Extraction System (Viogene, GF2002) according to manufacturer protocol and sent to sequencing (Genomics, New Taipei City, Taiwan).

doi:10.6342/NTU201901184

Validation for the expression pattern of the dorsal-expressed TFs of *S. speciosa* 'Espirito Santo'

Quantitative real time PCR (qRT-PCR) analysis was done to validate the RNA-seq data of the dorsal-expressed TFs that have been predicted to contain TCP binding sites at their 5' regulatory regions. qRT-PCR analysis was performed in Bio-Rad PCR machine (CFX-384) using KAPA SYBR® FAST qPCR Master Mix (2X) Kit (KAPA Biosystem, KR0389) (**Supplementary Table S3**). The recipe and program were listed below:

qRT-PCR mixture

Reagent	Volume
ddH_2O	1.0 μL
2x Master Mix	5.0 μL
Forward Primer (1μM)	1.0 μL
Reverse Primer (1μM)	1.0 μL
cDNA (5ng/μL)	2.0 μL
Total Volume	10 μL

Thermal cycle program:

Step	Temperature	Time
1	95 ℃	3 min
2	95 ℃	10 s
3	55-57 ℃	30 s
	Plate read	
4	Go to step 2, 39 cycles	
5	95 ℃	10 s
6	Melt curve 65 to 95℃, increment 0.5	5 s
	Plate read	

doi:10.6342/NTU201901184

After the running of PCR, the obtained data was analyzed using CFX Maestro™ Software for CFX Real-Time PCR Instruments (Version 1.1; Bio-Rad Laboratories Inc, 2017). The expression level of each TF was quantified as relative fold gene expression level ($2^{-\Delta\Delta CT}$), using *18s* as reference gene and ventral petals as the control. The ΔCt was calculated as Ct (dorsal/ventral) – Ct (reference gene) and the ΔΔCt was calculated ΔCt (dorsal petals) – ΔCt (ventral petals).

Genomic DNA Extraction

Genomic DNA (gDNA) extraction was performed with Hexadecyl trimethyl-ammonium bromide (CTAB) method (Doyle, 1990). The collected leaves were homogenized in liquid nitrogen using mortar and pestle. The homogenized tissue was added with 1 mL of CTAB, 20 mg of PVPP and 5 μL of β-mercaptoethanol, proceed by incubation at 65℃ for 30 minutes. Next, the mixture was added with 500 μL of PCI (phenol : choloroform : isoamyl alcohol, 25:24:1, pH = 8.0) and inverted for 15 minutes, followed by centrifugation at 13.000 rpm for 10 minutes. The upper layer of the solution was transferred to the new tube, added with 1 μL RNase A and incubated at 37℃ for 20-30 minutes. The solution was added with 500 μL of C:I (choloform: isoamyl alcohol, 24:1) and inverted for 15 minutes, followed by centrifugation at 13.000 rpm for 10 minutes. The upper layer was transferred to the new tube and added with one to tenth volume of 3

doi:10.6342/NTU201901184

M NaOAc (pH = 5.5), then precipitated with 0.7 volume of isopropanol. The mixture was incubated at -20 ℃ for 1 hour, proceed by centrifugation at 13.000 rpm for 10 minutes. The supernatant was discarded and the pellet was washed by the addition of 1 mL of 70% ethanol and centrifugation at 13.000 for 5 minutes. The supernatant was discarded and the pellet was air dried. Finally, 30-50 μL of ddH_2O was added to dissolve the pellet. The quantity and quality of extracted gDNA was measured with Nanodrop Spectrophotometer. The gDNA was stored at -20 ℃.

CTAB buffer (100 mL)

Reagent	per reaction
Hexadecyl trimethyl-ammonium bromide (CTAB)	2.0 g
1M Tris (pH = 8.0)	10.0 mL
0.5 Ethylenediaminetetraacetic acid (EDTA, pH = 8.0)	4.0 mL
5 M NaCl	28.0 mL
ddH_2O	56.0 mL

The pH was adjusted to 8.0 using NaOH and stored at room temperature

Isolation of the 5' regulatory region of *Sinningia speciosa* 'Espirito Santo'

There were several PCR based approaches used for isolating the 5' regulatory region of each dorsal-high expressed TF. The regulatory region of *Sispe038Scf1202g12026* (*SsOFP6*) was isolated using pair of primers designed directly from the predicted regulatory region of *S. speciosa* 'Avenida Nieyemer' (SsAN). The regulatory region of *Sispe038Scf1400g01001* (*SsCYC*) was isolated with forward primer designed directly from the predicted regulatory region of SsAN and reverse primer

doi:10.6342/NTU201901184

designed at the known coding sequence (CDS) of *S. speciosa* 'Espirito Santo' (SsES). Another two regulatory regions, *Sispe038Scf1061g02075* (*SsERF3*) and *Sispe038Scf2159g01072* (*SsCIB2*), were isolated by nested PCR using two sets of primers. The first set of primers contained the forward primer designed directly from the predicted regulatory region of SsAN and reverse primer designed on the known CDS of SsES. The second set of primers was design to amplify a secondary target within the first run product, thus reducing the non-specific binding in products. For the second round of the nested PCR, the product of the first PCR was diluted to 100 times. All of these three approaches were done with Phusion® High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA) according to manufacturer protocol (**Supplementary Table S4**). The amplified products were continued to cloning, using the same procedure described for transcription factor isolation and then sent to sequencing (Genomics, New Taipei City, Taiwan). Last, the regulatory region of *Sispe038Scf0228g08027* (*SsERF17*) was isolated with Thermal Asymmetric Interlaced PCR (TAIL-PCR), whereas for *Sispe038Scf0170g01016* (*SsRL2*), *Sispe038Scf2996g00029* (*SsNGAL1*) and *Sispe038Scf5680g00016* (*SsERF1*), the regulatory regions were isolated with Self-Formed Adaptor PCR (SEFA-PCR). All the isolated regulatory regions were screened for the presence of TCP binding sites.

doi:10.6342/NTU201901184

Thermal Asymmetric Interlaced PCR (TAIL-PCR)

Thermal Asymmetric Interlaced PCR (TAIL-PCR) is used to amplify the unknown sequence, in this case the regulatory region that is adjacent to the known CDS. It uses two sets of primers which are the gene-specific primers (GS primers) that usually have high melting temperatures and arbitrary degenerate primers (AD primers) (**Supplementary Table S5**) that usually have low melting temperatures. By using the combination of these primers, amplification of the expected sequence could be done from the known end and the unknown end, respectively. Specificity is obtained through subsequent rounds of TAIL-PCR, using nested gene-specific primers and alternate of high and low annealing temperatures cycles (**Supplementary Fig. S3a**). The TAIL-PCR used in this study was referred from Liu et al. (1995) and Liu and Whittier (1995) with modifications. The AD primers were adopted from Singer and Burke (2003). The recipe and program of TAIL-PCR were listed below:

Single reaction for primary TAIL-PCR

Reagent	Volume
Phusion DNA polymerase (0.02 units/μL)	0.1 μL
5X Phusion HF or GC Buffer	2.0 μL
10 mM dNTPs	0.2 μL
6 x AD primer	2.0 μL
10 μM GS1 primer	0.5 μL
gDNA (20 ng/μL)	0.5 μL
ddH_2O	add to 10 μL

doi:10.6342/NTU201901184

Thermal cycle for primary TAIL-PCR

Step	Temperature	Time
1	94 ℃	2 min
2	94 ℃	30 s
3	62 ℃	1 min
4	72 ℃	2.5 min
5	Go to step 2 for 4 cycles	
6	94 ℃	30 s
7	25 ℃	3 min
8	Ramping from 25 to 72 ℃ (rate = 0.3℃/sec)	
9	72 ℃	2.5 min
10	94 ℃	10 s
11	68 ℃	1 min
12	72 ℃	2.5 min
13	94 ℃	10 s
14	68 ℃	1 min
15	72 ℃	2.5 min
16	94 ℃	10 s
17	44 ℃	1 min
18	72 ℃	2.5 min
19	Go to step 10 for 14 cycles	
20	72 ℃	2.5 min

Single reaction for secondary TAIL-PCR

Reagent	Volume
Phusion DNA polymerase (0.02 units/μL)	0.1 μL
5X Phusion HF or GC Buffer	2.0 μL
10 mM dNTPs	0.2 μL
6 x AD primer	2.0 μL
10 μM GS1 primer	0.5 μL
1:1000 diluted 1st reaction	0.5 μL
ddH_2O	add to 10 μL

doi:10.6342/NTU201901184

Thermal cycle for secondary TAIL-PCR

Step	Temperature	Time
1	94 ℃	10 s
2	68 ℃	1 min
3	72 ℃	2.5 min
4	94 ℃	10 s
5	68 ℃	1 min
6	72 ℃	2.5 min
7	94 ℃	10 s
8	44 ℃	1 min
9	72 ℃	2.5 min
10	Go to step 1 for 11 cycles	
11	72 ℃	5 min

Single reaction for tertiary TAIL-PCR

Reagent	Volume
Phusion DNA polymerase (0.02 units/μL)	0.1 μL
5X Phusion HF or GC Buffer	2.0 μL
10 mM dNTPs	0.2 μL
6 x AD primer	2.0 μL
10 μM GS1 primer	0.5 μL
1:1000 diluted 3rd reaction	0.5 μL
ddH_2O	add to 10 μL

Thermal cycle for tertiary TAIL-PCR

Step	Temperature	Time
1	94 ℃	15 s
2	44 ℃	1 min
3	72 ℃	2.5 min
4	Go to step 1 for 19 cycles	
5	72 ℃	5 min

After the 3rd round of PCR, the products from 1st, 2nd and 3rd PCR were run together in gel electrophoresis. The product from the 1st round might contain the non-specific

doi:10.6342/NTU201901184

products which could be seen by the smear appearance on the gel. The expected specific products could usually be observed from the product of 2nd and 3rd round, with the 3rd round product having slight decreased in size. The largest band from the 3rd round product was isolated and continued to cloning, using the same procedure described for transcription factor isolation, then sent to sequencing (Genomics, New Taipei City, Taiwan) (**Supplementary Fig. S3b**).

Self-Formed Adaptor PCR (SEFA-PCR)

Self-Formed Adaptor PCR (SEFA-PCR) is developed to overcome the drawbacks of TAIL-PCR, which is the product is usually less than 1.0 kb. It combines the advantages of ligation-mediated PCR in its specificity and of TAIL PCR in its simplicity. It uses four primers that are located sequentially on the known DNA sequences. SP1, SP2, and SP4 are the specific primers designed from the known region and have relatively high annealing temperatures (e.g., 70°C), whereas SP3 (e.g., 5′-TACCCAAAGAAGCAGGAANNNNNNNGTGAAA-3′) is a partially degenerate primer which plays the key role in the process. First, a single cycle of PCR was carried out at a low annealing temperature (e.g., 35°C) with only primer SP3. At this low annealing temperature, SP3 can prime and elongate at many positions on the DNA template. A position probably exists somewhere downstream of the known DNA sequence

doi:10.6342/NTU201901184

where SP3 primes and extends, thus creating a nascent single strand which has a binding site for SP1. After a single cycle of PCR, the annealing temperature is increased to the point (e.g., 70°C) corresponding to the annealing temperature of SP1. Then, SP1 is added to the reaction mixture. At this high annealing temperature, only SP1 can prime the target site efficiently, thus creating a pool of single-stranded DNA with the SP1 sequence at the 3' end and the SP3 complementary sequence at the 5' end. Finally, several cycles of a low annealing temperature (e.g., 55°C) are performed to facilitate the loop-back extension, thus creating an adaptor which contains binding sites for SP1 and SP2. Once the adaptor has been created, the target sequences can be amplified efficiently by SP1. After SEFA PCR, a second round of nested PCR was run with the single primer SP2. A third round of thermally asymmetric PCR was run to improve the specificity with primer SP4 (e.g., annealing at 70°C) and the other short primer, SP5 (e.g., annealing at 60°C), positioned between SP2 and SP3 (**Supplementary Table S6; Supplementary Fig. S4a**). The SEFA-PCR used in this study was adopted from Wang et al. (2007) with modifications. The recipe and program were listed below:

doi:10.6342/NTU201901184

Single reaction for primary SEFA-PCR

Reagent	Volume
Phusion DNA polymerase (0.02 units/μL)	0.2 μL
5X Phusion HF or GC Buffer	4.0 μL
10 mM dNTPs	0.4 μL
5 μM SP3	1.0 μL
gDNA (1000 ng/μL)	1.0 μL
ddH_2O	add to 20 μL

Thermal cycle for primary SEFA-PCR

Step	Temperature	Time
1	98 ℃	30 s
2	35 ℃	3 min
3	Ramping from 35 to 70 ℃(rate = 0.2℃/sec)	
4	Add 3 μl of 5 μM SP1	
5	98 ℃	10 s
6	70 ℃	3 min
7	Go to step 5 for 24 cycles	
8	98 ℃	10 s
9	70 ℃	3 min
10	98 ℃	10 s
11	70 ℃	3 min
12	98 ℃	10 s
13	65 ℃	30 s
14	70 ℃	3 min
15	Go to step 8 for 10 cycles	
16	25 ℃	10 s

Single reaction for secondary SEFA-PCR

Reagent	Volume
Phusion DNA polymerase (0.02 units/μL)	0.1 μL
5X Phusion HF or GC Buffer	2.0 μL
10 mM dNTPs	0.2 μL
5 uM SP2	3.0 μL
1:10 diluted 1^{st} reaction	0.5 μL
ddH_2O	add to 10 μL

doi:10.6342/NTU201901184

Thermal cycle for secondary SEFA-PCR

Step	Temperature	Time
1	98 ℃	30 s
2	98 ℃	10 s
3	70 ℃	3 min
4	Go to step 2 for 29 cycles	
5	25 ℃	10 s

Single reaction for tertiary SEFA-PCR

Reagent	Volume
Phusion DNA polymerase (0.02 units/μL)	0.1 μL
5X Phusion HF or GC Buffer	2.0 μL
10 mM dNTPs	0.2 μL
5 uM SP4	3.0 μL
5 uM SP5	0.3 μL
1:10 diluted 1st reaction	0.5 μL
ddH_2O	add to 20 μL

Thermal cycle for tertiary SEFA-PCR

Step	Temperature	Time
1	98 ℃	10 s
2	70 ℃	3 min
3	98 ℃	10 s
4	70 ℃	3 min
5	98 ℃	10 s
6	65 ℃	30 s
7	70 ℃	3 min
8	Go to step 1 for 9 cycles	
9	25 ℃	10 s

After the 3rd round of PCR the products from 1st, 2nd and 3rd PCR were run together in gel electrophoresis. The 1st and 2nd might contain some non-specific products with low molecular weight, and the desired product is usually expected to be seen in the 3rd product.

doi:10.6342/NTU201901184

Therefore, the largest band from 3rd product was isolated and sent to sequencing (Genomics, New Taipei City, Taiwan) (**Supplementary Fig. S4b**). New forward primers were design to amplify the desired regulatory region paired with SP4 primers, using the same procedure as described in the transcription factor isolation (**Supplementary Table S4)**.

Vector construction for dual-luciferase assay

The PJD301-firefly driven by the 5' regulatory region of interest was used as the reporter (**Supplementary Fig. S5a**), whereas PJD301-renilla driven by 35s promoter was used as the internal control to normalized the transfection variability (**Supplementary Fig. S5b**) (Luehresen et al., 1995). The vector expressing *SsCYC* tagged with *GFP* was served as the effector for the tested group (**Supplementary Fig. S6**), whereas vector expressing only *GFP* without *SsCYC* was used as effector for the control group (**Supplementary Fig. S7**).The isolated regulatory region sequence of *SsRL1*, *SsERF17*, *SsOFP6*, *SsCYC*, *SsCIB2*, and *SsNGAL1* were amplified using PCR and cloned into the BamHI and SalI restriction sites of the PJD301-firefly, whereas *SsERF3* and *SsERF1* were amplified by PCR to add HincII and NCO1 restriction site and cloned into the AfeI and NcoI restriction site of the vector (**Supplementary Table S7**). The general recipe for enzyme digestion was described as below:

doi:10.6342/NTU201901184

Recipe for BamHI and Sal1 digestion

Reagent	Volume
BamHI buffer	10.0 μL
BamHI (10 U/μL)	2.5 μL
SalI (10 U/μL)	5.0 μL
DNA	400-500 ng
ddH_2O	add to 100 μL

Recipe for HincII and NcoI digestion

Reagent	Volume
1X Tango Buffer	10.0 μL
HincII (10 U/μL)	5.0 μL
NcoI (10 U/μL)	5.0 μL
DNA	400-500 ng
ddH_2O	add to 100 μL

Recipe for Afe1 and NcoI digestion

Reagent	Volume
2X Tango Buffer	20.0 μL
HincII (10 U/μL)	2.5 μL
NcoI (10 U/μL)	2.5 μL
DNA	400-500 ng
ddH_2O	add to 100 μL

The reaction mixtures were incubated overnight and the desired digestion products were purified by gel purification (Viogene, GP1002), following the manufacturer's protocol. The purified products were ligated to the PJD301-firefly vector following the recipe described below:

doi:10.6342/NTU201901184

Ligation reaction of PJD-firefly with the desired digestion product

PJD301: insert molar ratio	1:3
PJD301 fragments end conc.	3-30 fmol
Insert fragments end conc.	9-90 fmol
10x Ligation Buffer A	2.0 μL
10x Ligation Buffer B	2.0 μL
yT4 DNA ligase	1.0 μL
ddH_2O to final volume of	20 μL

The ligation mixture was incubated overnight. The next day, transformation was done using the heat shock method into the *Escherichia coli* HIT-DH5α (Real Biotech Corporation, Taipei, Taiwan). Amipicillin (100 μg/mL) plate was used as the selection medium. After 16-18 hours of incubation colony PCR was done to select the colony carrying the vector of interest. The colony that has been confirmed to carry the desired vector was cultured into LB contained Ampicillin (100 μg/mL) for maxi plasmid extraction (Viogene, GMV2002).

Protoplast isolation

Protoplast isolation was done according to '*Arabidopsis* mesophyll protoplasts protocol' (Yoo et al., 2007) with modifications. *Nicotiana benthamiana* leaves were used as the source of protoplasts instead of *Arabidospsis*. The plants were grown under 16/8 hours (day/night) cycle at 27°C with 70% relative humidity. The leaves from 4-5 weeks-old-plant were chosen and cut into 0.5–1-mm strips from the middle part of a leaf using

doi:10.6342/NTU201901184

a fresh sharp razor blade. The cut leaves were transferred into the prepared enzyme solution and digested in the dark for 3 hours at room temperature. After digestion, the solution was diluted with an equal volume of W5 solution. The enzyme solution containing the protoplasts was filtered through 75-μm nylon mesh into round-bottom tube. The filtered solution was then centrifuged at 100 g for 2 minutes. The supernatant was removed and the protoplasts were re-suspended with W5 solution at 2×10^5 ml^{-1} after counting cells under the microscope (× 100) using a hemacytometer. The protoplasts were rested on ice for 30 minutes. After 30 minutes, the W5 solution was removed and the protoplasts were re-suspended in MMG solution 2×10^5 ml^{-1}.

Protoplast DNA-PEG–calcium transfection

The protoplast transfection of vector mixture containing effector, reporter and internal control was also performed following the method described in '*Arabidopsis* mesophyll protoplasts protocol' (Yoo *et al*., 2007). About 10 μL of vector mixture (the amount of each vector was 10 μg in 10 μL) was added into a 2-ml microfuge tube, followed by 100 μl protoplasts (2×10^4 protoplasts), then the mixture was mixed gently. About 110 μl of PEG solution was added and the mixture was mixed gently by tapping the tube. The transfection mixture was incubated for 15 minutes at room temperature. After incubation, the mixture was diluted with 400 μl W5 solution and mixed gently

doi:10.6342/NTU201901184

by inverting. The mixture was centrifuge at 100 g for 2 minutes and the supernatant was removed. The protoplasts were re-suspended in 0.5 mL of WI solution in each well of a 12-well tissue culture plate. Incubation was done for 16 hours.

Experimental design for SsCYC and dorsal-expressed TFs interaction analysis

	Test	Control
Vector Mixture	SsCYC-GFP effector 5'regulatory region-PJD301 Firefly PJD301 Renilla	GFP effector 5'regulatory region-PJD301 Firefly PJD301 Renilla

Dual-luciferase assay.

After 16 hours of incubation, the transfected protoplasts were collected by moving them to 2 mL microfuge tube, followed by centrifugation at 100 g for 2 minutes and the supernatant was removed. The dual-luciferase assay was done in 96-well white flat bottom plate according to the instruction of Dual-Luciferase® Reporter Assay System for product E1960 (Promega Corporation, USA). About 20 µl passive lysis buffer was added into the protoplasts and the mixture was transferred into the well of the plate. After 5 minutes, 100 µl of LAR II reagent was added into the mixture and the firefly luciferase activity was measured by luminometer by 10s measurement using i-control™ Microplate Reader Software (Version 1.8; Tecan, 2011). Then, 100 µL of Stop & Glo® was added and the renilla luciferase activity was measured by 10s

doi:10.6342/NTU201901184

measurement. All reactions were run triplicate. The interaction of SsCYC and its downstream target was determined as normalized fold change (Δ fold activity) by calculating the firefly to renilla activity ratio of the tested group divided to the control group. One-Way Analysis of Variance was used to assess the up/down-regulation significance level, using One-Way Analysis of Variance Calculator (https://goodcalculators.com/one-way-anova-calculator/).

doi:10.6342/NTU201901184

Results

34 Transcription factors were predicted among 630 dorsi-ventral DEGs

The RNA-seq data has shown that there were 630 dorsi-ventral DEGs of *S. speciosa* 'Espirito Santo'. In order to screen for the TFs among these DEGs, iTAK was used as the identification and classification tool. Around 34 TFs were identified; 17 of them were the dorsal-expressed TFs and the others 17 were the ventral expressed TFs. Based on NCBI BLASTX analysis, *CYC* (*SsCYC*) which was previously known as the major regulator of floral zygomorphy of *A. majus* was identified in the dorsal-expressed TF group (**Table 1**).

doi:10.6342/NTU201901184

Table 1 The list of 34 TFs identified from 630 DEGs and their BLASTX annotations

Gene ID	Family Name	BLASTX Annotation	Expression
Sispe038Scf0044g00001	bZIP	*bZIP transcription factor 46-like*	ventral
Sispe038Scf0056g05037	MADS->MADS-MIKC	*SEPALLATA 1*	ventral
Sispe038Scf0116g02021	Tify	*protein TIFY 10A*	dorsal
Sispe038Scf0146g00043	WRKY	*probable WRKY transcription factor 14*	ventral
Sispe038Scf0163g00025	B3->B3	*transcription repressor VAL1-like*	ventral
Sispe038Scf0170g01016	MYB->MYB-related	*protein RADIALIS-like 3*	dorsal
Sispe038Scf0228g08007	Tify	*protein TIFY 10A-like*	dorsal
Sispe038Scf0228g08027	AP2/ERF->AP2/ERF-ERF	*ethylene-responsive transcription factor ERF017-like*	dorsal
Sispe038Scf0247g02018	C2C2->C2C2-CO-like	*zinc finger protein CONSTANS-LIKE 16*	ventral
Sispe038Scf0266g00013	HB->HB-HD-ZIP	*homeobox-leucine zipper protein ATHB-13*	ventral
Sispe038Scf0367g01001	MYB->MYB	*transcription factor MYBS1*	ventral
Sispe038Scf0439g00009	EIL	*ethylene-insensitive protein 3*	ventral
Sispe038Scf0608g04048	B3->B3-ARF	*auxin response factor 18-like*	ventral
Sispe038Scf0757g01046	MYB->MYB-related	*protein REVEILLE 1*	ventral
Sispe038Scf1061g02075	AP2/ERF->AP2/ERF-ERF	*ethylene-responsive transcription factor 14*	dorsal
Sispe038Scf1077g00028	MADS->MADS-MIKC	*MADS-box transcription factor 6*	ventral
Sispe038Scf1202g12026	OFP	*transcription repressor OFP6-like*	dorsal
Sispe038Scf1202g13005	B3->B3	*transcription repressor VAL1-like*	dorsal
Sispe038Scf1393g02049	WRKY	*WRKY transcription factor 26*	dorsal
Sispe038Scf1400g01001	TCP	*CYC*	dorsal
Sispe038Scf1614g02066	MADS->MADS-MIKC	*APETALA 1*	ventral
Sispe038Scf1651g00049	MYB->MYB	*transcription factor MYB14*	dorsal
Sispe038Scf1783g02026	AP2/ERF->AP2/ERF-ERF	*ethylene-responsive transcription factor 1A-like*	dorsal
Sispe038Scf1947g02019	HB->HB-HD-ZIP	*homeobox-leucine zipper protein ATHB-13*	ventral
Sispe038Scf1948g00046	AP2/ERF->AP2/ERF-ERF	*ethylene-responsive transcription factor ABR1-like*	dorsal
Sispe038Scf2159g01072	bHLH	*transcription factor bHLH62*	dorsal
Sispe038Scf2358g01033	WRKY	*probable WRKY transcription factor 14*	ventral
Sispe038Scf2515g00020	HB->HB-WOX	*WUSCHEL-related homeobox 1*	ventral
Sispe038Scf2996g00029	B3->B3	*B3 domain-containing protein At2g36080-like*	dorsal
Sispe038Scf3275g05006	WRKY	*WRKY transcription factor 13*	ventral
Sispe038Scf3458g00018	bHLH	*transcription factor bHLH62-like*	ventral
Sispe038Scf5680g00016	AP2/ERF->AP2/ERF-ERF	*ethylene-responsive transcription factor 2-like*	dorsal
Sispe038Scf6188g00023	B3->B3-ARF	*auxin response factor 2-like*	dorsal
Sispe038Scf6299g00006	Tify	*protein TIFY 9-like*	dorsal

19 out of 34 TFs were enriched with TCP binding sites at their predicted 5' regulatory regions

TCP binding site has been known as the important element that mediates gene regulation of TCP TF family. Basically, TCP binding site is classified into two classes

doi:10.6342/NTU201901184

with the consensus of GGNCCCAC for class I and GTGGNCCC for class II. Most of genes that are regulated by TCP-TFs are usually enriched with these binding sites. Moreover, regulation of these genes through the binding of TCP TFs at these binding sites has also been confirmed either in the *in vitro* or *in vivo* analysis. It is also important to be noted that each TCP TF might have different preference of TCP binding sites. There are also some evidences that the recognized binding site motifs of TCP TF might not always follow the exact common consensus of GGNCCCAC or GTGGNCCC. For instance, some TCP-TFs have been found to bind to the motif GAGGGACCCT, TTGGGACCTC, GTGGGAACCA (classified as class I), tGGKMCCa, GGACCA, tGGGtCCAC, and TGGKGCC (classified as class II) which actually do not resemble class I or class II consensus. Another case is that some TCP TFs have also been reported to recognize the binding motif TGGGC(C/T) or GGNCCCNC which is the combination of both class I and class II consensus, thus classified as class I&II (González-Grandío & Cubas, 2016).

Since *SsCYC* belongs to TCP TF family, SsCYC downstream regulation might also be facilitated by the presence of TCP binding sites at the regulatory region of its downstream, suggesting that the presence of TCP binding sites at the 5'regulatory region is the important indicator to determine SsCYC downstream among the dorsi-ventral DE-TFs. Therefore, the 2 kb sequences of 5' regulatory region of each TF were retrieved from the draft genome of *S. speciosa* 'Avenida Niemeyer' as it was the only available genome

doi:10.6342/NTU201901184

data. The retrieved sequences were then screened for the presence of TCP binding consensus, summarized from the paper 'TCP Transcription Factors: Evolution, Structure, and Biochemical Function' (González-Grandío & Cubas, 2016) which included TCP class I, TCP class II, combination of both class I and class II, as well as the unique sequences (the ones that not resemble both classes) that have been proved to be bound by TFs of TCP family.

Among 34 DE-TFs, there were 19 TFs that were predicted to contain TCP binding sites at their 5' regulatory regions; 9 of them, including *SsCYC* were the dorsal-expressed TFs and 10 of them were the ventral expressed TFs. Most of these TFs were enriched with TCP class I&II and class II binding sites (**Table 2**). This result suggested that these TFs might have the possibility as SsCYC downstream target. However, it is also possible that they might be regulated by other TCP TFs.

doi:10.6342/NTU201901184

Table 2 The list of 19 TFs predicted to contain TCP binding sites at their 5' regulatory region

Gene ID	Family	Gene Name	Expression	Total of TCP binding sites*		
				Class I & II	Class I	Class II
Sispe038Scf0044g00001	bZIP	*SsABF2*	ventral	7	2	5
Sispe038Scf0146g00043	WRKY	*SsWRKY35*	ventral	1		1
Sispe038Scf0170g01016	MYB-Related	*SsRL2*	dorsal	2		
Sispe038Scf0228g08027	AP2/ERF-ERF	*SsERF17*	dorsal	1		2
Sispe038Scf0247g02018	C2C2-CO-like	*SsBBX15*	ventral	4		2
Sispe038Scf0266g00013	HB-HD-ZIP	*SsHB13*	ventral	1		
Sispe038Scf0367g01001	MYB	*SsMYBS1*	ventral			1
Sispe038Scf0757g01046	MYB-Related	*SsRVE1*	dorsal			2
Sispe038Scf1061g02075	AP2/ERF-ERF	*SsERF3*	dorsal	1		
Sispe038Scf1077g00028	MADS-MIKC	*SsAGL6*	ventral	1		1
Sispe038Scf1202g12026	OFP	*SsOFP6*	dorsal	1		1
Sispe038Scf1400g01001	TCP	*SsCYC*	dorsal			3
Sispe038Scf1651g00049	MYB	*SsMYB14*	dorsal	2		
Sispe038Scf1947g02019	HB-HD-ZIP	*SsHB13*	ventral			1
Sispe038Scf2159g01072	bHLH	*SsCIB2*	dorsal	2		
Sispe038Scf2358g01033	WRKY	*SsWRKY14*	ventral			1
Sispe038Scf2515g00020	HB-WOX	*SsWOX1*	ventral			1
Sispe038Scf2996g00029	B3	*SsNGAL1*	dorsal	2		
Sispe038Scf5680g00016	AP2/ERF-ERF	*SsERF1*	dorsal	1		

*TCP binding consensus found:

Class I&II: GGNCCCNC and TGGGC(C/T)

Class I: GTGGGNCC

Class II: tGGKMCCa, GGACCA, and TGGKGCC

The expression pattern of dorsal-expressed TFs was consistent with the RNA-seq

In order to narrow down the possible SsCYC downstream TFs, this study focused on those TFs that might be the activation targets of SsCYC. These TFs should be those that have the similar expression pattern with *SsCYC*, which then should be the dorsal-expressed TFs. The qRT-PCR result showed that the 9 dorsal-expressed TFs expression pattern was consistent with the RNA-seq data, confirming their possibility as SsCYC activation targets (**Fig. 1**).

doi:10.6342/NTU201901184

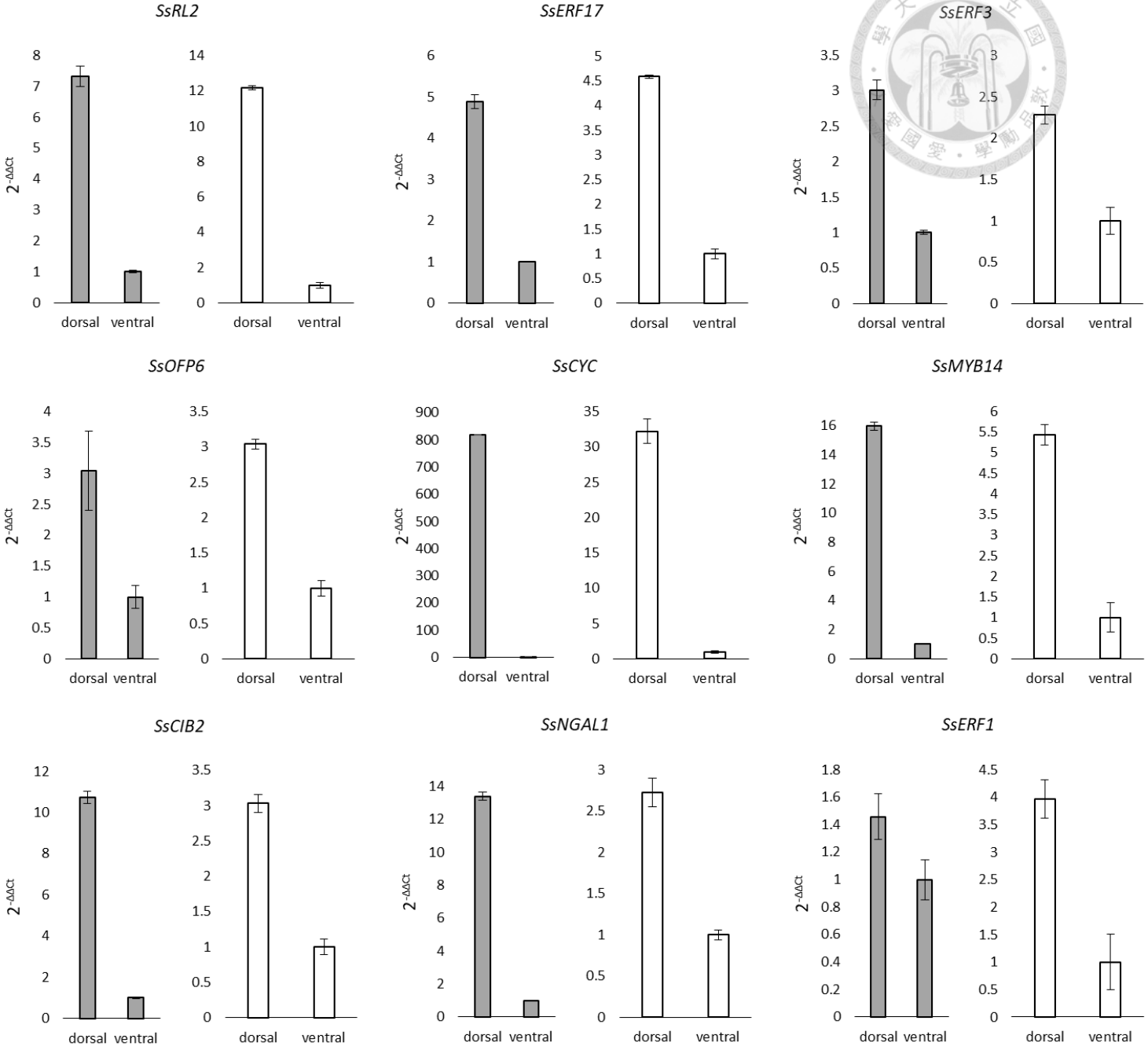


Figure 1　qRT-PCR confirmation of dorsal-expressed TFs that have been predicted to have TCP binding sites at their 5' regulatory region

The grey bars represent the qRT-PCR results and the white bars represent the RNA-seq results, expressed as the mean of relative fold gene expression level ($2^{-\Delta\Delta Ct}$) ± standard error of mean. *18s* was used as reference gene and ventral expression level was used as control.

doi:10.6342/NTU201901184

All the isolated 5' regulatory regions of *S. speciosa* 'Espirito Santo' dorsal-expressed TFs contained TCP binding sites

Since the 9 dorsal-expressed TFs have the consistent expression pattern with the RNA-seq data, then isolation of the 5' regulatory region of these TFs from *S. speciosa* 'Espirito Santo' (SsES) was conducted by PCR based methods. The reverse primer of each regulatory region was designed to facilitate overlap at the 3' with the beginning of the coding sequence (CDS) of the corresponding TF, except for *SsOFP6*.

The 5' regulatory regions that were successfully isolated were those belong to *SsCYC*, *SsRL2*, *SsERF17*, *SsERF3*, *SsOFP6*, *SsCIB2*, *SsNGAL1* and *SsERF1*. Their lengths were varied between almost 1 to 2 kb (**Table 4**). Each isolated regulatory region of SsES showed similarity ranging from 75% (*SsERF1*) to ~98% (*SsCYC*) when aligned with the predicted sequence of *S. speciosa* 'Avenida Niemeyer' (SsAN) (**Table 3**), indicating that regulatory sequence variations might appear within cultivars. All the isolated regulatory regions, with the exception of *SsOFP6* also have their 3' sequences overlap with the beginning of the CDS of their respective TFs. These results confirmed that all the obtained regulatory sequences were belong to their respective TFs. The differences between SsAN and SsES regulatory regions were due to several point mutations and indels. Comparing to the other regulatory regions, *SsERF1* showed significant differences between SsES and SsAN, which was characterized with frequent

doi:10.6342/NTU201901184

large gaps, caused due to large insertions or deletions and point mutations. Unfortunately, the regulatory region of *SsMYB14* was failed to be isolated using all the approaches, thus it did not continue to the remaining analysis. Moreover, the regulatory region of *SsCYC* in SsES showed two different alleles, also due to indels. The length of these two alleles only differed in 2 bp.

The isolated 5' regulatory regions were also screened for TCP binding sites using the same method described previously. All the isolated 5' regulatory regions contain TCP binding sites. Most of them were enriched with either TCP class I&II or/and class II. Since the regulatory regions of most TFs were quite similar to the predicted ones, they also shared similar binding consensus at almost similar position, except for *SsERF1*. However, the regulatory region of *SsRL2* in SsES was lack of 1 binding site that caused due to the shorter length comparing to *SsRL2* in SsAN, so that it could not cover the binding site found at the position between (-1934) and (-1939) of SsAN. Similar to *SsRL2*, *SsERF17* in SsES was also lack of 1 binding site which was caused by long deletion so that it missed the binding site found at the region between (-1544) and (-1549) of SsAN. In the case of *SsOFP6* of SsES, the lacking of 1 binding site was caused due to the change of one base from T to C at position -462, which eliminated this binding site in SsES. As the regulatory sequence of *SsERF1* of SsES had pretty low percentage of similarity to SsAN, thus it also had different binding site compared to the predicted one in the term of

doi:10.6342/NTU201901184

sequence and position (**Table 3, 4 & 5; Supplementary Fig. S8**). Yet, these results still suggested that these TFs might be the target of SsCYC or other TCP TFs.

Table 3 Percentage of identity between the 5' regulatory region of *S. speciosa* 'Espirito Santo' (SsES) and *S. speciosa* 'Avenida Niemeyer' (SsAN)

Gene Name	Identity (%)
SsRL2	97.86
SsERF17	95.56
SsERF3	94.75
SsOFP6	97.02
SsCYC_A	98.28
SsCYC_B	98.23
SsCIB2	97.72
SsNGAL1	96.56
SsERF1	75.48

Analysis was done by Clustal MUSCLE tool (http://www.ebi.ac.uk/Tools/msa/muscle/)

Table 4 Summary of TCP binding sites found at the 5' regulatory regions isolated from *S. speciosa* 'Espirito Santo' (SsES)

Gene Name	Pattern	Strand	Start	End	Sequence	Length (bp)	TCP binding site class
SsRL2	TGGGC(C/T)	+	-848	-853	TGGGCC	1444	I&II
SsERF17	TGGGC(C/T)	-	-247	-252	TGGGCC	997	I&II
	GGACCA	+	-472	-477	GGACCA		II
SsERF3	TGGGC(C/T)	-	-1303	-1308	TGGGCT	1338	I&II
SsOFP6	GGNCCCNC	+	-116	-123	GGTCCCTC	1423	I&II
*SsCYC_A**	TGGKGCC	-	-1105	-1111	TGGGGCC	1998	II
	TGGKGCC	+	-1108	-1114	TGGGGCC		II
	GGACCA	+	-1208	-1213	GGACCA		II
*SsCYC_B**	TGGKGCC	-	-1105	-1111	TGGGGCC	2000	II
	TGGKGCC	+	-1108	-1114	TGGGGCC		II
	GGACCA	+	-1208	-1213	GGACCA		II
SsCIB2	TGGGC(C/T)	-	-1370	-1375	TGGGCC	1861	I&II
	TGGGC(C/T)	+	-1405	-1410	TGGGCT		I&II
SsNGAL1	GGNCCCNC	+	-125	-132	GGCCCCCC	1697	I&II
	TGGGC(C/T)	-	-1584	-1589	TGGGCT		I&II
SsERF1	TGGGC(C/T)	+	-1912	-1917	TGGGCC	1953	I&II

**SsCYC*_A and *SsCYC*_B refer to *SsCYC* 5' regulatory region

doi:10.6342/NTU201901184

Table 5 Summary of TCP binding sites predicted from the 5' regulatory region of *S. speciosa* 'Avenida Niemeyer' (SsAN)

Gene Name	Pattern	Strand	Start	End	Sequence	TCP binding site class
SsRL2	TGGGC(C/T)	+	-863	-868	TGGGCC	I & II
	TGGGC(C/T)	+	-1934	-1939	TGGGCC	I & II
SsERF17	TGGGC(C/T)	-	-246	-251	TGGGCC	I & II
	GGACCA	+	-471	-476	GGACCA	II
	GGACCA	-	-1544	-1549	GGACCA	II
SsERF3	TGGGC(C/T)	-	-1306	-1311	TGGGCT	I & II
SsOFP6	GGNCCCNC	+	-115	-122	GGTCCCCC	I & II
	GGACCA	-	-496	-501	GGACCA	II
SsCYC	TGGKGCC	+	-1108	-1114	TGGGGCC	II
	TGGKGCC	+	-1111	-1117	TGGGGCC	II
	GGACCA	-	-1211	-1216	GGACCA	II
SsCIB2	TGGGC(C/T)	-	-1370	-1375	TGGGCC	I & II
	TGGGC(C/T)	+	-1405	-1410	TGGGCT	I & II
SsNGAL1	GGNCCCNC	+	-123	-130	GGCCCCCC	I & II
	TGGGC(C/T)	-	-1617	-1622	TGGGCT	I & II
SsERF1	TGGGC(C/T)	-	-347	-352	TGGGCT	I & II

SsCYC might have the ability to autoregulate itself and regulate other TFs

The interaction of SsCYC with its possible downstream targets was checked by dual-luciferase assay by co-transfecting effector, reporter and internal control into the same protoplasts of *Nicotiana benthamiana* leaves. The SsCYC-GFP was used as the effector to regulate the firefly luciferase activity driven by the 5' regulatory region of interest. The detected firefly luciferase signal of each tested regulatory region was normalized by renilla luciferase signal to encounter the transfection variability. The firefly/renila luciferase signal ratio obtained using SsCYC-GFP effector was compared to GFP effector (control) to specify the interaction of SsCYC with the corresponding TFs,

doi:10.6342/NTU201901184

expressed as normalized fold change (△fold activity).

Some of the dorsal expressed TFs showed response to SsCYC effector. Significant up-regulation by SsCYC was observed in the *SsCYC*, *SsRL2* and *SsERF1* regulatory region construct, indicating that they might be the activation target of SsCYC. The ability of SsCYC to activate itself might be considered as a positive autoregulation. In contrast, significant down-regulation by SsCYC was also observed in the *SsOFP6* and *SsERF3* regulatory region construct, indicating SsCYC might repress these TFs expression. The remaining construct did not show neither activation nor repression by SsCYC (**Fig. 2**).

doi:10.6342/NTU201901184

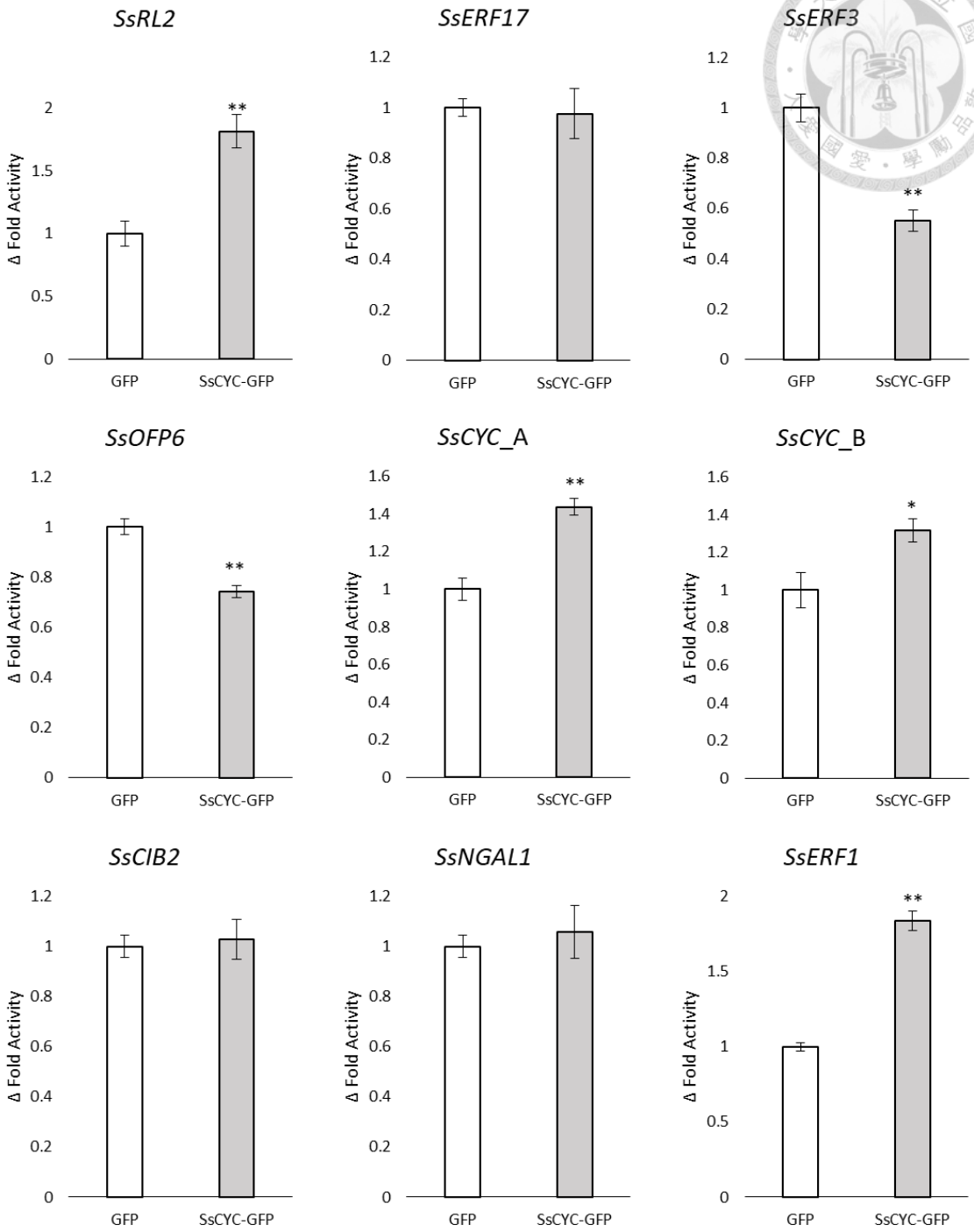


Figure 2 Dual-luciferase assay result

SsCYC regulation of the target TFs was expressed as mean of normalized fold change (△fold activity) ± standard error of mean, determined by calculating the firefly to renilla luciferase activity ratio of the tested group (SsCYC-GFP effector) divided to the control group (GFP effector). The results were analyzed using One-Way Analysis of Variance; *P-value<0.05, **P-value<0.01.

doi:10.6342/NTU201901184

Discussion

Floral zygomorphy study has mainly focused on the dorsi-ventral asymmetric expression of transcription factor (TF) *CYCLOIDEA* (*CYC*) which belongs to TCP TF family class II. The role of *CYC* in floral zygomorphy is early discovered in *A. majus*, where *CYC* is expressed in the dorsal petal of the flower, patterning the dorsal petal identity by affecting its size and shape so that it could be distinguished from the ventral petal (Costa et al., 2005; Hileman, 2014; Spencer and Kim, 2018). The phenomenon of *CYC* regulation of floral zygomorphy has also been observed in *S. speciosa* (Dong et al., 2018). This study showed the discovery of TFs that were responsive to CYC in *S. speciosa*. SsCYC was able to regulate certain dorsal-expressed TFs whose 5' regulatory regions were enriched with TCP binding sites, elements that have been known to mediate TCP TF family gene regulation. SsCYC regulation of these TFs could be linked to their function as SsCYC downstream in patterning the dorsal identity of *S. speciosa*.

The floral zygomorphy establishment in *S. speciosa* 'Espirito Santo' might involve another *RAD*-like gene

It has been well-known that the floral zygomorphy regulation in *A. majus* relies on AmCYC and AmDICH regulation of *AmRAD* in dorsal petal. The fact that AmCYC is

doi:10.6342/NTU201901184

able to bind to the TCP binding sites found at the promoter and intron of *AmRAD* suggests that TCP binding sites are also the important elements that provide AmCYC regulation of *AmRAD* (Costa et al., 2005). The classic pattern of *CYC-RAD-DIV* regulation is generally thought to be conserved in Lamiales, and even has been reported outside Lamiales which is in Dispacales (Pretson and Hileman, 2009; Boyden et al., 2013). For instance, the *CYC-RAD* regulation is found in *Bournea leiophylla* (Gesneriaceae) (Zhou et al., 2008), *Veronica montana* and *Gratiola officinalis* (*Antirrhinum* close relatives) (Preston et al., 2009), since their *RAD* genes are expressed in the similar manner with their *CYC* gene counterparts. Furthermore, study in *Chirita heterotricha* (Gesneriaceae) signifies CYC binding site enrichment at *RAD* promoter outside the *Antirrhinum* (Yang et al., 2010), which supports the hypothesis of *CYC-RAD* model conservation in establishing floral zygomorphy.

However, there are also evidences that the *CYC-RAD* model is actually not conserved. In both *Antirrhinum* and *Arabidopsis*, it has been found that there are some *RAD*-like genes. Observation of the 5 *RAD*-like genes in *Antirrhinum* shows that none of them are expressed like *AmRAD* in dorsal regions of the flower. The same phenomenon is also occurred in the 6 *RAD*-like genes of *Arabidopsis*, which they are not expressed at the same region with TCP1, the *Arabidopsis AmCYC* orthologue. Moreover, when *AmCYC* is overexpressed in *Arabidopsis*, it is also unable to increase the expression of

doi:10.6342/NTU201901184

endogenous *RAD*-like genes of *Arabidopsis*. Together, these studies suggest that there might be changes have occurred in the cis-regulatory elements of these *RAD*-like genes during the duplication which raise the possibility that the control of floral zyomorphy in other species does not always follow the *CYC-RAD-DIV* model of *Antirrhinum* (Baxter et al., 2007; Costa et al., 2005). In addition, it has also been reported that in *Saintpaulia ionantha* (Gesneraiceae), the *RAD* expression does not correlate with *CYC* (Hsu et al., 2018). This evidence supports the suggestion that even in Gesneriaceae, *CYC* might co-opt other pathways in regulating floral zygomorphy.

Although *S. speciosa* is belong to Lamiales, the RNA-seq data of *S. speciosa* 'Espirito Santo' (SsES) indeed only showed that *SsCYC* was differentially expressed while no *SsRAD* and *SsDIV* (**Table 1**), homologous of *AmRAD* and *AmDIV* were found to be dorsi-ventral differentially-expressed. Instead of *RAD*, other *RAD-like* gene (*SsRL2*) which has more similarity to the *Antirrhinum RAD-like 2* (*AmRL2*) was found to be dorsi-ventral differentially expressed. In the case of SsES, the *SsRL2* seemed to be activated by SsCYC since its 5' regulatory region was enriched with the TCP binding site and showed up-regulation by SsCYC in the dual-luciferase assay (**Table 4**; **Fig. 2**). Although the function of this TF is still unknown, but it might have some influences in the floral zygomorphy regulation of SsES.

doi:10.6342/NTU201901184

Based on this data, it was originally thought that the control of floral zygomorphy in *S. speciosa* might not mimic the model of *A. majus*, but later it was found that there were two other *RAD-like* genes (*SsRL1* and *SsRL3*) in SsES that surprisingly did not pass the dorsi-ventral differential expression filter in the RNA-seq. In order to confirm the existence of *CYC-RAD-DIV* model in *S. speciosa*, the expression of these *RAD*-like genes and *SsCYC* was compared within the dorsal and ventral petals of SsES and also with the whole petals of *S. speciosa* 'Avanti' (SsA), the other cultivar that has ventralized actinomorphic symmetry caused due to 10 bp deletion of *SsCYC*. As expected, *SsCYC* exhibited high expression pattern at the dorsal petals of SsES but was expressed in low level at the ventral petals of SsES and the whole petals of SsA, supporting the previous suggestion that *SsCYC* is the major role of floral zygomorphy in *S. speciosa*. Interestingly, the *SsRL1* was the only *RAD-like* gene that expressed almost in the similar manner with *SsCYC*, confirming that the establishment of floral zygomorhy in *S. speciosa* might follows the model of *Antirrhinum*. The high expression pattern of *SsRL2* in SsA indicated that there might be other regulation of this TF besides by SsCYC (**Fig. 3**).

doi:10.6342/NTU201901184

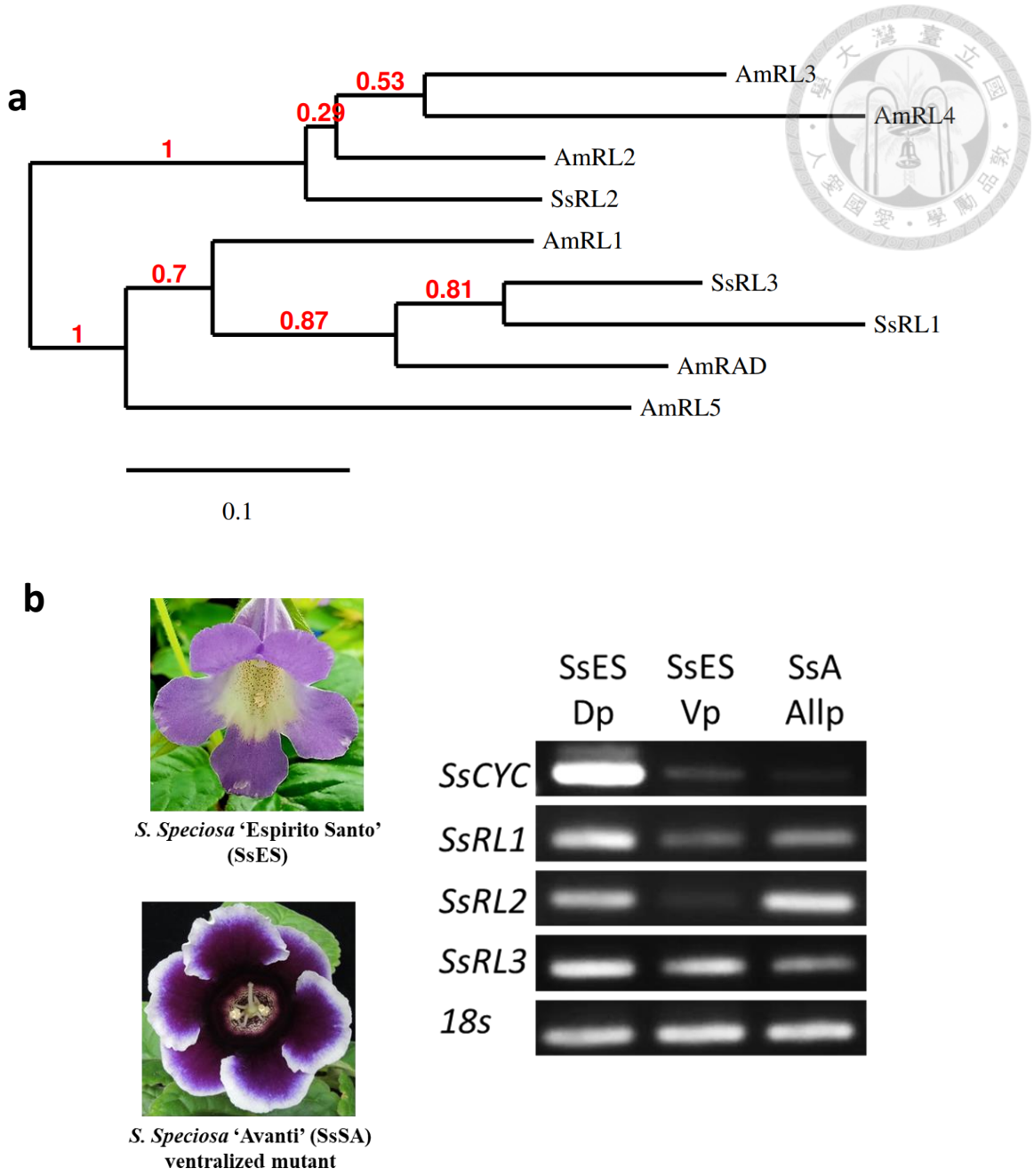


Figure 3 *SsCYC* and *SsRAD*-like genes expression profile in *S. speciosa* 'Espirito santo' and *S. speciosa* 'Avanti'

(a.) Phylogenetic tree of *S. speciosa* and *Antirrhinum RAD* and *RAD-like* genes. The tree was produced using Phylogeny.fr (http://www.phylogeny.fr/alacarte.cgi). Bootstrap values (red number) are based on 100 replicates. GenBank references: AmRL1, AJ791699; AmRL2, DQ375230; AmRL3, DQ375227; AmRL4, DQ375228; AmRL5, AJ793240. (b.) Real-time PCR analysis to compare the expression pattern of SsCYC and SsRAD-like genes within the dorsal (Dp) and ventral (Vp) petals of *S. speciosa* 'Espirito santo' (SsES), and the whole petals (AllP) of *S. speciosa* 'Avanti' (SsA). SsA is the ventralized actinomorphic mutant of *S. speciosa* caused due to 10 bp deletion of *SsCYC*.

doi:10.6342/NTU201901184

Regarding to the expression of *SsDIV*, whose ventral expression is important to complete the *CYC-RAD-DIV* model, the non-existence of this TF within the DE-TFs is actually explicable. In *A. majus*, *AmDIV* regulation by *AmCYC* and *AmDICH* appears to be post-transcriptional since it is expressed throughout the wild-type flower at the early stages of development and its expression is also not affected by *cyc* or *dich* mutation. Even in the later stage, *AmDIV* is still expressed in all petals although expression is enhanced in some ventral regions in a manner that depends on *AmDIV* itself (Galego and Almeida, 2002). In the previous data, the expression of *SsDIV* also followed the similar pattern of *AmDIV*; it showed similar dorsi-ventral expression pattern at flower bud stage 5 and it was also expressed throughout all the petals during the whole developmental stages of SsES (**Supplementary Fig. S9**), thus explain the absence of this TF in the DE-TFs. Since *S. speciosa* floral zygomorphy might follow the pattern of *CYC-RAD-DIV*, then further confirmation of SsCYC regulation of *SsRL1* might be needed.

SsCYC might work through *ERF*-mediated hormone pathway to affect the dorsal petal cell size

SsCYC control of the dorsal petal identity has been linked to its effect on petal size. For instance, the expression of *AmCYC* in *Arabidopsis* affects its petals and leaves in the different way. In leaves, *AmCYC* reduces the leaf size by inhibiting the cell proliferation

doi:10.6342/NTU201901184

at the early stage and reduces the cell expansion at the later stage. In petals, *AmCYC* expression leads to an increase of size in all petals because it promotes cell expansion at the later stage (Costa et al., 2005). Similar function of *CYC* in petal development has also been observed in *S. speciosa*. Recent study of the wild type *S. speciosa* 'pink flower' (SsPF) reveals the unique asymmetric expression of *SsCYC* at the dorsal petal of the flower, examined using *in situ* hybridization. In the dorsal petal of SsPF, *SsCYC* is expressed significantly high at the inner part of the gibbous structure while almost no expression of *SsCYC* is able to be detected at the outer part, and with *SsCYC* expression is slightly higher at the dorsal tube compared to the ventral. Examination of the cell size leads to the suggestion that *SsCYC* might repress cell expansion since the cell growth of the inner part of the gibbous structure is repressed whereas the outer gibbous structure is similar to the ventral petal in the way that cell growth is expanded in both parts. This suggestion is also supported by the morphology of the overexpressed *Arabidopsis* transformant, in which the plant is reduced in size, followed by the reduced of flower petals due to suppression of cell expansion (Dong et al., 2018). Overexpression of *SsCYC* in *Nicotiana* also shows the same result (Kuo, 2014, unpublished work). Unfortunately, recent comparison data of the inner dorsal and ventral petals of *S. speciosa* 'Espirito Santo' (SsES) using scanning electron microscopy revealed that the dorsal petals of SsES have larger cell area compared to the ventral petals, which also reflected by the fact that the

doi:10.6342/NTU201901184

dorsal petals of SsES are longer compared to the ventral petals (**Supplementary Fig. S10**). Although the cell observation of SsES and SsPF has opposite result, but both of them still indicate that *SsCYC* regulation of floral zygomorphy might be correlated with cell expansion of the dorsal petals.

In addition, some studies have mentioned that hormones might involve in *CYC* pathway of floral symmetry regulation (Spencer and Kim, 2018). Some TCP TFs are able to regulate the hormone pathway or be regulated by hormone (Danisman 2016). Recent study in *Arabidopsis* is one that links the effect of hormone regulation by TCP TF to the petal growth, showing that *TCP5* controls the cell elongation of petal by altering ethylene biosynthesis and response pathway (van Es, 2018). Besides the *TCP5*, another study in *Chrysanthemum morifolium* also reveals that *TCP20* could interfere the jasmonate signaling pathway to alter the petal elongation by interacting with *CmJAZ1-like* and down-regulating *CmBPE2* expression (Wang et al., 2019). In this study, SsCYC involvement in hormone pathway was revealed by its ability to regulate the ethylene-signaling regulators, *ethylene response factors* (ERFs) in the dual-luciferase assay (**Fig. 2**), in which SsCYC activated *SsERF1* and repressed *SsERF3*.

Ethylene has been known to control the ability of the cell to expand by directly acting on microtubule orientation and *EXPANSIN* family gene or genes that encode xyloglucan endotransglucosylase/hydrolases (XTHs) and endo-1,4-β-D-glucanases

doi:10.6342/NTU201901184

(EGases) (Dubois et al., 2018; Pierik et al., 2007). Supporting the mechanism of ethylene in cell expansion, some of *EXPANSIN* genes, XTH encoding gene and EGase encoding gene were indeed found within the dorsi-ventral differentially expressed gene, with higher expression in the dorsal petal (**Supplementary Table S8**). These three plant cell wall loosening agents act majorly on the cellulose:hemicellulose network. Expansins are a family of cell wall proteins that mediate the acid-induced extension of plant walls. These proteins act via disruption of the hemicellulose-cellulose noncovalent interactions, which allows slippage of the load-bearing polymers and thus, expansion. XTH can modify the substrate hemicellulose (xyloglucan) via hydrolysis or transglucosylation, while EGase causes the endo-hydrolysis of β1,4 linkages of cell wall glucans, which thereby alter the wall composition (Pierik et al., 2007).

Therefore, it can be concluded that *SsCYC* control of dorsal petal cell expansion involves hormone pathways mediated by *SsERF1* and *SsERF3* asymmetric regulation to control the expression of those cell wall loosening agents. Besides *ERF*s, an ovate domain containing transcriptional repressor, *SsOFP6* was also repressed. The cellular function of this TF is still unknown but as its overexpression changes the *Arabidopsis* leaves phenotype into having flat, thick and cyan appearance, then it is led to the suggestion that *SsOFP6* might have some effects to SsES petal phenotype which could be further investigated (Wang et al., 2011).

doi:10.6342/NTU201901184

SsCYC positive autoregulation is mediated by TCP binding sites at its 5' regulatory region

The idea of CYC ability to activate itself is observed in the study of *Primulina heterotricha*, which both CYC1C and CYC1D are able to form homo- or heterodimer to maintain each other expression at the dorsal petal of the flower, creating a double positive autoregulatory feedback loop (Yang et al., 2012). Recent study in *S. speciosa* 'Pink Flower' (SsPF) has also indicates that the positive autoregulation might be co-opted by SsCYC. They find that the promoter of *SsCYC* is enriched with CYC binding site. The Electrophoresis Mobility Shift Assay (EMSA) analysis also confirms the ability of SsCYC to interact with this element. Although it has been showed that SsCYC could bind to the cis-element at its promoter, this study in *S. speciosa* 'Espirito Santo' (SsES) provides a double confirmation that the interaction of SsCYC with its promoter is indeed a positive regulation (**Fig. 2**), which is not provided in the previous study of SsPF. The positive autoregulation of SsCYC in SsPF is based on the consensus of GGGGCCC found at its promoter, whereas in this study, the positive autoregulation was based on the enrichment of the GGACCA sequence and two TGGGGCC sequences. Although the hypothesis of interaction between the two studies are based on different binding consensus, but the TGGGGCC sequence found at the position between (-1105) and (-

doi:10.6342/NTU201901184

1111) of SsES (**Table 4**) is actually the same sequence of GGGGCCC found in SsPF (**Supplementary Fig. S8**), if it is extended to another 1 bp (Dong et al., 2018).

The fact that there were two alleles of *SsCYC* 5' regulatory region is another interesting discovery (**Table 4**). Although the differences between these two alleles did not alter the TCP binding site sequences, thus should not have any effect on SsCYC regulation, but the positive autoregulation did show difference in its strength of activation level. There are cases reported that allelic variations in the promoter region might affect its gene regulation, depending on the types of variations that differ the promoter of each allele. The variation in sequences of promoter could alter the cis-regulatory element and also DNA flexibility and curvature (Muterko et al., 2016; de Meaux 2005; Schwartz, et al. 2009). Alteration of cis-regulatory element could result in the change of the types of TFs that interact to it, whereas alteration of DNA flexibility and curvature could affect the protein-DNA interaction efficiency, protein-protein interaction and the interaction between TF and general transcription machinery. These in further might influence the gene expression level; as in TFs will alter the later downstream regulation. In the case of *SsCYC* regulatory region, the variation might cause enhancement of SsCYC_A flexibility and curvature which improved its performance in transcriptional activation upon the interaction with SsCYC protein (Muterko et al., 2016; Kanhere and Bansal, 2004; van der Vliet and Verrijzer, 1993).

doi:10.6342/NTU201901184

The effect of allelic variation at promoter region has been reported in wheat. The minor different within only 1 bp at the VRN-box of *VERNALIZATION1* can modulate vernalization sensitivity and flowering time of wheat, which is associated with the modulation of DNA curvature and flexibility in the promoter region (Muterko et al., 2016). In addition, study in *Chalcone Synthase* promoter of *Arabidopsis thaliana* reveals that allelic variation of the promoter can cause functional variation of this gene due to change at the cis-regulatory region (de Meaux 2005). Not only that, similar effect of allelic variation with change of cis-regulatory region also has been observed in *FLOWER LOCUS T* promoter of *A. thaliana* which influences the flowering response of the plant (Schwartz, et al. 2009). Despite that there was no change of the TCP binding sites at both alleles of *SsCYC* regulatory region, but it cannot rule out the possibility that changes could occur at other cis-regulatory elements. If these changes happen, then they will vary *SsCYC* upstream regulators, thus affecting *SsCYC* expression level as the consequence.

Positive autoregulation has been thought as *CYC* strategy to amplify and maintain its gene expression, which then should be important for the conservation of zygomorphic lineage during the evolution of flowering plant. The result in this study showing that how different *SsCYC* regulatory region alleles could alter its positive autoregulation efficiency could be an indication that variations at the regulatory region sequence do matter.

doi:10.6342/NTU201901184

Therefore, study towards *SsCYC* regulatory region which include the interaction between *SsCYC* upstream and its regulatory region may deserve further attention.

Limitation of the study

In this study, the determination of the potential *SsCYC* downstream targets was based on the presence of TCP binding sites at the 5' regulatory region of the TFs that was predicted from the draft genome of *S. speciosa* 'Avenida Niemeyer' (SsAN). After obtaining each TF regulatory region of *S. speciosa* 'Espirito Santo' (SsES) and comparing each sequence to SsAN, the differences of their sequences revealed cultivar variation of regulatory region. Similar to the effect of allelic variation, the cultivar variation of regulatory region could also be the reason of phenotypic different between cultivars, making them could be distinguished from each other, which has been mentioned in several studies (Wang, et al. 2013; Boccacci, et al. 2017; Ye, et al. 2018). Nevertheless, it is also important to be noted that the regulatory region retrieved from SsAN came from the draft genome which mistakes during the assembly process could occur.

In spite of having their regulatory region enriched by TCP binding sites, not all of the possible target TFs could be regulated by SsCYC (**Fig. 2)**. This indicates that SsCYC might have certain binding preferences. According to the recent study of SsCYC, the only binding site that has been confirmed could be bound by SsCYC is GGGGCCC which

doi:10.6342/NTU201901184

actually refers to TGGGGCC observed in this study (Dong et al., 2018). This motif is only present in *SsCYC* regulatory region and it represents more of the TCP class II binding site, suggesting that *SsCYC* might have more preference to TCP class II rather than class I (**Table 4**) (González-Grandío and Cubas, 2016), which is different to AmCYC of *Antirrhinum majus* (Costa et al., 2005) and ChCYC of *Chirita heterotricha* (Yang et al., 2010). However, as SsCYC was still able to regulate other TFs, it also means that there might be other binding consensus that could be accommodate by SsCYC. Some studies have reported that TCP TFs have the capability to bind to overlap consensus of class I and II, the ability that might be co-opted by SsCYC as it could regulate those TFs that only enriched by the overlap consensus of class I and II (**Fig. 2**; **Table 4**) (González-Grandío and Cubas, 2016).

TCP TFs binding to DNA usually depends on the type of basic residues and helix-loop-helix motif. The basic residues involve in determining TCP TF binding preference of class I and class II, in which they affect DNA recognition and amino acid positioning. HLH motif also influences the selectivity of TCP TFs, allowing more or less efficient discrimination among related sequences. These properties of SsCYC could be further studied in order to determine its binding preference, selectivity and flexibility to its DNA binding sites (Viola et al., 2012). In addition, it is still unknown whether SsCYC did bind to the TCP binding sites of its identified target, and also which binding sites are crucial to

doi:10.6342/NTU201901184

facilitate SsCYC regulation. Analysis using EMSA may help to determine the capability of SsCYC to bind to the TCP binding site sequences and serial deletion of the 5' regulatory region could be considered as the way to determine the important region for SsCYC regulation.

One cannot argue the fact that the interaction of SsCYC with its downstream in this study did not represent the actual condition, as the experimental was designed to allow SsCYC to have better access to the TF regulatory regions which might not be the case in real situation. This raises the contradictory question of whether the TFs regulated by SsCYC are its real direct target or if they are actually the indirect target of SsCYC. Especially for *SsERF3* and *SsOFP6*, their role as SsCYC repression targets was not synchronized with their dorsal-high expression pattern. In this case, there is possibility that *SsERF3* and *SsOFP6* regulation in the actual condition is not related to SsCYC, means that they are probably not SsCYC targets. As the experiment was performed in the leaves protoplasts of *N. benthamiana* rather than *S. speciosa* itself, it also leads to the another argument if those TFs that showed no regulation by SsCYC are certainly not its target.

TF binding is affected by intra- and intermolecular TF interactions, which include the interaction between SsCYC with other TF or with the non-DNA-binding cofactors. As the TCP TF family, SsCYC binding needs to be facilitated with either homo- or

doi:10.6342/NTU201901184

heterodimer (Inukai et al., 2017; Atchley and Fitch, 1997). When *SsCYC* was expressed in the protoplasts of *N. benthamiana* leaves, there was a consequence that SsCYC was lacking of its binding partners leading to inability to regulate its actual targets, thus they were missed to be identified as SsCYC targets. In contrast, there is also a possibility that SsCYC regulation might be interrupted by the presence of other endogenous transcriptional factors and regulators (TFs and TRs) in *N. benthamiana*, that caused certain non-target TFs being mistaken as SsCYC targets (**Fig. 4**). Therefore, it is suggested to validate these TFs response to SsCYC using the protoplast of *S. speciosa* 'Avanti' flowers, so that the SsCYC regulation will not be affected by other species endogenous TFs and transcriptional regulators (TRs), as well as its endogenous SsCYC (considering if the experiment is performed in *S. speciosa* 'Espirito Santo'). The reviewers also recommended to calculate the transformation efficiency of each of the co-transfected vectors to avoid false result and to do three times reading of the firefly and renilla signal to make sure signal stability within the experiment. Moreover, further *in vivo* confirmation, such as CHIP-qPCR assay may be needed to confirm whether SsCYC has the capability to bind to these responsive TFs in actual condition.

doi:10.6342/NTU201901184

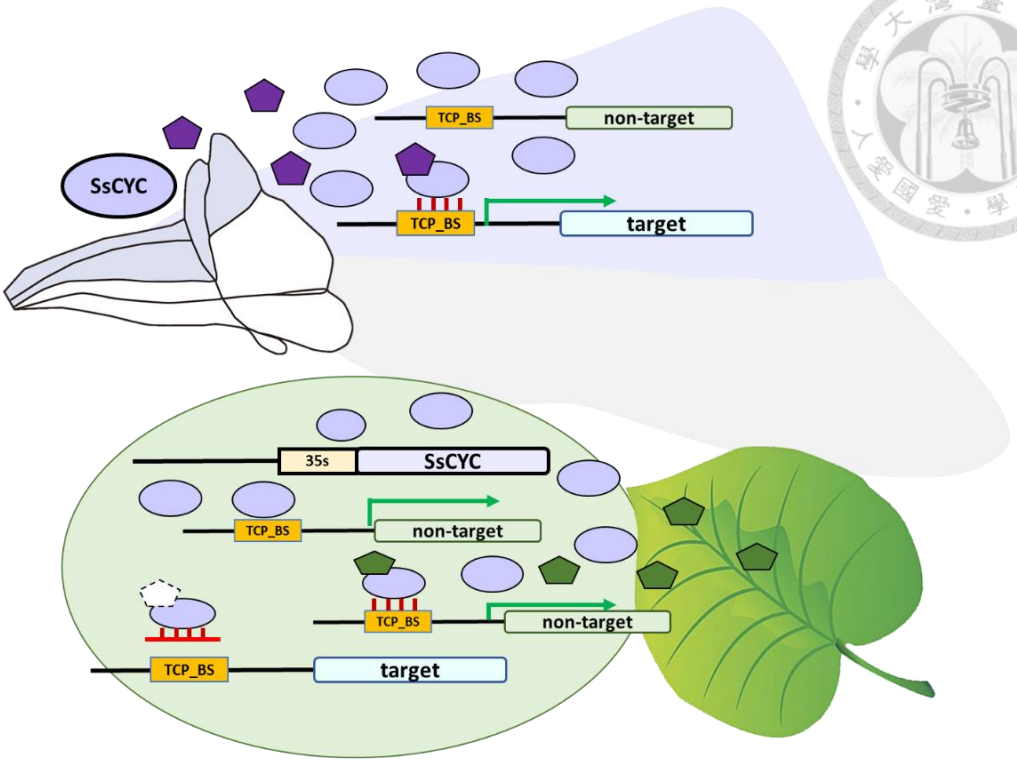


Figure 4 Possible bias that might occur in the dual-luciferase assay

SsCYC might need certain protein or co-factor to regulate its targets, and these elements were lacking in the *N. benthamiana* leaves protoplasts. As the consequence, this targets of SsCYC was missed to be identified. In contrast, the endogenous transcription factor or regulators in N. benthamiana leaves might interrupt with SsCYC regulation, causing certain non-target TFs, being mistaken as SsCYC targets.

doi:10.6342/NTU201901184

Conclusion

SsCYC might regulate the floral zygomorphy of *S. speciosa* through the *CYC-RAD-DIV* model similar to *Antirrhinum*, as well as positive autoregulation and interaction with other transcription factors (TFs), facilitated by the TCP binding site found at their 5' regulatory regions (**Fig. 5**). The TFs upregulated by *SsCYC* include *RADIALIS-like* ortholog (*SsRL2*) whose function is unknown and ethylene response transcriptional activator (*SsERF1*). In contrast, *SsCYC* also downregulated ethylene response transcriptional repressor (*SsERF3*) and an ovate family transcriptional repressor, *SsOFP6* whose function is unknown.

The finding of *SsERF1* and *SsERF3* as SsCYC responsive TFs could be linked to their function as downstream regulators of ethylene signaling pathway. They might alter dorsal cell expansion via regulation of *EXPANXIN* (*EXPA*) genes, xyloglucan endotransglucosylase/hydrolase (XTH) encoding gene and endo-1,4-β-D-glucanase (EGase) encoding gene to loosen the cell wall, since these three genes were identified as the dorsal expressed genes in the RNA-seq data of SsES. This suggestion is also reflected by the observation that the dorsal petals of SsES have larger cell area, thus are longer in length compared to the ventral petals, which is considered as one of the factors that generates floral zygomophy in this flower.

doi:10.6342/NTU201901184

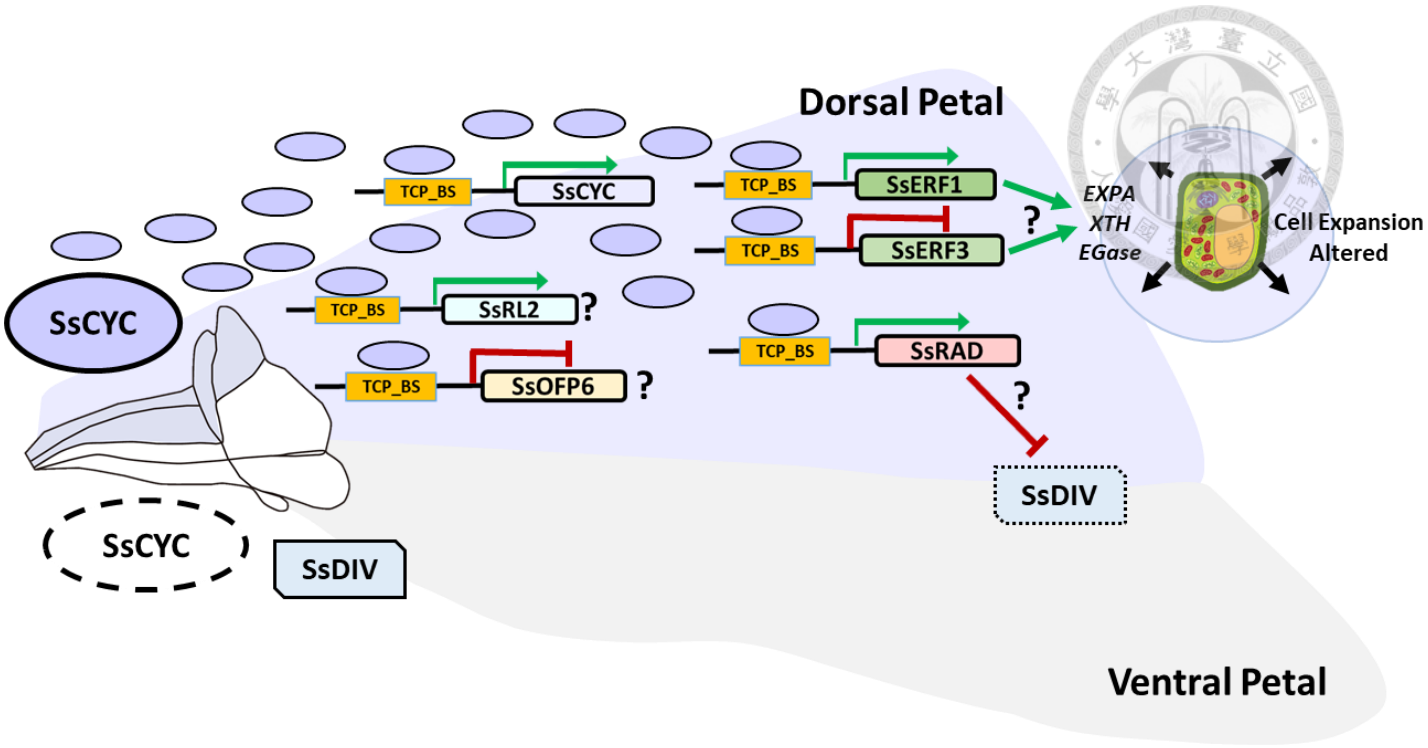


Figure 5 Hypothesis of *S. speciosa* floral zygomorphy regulation by SsCYC

SsCYC expressed highly in dorsal and maintain its expression through positive autoregulation. SsCYC protein activates *SsRAD*, producing SsRAD protein that inhibits SsDIV to the dorsal petal (further confirmation is needed). The dorsal petal size might be altered through SsCYC regulation of *SsERF1* and *SsERF3*. Both ERFs act as the downstream regulator of ethylene pathway and might alter dorsal cell expansion via regulation of *EXPANXIN* (*EXPA*) genes, xyloglucan endotransglucosylase/hydrolase (XTH) encoding gene and endo-1,4-β-D-glucanase (EGases) encoding gene to loosen the cell wall. *SsRL2* and *SsOFP6* also responded to SsCYC, yet their functions are still unknown.

doi:10.6342/NTU201901184

References

Atchley, WR., and Fitch, WM. (1997). A natural classification of the basic helix-loop-helix class of transcription factors. *Proc Natl Acad Sci* USA 94(10):5172–5176.

Baxter, CE., Costa, MM., and Coen, ES. (2007). Diversification and co-option of *RAD*-like genes in the evolution of floral asymmetry. *Plant J* 52:105-113. doi: 10.1111/j.1365-313X.2007.03222.x.

Boccacci, P., Mela, A., Pavez Mina, C., Chitarra, W., Perrone, I., Gribaudo, I., and Gambino, G. (2017). Cultivar-specific gene modulation in *Vitis vinifera*: analysis of the promoters regulating the expression of *WOX* transcription factors. *Scientific reports* 7:45670. doi:10.1038/srep45670.

Boyden, GS., Donoghue, MJ., and Howarth, DG. (2013). Duplications and expression of *RADIALIS*-like genes in Dipsacales. *Int J Plant Sci* 173(6):971–983. doi: 10.1093/molbev/msp051.

Brioudes, F., Joly, C., Szécsi, J., Varaud, E., Leroux, J., Bellvert, F., Bertrand, C., and Bendahmane M. (2009). Jasmonate controls late development stages of petal growth in *Arabidopsis thaliana*. *Plant J* 60(6):1070-80. doi: 10.1111/j.1365-313X.2009.04023.x.

Chandler, JW. (2011). The hormonal regulation of flower development. *J Plant Growth Regul* 30(2):242–254. doi:10.1007/s00344-010-9180-x.

Citerne, HL., Moèller, M., and Cronk, QCB. 2000. Diversity of *CYCLOIDEA*-like genes in Gesneriaceae in relation to floral symmetry. *Annals of Botany* 86:167-176. doi: 10.1006/anbo.2000.1178.

Corley, SB., Carpenter, R., Copsey, L., and Coen, E. (2005). Floral asymmetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum*. *Proc Natl Acad Sci USA* 102(14): 5068–5073. doi: 10.1073/pnas.0501340102.

Costa, MM., Fox, S., Hanna, AI., Baxter, C., and Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. *Development* 132(22):5093-101. doi: 10.1242/dev.02085.

Cubas, P. (2004). Floral zygomorphy, the recurring evolution of a successful trait. *Bioessays* 26(11):1175-84. doi: 10.1002/bies.20119.

Danisman S. (2016). TCP transcription factors at the interface between environmental

doi:10.6342/NTU201901184

challenges and the plant's growth responses. *Frontiers in plant science* 7:1930. doi:10.3389/fpls.2016.01930.

de Meaux, J., Goebel U., Pop, A., and Mitchell-Olds, T. (2005). Allele-specific assay reveals functional variation in the *Chalcone Synthase* promoter of *Arabidopsis thaliana* that is compatible with neutral evolution. *Plant Cell* 17(3): 676–690. doi: 10.1105/tpc.104.027839.

Dong, Y., Liu, J., Li, PW., Li, CQ., Lü, TF., Yang, X., and Wang, YZ. (2018). Evolution of Darwin's *Peloric* Gloxinia (*Sinningia speciosa*) is caused by a null mutation in a pleiotropic TCP Gene. *Mol Biol Evol* 35(8):1901-1915. doi: 10.1093/molbev/msy090.

Dubois, M., Van den Broeck, L., and Inzé, D. (2018). The Pivotal Role of Ethylene in Plant Growth. *Trends in plant science* 23(4):311–323. doi:10.1016/j.tplants.2018.01.003.

Endress, PK. (1997). *Antirrhinum* and *Asteridae*--evolutionary changes of floral symmetry. *Symp Soc Exp Biol* 998(51):133-40.

Galego, L., and Almeida, J. (2002). Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers. *Genes & development* 16(7):880–891. doi:10.1101/gad.221002.

González-Grandío, E., Cubas, P. (2015). TCP transcription factors: evolution, structure, and biochemical function. In D.H. Gonzalez (Ed.), Plant *transcription factors evolutionary, structural and functional aspects* (pp. 139-151). Oxford, UK: Academic Press.

Hileman, LC. (2014). Trends in flower symmetry evolution revealed through phylogenetic and developmental genetic advances. *Philos Trans R Soc Lond B Biol Sci* 369(1648): 20130348. doi: 10.1098/rstb.2013.0348.

Hsu, HJ., He, CW., Kuo, WH., Hsin, KT., Lu, JY., Pan, ZJ., and Wang, CN. (2018). Genetic analysis of floral symmetry transition in African Violet suggests the involvement of trans-acting factor for *CYCLOIDEA* expression shifts. *Front Plant Sci* 9:1008. doi: 10.3389/fpls.2018.01008.

Inukai, S., Kock, K. H., and Bulyk, M. L. (2017). Transcription factor-DNA binding: beyond binding site motifs. *Current opinion in genetics & development* 43:110–119. doi:10.1016/j.gde.2017.02.007.

Irish, FV. (2008). The *Arabidopsis* petal: a model for plant organogenesis. *Trends Plant Sci* 13(8):430-436. doi: 10.1016/j.tplants.2008.05.006.

doi:10.6342/NTU201901184

Kanhere, A., and Bansal, M. (2004). DNA bending and curvature: a 'turning' point in dna function?. *Proc Indian natn Sci Acad* B70(2):239-254.

Kosugi, S., and Ohashi, Y. (1997). PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell nuclear antigen gene. *The Plant cell* 9(9):1607–1619. doi:10.1105/tpc.9.9.1607

Kosugi, S., Ohashi, Y. (2002). DNA binding and dimerization specificity and potential targets for the TCP protein family. *Plant J* 30(3):337-48.

Kuo, WH. (2014). Ectopic expression of *SsCYC* in *Nicotiana benthamiana* and optimizing regeneration system of *Sinningia speciosa*. (Unpublished master's thesis). National Taiwan University, Taipei, Taiwan.

Li, C., Potuschak, T., Colón-Carmona, A., Gutiérrez, R.A., and Doerner, P. (2005). Arabidopsis TCP20 links regulation of growth and cell division control pathways. *Proc Natl Acad Sci USA* 102(36): 12978–12983. doi: 10.1073/pnas.0504039102.

Liu, YG., Mitsukawa, N., Oosumi, T., and Whittier, RF. (1995). Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. *Plant J* 8(3):457-63.

Liu, YG., Whittier, RF. (1995) Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics* 10;25(3):674-81.

Luehrsen, KR., de Wet, JR., and Walbot, V. (1995). Transient expression analysis in plants using firefly luciferase reporter gene. *Methods Enzymol* 216:397-414.

Luo, D., Carpenter, R., Copsey, L., Vincent, C., Clark, J., and Coen, E. (1999). Control of organ asymmetry in flowers of *Antirrhinum*. *Cell* 99(4):367-76.

Martín-Trillo, M., and Cubas, P. (2010). TCP genes: a family snapshot ten years later. *Trends Plant Sci* 15(1):31-9. doi: 10.1016/j.tplants.2009.11.003.

Müller, M., and Munné-Bosch, S. (2015). Ethylene response factors: a key regulatory hub in hormone and stress signaling. *Plant physiology* 169(1), 32–41. doi:10.1104/pp.15.00677.

Muterko, A., Kalendar, R., Salina, E. (2016). Novel alleles of the *VERNALIZATION1* genes in wheat are associated with modulation of DNA curvature and flexibility in the promoter region. *BMC Plant Biol* 16(Suppl 1): 9. doi: 10.1186/s12870-015-0691-2.

Neil, PR., Dafni, A., and Giurfa M. (1998). Floral symmetry and its role in plant-pollinator systems: terminology, distribution, and hypotheses. *Annu Rev Ecol Syst*

doi:10.6342/NTU201901184

29:345–373. doi: https://doi.org/10.1146/annurev.ecolsys.29.1.345.

Pierik, R., Sasidharan, R., and Voesenek, LACJ. (2007). Growth control by ethylene: adjusting phenotypes to the environment. *J Plant Growth Regul* 26:188–200. doi: 10.1007/s00344-006-0124-4.

Preston, JC., and Hileman, LC. (2009). Developmental genetics of floral symmetry evolution. *Trends Plant Sci* 14(3):147–154. doi: 10.1016/j.tplants.2008.12.005.

Preston, JC., Kost, MA., and Hileman, LC. (2009). Conservation and diversification of the symmetry developmental program among close relatives of snapdragon with divergent floral morphologies. *New Phytol* 182(3):751-762. doi: 10.1111/j.1469-8137.2009.02794.x.

Pruneda-Paz, JL., Breton, G., Para, A., Kay, SA. (2009). A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock. *Science* 323(5920): 1481–1485.

Schommer, C., Palatnik, JF., Aggarwal, P., Chételat, A., Cubas, P., Farmer, EE., Nath, U., and Weigel, D. (2008). Control of jasmonate biosynthesis and senescence by miR319 targets. *PLoS Biol* 6(9):e230. doi: 10.1371/journal.pbio.0060230.

Schwartz, C., Balasubramanian, S., Warthmann, N., Michael, T. P., Lempe, J., Sureshkumar, S., Kobayashi, Y., Maloof, JN., Maloof, JN., Borevitz, JO., Chory J, and Weigel, D. (2009). Cis-regulatory changes at *FLOWERING LOCUS T* mediate natural variation in flowering responses of *Arabidopsis thaliana*. *Genetics* 183(2), 723–732. doi:10.1534/genetics.109.104984.

Singer, T., Burke, E. (2003). High-throughput TAIL-PCR as a tool to identify DNA flanking insertions. *Methods Mol Biol* 236:241-72. doi: 10.1385/1-59259-413-1:241

Spencer, V., and Kim, M. (2018). Re"*CYC*"ling molecular regulators in the evolution and development of flower symmetry. *Semin Cell Dev Biol* 79:16-26. doi: 10.1016/j.semcdb.2017.08.052.

van der Vliet, PC., Verrijzer, CP. (1993). Bending of DNA by transcription factors. *Bioessays* 5(1):25-32. doi: 10.1002/bies.950150105.

van Es, SW., Silveira, SR., Rocha, DI., Bimbo, A., Martinelli, AP., Dornelas, MC., Angenent, GC., and Immink, R. (2018). Novel functions of the *Arabidopsis* transcription factor TCP5 in petal development and ethylene biosynthesis. *The Plant journal : for cell and molecular biology* 94(5):867–879. doi:10.1111/tpj.13904.

Viola, I. L., Reinheimer, R., Ripoll, R., Manassero, N. G., and Gonzalez, D. H. (2011).

doi:10.6342/NTU201901184

Determinants of the DNA binding specificity of class I and class II TCP transcription factors. *The Journal of biological chemistry* 287(1):347–356. doi:10.1074/jbc.M111.256271.

Wang, S., Chang, Y., Guo, J., Zeng, Q., Ellis, BE., and Chen, JG. (2011). *Arabidopsis* ovate family proteins, a novel transcriptional repressor family, control multiple aspects of plant growth and development. *PloS one* 6(8):e23896. doi:10.1371/journal.pone.0023896.

Wang, J., Guan, Y., Ding, L., Li, P., Zhao, W., Jiang, J., Chen, S., and Chen, F. (2019). The *CmTCP20* gene regulates petal elongation growth in *Chrysanthemum morifolium*. *Plant Science* 280:248-257. doi: https://doi.org/10.1016/j.plantsci.2018.12.008.

Wang, S., He, J., Cui, Z., and Li, S. (2007). Self-formed adaptor PCR: a simple and efficient method for chromosome walking. *Applied and environmental microbiology* 73(15):5048–5051. doi:10.1128/AEM.02973-06.

Wang, K., Zhang, X., Zhao, Y., Chen, F., and Xia, G. (2013). Structure, variation and expression analysis of glutenin gene promoters from *Triticum aestivum* cultivar Chinese Spring shows the distal region of promoter 1Bx7 is key regulatory sequence. *Gene* 527(2):484-90. doi: 10.1016/j.gene.2013.06.068.

Yang, X., Cui, H., Yuan, ZL., and Wang, YZ. (2010). Significance of consensus CYC-binding sites found in the promoters of both *ChCYC* and *ChRAD* genes in *Chirita heterotricha* (Gesneriaceae). *Journal of Systematics and Evolution* 48(4):249–256. doi: doi.org/10.1111/j.1759-6831.2010.00086.x.

Yang, X., Pang, HB., Liu, BL., Qiu, ZJ., Gao, Q., Wei, L., Dong, Y., and Wang, Y Z. (2012). Evolution of double positive autoregulatory feedback loops in *CYCLOIDEA2* clade genes is associated with the origin of floral zygomorphy. *The Plant cell* 24(5):1834–1847. doi:10.1105/tpc.112.099457.

Ye, NH., Wang, FZ., Shi L., Chen, MX., Cao, YY., Zhu, FY., Wu, YZ., Xie, L.J, Liu, TY., Su, ZZ., Xiao, S., Zhang, H., Yang, J., Gu, HY., Hou, XX., Hu, QJ., Yi, HJ., Zhu, CX., Zhang, J., and Liu, YG. (2018). Natural variation in the promoter of rice *calcineurin B-like protein10* (OsCBL10) affects flooding tolerance during seed germination among rice subspecies. *Plant J* 94(4):612-625. doi: 10.1111/tpj.13881.

Ye, BH. (2018). Identification, expression profiles and characterization of the *TCP* genes in *Sinningia speciosa*. (Unpublished master's thesis). National Taiwan University, Taipei, Taiwan.

doi:10.6342/NTU201901184

Yoo, SD., Cho, YH., and Sheen, J. (2007). *Arabidopsis* mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat Protoc* 2(7):1565-72. doi: 10.1038/nprot.2007.199.

Zhou, XR., Wang, YZ., Smith, JF., and Chen, R. (2008). Altered expression patterns of *TCP* and *MYB* genes relating to the floral developmental transition from initial zygomorphy to actinomorphy in *Bournea* (Gesneriaceae). *New Pyhtol* 178(3):532-543. doi: 10.1111/j.1469-8137.2008.02384.x.

Zaitlin, D., Pierce, AJ. (2010). Nuclear DNA content in *Sinningia* (Gesneriaceae); intraspecific genome size variation and genome characterization in *S. speciosa*. *Genome* 2010 53(12):1066-82. doi: 10.1139/G10-077.

doi:10.6342/NTU201901184

Supplementary Data

Supplementary Table S1 Summarized TCP binding consensus

Binding Consensus	Class
GGNCCCNC	I & II
TGGGC(C/T)	I & II
GAGGGACCCT	I
TTGGGACCTC	I
GTGGGAACCA	I
GTGGGNCC	I
GTGGNCCC	II
tGGKMCCa	II
GGACCA	II
tGGGtCCAC	II
TGGKGCC	II

The consensus was summarized from the TCP binding sequences found in 'TCP Transcription Factors: Evolution, Structure, and Biochemical Function' (González-Grandío and Cubas, 2016)

doi:10.6342/NTU201901184

Supplementary Table S2 Primer list for the isolation of *S. speciosa* 'Espirito Santo' transcription factor coding sequence

Primer	Sequence
SsABF2_F	AGGTGTTTGGATTTGCTTTGCAC
SsABF2_R	GATTTGTGGTAAGGAGCATTCTACTGC
SsRL2_F	CGTCTTCGTTTCTCGGTTTCTTGG
SsRL2_R	CCATATTCCAATGATAATGGACTGAGGTTT
SsERF17_F	AATGGTGAAACCGCAATCGAGAAAG
SsERF17_R	CAATAATTCCATATGCGTGATGTATTCAAAACC
SsHB13_F	GCAGGTGGCAACAGTTTCATAGG
SsHB13_R	CGATTCGTGTTCCCATCTTGTTCT
SsMYBS1_F	AGTATGGGAGAGGAAATAGGAGTGG
SsMYBS1_R	CAAGACAGTTCAATGTAACAGCCTCTAAT
SsRVE1_F	AGAACTGATAGGTTCTGAGGCTATGG
SsRVE1_R	GGTAAGCCAGATACCCTGCTTCAA
SsERF3_F	AAACCTCATTCCTACAGACCAACC
SsERF3_R	GTGGTGCTGGAAGATTCAGGTC
SsAGL6_F	AGAATGGGGAGAGGAAGAGTGGAGTTG
SsAGL6_R	GGCTTAGAGCAAATTAAAGTGTCCATCCTTCG
SsOFP6_F	CCTGGTTTGCCAATGTCTAGCATTAAG
SsOFP6_R	GTGCAGTCCCTAGAAGTCACGT
SsCYC_F	ATGTTTAGCAAGAGCACATACCTTCATG
SsCYC_R	CCACAGAAACCACGCAGAATTACA
SsMYB14_F	AAATGGGTCGGGCTCCATGTTG
SsMYB14_R	GTCTACATGTTACAGGAGTGACGGT
SsCIB2_F	TTTGGAATCTTGATGATGGATAAGGAGTAC
SsCIB2_R	GCTTCAAATCTGGCATAAGTGACTAGTT
SsNGAL1_F	ACACGCACTGAAATGTCAATAAACCAC
SsNGAL1_R	CCATCCCATGTTATAATTCATACAAACAGGAT
SsERF1_F	TCATGTACCAGCCAATTTTCAGTGAG
SsERF1_R	CAATAGAGCCTTTGATCCACGCATTC

doi:10.6342/NTU201901184

Supplementary Table S3 Primer list for qRT-PCR confirmation of dorsal-expressed transcription factor of *S. speciosa* 'Espirito Santo'

Primer	Sequence
SsRL2_qp_F	TCCTCTTCATGGACACCTAAGC
SsRL2_qp_R	TCGGGTGTATCCTTGTCGTAC
SsERF17_qp_F	TAGCGGCGGATGAATTGTCTCG
SsERF17_qp_R	TTCCACCACCGGTTGTTTCGAC
SsERF3_qp_F	ACAGCGACGTTTCCTCAGTAGC
SsERF3_qp_R	CTCCGAACGCAGAGCTGTAGAC
SsOFP6_qp_F	GGAAACCACCACCACCACCTTC
SsOFP6_qp_R	CCTTGAACGGCCCTCAGAGTTG
SsCYC_qp_F	ACCTCACAATCCAACCTGTGTGAC
SsCYC_qp_R	CCACAGAAACCACGCAGAATTAC
SsMYB14_qp_F	ACAACCCAAACCCGAATTCGAC
SsMYB14_qp_R	AATACCTCCGACCAGAAGCTTTCG
SsCIB2_qp_F	GCGTTCGAAACCAACAGAAAGTGG
SsCIB2_qp_R	CGGACTCTTTCTGCTAAACTGTGG
SsNGAL1_qp_F	TCCAAGTCAACAGCATCAAGGG
SsNGAL1_qp_R	ACCTCTTTGCATTCCCACTTGC
SsERF1_qp_F	TGGTGCAAGAGTTTGGCTTG
SsERF1_qp_R	AGCCTTTGATCCACGCATTC

doi:10.6342/NTU201901184

Supplementary Table S4 Primer list for the isolation of *S. speciosa* 'Espirito Santo'5' regulatory region

Primer	Sequence
SsRL2_reg_F	TTAGATCACAGGTTATAACCCGATCTAATTTC
SsERF3_reg_F1	GAGGAAGTAAAACGTGTGGGGTTCTC
SsERF3_reg_F2	GTTTGGACGACTTTAAACCACCAG
SsERF3_reg_R	CTGCGGTCCCTTCTTTACGTAAA
SsOFP6_reg_F	AGAGCATCGCTATATTTGTGGCT
SsOFP6_reg_R	GTGGCAATAAGGGAAACTGAAGTC
SsCYC_reg_F	TCCCTGCAAGAACGTATAGGAATC
SsCYC_reg_R	GGTCAACCAAAGAAGTAGAGGCA
SsCIB2_reg_F1	ATCGCGCGTCACGTTTACTTA
SsCIB2_reg_R1	CTTCACTACAATTCAGTCCATTCC
SsCIB2_reg_F2	ATTTCCACCACAAAGCTGGGAAG
SsCIB2_reg_R2	CTCCTTATCCATCATCAAGATTCC
SsNGAL1_reg_F	CAGAGATGGTGTGGTCACAGGGAATC
SsERF1_reg_F	GAATCACTGGGATAACTCAGCCATCTGCAG

doi:10.6342/NTU201901184

Supplementary Table S5 Thermal Asymmetric Interlaced PCR (TAIL-PCR) primer list for the isolation of *S. speciosa* 'Espirito Santo' 5' regulatory region

Primer	Sequence
SsERF17_TAIL1	CGTGCTTCGACGCAGCAACCTGAATCTGC
SsERF17_TAIL2	TAAGTGTCACCGAAGGTCCACGTAAGC
SsERF17_TAIL3	CTGTTGGGTTGTCGAACCTCTG
SsERF17_TAIL4	GCGATTCCTTTCTCGATTGCGGTTTCACC
AD1	NGTCGASWGANAWGAA
AD2	TGWGNAGSANCASAGA
AD3	AGWGNAGWANCAWAGG
AD4	STTGNTASTNCTNTGC
AD5	NTCGASTWTSGWGTT
AD6	WGTGNAGWANCANAGA

doi:10.6342/NTU201901184

Supplementary Table S6 Self-Formed Adaptor PCR (SEFA-PCR) primer list for the isolation of *S. speciosa* 'Espirito Santo' 5' regulatory region

Primer	Sequence
SsRL2_SP1	CCTTCTCACTTCCTCTGCTGATTTACCAGTAACC
SsRL2_SP2	CCTTGTCGTACATAGCCAGAGCTTCTTCG
SsRL2_SP3	ACGTGAAGATGACATGGANNNNNNNNGGTGAC
SsRL2_SP4	GAGATCGACCAAGAAACCGAGAAACGAAGACG
SsRL2_SP5	CTTAGGTGTCCATGAAGAGGACG
SsNGAL1_SP1	GCAGTCAACTGTTCATCGACCACTAATTCTTCCTC
SsNGAL1_SP2	GGAACCACCAGCTCCACCTCCTTCTACGG
SsNGAL1_SP3	AGGATTCCATCATCATATNNNNNNNNGGCCAG
SsNGAL1_SP4	GTGGGCTTCTGGAATCTGGTCTGAAGAGTAGTGG
SsNGAL1_SP5	CCCAGAAACTAGAGTTCGTAGTATTAG
SsERF1_SP1	GCGGAGAACTCACGTATTCCGTTTTCACCG
SsERF1_SP2	GGCTCAATTTTACATCTCGTCTTGGGTTTCACG
SsERF1_SP3	GACCGCAAATTACCATANNNNNNNNATCATC
SsERF1_SP4	TTTAACGGCAAGTCTCCCCAAGTTTCCGTC
SsERF1_SP5	AAACGGCGTCCATCCATCATTAACC

doi:10.6342/NTU201901184

Supplementary Table S7 Primer list for the construction of dual-luciferase vector

Primer	Sequence
SsRL2_F_BamHI	GGATCCTTAGATCACAGGTTATAACCCGATC
SsRL2_R_Sal1	GTCGACGAGATCGACCAAGAAACCGAG
SsERF17_F_BamHI	GGATCCAGCTTGGTCGAGTGA
SsERF17_R_SalI	GTCGACTTTTTCATAGCCTCTGCAAA
SsERF3_F_HincII	GCAGTCGACTCTAGAGGGGAT
SsERF3_R_NcoI	CCATGGTTGGTACCTTTTGCTGAGC
SsCYCA_F_BamHI	GGATCCTCCCTGCAAGAACGTATAG
SsCYCB_F_BamHI	GGATCCTTCCTGCAAGAACGTATAG
SsCYC_R_SalI	GTCGACTTTTCTTTTTTGGGAGAGGG
SsCIB2_F_BamHI	GGATCCATTTCCACCACAAAGCT
SsCIB2_R_Sal1	GTCGACCATCAAGATTCCAAAAAACAA
SsNAGL1_F_BamH1	GGATCCCAGAGATGGTGTGGTCAC
SsNGAL1_R_Sal1	GTCGACCTCAGTGCGTGTAGTGTG
SsERF1_F_HincII	GTCGACGAATCACTGGGATAACT
SsERF1_R_NcoI	CCATGGGAAGAATTGATCAATTGAAGTAA

doi:10.6342/NTU201901184

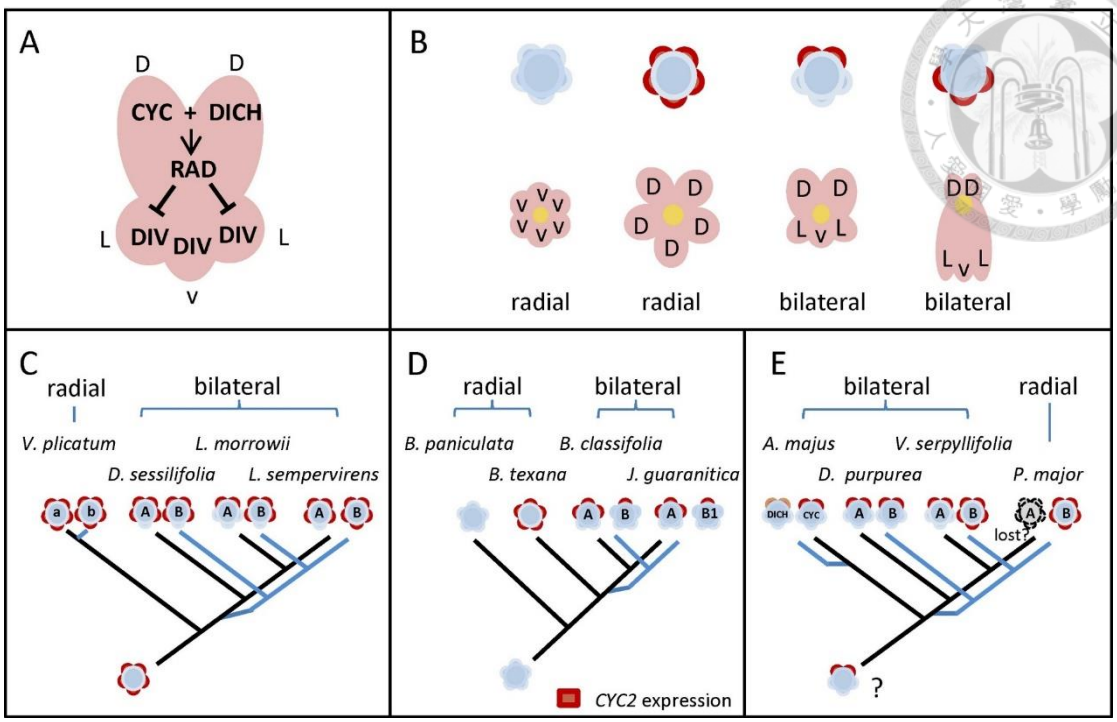
Supplementary Table S8 RNA-seq and BLASTx annotation result of dorsal expressed genes encoding cell wall loosening agent

Gene ID	dorsal replicate 1 (RPKM)	dorsal replicate 2 (RPKM)	ventral replicate 1 (RPKM)	ventral replicate 2 (RPKM)	Blastx Annotation
Sispe038Scf2587g00004	39.97	38.14	13.12	17.76	expansin-A10 isoform X1
Sispe038Scf1947g01015	25.66	29.69	11.82	14.44	expansin-A10
Sispe038Scf0517g00048	8.10	7.12	2.51	2.83	Expansin A1, ALPHA 1.2,EXPA1
Sispe038Scf0224g00031	75.27	87.38	30.65	35.95	expansin-A6-like
Sispe038Scf7657g00015	91.96	90.44	38.81	43.53	endoglucanase 17
Sispe038Scf1008g10041	1.89	1.81	0.81	0.94	xyloglucan endotransglucosylase/hydrolase protein 23

RPKM = Reads Per Kilobase per Million mapped reads

(Source: Pan, Z.J., unpublished work)

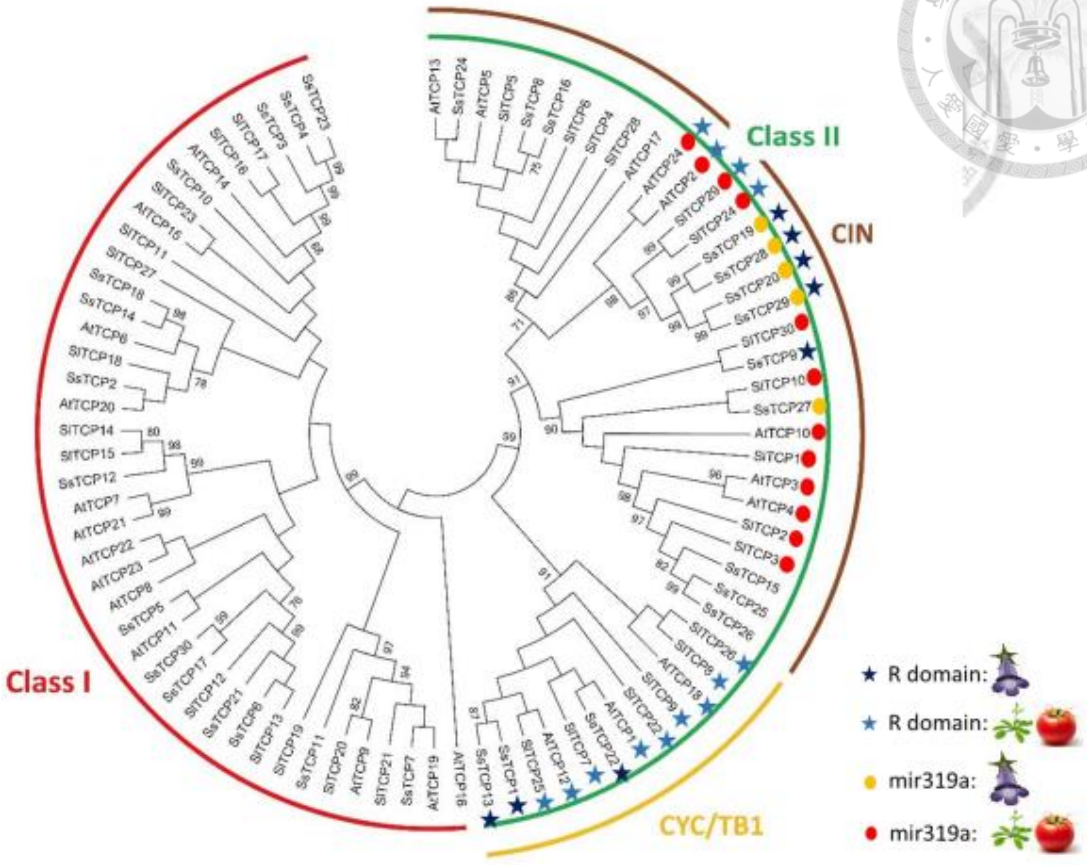
doi:10.6342/NTU201901184



Supplementary Figure S1 The role of CYC in floral zygomorphy

(A) Genetic regulation of floral symmetry in *Antirrhinum*. *CYCLOIDEA* (CYC*),* *DICHOTOMA* (*DICH*) and *RADIALIS* (*RAD*) specify dorsal petal identity, whereas *DIVARICATA* (*DIV*) determines ventral petal identity. *CYC* and *DICH* activate *RAD*, which in turn represses *DIV* activity in the dorsal and lateral regions. (B) Alteration in spatial CYC expression patterns can generate different flower symmetry. No *CYC* expression in any petals can create radial flowers, and ubiquitous *CYC* expression in all petals can also make radial flowers. Dorsal or ventral expression of *CYC* generates bilateral flowers. (C–E) Changes in *CYC2* expression coincide with flower symmetry evolution. (C) In Dipsacales, *CYC2* duplication and its expression change from ubiquitous to differential (only dorsal/lateral), coinciding with the radial to bilateral transition. (D) Similarly, in Malpighiaceae and its related basal families such Centroplacaceae and Elatinaceae, a progression of *CYC2* expression (absent- ubiquitous − dorsal/lateral − dorsal) plays an important role in the radial to bilateral transition. (E) When dorsal *CYC2* expression was lost or reverted to the ubiquitous status, radial symmetry was regained in Plantago major flowers. [*CYC2* expression in red, D; dorsal petal, L; lateral petal, v; ventral petal] (Spencer and Kim, 2018).

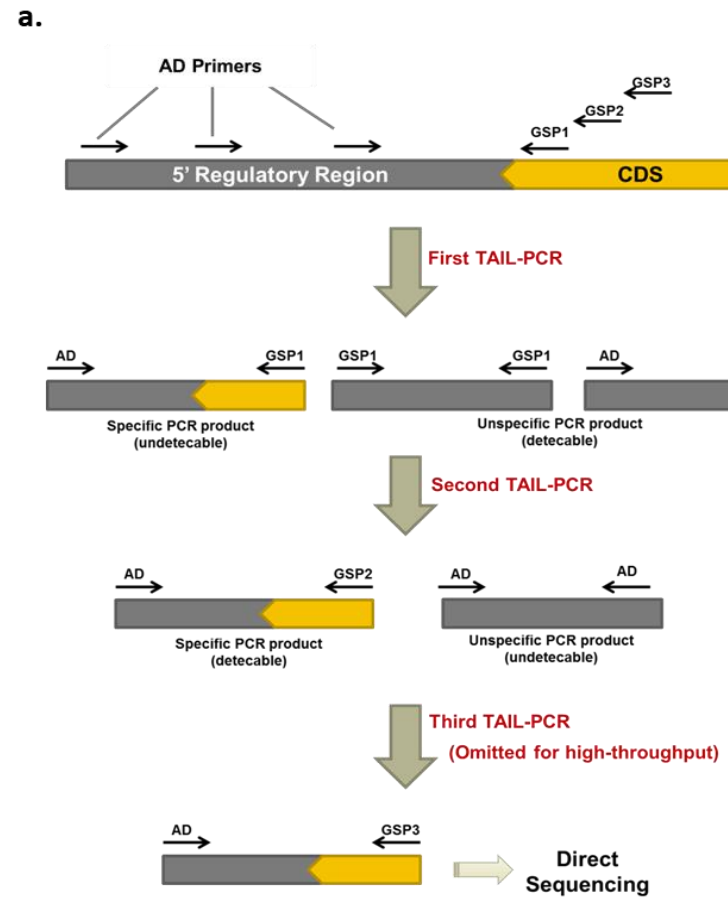
doi:10.6342/NTU201901184

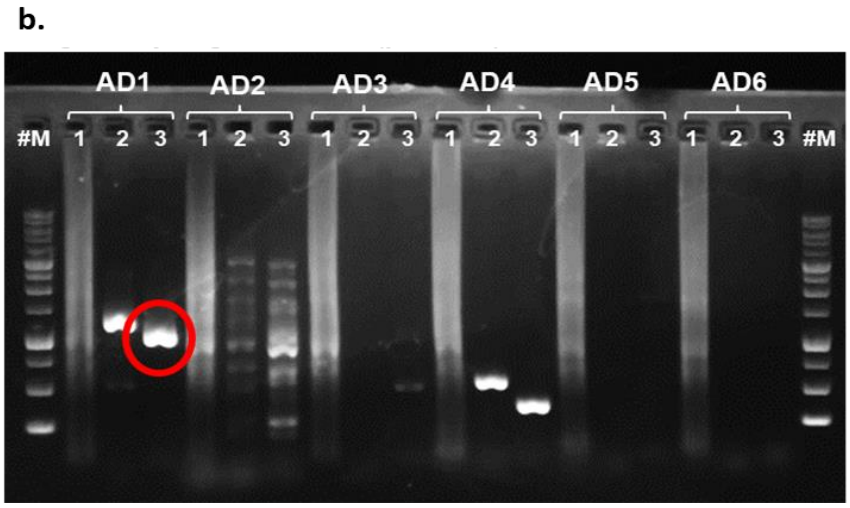


Supplementary Figure S2 Phylogeny of TCP proteins from *Sinningia speciosa*, *Solanum lycopersicum* and *Arabidopsis thaliana*

The phylogenetic tree was constructed using Neighbor-Joining method with 1000 bootstrap support indicated at each node. SsTCP: TCPs of *S. speciosa*; SlTCP: TCPs of *S. lycopersicum*; AtTCP: TCPs of *A. thaliana*. Proteins marked with stars have mRNA's containing an R domain. Those with circles containing putative target sites for miR319. *SsTCP22* is the *SsCYC*, *Antirrhinum majus CYC* orthologue which in the tree showed to have only a single copy. (Source: Ye, 2018, unpublished work).

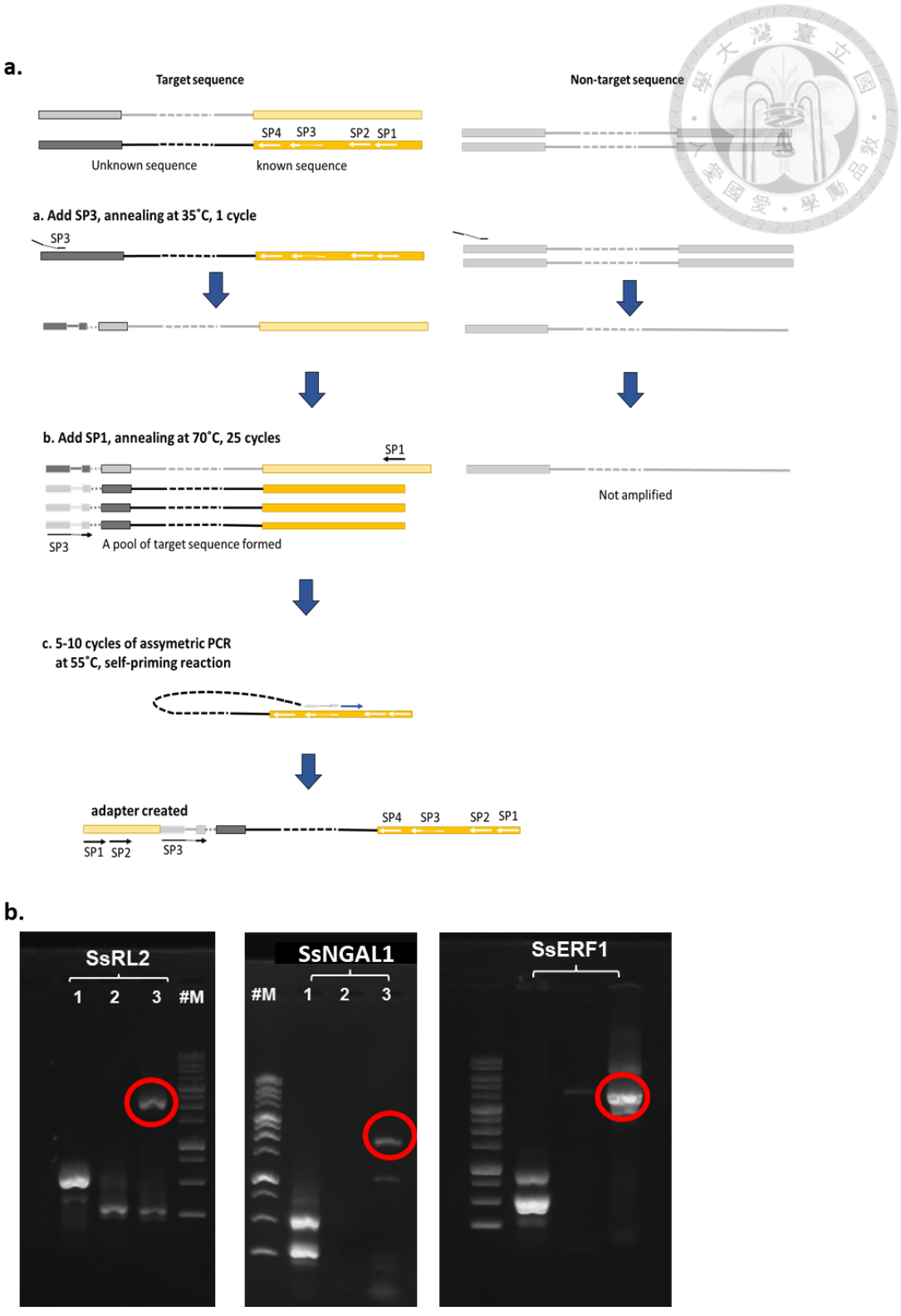
doi:10.6342/NTU201901184





Supplementary Figure S3 Thermal Asymmetric Interlaced PCR (TAIL-PCR)
(a.) Schematic representation of primer binding sites in TAIL-PCR and TAIL-PCR products. (b.) The TAIL-PCR result of *SsERF17* 5' regulatory region from first to third PCR reaction using each of 6 AD primers. The red circle shows the isolated band.

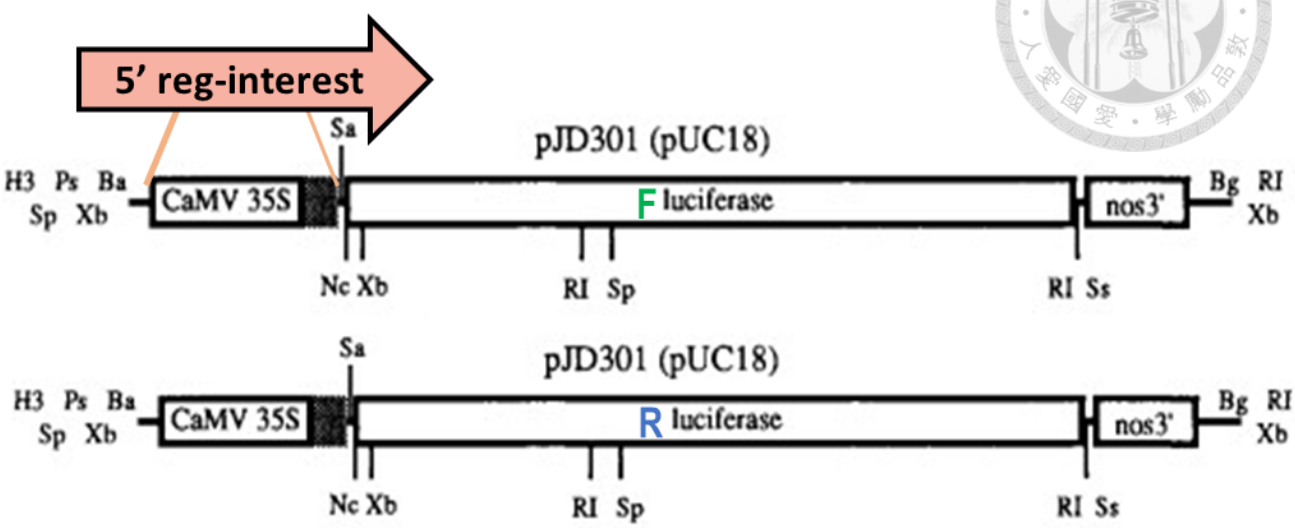
doi:10.6342/NTU201901184



Supplementary Figure S4 Self-Formed Adaptor PCR (SEFA-PCR)

(a.) Schematic representation of primer binding sites in SEFA-PCR and SEFA-PCR products. (b.) The SEFA-PCR result of *SsRL*, *SsNGAL1*, and *SsERF1* 5' regulatory region from the first to third reaction. The red circle shows the isolated band.

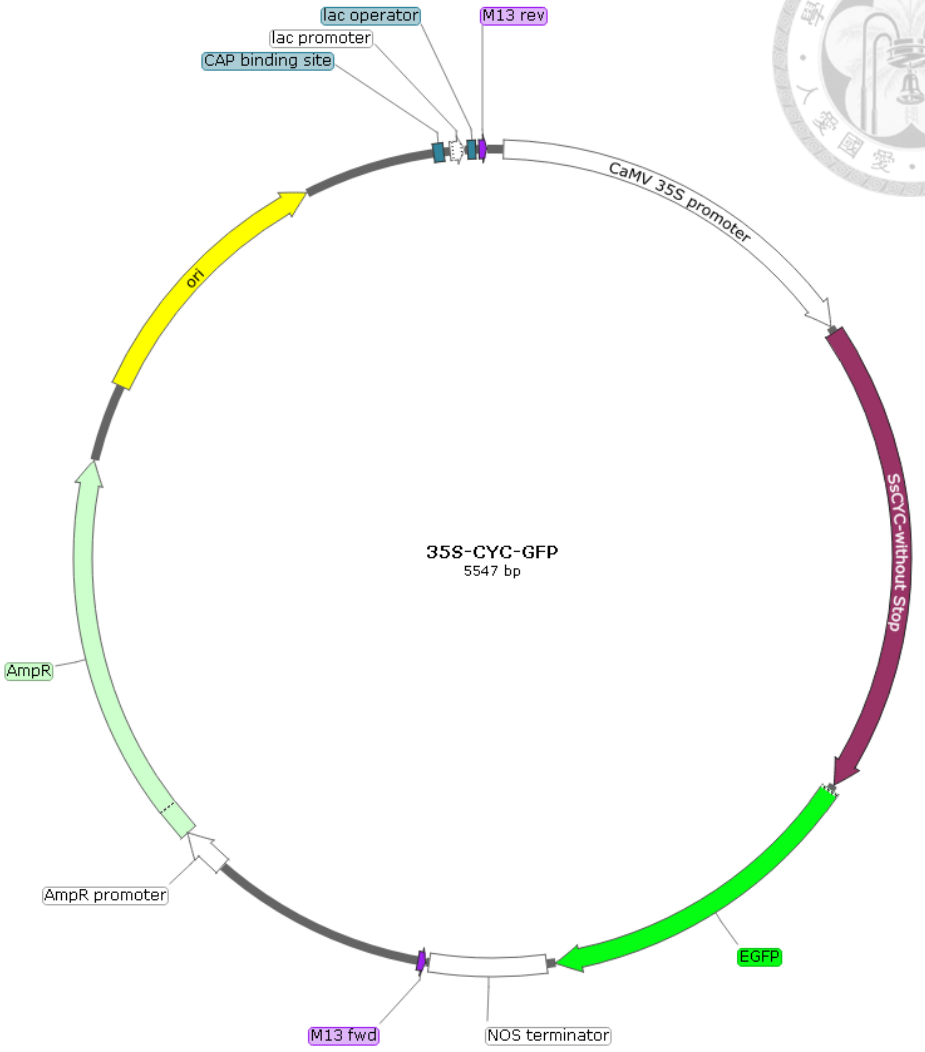
doi:10.6342/NTU201901184



Supplementary Figure S5 Reporter and internal control construct for dual-luciferase assay

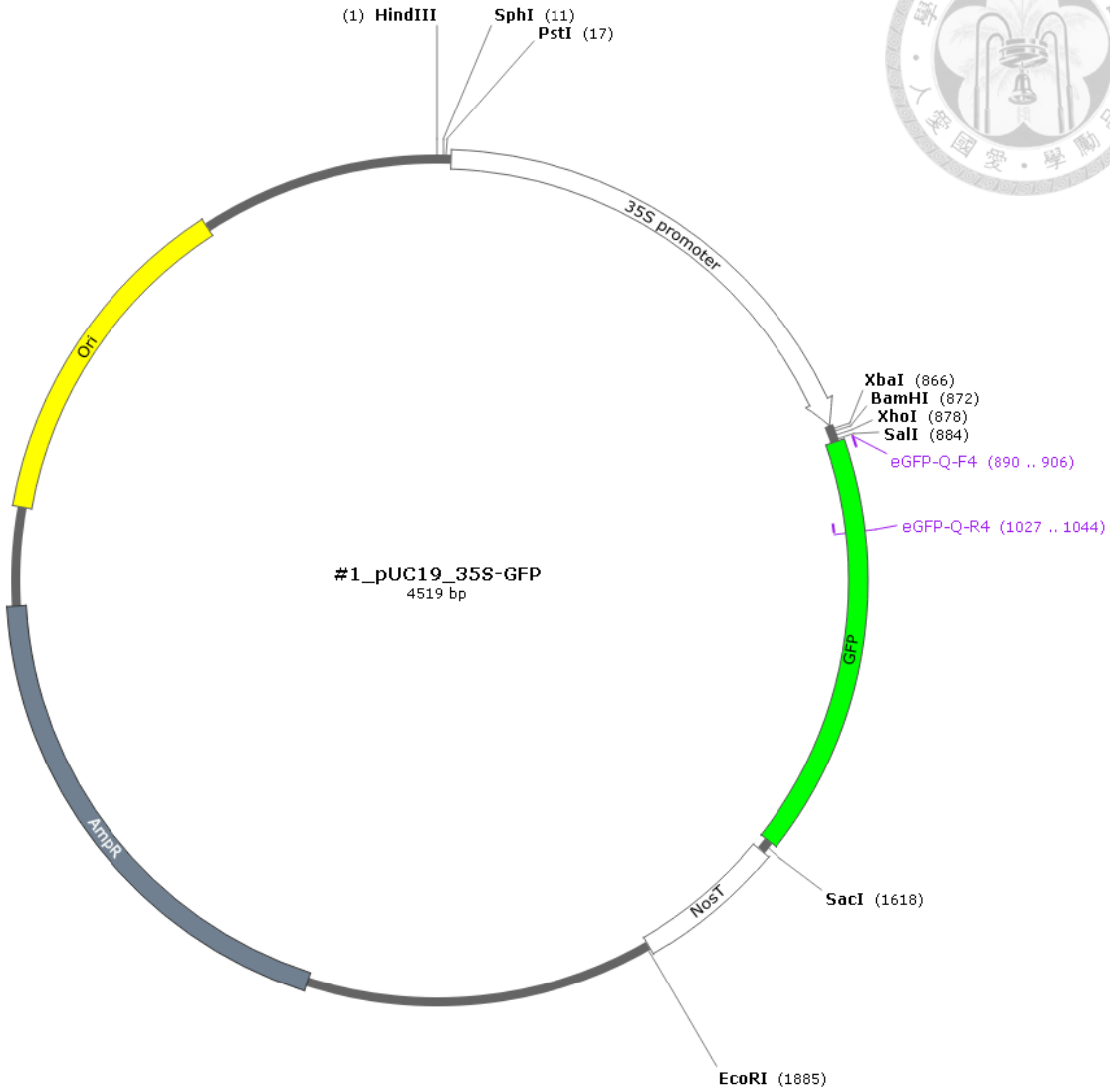
The PJD301-luciferase was constructed following Luehresen et al. (1995). (a.) Reporter construct in which the 5' regulatory region of interest was used to replace the CAMV 35s promoter to drive the firefly (F)-luciferase transcription. (b.) Internal control construct using renilla (R)-luciferase as the signal.

doi:10.6342/NTU201901184



Supplementary Figure S6　Construct of *SsCYC* tag *GFP* (by Yu-An, Shi)

doi:10.6342/NTU201901184



Supplementary Figure S7　Construct of *35s-GFP* (by Yu-An, Shi)

doi:10.6342/NTU201901184

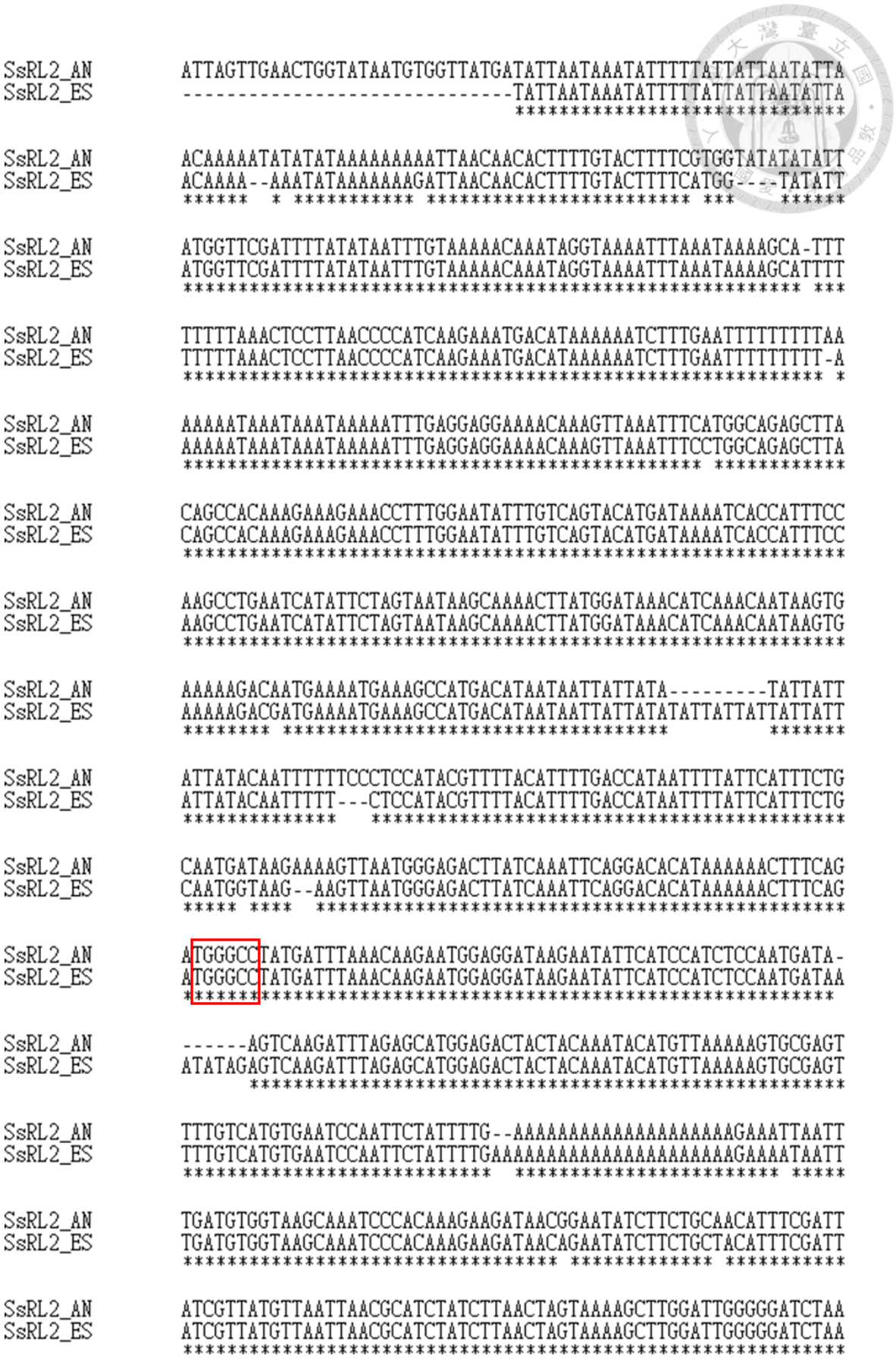
Supplementary Figure S8 Sequence alignment for the 5'regulatory region of *S. speciosa* 'Espirito Santo' and "Avenida Niemeyer'

Sequence Alignment was done by Clustal MUSCLE tool (http://www.ebi.ac.uk/Tools/msa/muscle/) to compare the 5'regulatory region of *S. speciosa* 'Espirito Santo' (SsES) and 'Avenida Niemeyer' (SsAN) of the TFs *SsRL2*, *SsERF17*, *SsERF3*, *SsOFP6*, *SsCYC*, *SsCIB2*, *SsNGAL1*, and *SsERF1*. The TCP binding sites found at the positive strand were marked with red box. The TCP binding sites found at the negative strand were marked with green box. The allelic variations of *SsCYC* regulatory region were marked with blue box.

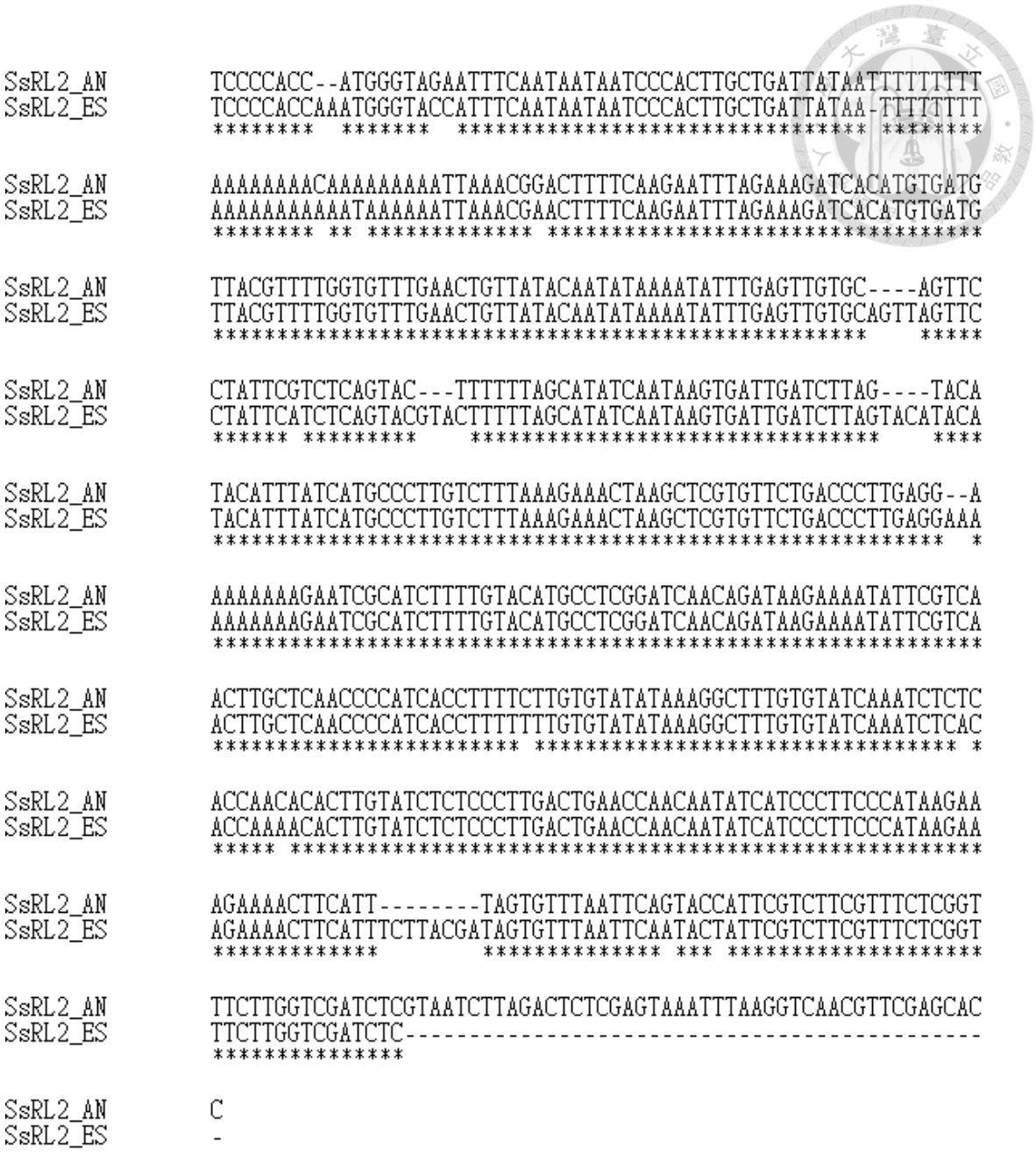
SsRL2

```
SsRL2_AN        CAAATCAAATGGTGAAAAATCAAATGGATACAAGATTGTTTTTTAAATGAGAGTTGGCTG
SsRL2_ES        ------------------------------------------------------------

SsRL2_AN        TGGGCCTGTACAATTTTTTTTTAAATGAGAGTGGACACAAATGAATGAGGCTCTACTATT
SsRL2_ES        ------------------------------------------------------------

SsRL2_AN        GAGTATGGACCTGTATAATTTTTTTTTAAATGAGGCACACCCTCAATTAAATATGGCTCT
SsRL2_ES        ------------------------------------------------------------

SsRL2_AN        ACCACTATATATATATATATATATATTTGGTCATAAAATAGAGTAATATTAAGTCTTT
SsRL2_ES        ------------------------------------------------------------

SsRL2_AN        TGTATGTGCTAAAACTATTTTCTACCAATCTTAAAATGTAATAAACGATTTAATTTGTGA
SsRL2_ES        ------------------------------------------------------------

SsRL2_AN        AATGAACTCAGATTAATTAGGAAGTTTAGATGTAAACTATGGATTATTACTCGAAATTAT
SsRL2_ES        -------------------------TTAGAT------CACAGGTTATAACCCGA------
                                         ******       *  * **** ** ***

SsRL2_AN        AACTAATGTTAACTTTATACTTTAAATATTTTAAAAGTCGTATTCGAATAACAACTTAAA
SsRL2_ES        ------------------------------------------------------------

SsRL2_AN        AGTTATAGTTTGACTATAACTCTACCTGCTAACTTTAGTCACGAGAATATTGATGCTTAT
SsRL2_ES        ----------------------------TCTAATTTC----------------------
                                            **** **

SsRL2_AN        ATCGAACCATTTAAAGTTGAATAGATAATTGAATCAATGGCATATTTGTTTATCGTTTTT
SsRL2_ES        ------------------------------------------------------------
```


doi:10.6342/NTU201901184



doi:10.6342/NTU201901184

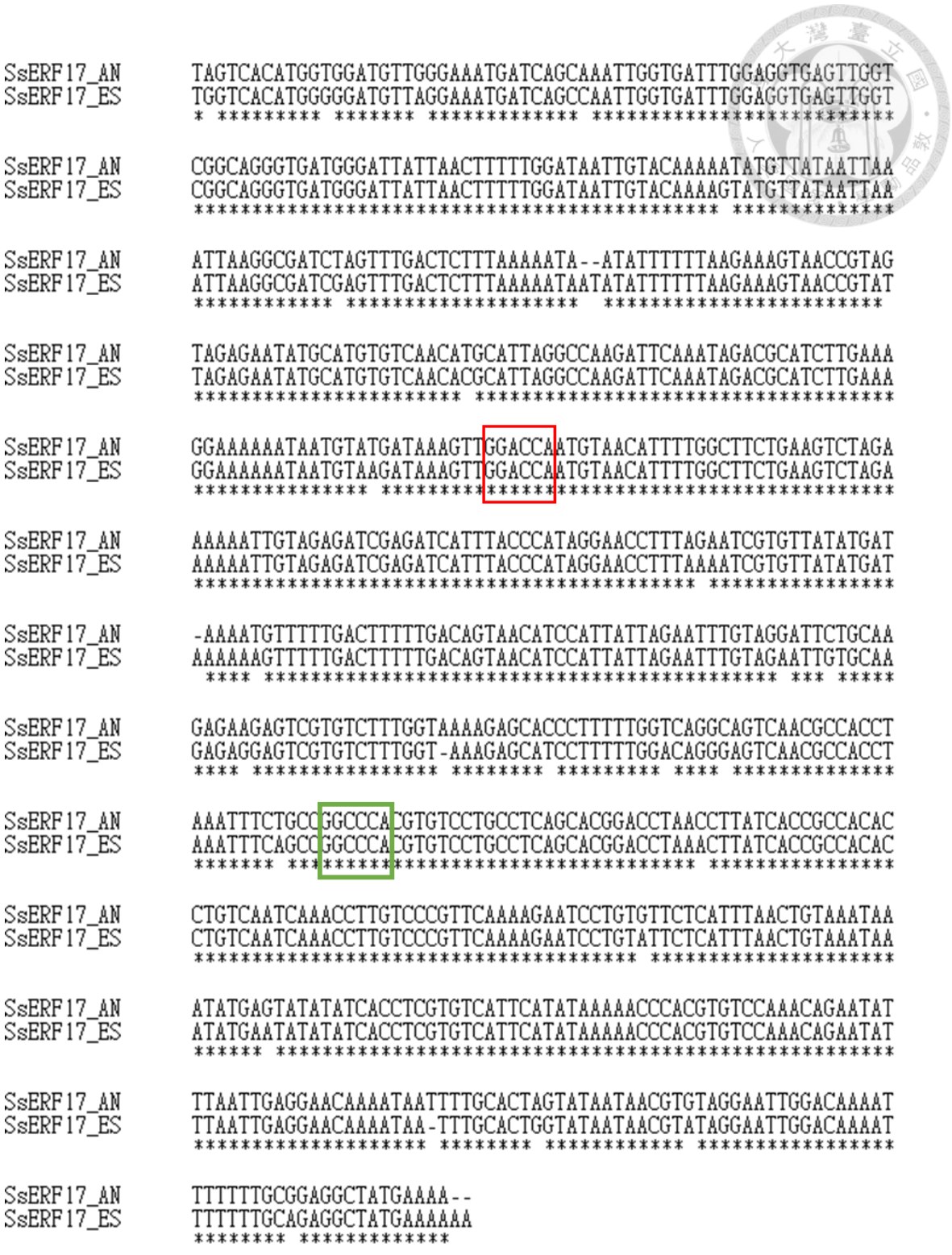


doi:10.6342/NTU201901184

SsERF17

```
SsERF17_AN      ATCTTCCAGGCTTGATAGGGTGTCGTGGCAAGTGAAAATATTCTCGAGCATCAAGTCATCC
SsERF17_ES      --------AGCTTGGTCGAGTG----------------------------------------
                        ***** * * ***

SsERF17_AN      ATGCTCGACTGGATACAGGATAGAGCCCTTTGGTCCCTTGCTCTAGCTTCTCTTGCAGCC
SsERF17_ES      ------------------------------------------------------------

SsERF17_AN      TTCTTCACTGCCTGCGAGGCTGCTGCCAGATCTTTTGGATCCTTGTACTCTTCCTCGATA
SsERF17_ES      ------------------------------------------------------------

SsERF17_AN      GCATCCCAGAGATCCTGTCCTCCGAGGAACGACTTCATCTGAAGGCTCCAGGTAGAGTAG
SsERF17_ES      ------------------------------------------------------------

SsERF17_AN      TTGTCATTAGTGAGTTTTGGAATTGGACATACTCCACTCCTCATTGTATTGTTTTGTGTA
SsERF17_ES      ------------------------------------------------------------

SsERF17_AN      AGACCTTAGCTCTGATACCACTTTGTTGGATCGAATTAATAAAGAAACACACACAAACAA
SsERF17_ES      ------------------------------------------------------------A
                                                                             *

SsERF17_AN      GACAAATATGGAGAAAAACAGAAAATCTTTATTCTGCACAAAAAACTCACGACTCACAGA
SsERF17_ES      GATGAA-----AGAAAACCAG----------------------------------------
                **  **     ****** ***

SsERF17_AN      CTCTCACACAGAGGATCACTCTCGTTTTGATCTCTCTATGTTCTCTCTCTTTTCTGAATA
SsERF17_ES      ------------------------------------------------------------

SsERF17_AN      CAAAAAACCAATGCCTACTGACCCTTATTTATAGGCTAACAATACAAGGAGGTTGAATCA
SsERF17_ES      ----------------------------------------------GAGGAGGATGGATCA
                                                              ****** ** ****

SsERF17_AN      TACTAGGAAATAAATCTAAAATAGATTTTTATCAAATCTAATCTTATCTTATGTAGATAA
SsERF17_ES      AAC-------------------------------------------------------
                 **

SsERF17_AN      TCTTTATCAAATCTAATCATAATTATCTAAATTTTATTTTTTTAAACAAATTCTAATTTA
SsERF17_ES      ------------------------------------------------------------

SsERF17_AN      ATAAGGACTCTGAATTTTGAATCAAAATTACTTTTTAAATAGTGATGGCTGTGGCAAAGG
SsERF17_ES      --------------------------------TCTCAACAGTGATGGCTGTGGCAAAGG
                                               * * ** **********************

SsERF17_AN      AATCGCAGGTCTGTCTACTGCAAGAGGAATGATTAAAGATCATATGGGCAATGAGCTTAT
SsERF17_ES      AATCCCAGTTCTGTCTACTGCAAGAGGAATGATTAAAGATCATATGGGCAATGAGCTTAT
                **** *** ************************************************

SsERF17_AN      ACTAGCCGACATGTCAAGCATCCAATTGGATTGAATTAATACCTTGAAACTGAACTTACA
SsERF17_ES      ACTAGCCGACATGTTTGGCATCCAATTGGATTGAATTAATACCTTGAAACTGAACTTACA
                **************   *********************************************

SsERF17_AN      AGCACAACTAAGTCATTGATTTTTCTTCTTAACCATTCATCAAATATCTGAGGCAGTTAA
SsERF17_ES      AGCACAACTAAGTCATTGAGTTTTCTTCTTAACCATTCATC-AATATCTGAGGCAGTTAA
                ******************* ********************** ******************
```


doi:10.6342/NTU201901184



doi:10.6342/NTU201901184

SsERF3

```
SsERF3_AN       TACTGTGCGGAAAAAATTGTGTACTGTCAGCCCACTTATCATTCCCCGATACTGAAAGAA
SsERF3_ES       -------------------------TCAGCCCACTTATCATTACCCGATACTGAAAGAA
                                         ********** ********** **********

SsERF3_AN       GTGCGTCGTAATTCGCCGAATATTTAACTATTCAATGTACCCAAAATAATATCGACGCGG
SsERF3_ES       GTGCGTCCTAATTCGCCGAATATTTTACTATTCAATGTACCCAAAATAATATCGACGCGG
                ******* ********** ********** **********************************

SsERF3_AN       ATCCCAACCTTGACTGCTAAATGATTATCCTTTTCTAATCAAA-ATATCATTAATTGGTT
SsERF3_ES       ATCCCAATCTTGACTGCTAAATGATTATCCTTTTCTAATCAAATATATCATTAATTGGTT
                ******* ********** ********** **********************************

SsERF3_AN       CAAATATTTGTCAATTCATTTGATTTATCAAATCAATCATAATCAAAATTCAAAATTAAA
SsERF3_ES       CAAATATTTGTCAATTCATTTGATTTATCAAATCAATCATAATCAAAATTCAAAATTAAA
                ************************************************************

SsERF3_AN       AACAAATCGAGTCAAAATATTAAATTGATTTAAATTATCGAACTGTAAAAAATATTCAAA
SsERF3_ES       AACAAATCGAGTCAAAATATTAAATTGATTTAAATTATCGAATTTTAAAAAATATTCAAA
                ******************************************** * **********

SsERF3_AN       GTATTTGGAATTAAATTAATAAAATTCAATCATTTTTTTTTATCACAGTCAAATATTTTT
SsERF3_ES       GTATTTGGAATTAAATTAATGAAATTTAATCA-TTTTTTTTATCA-----------TTTT
                ********** ***** ***** **********           ****

SsERF3_AN       TTCTTAAAAAAATTAAAGTTATGTTTTAAATTTGTTAGGAGAAATGAAAATTTGAGAATA
SsERF3_ES       TTCTTAAAAACATTAAAGTTATGTTTTAAATTTGTTACGAGAAATGAAGAATTGAGAATA
                ********** ********** ********** * **********

SsERF3_AN       CAACTTTTTATGGAATAAGGTTTCTTTGACTATCAAAATGATACGTGTAATTATCATGTT
SsERF3_ES       CAACTTTTTATGGAACAAGGTTTATTTGACTATCAAAATGATACGTGTGATTACCAAGTT
                ********** ******* ********** **** ** ***

SsERF3_AN       TATTATAATAAGTTTTTTTT-TTTTCAAAGTCAATATTTTTTATTGTCACACGTTAATTT
SsERF3_ES       TATTATAATAAGTTTTTTTTAAAAAAAAAGTCAATATTTTTTATTGTCACACATTAATTT
                ********** **********  ********** **********

SsERF3_AN       ATTATTATTTTTGTATGGATTATAAGAAAACAACTTACATGGATTATGCATCCAAATTCA
SsERF3_ES       ATTATTATTTTTGTATGGATTATAAGAAAACAACATACATGGATTATGCATCCAAATTCA
                ********** ********** ********** **********

SsERF3_AN       TTGAATGAAATGAAATGTTGTTTTTAGTAAAGAAGAAGAT-AATGTGACATGTGAAACAT
SsERF3_ES       TTGAATCAAATGAAATGTTGTTTTTAGTAAAGAATAAGATAAATGTGACATGTGAAACAT
                ********** ********** ********** **********

SsERF3_AN       TAGTCCAATTAAATGGATCCTTAAGTATGGATATAAATTATATAAATATCGTTACACTTC
SsERF3_ES       TAGTCCAATTTACTGGATCCTTAAGTACGGATATGAATTATATAAATATCGTTACACTTT
                ********** * ********** ********** **********

SsERF3_AN       AACTC-TTGTTTAAAAATTATAATTTTAATATATTTAATTTTTTTTTTTTCGTATCTTATG
SsERF3_ES       AACTCTTTGTTTAAAAATTATAATTTTAATATATTTAA-TTTTTTTTTTCGTATCTTATG
                ********** ********** ********** **********

SsERF3_AN       AAATGAAAAAAGAATAATTAACTGTAACTTGATGGTCGAGTAAAATGATAAAAAGTTAAT
SsERF3_ES       AAATTAAAAAAGAATAATTAACGGTAACTTAATGATCGAGTAAAATGATGAAAGATTAAC
                **** ********** ******* *** ********** ***  ****

SsERF3_AN       GTATTTTATGGCATGTATTGTATTGTATGTTCATAGTAGGCATGAGTTTTAAAATATAAT
SsERF3_ES       GTATCTTATGGCA-----CGTATTGTATGTCCGTGATACATGCGAGCTCTAAAATAAAAT
                **** ********      ********** * *  **     *** * ******* ***
```


doi:10.6342/NTU201901184

```
SsERF3_AN       TAT-----------------TTTTATTATTTTTAAAATTAATATAAAAAGACAGTTATTT
SsERF3_ES       TATTTCTGAAAAAAAATTAATTTTATTGTTTTTAAAATTAATATAAAAAGACAGTTATTT
                ***                 ******* ********************************

SsERF3_AN       TAAATTCTTGTACTTTAATTTAATGATATTAGATGAGACTAAATGTTTGAAGTTAAAAAA
SsERF3_ES       TAAACTCTTGTACTTTAATTTAATGATATTAGATGAGACTAAATGTTTGAAGTTAAAAAA
                **** ************************************************************

SsERF3_AN       CGTATAATTAAAGTCCTAAAAATGGCGTTACTAAAAAAAAAAAAAAAATTAATTTTCCCAA
SsERF3_ES       CGTATAATTAAAGTCCTAAAAATGGCGTTACT-------AAAAAAAATTAATTTTCCCAA
                ********************************       **********************

SsERF3_AN       AGATTTGCTAATTATTAATATTAAAGATTTTTTCACATTGTAAATCTTGATTTACACGCC
SsERF3_ES       AGATTTGCTAATTATTAATACTAAAGATTTTTTCACATTGTAAATCTTGATTTACACGCC
                ******************** ****************************************

SsERF3_AN       ATGCATTTAATGCCTCTCTCACTCCTTATCCTGCGCCCAAAACTTATACTGTAAATAAAT
SsERF3_ES       ATGCATTTAATGCCTCTCACACTCCTTATCCTGCGCCCAAAACTTATACTGTAAATAAAT
                ******************* ****************************************

SsERF3_AN       AAAATAATGGAGTGGAGAAATTAATAATAATAAATAAATAAAAAGAGAATAATCACAAAG
SsERF3_ES       AAAATAATGGAGTGG-GAAATTAATAATAATAAATAAATAAAAAGAGAATAATCACAAAG
                *************** ****************************************

SsERF3_AN       AGTCGAACCTGTCGGCGGAAGCAGAGGCAGCTATAAATAACCCTCAAAACCTCATTCCTA
SsERF3_ES       AGTCGAACCTGTCGGCGGAAGCAGAGGCAGCTATAAATAACCCTCAAAACCTCATTTCTA
                ************************************************************ ***

SsERF3_AN       CAGACCAACCAAAATATCCACGCGCTCAGCAAAAGGTAC--A
SsERF3_ES       CAGACCAACCAAAATATCCACGCGCTCAGCAAAAGGTACCAA
                ****************************************  *
```


doi:10.6342/NTU201901184

SsOFP6

```
SsOFP6_AN       AAAGAGAGCATCGCTATATTTGTGGCTGGTTTAGCAAATTCCTTTATACTTTGATTGGTA
SsOFP6_ES       ----AGAGCATCGCTATATTTGTGGCTGGTTTAGCAAATTCCTTTATACTTTGATTGGTA
                    ************************************************************

SsOFP6_AN       CAGTGTGGTTATTGTATTCAGGCAAT--------------------------------
SsOFP6_ES       CAGTGTGGTTATTGTATTCAGGCAATGAATACCAGATCAAAAGATGTGGTTATTGTATTC
                *************************

SsOFP6_AN       --------GAATACCAGATCAAAAGATGGATGGATTCAGCTGCGCTG----CATTTCCAA
SsOFP6_ES       AGGCAATGGAATACCAGATCAAAAGATGGATGGATTCAGCTGCGCTGCATACATTTCCAA
                        ****************************************    *********

SsOFP6_AN       AATTGTTGTGAAATTACCAGAATTGGCAGATCCTCCAGCCGGCATGAATGC--TACATTG
SsOFP6_ES       AATTGTTGTGAAATTACCAGAATTGGCAGATCCTCCAGCCGGCATGAATGCTATACATTG
                ************************************************  *******

SsOFP6_AN       AGTGTTTATAGAAACAAGATTTTACGGATCATGAAGAAATTAAGTTCTGAAAAAAAGAAG
SsOFP6_ES       AGTGTTTATAGAAACAAGATTTTACGGATCATGAAGAAATTAAGTTCTG-AAAAATGAAG
                ************************************************ ***** ****

SsOFP6_AN       ACTTTATTAAAAGATCTTGAAAAAGTCCCGATTCAAATACTAGACAGTGAAAGCTCAAAT
SsOFP6_ES       ACT---TTAAAAGATCTTGAAAAAGTCCCGATTCAAATACTAGACAGTGAAAGATCAAAT
                ***   ********************************************** ******

SsOFP6_AN       CAGCAAGGGATGGAACTATCAACGCTCAGAATCAAAGGGAAAACAAGGGGATTGCTGCAT
SsOFP6_ES       CAGCAAGGGATGGAACTATCAACGCTCAGAATCAGAGGGAAAACAAGGGGATTGCTGCAT
                ********************************** ************************

SsOFP6_AN       TAATCTCTGAGAGTATATGTGCAAGA--ATGTATCAAGTTTATTTACACC--TATTATTT
SsOFP6_ES       TAATCTCTGAGAGTATATGTGCAAGAATATGTATCTAGTTTATTTACACCTATATTATTT
                *************************  ******* **************  ********

SsOFP6_AN       CTTCATTAGTTAATAAAAATACAATGTGTCAAAAGGTATATATACCAATTTATATATATT
SsOFP6_ES       CTTCATTAGTTAATAAAAAAACAATGTGTCCAAAGG----TATACCAATTTATATATATT
                ******************* ********** *****    **********************

SsOFP6_AN       ATGAGATTTCACCTTTTGCTCAACTTTAATTTTCTCGACTTCAATCCTGAGTCAAAAAAC
SsOFP6_ES       ATGAGATTTCACCTTTTGCTCAACTTTAATTTTCTCGACTTCAATCCTCAGTCAGAAAAC
                ************************************************ ***** *****

SsOFP6_AN       GGCATTGTGAGATTCAAATTAAACTGGAAATATATATTTACACTATACTCGAATAGCCCC
SsOFP6_ES       GACATTGTGGGATTCAAATTAAACTGGAAATATATACTTACACTGTACTCGAATAGCCCC
                * ******* ************************** ******* ***************

SsOFP6_AN       GTATCAAATTAAACTTGTATAAAATCATGAGATTATAATAACCACACTTTCATGCTTTAA
SsOFP6_ES       GTATCAAATTAGAC---TATAAAATCATGAGATTATAATAACCACACTTTCATGCTTTAA
                *********** **   ********************************************

SsOFP6_AN       TTTTAAAAATTAAAATATCTTGAACTGACTTTTACAGATATAATGTGCATAGGTGACATA
SsOFP6_ES       TTTTAAAAATTAAAATATCTTGAACTGACTTTTACAGATATAATGTGCATAGGTGACATA
                ************************************************************

SsOFP6_AN       AAGAAGATTTCCATTTTCGATCCAAATCATAGCCATTGCTTAATAAATTATATCTTAATG
SsOFP6_ES       AAGAAGATTTCCATTTTTGATCCAAATCATACCCATTGCTTAATAAATTATATCTTAGTG
                ***************** ************* ************************ **

SsOFP6_AN       TAAAAGATTCGCATCTAATCATACCCCGCC--------------ATTGTCAATTGTCATC
SsOFP6_ES       TAATTGATTCGCATGTAATCATACCCCGCCATTGTCAATTGTCAATTGTCAATTGTCATC
                ***  ********* ***************             *****************
```

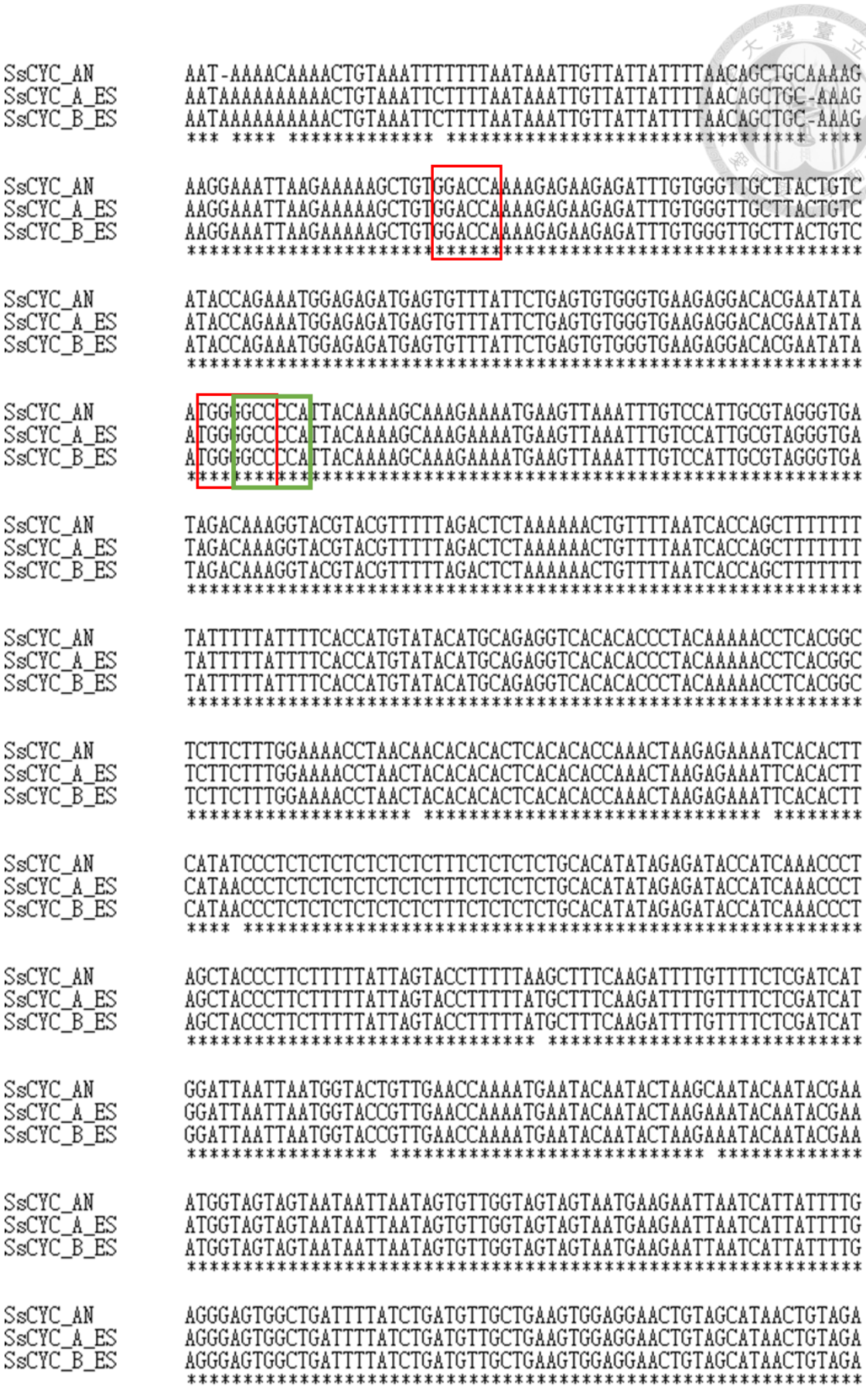

doi:10.6342/NTU201901184

doi:10.6342/NTU201901184

SsCYC

```
SsCYC_AN        TCCCTGCAAGAACGTATAGG-AATCTTTTAACACACATTCTAGACTCATAAGATCCATTA
SsCYC_A_ES      -CCCTGCAAGAACGTATAGG-AATCTATTAACACACATTCTAGACTCATAAGATCCATTA
SsCYC_B_ES      TCCCTGCAAGAACGTATAGGCAATCTATTAACACACATTCTAGACTCATAAGATCCATTA
                 ******************* ***** ***********************************

SsCYC_AN        AGAGGAAAATTACTGTACCAAGTATTTCAATTTCTTTCAATCTTACCCTTGAATTTTTTA
SsCYC_A_ES      AGAGGAAAATTACTGTACCAAGTATTTCAATTTCTTTCAATCTTACACTTGAATTTTT--
SsCYC_B_ES      AGAGGAAAATTACTGTACCAAGTATTTCAATTTCTTTCAATCTTACACTTGAATTTTT--
                ************************************************ ***********

SsCYC_AN        TTTTTTTTTTGAAGAAAAAAAATAATTGTATTTACTGGAAAAATG-AAGTAAATTAGGGT
SsCYC_A_ES      TTATTTTTTTGAAGAAAAAAAAAAATTGTATTTACTGGAAAAGTGAAAGTAAATTAGGGT
SsCYC_B_ES      TTATTTTTTTGAAGAAAAAAAAAAATTGTATTTACTGGAAAAGTGAAAGTAAATTAGGGT
                ** ****************** ****************** ** **************

SsCYC_AN        TCATTTATTCATACACAAATTTAATTTTCTTTCACCGAATCTGTATAAATGTATTATTCA
SsCYC_A_ES      TCATTTATTCACACACAAATTTAATTTTCTTTCACCGAATCTGTATAAATGTATTATTCA
SsCYC_B_ES      TCATTTATTCACACACAAATTTAATTTTCTTTCACCGAATCTGTATAAATGTATTATTCA
                *********** ************************************************

SsCYC_AN        CTTACCATCACCGGAAACCATGTCTTTTTTCTTCAACTGAAGTAACCCTTTTTTTTGT--
SsCYC_A_ES      CTTACCATCACCGGAAACCATGTCTTTTTTCTTCAACTGAAGTAACCCTGTTTTTTGTTT
SsCYC_B_ES      CTTACCATCACCGGAAACCATGTCTTTTTTCTTCAACTGAAGTAACCCTGTTTTTTGT-T
                ************************************************ ********

SsCYC_AN        ------------TTTTTCCATTCCACCAACTGATTGAAGCTGATCTGATCTGGTGTTCTT
SsCYC_A_ES      TTTTTTTTTTTGTTTTTCCTTTCCACCAACTGATTGAAGCTGATCTGATCTGGTGTTCTT
SsCYC_B_ES      TTTTTTTTTTTGTTTTTCCTTTCCACCAACTGATTGAAGCTGATCTGATCTGGTGTTCTT
                        ******* ******************************************

SsCYC_AN        ATAATAGAAGAGATGAAATTAAAGCAAACCTTAAACAAACAGTTCAACAAGAACTCAAAA
SsCYC_A_ES      ATAATAGAAGAGATGAAATTAAAGCAAACCTTCAACAAACAGTTCAACGAGAACTCAAAA
SsCYC_B_ES      ATAATAGAAGAGATGAAATTAAAGCAAACCTTCAACAAACAGTTCAACGAGAACTCAAAA
                ********************************** *************** ***********

SsCYC_AN        ACCCAGTCTTTAAATTGTATTTAAAAATAATTACCCACCAAAAAAAAAAAAAAGAAAAAGA
SsCYC_A_ES      ACCCAGTCTTTAAATTGTATTTAAAAATAATTACCCACC-AAAAAAAA-----AAAAAAA
SsCYC_B_ES      ACCCAGTCTTTAAATTGTATTTAAAAATAATTACCCACCAAAAAAAAA-----AAAAAAA
                ************************************************     ***** *

SsCYC_AN        AAAAGAATGCATCTGAAAATTAGTAATCATACATAGGGCAAGAAACCCACTCTTCAAGAA
SsCYC_A_ES      AAAAGAATGTATCTGAAAATTACTATTCATACATAGGGCAAGGAACCCACTCTTCAAGAA
SsCYC_B_ES      AAAAGAATGTATCTGAAAATTACTATTCATACATAGGGCAAGGAACCCACTCTTCAAGAA
                ********* ************ ** ***************** ******************

SsCYC_AN        TTAGGGTTTTCAGAAGTTTTC-TTAAAACAGGACGAGCACTAAACGTTTACAAACCCAAA
SsCYC_A_ES      TTAGGGTTTTCAGAAGTTTTCTTTAAATCAGGACGAGCACTAAATGTTTACAAACCCAAA
SsCYC_B_ES      TTAGGGTTTTCAGAAGTTTTCTTTAAATCAGGACGAGCACTAAATGTTTACAAACCCAAA
                ********************** ***** ***************** ***************

SsCYC_AN        TTATTCGCTCTTAATTCACACTTCAAAATGGTTCAAACCCAGAAACAAAAAAAAAAAAAA
SsCYC_A_ES      TTCTTCACTCTTAATTCACACTTCAAAATGGTTCAAACACAGAAACAAAAAAAAAAAA--
SsCYC_B_ES      TTCTTCACTCTTAATTCACACTTCAAAATGGTTCAAACACAGAAACAAAAAAAAAAAA--
                ** *** ******************************** ********************

SsCYC_AN        AATGGGGTTTCAAAAAAATTACACAAACATTCAGACTAAAAGCC-AAAAGAAAAAGGAAA
SsCYC_A_ES      ----GGGTTTCAAAAAAATTACACAAACATTCAAACTAAAAGCCAAAAAGAAAAAGGAAA
SsCYC_B_ES      ----GGGTTTCAAAAAAATTACACAAACATTCAAACTAAAAGCCAAAAAGAAAAAGGAAA
                    **************************** ********** ***************
```


doi:10.6342/NTU201901184



doi:10.6342/NTU201901184

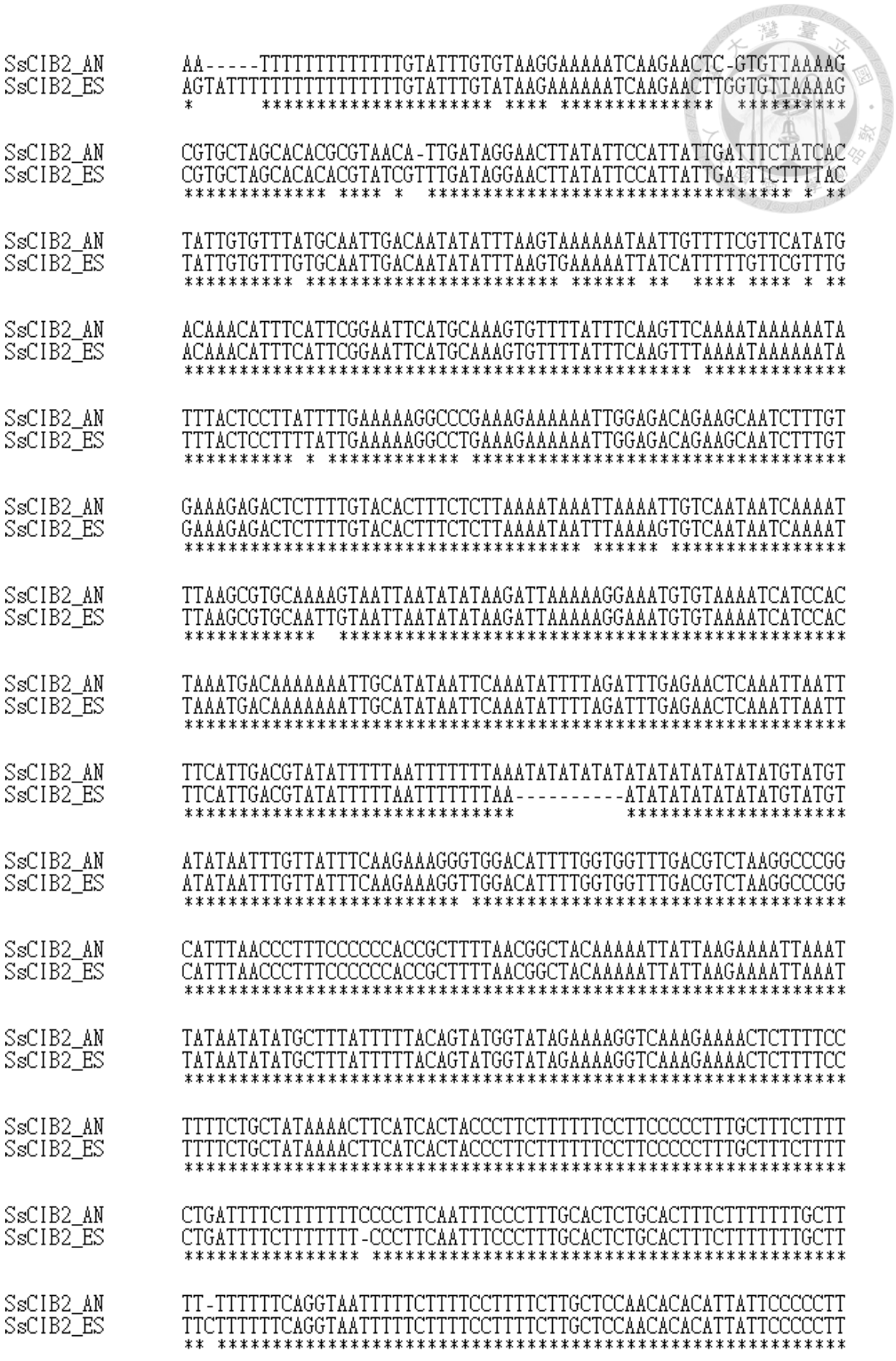
```
SsCYC_AN        AGGGAGTGGCTGATTTTATCTGATGTTGCTGAAGTGGAGGAACTGTAGCATAACTGTAGA
SsCYC_A_ES      AGGGAGTGGCTGATTTTATCTGATGTTGCTGAAGTGGAGGAACTGTAGCATAACTGTAGA
SsCYC_B_ES      AGGGAGTGGCTGATTTTATCTGATGTTGCTGAAGTGGAGGAACTGTAGCATAACTGTAGA
                ************************************************************

SsCYC_AN        TTACATTTTGAATTGACAATAAATTTTTGTACTGCGCTAAAAGTGAAGAAGATAAAGTTA
SsCYC_A_ES      TTACATTTTGAATTGACAATAAATTTTTGTACTGCGCTAAAAGTGAAGAAGATAAAGTTA
SsCYC_B_ES      TTACATTTTGAATTGACAATAAATTTTTGTACTGCGCTAAAAGTGAAGAAGATAAAGTTA
                ************************************************************

SsCYC_AN        AACTAGGTAGTTTTTTTTTATTATTATTATCACCAATTTAATACCCTATTCAGTGCATCTG
SsCYC_A_ES      AACTAGGTAGTTTTTTTTTATTATTATTATCACCAATTTAATACCCTATTCAGTGCATCTG
SsCYC_B_ES      AACTAGGTAGTTTTTTTTTATTATTATTATCACCAATTTAATACCCTATTCAGTGCATCTG
                ************************************************************

SsCYC_AN        AACAAATTTTATTTGGAGATTAAAGAAAGGGTACAATACTTTTACTCCTGAAAACCCAAA
SsCYC_A_ES      AACAAATTTTATTTGGAGATTAAAGAAAGGGTACAATACTTTTACTCCTGAAAACCCAAA
SsCYC_B_ES      AACAAATTTTATTTGGAGATTAAAGAAAGGGTACAATACTTTTACTCCTGAAAACCCAAA
                ************************************************************

SsCYC_AN        ATTTTTCCCAATTCATCATATCTTCGTCCTCCATTTTTCACCTACACGCTAGCCTTCCAG
SsCYC_A_ES      ATTTTTCCCAATTCATCATATCTTCGTCCTCCATTTTTCACCTACACGCTAGCCTTCCAG
SsCYC_B_ES      ATTTTTCCCAATTCATCATATCTTCGTCCTCCATTTTTCACCTACACGCTAGCCTTCCAG
                ************************************************************

SsCYC_AN        TCTTTCTCAGGCAAAGTCATTTTCTTTGGTGTAATATAAAGCAAAGACAAGAAAAATTTG
SsCYC_A_ES      TCTTTCTCAGGCAAAGTCATTTTCTTTGGTGTAATATAAAGCAAAGACAAGAAAAATTTG
SsCYC_B_ES      TCTTTCTCAGGCAAAGTCATTTTCTTTGGTGTAATATAAAGCAAAGACAAGAAAAATTTG
                ************************************************************

SsCYC_AN        CATATAACTATATATACACACACACATTTATCATCAATAATAAATAAGTGATGCTAGAGT
SsCYC_A_ES      CATATAACTATATATATACACACACATTTATCATCAATAATAAATAAGTGATGCTAGAGT
SsCYC_B_ES      CATATAACTATATATATACACACACATTTATCATCAATAATAAATAAGTGATGCTAGAGT
                ********** ***************************************************

SsCYC_AN        TATTGATTTCTTGAGGAAAAAAAAGAAA-AAAAACCTTAGTTCTCATTCTGGAGAAACCT
SsCYC_A_ES      TA-TGATCTCTTGAGGAAAAAAAAAAGATAAAAACCTTAGTTCTCATTCTGGAGAAACCT
SsCYC_B_ES      TA-TGATCTCTTGAGGAAAAAAAAAAGATAAAAACCTTAGTTCTCATTCTGGAGAAACCT
                ** **** ********** * * ********************************

SsCYC_AN        TCAAACCAGCTCTCATGACAGGTTGATTGCATAAACAATAAATATGGTTAAACAATTCAA
SsCYC_A_ES      TCAAACCAGCTCTC---ACAGGTTGATTGCATAAACAATAAATATGGTTAAAAAATTCAA
SsCYC_B_ES      TCAAACCAGCTCTC---ACAGGTTGATTGCATAAACAATAAATATGGTTAAAAAATTCAA
                ********** *********************************** *******

SsCYC_AN        GAACTTAAGGGTTTCTTTCCTTCTTTTTTTCTTTTTTATGTAAGAAATTAATTAGGGTTT
SsCYC_A_ES      GAACTTAAGGGTTTCTTTCCTTCTTTTTTTCTTTTTTATGTAAGAAATTAATTAGGGTTT
SsCYC_B_ES      GAACTTAAGGGTTTCTTTCCTTCTTTTTTTCTTTTTTATGTAAGAAATTAATTAGGGTTT
                ************************************************************

SsCYC_AN        ATTAACCCTTCTTCCCCTCCCCTCTCCCAAAAAAGAAAA
SsCYC_A_ES      ATTAACCCTTCTTCCCCTCCCCTCTCCCAAAAAAGAAAA
SsCYC_B_ES      ATTAACCCTTCTTCCCCTCCCCTCTCCCAAAAAAGAAAA
                ****************************************
```


doi:10.6342/NTU201901184

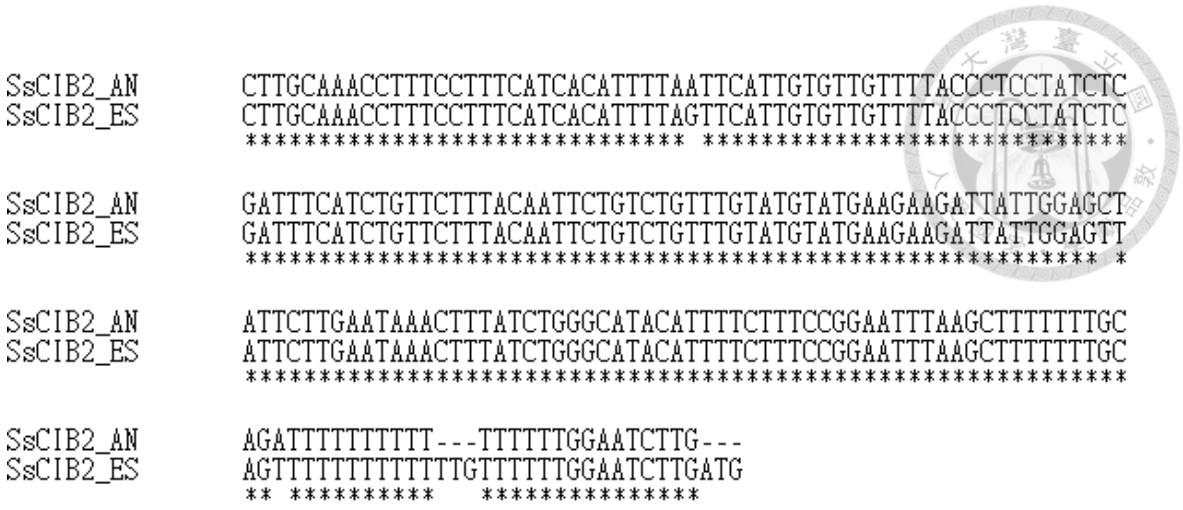
SsCIB2

```
SsCIB2_AN       TTAATAAAAATTTATCTCAAAAATTATTTAAATACAGTTTTCGTGATTCATTACCTATAT
SsCIB2_ES       ------------------------------------------------------------

SsCIB2_AN       TAATCTTGTTATTAAAAATTGAATTCCATCGCGCGTCACGTTTACTTAAATATTTATTAA
SsCIB2_ES       ------------------------------------------------------------

SsCIB2_AN       TCGTTATGCACAAAATTTTATTTCCACCACAAAGCTGGGAAGCATGTCTACAATATGTTA
SsCIB2_ES       ------------------ATTTCCACCACAAAGCTGGGAAGCATGTCTACAATATGTTA
                                   ****************************************

SsCIB2_AN       CAAATCTATTCTTGAAATTTACATTTTTA-TTATTTTTTATTAAAAAAAAAAAAAAAAAA
SsCIB2_ES       CAAATCTATTCTTGAAATTTACATTTTTATTTATTTTTTATTTAAAAAAAAAGAAAAAGA
                **************************** ************ ********* ***** *

SsCIB2_AN       AAAGAAAAACAAAGGGAAGTAAATTGTTATTTTCATTACTATGGGAATCCGACCAAAACC
SsCIB2_ES       AAAGAAAAACAAAGGGAAGTAAATTGTTATTTTCATTACTATGGGAATCCGACCAAAACC
                ************************************************************

SsCIB2_AN       CGATTTAATCGGTGAGCTTGGAATTTAATGCAATGATTGACTTACCCAAAACACGTTAAC
SsCIB2_ES       CGATTTAATCGGTGAGCTTGGAATTTAATGCAATGATTGACTTACCCAAAACACGTTAAC
                ************************************************************

SsCIB2_AN       GATTTTTTCCCTCCTCCCCAGAGAAAAATTTATATTTAAAGTCGCATTTAATTTATTCAT
SsCIB2_ES       GATTTTTTCCCTCCTCCCCAGAGAAAAATTTATATTTAAAGTCGCATTTAATTTATTAAC
                ************************************************************ *

SsCIB2_AN       ATTTGAGCCAAACGATAAATAAAAATAAAAAAAATTGAAAATTAACTGTCTACAATAAATTA
SsCIB2_ES       ATTTGAGCCAAACGAAAAATAAAAATAAAAAAAATTGAAAATTAACTGTCTATAATAAATTA
                *************** ********************************** *********

SsCIB2_AN       TCTTTTTATTTTATTATTGTTTGACTTTAACCTTAATGGGCAATTTTCAATTTGACCCAA
SsCIB2_ES       TCTTTTTATTTTATTATTGTTTGACTTTAACCTTAATGGGCAATTTTCAATTTGACCCAA
                ************************************************************

SsCIB2_AN       TCATTGAGGAATATAAAAATTGGAAAATAGATCTTACAATTTTAATTTAATGGGCTTTGG
SsCIB2_ES       TCATTGAGGAATATAAAAATTGGAAAATAGATCTTACAATTTTAATTTAATGGGCTTTGG
                ************************************************************

SsCIB2_AN       CCTATGATAATAATGTGAAATTCGAGGCCCAAAAATTAAGGTAAGTGTAAATTAAATTGT
SsCIB2_ES       CCTATGATAATAATGTGAAATTCGAGGCCCAAAAATTAAGGTAAGTGTAAATTAAA-TAT
                ****************************************************** * *

SsCIB2_AN       CAAGGAATAAAAATATAATATAAACCCTAGATTTAAAAAAAGGAATTAGGTAAGCATATC
SsCIB2_ES       CAAAGAATAAAAATATAATATAAACCCTAGATTTAGAAATAGGAATTAGGTAAGCATATC
                *** ********************************* *** ********************

SsCIB2_AN       ATCATTTATGAGATCATTAGCCTTATCTTTTTCCTCAAAATAGTATTTTTATAGTTTTGT
SsCIB2_ES       ATCATTTATGAGATCATTAGCCTTATCTTTTTCCTCAAAATAGTATTTTTATAGTTTTGT
                ************************************************************

SsCIB2_AN       TACGTACACTAATTAAAGTTTTTAGTATTTTGTTTAAGAAGAAATTTAAGGAAATTTTTT
SsCIB2_ES       TACGTACACTAATTAAAGTTTTTAGTATTTTGTTTAAGAAGAAATTTAAGGAAATTTTTT
                ************************************************************

SsCIB2_AN       TTATTATTTTTTTTTTGAATTTTACGGTCACATAATATTCGAATTTGATATTTTTATATAT
SsCIB2_ES       TT-TTA-TTTTTTTTGAATTTTACAGTCACGTAATATTCGAATTTGATATTTTTATATAT
                ** *** ***************** ***** ******************************
```


doi:10.6342/NTU201901184



doi:10.6342/NTU201901184

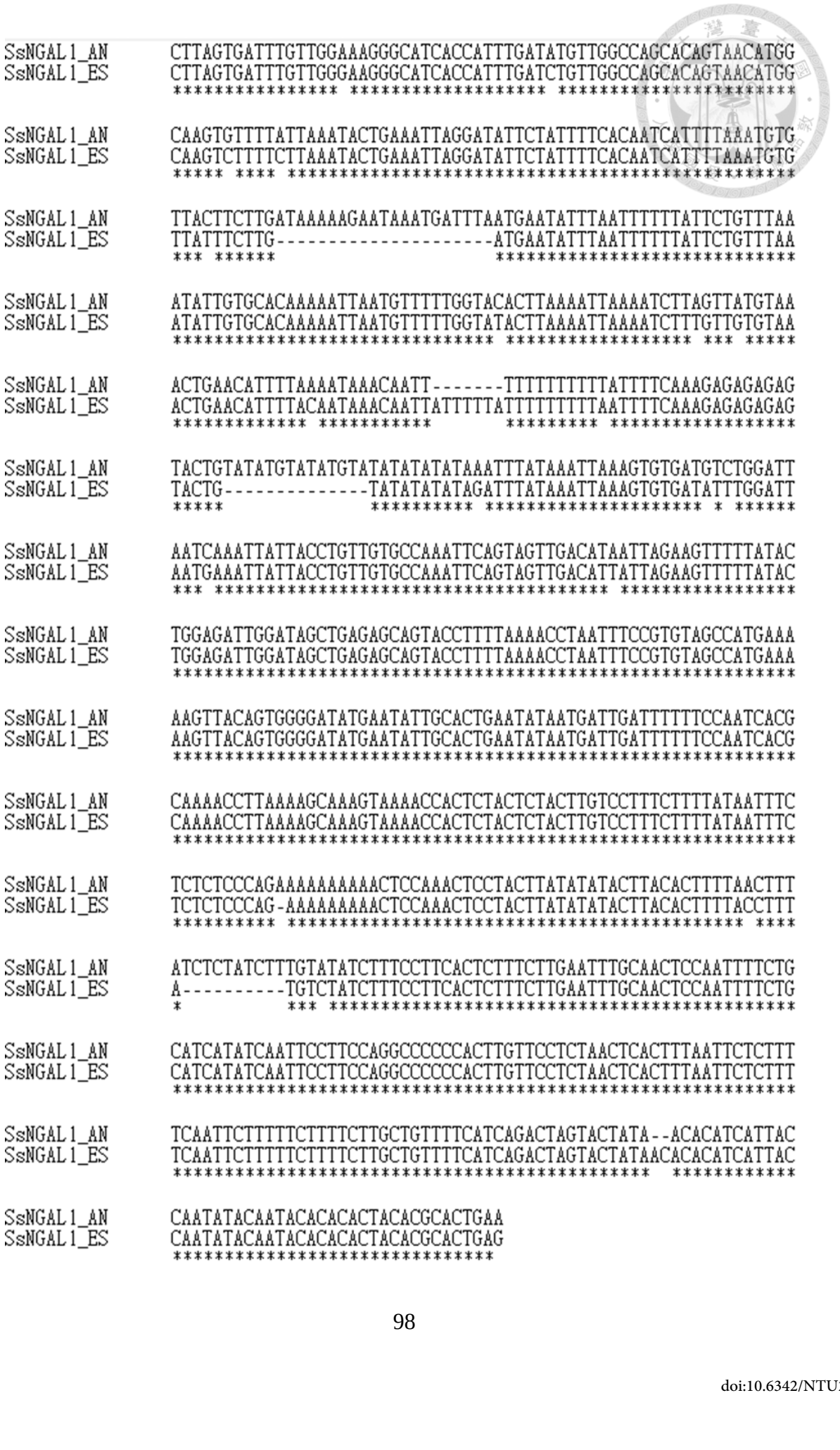


doi:10.6342/NTU201901184

SsNGAL1

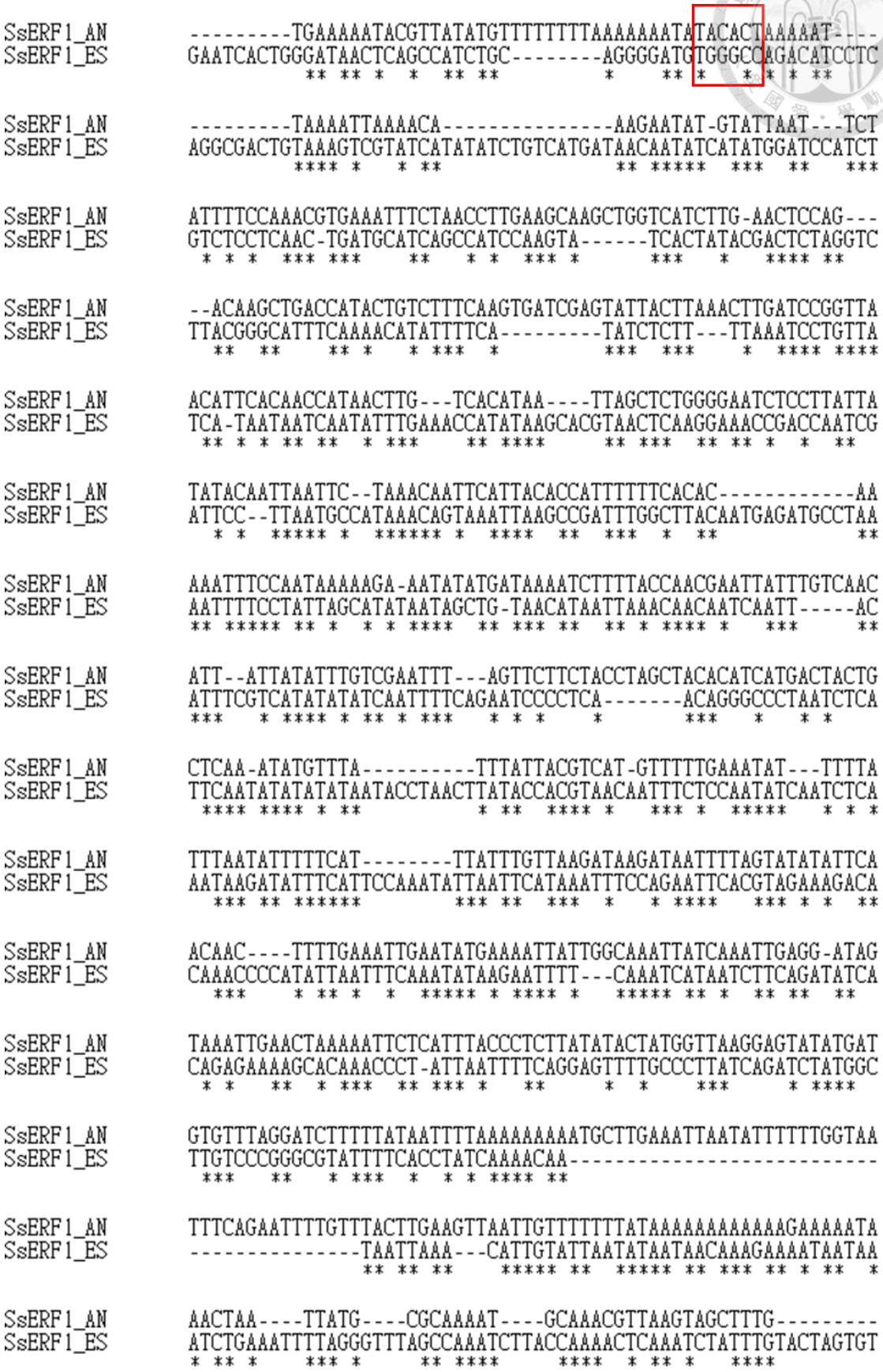
```
SsNGAL1_AN      AAATTAATTTTTAATTTATTAGATTTTTCCAGAGATGGTGTGGTCACAGGGAATCGGTTC
SsNGAL1_ES      ----------------------------CAGAGATGGTGTGGTCACAGGGAATCAGTTC
                                            *************************** ****

SsNGAL1_AN      CTAATATATAAAAAAATTAAGTACCCTCATGCAAAGGGAAACTACTCTTTAATCAAAGCT
SsNGAL1_ES      CTAATATATAAAAAAATTAAGTACCCTCATGCAAAGGGCAACTACTCTTTAATCAAAGCT
                ************************************ **********************

SsNGAL1_AN      TCAATTTTCTTCCATTTAGCCCACACCGATAGGCATAGCCCCACATACCAACTTTTGACA
SsNGAL1_ES      TCAATTTTCTTCCATTTAGCCCACACCGATAGGCATAGCCCCACATACCAACTTTTGGCA
                ********************************************************* **

SsNGAL1_AN      ACAATTTCTAAATAATAATAAATTTTTTTTTTTTTCTTCTAATTAAACTAGTACGTATAA
SsNGAL1_ES      ACAATTTCTAAATAATAATAATTCTTTTTTTTCTTCTTCTAATTAAATTAGTACGTATAA
                ********************* * ******** ************** ************

SsNGAL1_AN      TTTTAAATGTATGTTTAGAAAATAAAGTTGAAAAGCCAGCAGAATTTGAATTTTTCAAGA
SsNGAL1_ES      TTTTAAATGTATGTTTAGAAAATAAAGTTGAAAAGCCAGCAAAATTTGAATATTTAAAGA
                ****************************************** ********* *** ****

SsNGAL1_AN      CACCATTTCAATAGTCTGTTATGTAATAATTTTTTATTTTTTAAAATAAAAAATATAGTG
SsNGAL1_ES      CACCATTTCAATAGTCTGTTACGTAATAATTTTTTATTTTTTAAAATAAAAAATATAATG
                ********************** ********************************** **

SsNGAL1_AN      ACACAGTTGTTTCGCAATTATTTTTAAAAAAATAAAAAAATTACAAAACAAACTTAAAAA
SsNGAL1_ES      ACACAGTTGTTTCGCAATTATTTTTTAAAAAATAAAAAATTTACAAAACAAACTTAAAAA
                ************************** ************* ********************

SsNGAL1_AN      TAACTTATTTTATAACAAAACTTAAAATCT-TTTTAAAACTTTTATG-CATGTTATTATT
SsNGAL1_ES      TAACTTATTTTATAACAAAACTCAAAATTTGTTTTTAAATTTTTATGTTACCCTATTATT
                ********************** ***** * **** *** *******  *   *******

SsNGAL1_AN      CAAAAATATAATTAAACAAAAATTATCAATTTTTTAAAATAAATTTTGATAAATAACTGT
SsNGAL1_ES      CAAGAATATATCTAAACAAAAATCATCAATTTTTCAAAATAAAATTTGATAAATAATTGT
                *** ******  *********** ********** ******** ************ ***

SsNGAL1_AN      TT-ATTACACAACTACGTAGATACATAATTTTGCATTCTAACCTTTTGAAAAATTTAATA
SsNGAL1_ES      TTAATTACACAACTACATAGATACATAACTTTGCATTCTAACTTTTTGAAAAATTTAATA
                ** ************* *********** ************* *****************

SsNGAL1_AN      AAAGATTAATTAAAAACCATTTTTAATTATTGTATTTTACATATTAAGCAAAAACACATT
SsNGAL1_ES      AAAGATTAATTAAAAACCATTTTTAATTATTGTATTTTACATATTAAGCAGAAATACATT
                ************************************************ *** *****

SsNGAL1_AN      TAAACATACCCTTAATCTAC-GTATAAGTACTGAAAATAAATAAACTATAGCCATTTTTA
SsNGAL1_ES      TAAACATACCCTTAATCCACGGTATAACTACTGAAAATAAATAAACTATAGCTATTTTTA
                ****************** ** ****** ************************ *******

SsNGAL1_AN      TGTATGGAGTAAAAATTAATTAATATCATGTACTTTGCCACATTAATTAATTTGAACTTA
SsNGAL1_ES      CGTATGGAGTAAAAATTAATTAATATCATGTACTTTGCCACATTAATTAATTTGAACTTA
                 ************************************************************

SsNGAL1_AN      TTTTGCATTACACATATTGGGAAATTTAAATTTGAAGTATTTAATTAATAGTATTTAAAA
SsNGAL1_ES      TTTTGCATTACACATATTGGGAAATTTAAATTTGAAGTATTTAAGTAATAGTATTTAAAA
                ******************************************** ***************

SsNGAL1_AN      AAAGTAGAGTTAGAGACCATAGACTAAAAATTAAGAATTAATTGTGAATTAATAGGCAGC
SsNGAL1_ES      AAAGTAGAGTTAGAGACCATAGACTAAAAATTAAGAATTAATTGTGAATTAATAAGCAGC
                ********************************************************* *****
```


doi:10.6342/NTU201901184

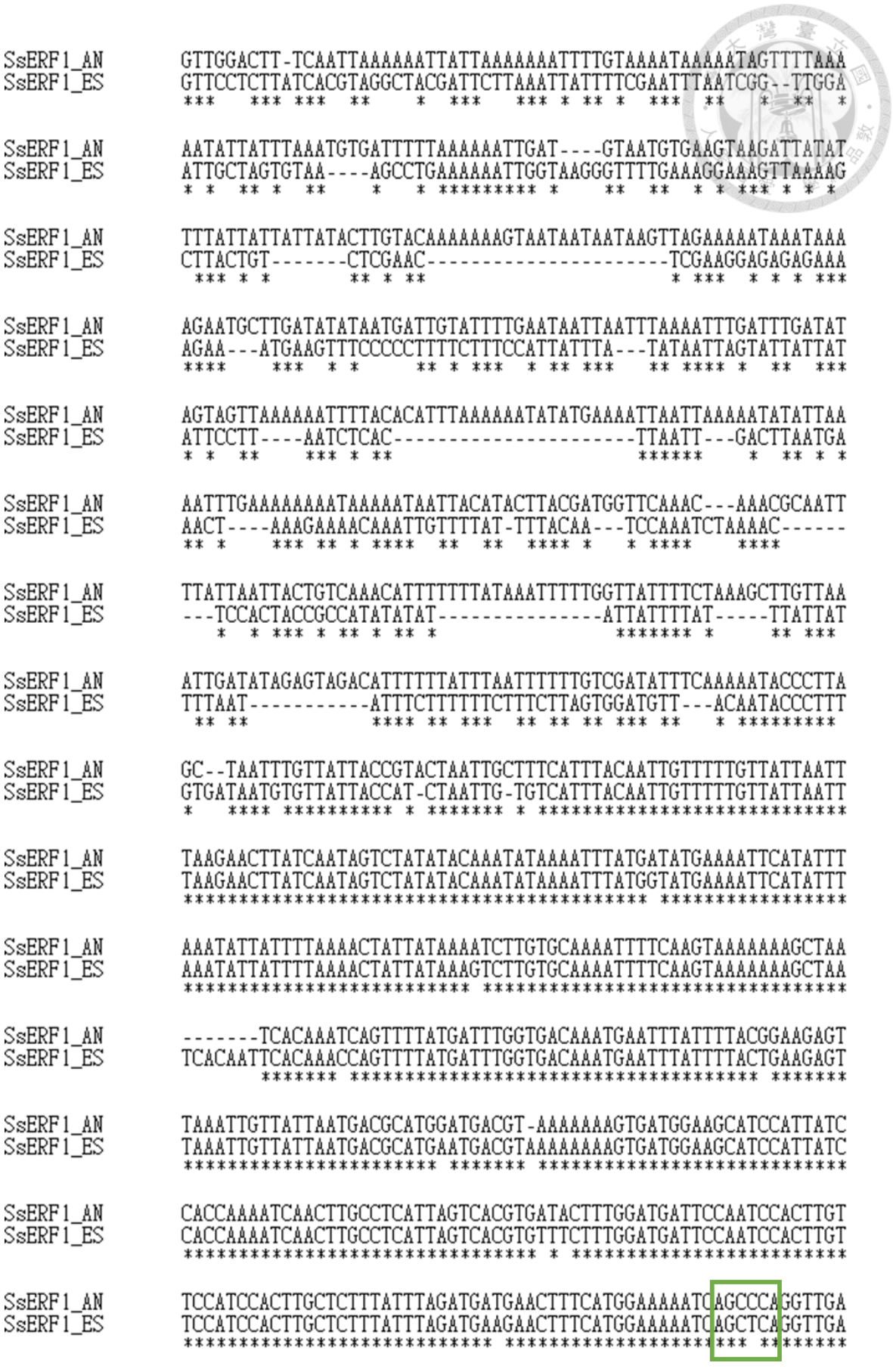


doi:10.6342/NTU201901184

SsERF1



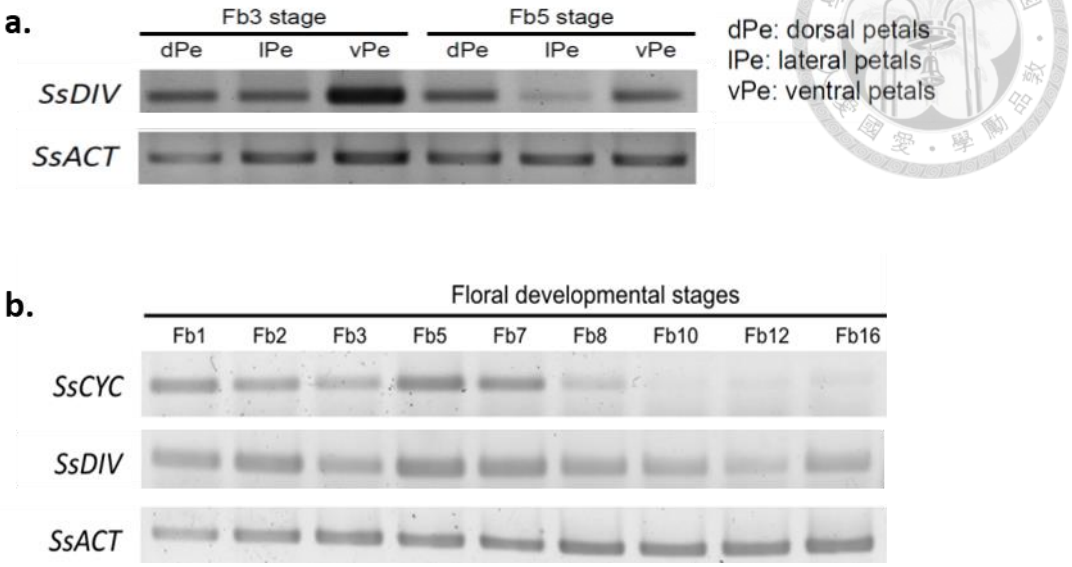
doi:10.6342/NTU201901184



doi:10.6342/NTU201901184

```
SsERF1_AN       TTATTATTAATTCACAATAATTTAATATTTTAAAATCTTTTTATGGTGCATGCACCCTTT
SsERF1_ES       TTATTATCAATTCAAAATAATTTAATATTTTAAAATC-TTTTATGGTGCATGCACCCTTT
                ******* ****** ********************** **********************

SsERF1_AN       TCTCACATCTTTTATCTTATATAAACCCCCTCCTAACATTTATCCCAAAAATTCAGAAAG
SsERF1_ES       TCTTACATCTTTTATCTTATATAATCCCCCTCCTAACTTCTATCCCAAAAATTCAGAAAG
                *** ******************** ************ * ********************

SsERF1_AN       TTCGATCAATTCATTTCTCTCTCTATAAATCTTACAAATTCAAGAAAATCCAAGTTTTTG
SsERF1_ES       TTCGATCAATTCATTTTTCTCTCTATAGGTCTTACAAATTCAAGAAAATCCTAGATTTTG
                ***************** **********  ********************** ** *****

SsERF1_AN       GAAAGCCCTTTAATTTTTTTGATCCGAAGCTTCTACAATTTCAAAAAAAAAAAAGAAAGA
SsERF1_ES       GAAAGCCCTTTAATTTTTTTGATCCGAAGCTTCTGCAATTTCCAAAAGACAAAAAAAAGA
                ********************************** ******* **** * **** *****

SsERF1_AN       AACCCAAGAA-------TTTTGTTGAATCCTTTTGGCTTAAAT-----------------
SsERF1_ES       AAAAAAAAAGCAAGTATTTTTGTTGAATCCTTTTGGCTTAAATTGATCAAAGACTGTCAC
                **   ** *        ***************************

SsERF1_AN       AGTTTTAAGAAAATCCAAGATTTTTGAAAAGCAAGCCCCTTTTGCTTGGATTGATCAATT
SsERF1_ES       GGTTTTAAGAAAATCCAAGATTTTTGGAAAG-AAGCCCCTTTTACTTCAATTGATCAATT
                 ************************** **** *********** ***  ***********

SsERF1_AN       CTTC
SsERF1_ES       CTTC
                ****
```

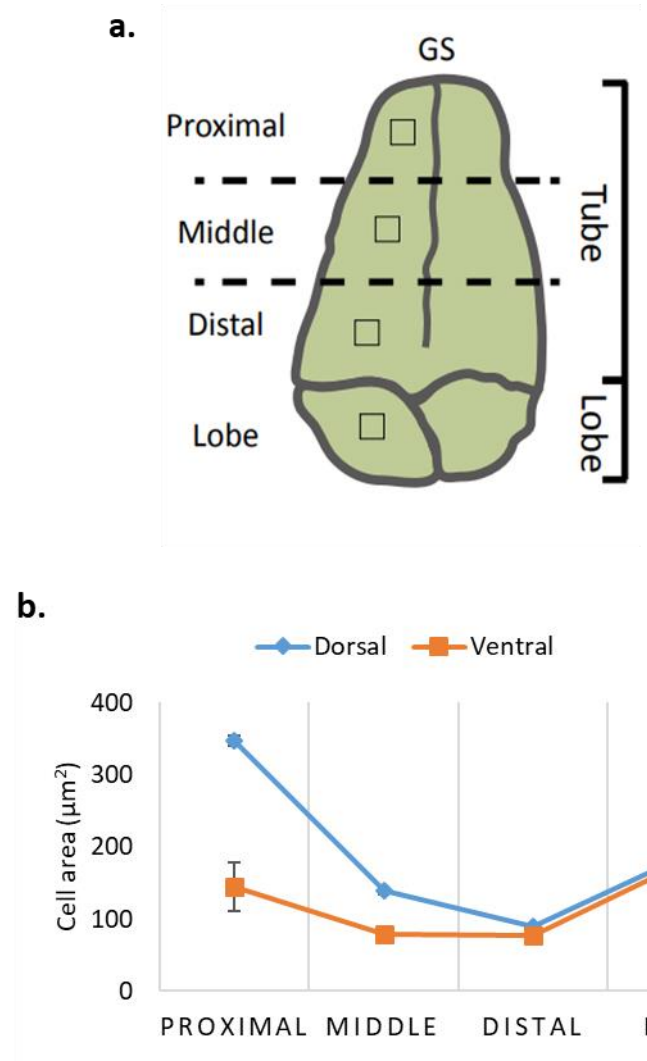

doi:10.6342/NTU201901184



Supplementary Figure S9 Expression profile of *SsDIV* in *S. speciosa* 'Espirito Santo' (SsES)

(a.) *SsDIV* expression in the floral bud stage 3 and 5 of SsES. (b.) *SsDIV* expression during the floral developmental stages of SsES. (Source: Pan Z.J., unpublished data).

doi:10.6342/NTU201901184



Supplementary Figure S10 Dorsal and ventral petals observation of *S. speciosa* 'Espirito Santo' (SsES)

(a.) Sectioning of SsES flower bud stage 5 for scanning electron microscope. (b.) Comparison of total cell area within the proximal, middle, central and lobe of SsES flower bud stage 5, observed using SEM. (Source: Pan Z.J. and Nien, Y.C., unpublished data).

doi:10.6342/NTU201901184

Supplementary Table S9 Full coding sequences of TFs containing TCP binding sites at their 5' regulatory regions

The sequences were isolated from *S. speciosa* 'Espirito Santo'

> SsABF2

AGGTGTTTGGATTTGCTTTGCACATATCTTTGTAAAGCTGCTTTGAAATGGGGAGTAATTTGAA
CTTCCAGAATATGGGGAATGACCTGCCAGTGGAGGGAGACGGTGGTGGAAGGCTGCCATTTA
GTTTTCCCTTGACACGGCAGCCCTCAATCTATGCTTTGACCTTTGATGAGTTCCAAAGTACAAT
GGGGGGACTTGGGAAGGATTTTGGGTCAATGAACATGGATGAGTTGCTGAAAAACATATGGA
GTGCTGAGGAGACTCAGAACATTGCATCCATTAGTGGCTGTGGTGGTGGACAAGAAGGAGGT
GGCTATTTGCAGAGGCAGGGGTCATTGACTATTCCTCGGACGCTGAGCCTAAAGACTGTTGAT
GAGGTTTGGCGGGATATGTCGAAGGAGTTTGGTGGGGGAACGGACACTAGTGGTTGCACTGG
TGTTTTTAGTATGTCTCAGAGGCAACCGACTTTAGGGGAGATTACACTTGAGGAATTCTTGGT
GAGAGCTGGGGTTGTGAGGGAAGAGAGTCAAGTGGCTGGAAAGCCTAATACTGTTGGATATC
TTGATAATTTACCACCCTCCTCAAATAATTGGGATTTTGGATTTGGAAATCAGCAGGCTAATGG
GACCGGAGGCTTGATCAATGGTAGGATTGCTGAAAGTAGTAATCAGATTGCTATGCAATCTGC
TAAGTTACCATTGAATGTAAATGGGGTAAGATCTTCTGCACATCAGTCTGTGAGTCAACAGCA
ATCCGTCCAATCAACACAGCAGCAACAGCTCCTTCCAAAGCAACCTGCCTTGGCATATGCAGC
TCCAATAGGAGTTCCAAACAATTGCCAGTTGAGTAGTCCGGAGATTAGGGGTGGAATTGTGG
GGATTACTGATGCAACAATGAATAATACATTTGTCCAGAATACGGCATTACAGGGTGGAGGATT
GGGGTTGCTTGGTTTAGGAGCTAATGGTGTTGGTGTTGCAACAGGGTCTCCGGCAGTTTCATC
AGACGGGCTCATAAAGTGTAATGGAGATACCTCTTCTGTGTCACCACTCCCTTATGTGTTTAAT
GGTGGTTTACGGGGGAGGAAAATTACTGCTGTAGAGAAGGTCGTTGAAAGGAGGCATAGGAG
AATGATTAAAAACAGAGAGTCTGCTGCAAGATCACGGGCTCGAAAGCAGGCATACACTATGG
AGTTGGAAGCAGAAGTTGCAAAATTGAAGGAGGAAAACCAAGAATTGCAGAAGAAACAGG
CAGCATTGGTGGAAATCCAGAAGAATCAGGTTCTGGAGATGATGAACCAGCAGAAGAATGGG
ACTAAGAGGCAATGCTTGAAAAGGACACATACAGGTCCATGGTAAAGGATGTTGCAGGAGAT
ACATATATAGGTGTACGTAAATAGTCTATAGGGACATGTTTCTACTGTATATGAAAGAGAAATTA
GACCGAGTTGTACTGCATTTTGCAGTAGAATGCTCCTTACCACAAATC

>SsRL2

CGTCTTCGTTTCTCGGTTTCTTGGTCGATCTCACAATCTTAAACGCTCGAACAAATTTAAGGTC
ACCGTTCGAGCTCCATGTCATCTTCACGTGGATCGTCCTCTTCATGGACACCTAAGCAAAACA
AGCAATTCGAAGAAGCTCTGGCTATGTACGACAAGGATACACCCGACCGCTGGCATAACATAG
CCAGGGCGGTTACTGGTAAATCAGCAGAGGAAGTGAGAAGGCATTATGAGGCATTAGTCAAA
GACATTATGCAGATAGAAACTGATCAAGTTCCCATACCTAATTACAGAGTCATTGCCAACAGTG
GCAGAGCTTATGTCAGTGATCAGAGGCTTTTGAAGAATCTGAAGCTGCAGTGA

doi:10.6342/NTU201901184

TTTGACTCCATCTGTAGAATAAATTAAAAAAAACCTCAGTCCATTATCATTGGAATATGG

>SsERF17

AATGGTGAAACCGCAATCGAGAAAGGAATCGCGTGACGGGCACTACAAGGGTGTGAGGATG
AGGAAATGGGGCAAGTGGGTGGCAGAGGTTCGACAACCCAACAGTCGTGATAGGATTTGGTT
AGGCTCTTATGAGACGGCGGAGGAGGCTGCTCGAGCTTATGACGCCGCCGTGTTTTGCTTACG
TGGACCTTCGGTGACACTTAATTTTCCTGATGATCCACCTCATATACTAGCGGCGGATGAATTG
TCTCGTTCGCAGATTCAGGTTGCTGCGTCGAAGCACGCTCGCCGCATGAGACACTCTGCTGTG
CCGGATGCAGCCTCGGCGGAGTCGAAACAACCGGTGGTGGAAAATATGTTCTTTGTGGAGCA
TACGGAATTGGGTTCTTCATCTTTGGATCAATTTTTGGGTAGAGAGAGCTTTTTGGCAGCCGAT
GGGAGAGGCGGCGGAGGAACTGATACATCACCATGGATACTGAGGGGGTTTTGAATACATCA
CGCATATGGAATTATTG

>SsHB13

GCAGGTGGCAACAGTTTCATAGGTCATGACTTGTACTGAAATGGCATTCTTCCACTCCAATTTT
ATGCTACAAAATTCTCATGAAGATGATCACAATCAACCCTCCACTTCTCTTGCTCCAATTCTTC
CTTCTTGTAGCCCCCAAGAATTTCATGCTTCGCTATTAGGGAAGAGATCTTCCATGTCATTCTC
CATGGGAATCGACGTTTGCGAAGAGATGAATAATCATGGAGAGGATGAATTATCTGATGATGG
TTCACAACTCGGGGAGAAGAAGAGGAGGCTTAACATGGAGCAAGTGAAGACACTTGAGAAA
AACTTTGAGCTAGGCAACAAGCTTGAGCCCGAAAGGAAATTGCAGCTGGCCCGAGCACTTG
GCCTGCAGCCTAGACAGATTGCTATTTGGTTTCAAAACAGGAGAGCAAGATGGAAGACTAAA
CAATTGGAGAAAGATTATGAACTTCTTAAGAGACAATTTGAAGCTGTTAAATCAGAAAATGAT
GCACTTCAACTCCAGAATCAGAAACTTCATGCTGAGATAATGGCACTAAAGACTAGGGAGCC
AACAGAATCCATCAATCTGAACAAAGAAACAGAAGGTTCTTGCAGCAATAGAAGTGAAAAC
AGCTCAGAAATAAAGCTTGATATTTCAAGAACTCCTGCAATTGATAGCCCATTATTAACAAATC
CCACTACAAGCAGACCATTTTTCCCATCTTCACTCAGGCCAAATGGAGTGTCACAGCTCTTCC
AAAATGCTTCAAGGCCAGAAATTCAGTGCCCCAAAATGGACCAAACTGTTAAGGAAGAAAG
CTTGTGCAACATGTTTTGTGGCATGGATGATCAGACAGGATTTTGGCCATGGTTAGAGCAACA
GCATTTCAATTAAATTGCCTCAAGTTTGAGTAAAGATTTGTCTGGAGAAATGGTTAAAAAAAG
AAAAGAAAAAGAAAAACCCATTTGGGTGAGATCAAGAACAAGATGGGAMCACGAATCG

>SsMYBS1

AGTATGGGAGAGGAAATAGGAGTGGAATATTGGAGCAGAGAAGAAGAGAAAGCATTTGAGA
ATGGAATTGCAAAGCATTGGATTATTAAGGAAGACGAAGACAAAAACGAAAATGAATGTATTA
TTTGGAATAAGATTGCTTCAATGGTTCCCACTAAAAGCATTCAACAGTTGAAACATCATTATCA
ACTTCTAGTGGAAGATGTTCAAGACATTGAAGCAGGAAATGTTCCATTGCCAAAATATTCTTC
CCATCACTTCCAATCTCAACCTCCATTAATATCCACTAAGGATTCTTATCTTGAAAAGGATAAA

doi:10.6342/NTU201901184

AGATTCATCACTAACTGTAATTTTGAATCTGTCACTTCCTCTGGTAAAGGAGTAGGATCCTCCA
CCAGGTCAGACCAAGAGCGTCGAAAAGGAATTCCATGGACAGAAGAAGAACACAGGTTATT
TTTGCTTGGTTTGGACAAGTTTGGGAAAGGGGATTGGAGAAGTATTTCAAGAAACTTTGTCAT
TTCAAGAACTCCAACACAAGTGGCTAGTCACGCTCAGAAATATTTCATTCGTTTGAATTCCATC
AACAGAGATAGAAGGAGATCTAGCATACATGACATTACCAGTATCAATGGTGGGGATGTCTCA
TCTTCCACTCATCAACCTCCTCCTATTACAGGTCAACAGACACCAACTGCGATCAAACATCAC
AGAGCAAACGTGCAAGGATTAGGAATAATCTATGGTGGCGCACCGATGGGCCATCCTGTTACC
CTCCTCATGGAGGTAGTCACATTGCACCTGCAGTTGGCACTCCAGTCATGATTCCTCCAGGAC
ATCATCATCCCTCATATGTTCTTCCCGTTGCATACCCTCTGCCGCCGCCGCCACCACCGCACCA
ATAACAAACAAAACTGAAATTGCTGGAACATCAAGTTTTAGTCTTTTAGAGACACTAGTTTTG
AACTCTAGTTTATATGCAAATGGCTCATCTTCAACTAAAATTAGAGGCTGTTACATTGAACTGT
CTTG

>SsRVE1

AGAACTGATAGGTTCTGAGGCTATGGCCGTTCAGGAACAACACGGAGGCACAGGGTCTGATA
TTTCTCTGCCTGCTAGCAACAAAATTTCATTAGACAGTGGGGCACCTTCCATTATGGGTATCCA
GTTGAAACATCAGTTTAACTCAGAAGACGAGTTTACTCCAAAGGTGAGGAAACCTTACACCA
TTTCTAAGCAAAGAGAAAGATGGACTGAGGAAGAGCATGAAAAATTTCTAGAGGCGTTAAAA
CTTTATGGTCGGGCATGGAGGAGAATTGAAGAGCATGTAGGTACAAAAACTGCAGTGCAAAT
TCGAAGCCATGCACAGAAGTTTTTCTCCAAGGTTGCTCGTGAATCTTATGGAGATGATGTCAG
CTCTGCAAAACCAGTTGAAATTCCACCTCCAAGGCCCAAAAGAAGACCTATATACCCTTATCC
TCGGAAGCTGGTTCCTCCAGTTAAAACTGTTCTGAATCTATCTACTTCAGAAAAGGAGAATCA
ATCTCCTACATCCATATTATCTGCAATGGCTTCAGATACATCTGGTGGGACAGTTTCTTGCGCGC
CTAATGGCTGCTCCTCACCCATTCCATCTGTCGTTGCTCAAAATGATTGTGCAGTTTTTTGTTCT
GCAAAAGATGCCAAATCTTCCTTGCATTGCCATGAAGATGAAAGTTCAAGTCCAGATGAACA
AGTTCCTCTGCAATTGGAGATTTCATCTCAAAAGAATGCTCTTGTTGAAGAAGATTTAAATGA
ATCAGCCGCTCAACGTTTGAAGCTTTTCGGCAAGACATTATTAGTCACTGATTCTCACAGACA
ATCACATTCGACCCTGGGGACCTATGAGCAGGAGTCGTTAGATAGAAGTGAGGTCCCTTGTTT
ATTTCCTTGGAGTGTCATGCCTATAAAATTCCCTCCAAGCAATTCAGAGTATGTCTGGAGTTTG
TTCTCTCAGAGGTCGACTTCAGCAGTCATTTGCACTGAACTAAGGGGTGATATCTACTGCAGT
GAAATGATGAGTGATAAGTTGAATCCTATAAATGGTTGTTGTTCAGTTCCTATGCCATGGTTGC
CACTTTGTGGTGGGGCATTGTTTTCAACCCCGGAATTGCATAATCCAACCCCAATTAAAGCTC
GACCATCGTATGATAAGAGAGAAAAGCTTGATGACGATGAACAAAATGAAGGGTCTTCGACA
GGTTCAAATACTGACATAGGTAGTGCATCAGGAGAGGGAGAACAAAGTTTGGATGTTGATAG
CTGTTGTCTTTCACTTGGAAGGAAATTGAATAGGGAAGAATCACTCTCCTTCAACGGTATAAC
TAAGAAAATTTCAGCAAACTCTGTAAGTTGTCGGAAGGGCTTTGTTCCCTACAAAAGATGCTT
AACAGAGAGGAATTCTACGCTTTCCTCAACAGAAACTCCTGAAGAGCGAGAAAAACAGCGG

doi:10.6342/NTU201901184

ATCCGGCTTTGCTTATAGTTAGTTGCAATATTTTGCAAGTCCTTGGTGAAGATCTGTTTAAGATG
CAGTGACTGTTATTAACCTGTAGTAATGCTTACAGTTGCATGGCGAGGTAGCTTGAAGCAGGG
TATCTGGCTTACC

>SsERF3

AAACCTCATTCCTACAGACCAACCAAAATATCCACGCGCTCAGCAAAAGGTACCAAATGGTG
GTTGAAAAACTTGAGAACGCCGCCGTGGCTGGTGGCGCTCCGCCGTCATTTTACGTAAAGAA
GGGACCGCAGTTCCGCGGCGTGAGGAAGCGTCCTTGGGGAAGGTACGCAGCGGAGATACGC
GATCCGTGGAAGAAGACGAGGAAGTGGCTCGGCACTTTCGACACGGCAGAGGAGGCGGCGT
TGGCTTACGATGAGGCAGCGAGGAGTCTTCGCGGTGCAAAAGCAAAGACGAATTTCCCGTAC
AGCGACGTTTCCTCAGTAGCGCCGCCGCCGCTTAACGTGAACATTTCGTGTTGGCGGTCGCCG
GAGTTTTTCCGTGATGATGGTGAGTCTACAGCTCTGCGTTCGGAGTACACCGGGTATAAGATA
GAGGAAGTAGGTGCGGTGGTTATGAATGAGCAAGAAAAGAAGATGAGGA
GTGAGAAGAAACCGTTTCTGTTTGACCTGAATCTTCCAGCACCAC

>SsAGL6

AGAATGGGGAGAGGAAGAGTGGAGTTGAAGAGAATAGAGAACAAGATTAACAGGCAGGTGA
CATTTTCTAAGAGAAGAAATGGACTGTTGAAGAAAGCTTATGAACTCTCTGTACTTTGTGATG
CTGAAGTTGCCCTTATCATCTTCTCTAGCAGGGGAAAGCTCTATGAATTTTGTAGCACTAGTAT
GATGAATGCCCTTGAGCGCTACCAGCGCTGCTGCTTCAATCCACATGACAACAGTGCTGATCA
AACTGAAACACAGAGCTGGTATCAAGAGGTTTCAAAATTAAAGGCCAAGTATGAATCTCTTC
AAAGGACTCAAAGGCATTTGCTTGGAGAAGATCTTGGACCACTGAGTGTTAAGGAGTTGCAG
AATCTTGAAAGACAACTAGAAGGAGCTCTTGCTCAAGCCAGGCAAAGGAAGACACAGATTAT
GATGGAACAAATGGAAGAGCTCCGCAGAAAGGAGCGCCAACTTGGAGACATGAACAAGCAG
CTGAAGATCAAGCTAGAGTCAGAAGGACAAGGCCTAAGATCCCTTCCCTGTCCATGGAATGC
TGGTACATCAGCAGCTGGAACAAGCAATTTTCTTGTCCAACTACCAGCTTCCAATCCAATGGA
AATTGAACAAGAACCTGTGCTCCAAATAGGGTATCATCACTATAATCTTGGACAAGGATCATCT
GTTCCAAGAAGCATGGGTGTTGAAAGTAATATTATCGAAGGATGGACACTTTAATTTGCTCTAA
GCC

>SsOFP6

GTGCAGTCCCTAGAAGTCACGTGACATCCACATTCCATGCGAGTTGGGAGAAGCCGGCGCGT
TAGGCCTAACAGAGTACACACCATTCCAGATCTCTGTAAAAGCTCTAACAATGATGCCATGGT
AGTACGGAGAATTCAGCTGCAAGAAACAATTCAGCAGCTCTTTCAAATCATCCTTGGAGTAAA
TCTGTTTCTCCAGAATCATCTGAAGCATGGAGTGGCGGAAATCAAGATACGGGTCGTCGGAAT
CTTTCTCCACCGCCACGCTCTCCCCGCCGATTCTTCCGAATCCTTGAACGGCCCTCAGAGTTG
TAGCACTCTTGATGTCTGAGTCAGTATCGGAGGAGTAACACGCCGGTGGGGTGCCGACAGCA

doi:10.6342/NTU201901184

GAGGAGAAGGTGGTGGTGGTGGTTTCCCATGAGCTGGATGAGGCGGGGGAGTGGTGGTGGT
AGTTGGGGTGGGTGTGTAATCGTGTTCTAAGTTTACTTTTAATTAGGTTTGGGTGGAATATGGC
TGAAATTCTTGGCCTTCTGCAACTACTGCTGCATCCTAAATCTACACTCTCTG
TATTCAAAGTGTATTTCTTCTTAATGCTAGACATTGGCAAACCAGG

>SsCYC

ATGTTTAGCAAGAGCACATACCTTCATGTTCCACAGGTTTCACCATCTCTTCAATCTCGTGCCT
CTACTTCTTTGGTTGACCTTAATGGAGGTGAAATCTTGCTTCATAACCACCACCACCATGACAT
GCTTTCCAGCCATTACTTAGCCGTGAATGCCCCGTTTCTTGAGGCTTCCTCCTTGTATAACCAA
GATGCTATTGTTGGTCTAAATGAAGATCCTTCTGCCATGGCCAACACGTTTCCAAGGAAGCAA
ACAGTGAAAAAAGATAGGCACAGTAAAATTGTTACAGCTCAAGGGCCGAGGGATCGGAGAG
TCAGGCTTTCTATTGGCATAGCAAGAAAGTTCTTTGATCTTCAAGAAATGCTAGGTTTTGACA
AGCCAAGTAAAACCCTTGACTGGTTGCTCACCAAATCTAAAGCAGCCATTAAGGAGCTAGTG
CAGGCTAAGAAAAGTGGGAGTGGGAGTGCTAAGAGCATTTCTTCCCCTTCTGAATGCGAGGT
AGTGTCTGCAGGAAATGGTGAAACTTTCGAAAATGGCAGCTATTTGGATGTGGAATCAAAGA
AGAAATCACTGCCCCTGAATCCTAATTACAAGTGTAAAGAATATTCAAAAGATCCACAGCAGT
CTGCATTAAATCTTGCAAAAGTATCAAGGGCTAAGGCAAGAGCAAGGGCCAGAGAAAGAAC
AAGAGAGAAAATGTGCATCAAGAAGCTTAATGAATCAAGAAGCATGGATCCTGATTTGAACC
CTTCAAACCAAATTCAGCCGACCCTCCACTGTCCCTTAACTAATAATGTACCTGCTGCAACAA
CTGAAGATTTAATTCAAGAATCCATTGTCATTAAAAGGATGTTGAAACAGTACCCTTCATTTTT
TGGATTTCAACAAAACCTTATCATTTCAAGGGATTTGAACTGCAATCTCCCTTCTCCTAATATC
AACGATAATTGGGATATCAATAGCTTAACCTCACAATCCAACCTGTGTGACATTTTGGATCAGC
ACAAGTTCATGAATAGCTCTTCAAATATATAGGAAACTTTTGGAATCTGCAACTAATTAAGGTT
CAGAATCATCGATGTAATTCTGCGTGGTTTCTGTGG

>SsMYB14

AAATGGGTCGGGCTCCATGTTGTGAGAAAATGGGATTGAAGAAAGGTCCCTGGACTACAGAA
GAAGATATCATTCTTGTCAATTACATTAACCAGAATGGTCATGGAAATTGGAGGGCTCTCCCAA
AACGGGCAGGGCTGTTAAGGTGTGGAAAGAGCTGCCGACTCCGGTGGATGAATTACTTGCGG
CCTGACATTAAGAGAGGAAATTTTACCAACGAAGAAGAGGACACAATCATTAAATTGCATCA
AGAATTGGGGAACAAATGGTCTGTAATTGCAGCAAGATTACCGGGCCGCACAGACAACGAAA
TTAAAAACGTGTGGCACACCCACTTGAAGAAGAAGCTCATTAAAAAATCACAACCCAAACCC
GAATTCGACCAAAAGGTAGACGACAATAATGACTCGGAGCCAAACTCTCCATCCAGCGAAAT
TTCATCAGTAACCACCACCAACGCCGGAGGCCCTGATGACATGATCTCTGTTTCCTTACAAAA
TTTCCCAGAAATGGACGAAAGCTTCTGGTCGGAGGTATTTTCCGCGGATCACTCCAGTGCTAC
TAGTGATTTCCAGCTTCAATTTTCAATGGAGAAGGAGAATTACTTGATTGGTCCTGAGTTATCC
ATGGAAAGTAGCATGGATTTTTGGCATGACTTGTTTAATCAGGCAGAGGAGCTAATAGAATTG

doi:10.6342/NTU201901184

CCAGATTTTTTATGAATTTTCAGTAATTTCTTACTACTCTTTTTAACTTTTTTACCTTGATTGATT
AGGCTTTTGAGGATTGAGACAGTAAAAAAAAAAAAAACTAAAAAGCTGTGATTACAAAATTTA
CCGTCACTCCTGTAACATGTAGAC

>SsCIB2

TTTGGAATCTTGATGATGGATAAGGAGTACTATATGAATGCTGGAATTCCAACACCCCATCCGC
TAGAATTTGAAACTATAATGCCAATTGGATGGAATGGACTGAATTGTAGTGAAGAACAATCGT
TCTTGGACCCAAATCCTTCTGTGGATCAATATTCACATTTTGAGTCAGCTTTGAGTTCAATGGT
GTCTTCCCCTGCTGCCTCCAGCTCAGGCTTGTCCAATGATGCTTTTATCATCCGTGAAT
TGATCGGAAAACTGGGTGGCATTGGCAACTCCATAGCTTTACCAACGGCAGCAACCACCGTC
GTGGCGACAGGGAGTAGTAATAATCCTACTAATGAATCTTGTTACAGCACACCTTTAAGTTCTC
CACCTAAGCTAAACTTGCCAATTCTTGATCATATTAAGATGCCTAATTTGGGGAATTCAGTTTC
GCTGACTCCTCTTCCTTCCTTTTCAACTGATCCGGGGTTTGCCGAAAGGGCTGCCAAGTTTTC
TTGCTTTGGCAGTTGGAGCTTCAATGGTAGGGGAAGTCCATTTGGGATGAACAATGCTGAATT
GGTACATAGATCCAGAAGTCAATTGATGGGTAATGGGAAGTTATCTCGAGTTTCGAGCAGCCC
TTCTCTTAAGCAAGATGGATCCCCTGTAAAAAACCAGAATTTAAGTCAACCCCAGATGAATAT
GACACCCATTGATCAAATGGTCGCAGGTTCTGACAAGAAATTGAGTAAATTGTCAGACTCATT
TGCCAATTCTAATGAAGAATCCTCTGTTTCTGAGCAAATTCCAAGTGCAGAAACAGGTTCAAA
AACGTTCAACGAGCTAAATTCTAGGAAAAGAAAACCGATATCCAGAGGAAAATCAAAACAAG
ATGGATCAACTTCAGCTAAGGGAGTCAATGGCGATGATGATGCAAATGCCAAGCGTTCGAAA
CCAACAGAAAGTGGCAAAATCGAAAACAATGGTGCTAAAACAGAGGAAGAAGCAAAGGGG
GCCTCGACTGATGAAAACGATAAACAAAAGACTAATCAAAAGCCACCTGAGCCACCAAAGG
ATTACATTCATGTCAGAGCAAGAAGGGGTCAAGCTACTGATAGCCACAGTTTAGCAGAAAGA
GTCCGACGAGAGAAAATCAGCGAAAGAATGAAACTTCTCCAGGATCTTGTACCAGGTTGTAA
TAAGGTGACTGGAAAAGCACTGATGCTTGATGAAATCATAAATTATGTACAATCATTGCAACG
ACAGGTTGAGTTTCTGTCAATGAAATTAGCCTCAGTAAATCCAGGGCTGGATTTCAACATGGA
AAATCTTCTCTCCAAGGAAACTTTTCAACAAAATCCGACTTTACCCCAACAAATGTACCCTTT
GGATTCCTCAACACCAGCATTCTTGAATCATCAGCCTCATCAAATTCCACAACAACTGCATAA
CAACAATGCTTCAAGCAGACCTTTAACCCAAAATTTAAGCCTTGATGGATTTGATTTCCCAGA
ATTTGGTGAAGGTGATTTGCACAGCATTTTCCAGATGGGTTTTGGCCAGAATCCCGTCAACTT
TCCAGTTCCAATTCCAAATCAAACACCTAACATGAAAATTGAGCGGTGAAAACTAGTCACTTA
TGCCAGATTTGAAGC

doi:10.6342/NTU201901184

>SsNGAL1

ACACGCACTGAAATGTCAATAAACCACTACTCTTCAGACCAGATTCCAGAAGCCCACTTGTAC
TGGCCTTCACAATATATGATGATGGAATCCTCGTCTTCTAATCAGAATAAATCTACCTTTTTTTC
ACATTTAATCCCGAATAATACTAATACTACGAACTCTAGTTTCTGGGGGCCCCGGAATCAGTTT
TACCACCACCACTCTAGCGCCGTAGAAGGAGGTGGAGCTGGTGGTTCCAGTAGTACGACTGC
AATGTTTAATCTGAACAATGAAGATGAGGAAGAATTAGTGGTCGATGAACAGTTGACTGCAG
ATGATACAGATAATATTAATAATAATTTGGAACATGAAGGAATGGAAATCCCCAAAGAACCCTT
GTTTGAGAAACCATTGACTCCCAGCGACGTGGGTAAGCTCAATCGTCTCGTGATACCGAAGC
AACACGCCGAGAAATACTTTCCATTAAGCGGCGGCGGAAGCGCTGGAGGTGACTCGGGGGA
AAAGGGATTTCTATTAAGTTTTGAGGATGAGATGGGAAAATCATGGAGGTTTCGTTACTCTTAT
TGGAACAGTAGCCAGAGCTATGTCTTGACAAAAGGGTGGAGCCGATTCGTGAAGGAAAAAA
GGCTAGACGCAGGCGATTTTGTCCTGTTTTCGAGACACCGTGCTGATGCCGGCCGCCTTTTCA
TAGGGTGGCGGCGGAGAAACTCCGGAGTAGAGAGTGGTGGTTCTCAGCCGGTGGGTAGCGG
CGGTTGGTATAACAGAGTATTTTATCCTGCAGGCAATCCTTATCCAAGTCAACAGCATCAAGGG
TCTTCTTCTTCTTCTCATCCACACCAACCTGACTGTCTTCATGCAGGATCAGTTTTACAAAACC
AAACATCAACAGCAGCTGTAACAGCAAGTGGGAATGCAAAGAGGTTAAGATTATTTGGTGTA
AATTTAGAATGCCAAGCAGATGAATCTGAGCCATCCACACCAACAGAAGGTTCGCCCATGTCC
AGCCACAACCACCACCACCACCACCACCACCACCAGAATCCATATCAACACCAATTTTACTCC
ACCCATCACAATCACATGGGAGGAGGACAAGGACAAGGACAAGGACAGGATATAAATTTCTC
AACAGGAGATCATGTATATCGCCAAGGATAAGATCTGTGAGATGTGACGTCCACATTCCACAA
TCAGAGGGCTGGCGAGGACGTGGGATAAATTCATTTTTCTTTATTGAAAAAGAAAAGAAAAA
AAAAGAAGGGTGAAAATCCTGTTTGTATGAATTATAACATGGGATGG

>SsERF1

TCATGTACCAGCCAATTTTCAGTGAGTTAACGCCGGTGGATTTGTCGCCGGTGGTTTACCAGA
GCTCGAGTTTCAGCAGTCTGGTGCCATTTTTGACGGAAACTTGGGGAGACTTGCCGTTAAAA
GTTGATGATTCGGAAGATATGGTAATTTGCGGTCTATTGCGTGACGCGGTTAATGATGGATGGA
CGCCGTTTAACAACGTGAAACCCAAGACGAGATGTAAAATTGAGCCGGAGCCGAGCCTATCC
GCGGTGAAAACGGAATACGTGAGTTCTCCGCCGGAGATGACCGCGCCGGCGTTGGCGCGGCC
TAAAGGAAGGCACTACAGAGGAGTGAGGCAGCGTCCGTGGGGGAAATTTGCAGCCGAGATA
AGAGACCCGGCTAAAAATGGTGCAAGAGTTTGGCTTGGAACATATGAAACGGCTGAAGAAGC
GGCTTCTGCTTACGACAGAGCGGCTTACAGAATGCGTGGATCAAAGGCTCTAT
TG

doi:10.6342/NTU201901184

Supplementary Table S10　5' Regulatory region sequence of dorsal-expressed TFs of *S. speciosa* 'Espirito Santo'

The 5' regulatory region sequences were referred to the sequence before translation start site (ATG)

>SsRL2

TTAGATCACAGGTTATAACCCGATCTAATTTCTATTAATAAATATTTTTATTATTAATATTAACAAA
AAAATATAAAAAAAGATTAACAACACTTTTGTACTTTTCATGGTATATTATGGTTCGATTTTATAT
AATTTGTAAAAACAAATAGGTAAAATTTAAATAAAAGCATTTTTTTTTAAACTCCTTAACCCCA
TCAAGAAATGACATAAAAAATCTTTGAATTTTTTTTTAAAAAATAAATAAATAAAAATTTGAGG
AGGAAAACAAAGTTAAATTTCCTGGCAGAGCTTACAGCCACAAAGAAAGAAACCTTTGGAAT
ATTTGTCAGTACATGATAAAATCACCATTTCCAAGCCTGAATCATATTCTAGTAATAAGCAAAA
CTTATGGATAAACATCAAACAATAAGTGAAAAAGACGATGAAAATGAAAGCCATGACATAATA
ATTATTATATATTATTATTATTATTATTATACAATTTTTCTCCATACGTTTTACATTTTGACCATAATT
TTATTCATTTCTGCAATGGTAAGAAGTTAATGGGAGACTTATCAAATTCAGGACACATAAAAAA
CTTTCAGATGGGCCTATGATTTAAACAAGAATGGAGGATAAGAATATTCATCCATCTCCAATGA
TAAATATAGAGTCAAGATTTAGAGCATGGAGACTACTACAAATACATGTTAAAAAGTGCGAGT
TTTGTCATGTGAATCCAATTCTATTTTGAAAAAAAAAAAAAAAAAAAAAAGAAAATAATTTGA
TGTGGTAAGCAAATCCCACAAAGAAGATAACAGAATATCTTCTGCTACATTTCGATTATCGTTA
TGTTAATTAACGCATCTATCTTAACTAGTAAAAGCTTGGATTGGGGGATCTAATCCCCACCAAA
TGGGTACCATTTCAATAATAATCCCACTTGCTGATTATAATTTTTTTTAAAAAAAAAAATAAAAA
ATTAAACGAACTTTTCAAGAATTTAGAAAGATCACATGTGATGTTACGTTTTGGTGTTTGAACT
GTTATACAATATAAAATATTTGAGTTGTGCAGTTAGTTCCTATTCATCTCAGTACGTACTTTTTAG
CATATCAATAAGTGATTGATCTTAGTACATACATACATTTATCATGCCCTTGTCTTTAAAGAAAC
TAAGCTCGTGTTCTGACCCTTGAGGAAAAAAAAAAGAATCGCATCTTTTGTACATGCCTCGGA
TCAACAGATAAGAAAATATTCGTCAACTTGCTCAACCCCATCACCTTTTTTTGTGTATATAAAG
GCTTTGTGTATCAAATCTCACACCAAAACACTTGTATCTCTCCCTTGACTGAACCAACAATATC
ATCCCTTCCCATAAGAAAGAAAACTTCATTTCTTACGATAGTGTTTAATTCAATACTATTCGTCT
TCGTTTCTCGGTTTCTTGGTCGATCTC

>SsERF17

AGCTTGGTCGAGTGAGATGAAAGAAAACCAGGAGGAGGATGGATCAAACTCTCAACAGTGA
TGGCTGTGGCAAAGGAATCCCAGTTCTGTCTACTGCAAGAGGAATGATTAAAGATCATATGGG
CAATGAGCTTATACTAGCCGACATGTTTGGCATCCAATTGGATTGAATTAATACCTTGAAACTG
AACTTACAAGCACAACTAAGTCATTGAGTTTTCTTCTTAACCATTCATCAATATCTGAGGCAGT
TAATGGTCACATGGGGGATGTTAGGAAATGATCAGCCAATTGGTGATTTGGAGGTGAGTTGGT
CGGCAGGGTGATGGGATTATTAACTTTTTGGATAATTGTACAAAAGTATGTTATAATTAAATTAA

doi:10.6342/NTU201901184

GGCGATCGAGTTTGACTCTTTAAAAATAATATATTTTTTAAGAAAGTAACCGTATTAGAGAATAT
GCATGTGTCAACACGCATTAGGCCAAGATTCAAATAGACGCATCTTGAAAGGAAAAAATAATG
TAAGATAAAGTTGGACCAATGTAACATTTTGGCTTCTGAAGTCTAGAAAAAATTGTAGAGATC
GAGATCATTTACCCATAGGAACCTTTAAAATCGTGTTATATGATAAAAAAGTTTTTGACTTTTTG
ACAGTAACATCCATTATTAGAATTTGTAGAATTGTGCAAGAGAGGAGTCGTGTCTTTGGTAAA
GAGCATCCTTTTTGGACAGGGAGTCAACGCCACCTAAATTTCAGCCGGCCCACGTGTCCTGCC
TCAGCACGGACCTAAACTTATCACCGCCACACCTGTCAATCAAACCTTGTCCCGTTCAAAAGA
ATCCTGTATTCTCATTTAACTGTAAATAAATATGAATATATATCACCTCGTGTCATTCATATAAAA
ACCCACGTGTCCAAACAGAATATTTAATTGAGGAACAAAATAATTTGCACTGGTATAATAACGT
ATAGGAATTGGACAAAATTTTTTTGCAGAGGCTATGAAAAAA

>SsERF3

GCAGTCGACTCTAGAGGGGATCCAGATCTCAGCCCACTTATCATTACCCGATACTGAAAGAAG
TGCGTCCTAATTCGCCGAATATTTTACTATTCAATGTACCCAAAATAATATCGACGCGGATCCCA
ATCTTGACTGCTAAATGATTATCCTTTTCTAATCAAATATATCATTAATTGGTTCAAATATTTGTC
AATTCATTTGATTTATCAAATCAATCATAATCAAAATTCAAAATTAAAAACAAATCGAGTCAAA
ATATTAAATTGATTTAAATTATCGAATTTTAAAAAATATTCAAAGTATTTGGAATTAAATTAATGA
AATTTAATCATTTTTTTTATCATTTTTTCTTAAAAACATTAAAGTTATGTTTTAAATTTGTTACGA
GAAATGAAGAATTGAGAATACAACTTTTTATGGAACAAGGTTTATTTGACTATCAAAATGATAC
GTGTGATTACCAAGTTTATTATAATAAGTTTTTTTTAAAAAAAAAGTCAATATTTTTTATTGTCA
CACATTAATTTATTATTATTTTTGTATGGATTATAAGAAAACAACATACATGGATTATGCATCCAA
ATTCATTGAATCAAATGAAATGTTGTTTTTAGTAAAGAATAAGATAAATGTGACATGTGAAACA
TTAGTCCAATTTACTGGATCCTTAAGTACGGATATGAATTATATAAATATCGTTACACTTTAACTC
TTTGTTTAAAAATTATAATTTTAATATATTTAATTTTTTTTTTTCGTATCTTATGAAATTAAAAAAGA
ATAATTAACGGTAACTTAATGATCGAGTAAAATGATGAAAGATTAACGTATCTTATGGCACGTAT
TGTATGTCCGTGATACATGCGAGCTCTAAAATAAAATTATTTCTGAAAAAAAATTAATTTTATTG
TTTTTAAAATTAATATAAAAAGACAGTTATTTTAAACTCTTGTACTTTAATTTAATGATATTAGAT
GAGACTAAATGTTTGAAGTTAAAAAACGTATAATTAAAGTCCTAAAAATGGCGTTACTAAAAA
AAATTAATTTTCCCAAAGATTTGCTAATTATTAATACTAAAGATTTTTTCACATTGTAAATCTTG
ATTTACACGCCATGCATTTAATGCCTCTCACACTCCTTATCCTGCGCCCAAAACTTATACTGTAA
ATAAATAAAATAATGGAGTGGGAAATTAATAATAATAAATAAATAAAAAGAGAATAATCACAAA
GAGTCGAACCTGTCGGCGGAAGCAGAGGCAGCTATAAATAACCCTCAAAACCTCATTTCTAC
AGACCAACCAAAATATCCACGCGCTCAGCAAAAGGTACCAA

>SsOFP6

AGAGCATCGCTATATTTGTGGCTGGTTTAGCAAATTCCTTTATACTTTGATTGGTACAGTGTGGT
TATTGTATTCAGGCAATGAATACCAGATCAAAAGATGTGGTTATTGTATTCAGGCAATGGAATA

doi:10.6342/NTU201901184

CCAGATCAAAAGATGGATGGATTCAGCTGCGCTGCATACATTTCCAAAATTGTTGTGAAATTAC
CAGAATTGGCAGATCCTCCAGCCGGCATGAATGCTATACATTGAGTGTTTATAGAAACAAGATT
TTACGGATCATGAAGAAATTAAGTTCTGAAAAATGAAGACTTTAAAAGATCTTGAAAAAGTCC
CGATTCAAATACTAGACAGTGAAAGATCAAATCAGCAAGGGATGGAACTATCAACGCTCAGA
ATCAGAGGGAAAACAAGGGGATTGCTGCATTAATCTCTGAGAGTATATGTGCAAGAATATGTA
TCTAGTTTATTTACACCTATATTATTTCTTCATTAGTTAATAAAAAAACAATGTGTCCAAAGGTAT
ACCAATTTATATATATTATGAGATTTCACCTTTTGCTCAACTTTAATTTTCTCGACTTCAATCCTC
AGTCAGAAAACGACATTGTGGGATTCAAATTAAACTGGAAATATATACTTACACTGTACTCGA
ATAGCCCCGTATCAAATTAGACTATAAAATCATGAGATTATAATAACCACACTTTCATGCTTTAA
TTTTAAAAATTAAAATATCTTGAACTGACTTTTACAGATATAATGTGCATAGGTGACATAAAGA
AGATTTCCATTTTTGATCCAAATCATACCCATTGCTTAATAAATTATATCTTAGTGTAATTGATTC
GCATGTAATCATACCCCGCCATTGTCAATTGTCAATTGTCAATTGTCATCTCAGCTTCCTCTTTC
CATAATATTTCTCTATTTAGACCTTATCATAATTATTTAATTTAAGCTAATAATTATTGATCACATAT
AAAGCAAATTTTTCCAAATTATATTAACCATTACTGGCATTTGGCCCTAAAAGTTTTCCCAAGG
AATGCGACTTATTATACATTGAAAGTAAAAATTTATATTGTTTTATGAGAAATAAGACTTAAAAC
TTTTTTAAAGACTATACATAAAATAATTTATTAATACAGCGACGACCAAGCTTTAATAATATCCC
TTTGCAAAAAAAAAAAATAAATAATAATGATATTCAATGCTTAAAGTAACAGGATTGATCTCACC
ATCATTCTATCACAACCAAATGATAAATCCCACTTGTCTTGAGAAAATATAATTATTAAAAAAAT
TAAATATTTATGTCCAAAATTTATCAAAAAAGCCAATTTGTCTGTGAGTTTAGCATGAATGGGT
CCCTCAAATCAAAAGCTCACATGTAGAATTTATATATATGAAGGGACTTCAGTTTCCCTTATTGC
CAC

>*SsCYC*_A

CCCTGCAAGAACGTATAGGAATCTATTAACACACATTCTAGACTCATAAGATCCATTAAGAGGA
AAATTACTGTACCAAGTATTTCAATTTCTTTCAATCTTACACTTGAATTTTTTTATTTTTTTGAAG
AAAAAAAAAAAATTGTATTTACTGGAAAAGTGAAAGTAAATTAGGGTTCATTTATTCACACACA
AATTTAATTTTCTTTCACCGAATCTGTATAAATGTATTATTCACTTACCATCACCGGAAACCATG
TCTTTTTTCTTCAACTGAAGTAACCCTGTTTTTTGTTTTTTTTTTTTTTGTTTTTCCTTTCCACCA
ACTGATTGAAGCTGATCTGATCTGGTGTTCTTATAATAGAAGAGATGAAATTAAAGCAAACCTT
CAACAAACAGTTCAACGAGAACTCAAAAACCCAGTCTTTAAATTGTATTTAAAAATAATTACC
CACCAAAAAAAAAAAAAAAAAAAGAATGTATCTGAAAATTACTATTCATACATAGGGCAAGG
AACCCACTCTTCAAGAATTAGGGTTTTCAGAAGTTTTCTTTAAATCAGGACGAGCACTAAATG
TTTACAAACCCAAATTCTTCACTCTTAATTCACACTTCAAAATGGTTCAAACACAGAAACAAA
AAAAAAAAAGGGTTTCAAAAAAATTACACAAACATTCAAACTAAAAGCCAAAAAGAAAAAG
GAAAAATAAAAAAAAAACTGTAAATTCTTTTAATAAATTGTTATTATTTTAACAGCTGCAAAGA
AGGAAATTAAGAAAAAGCTGTGGACCAAAAGAGAAGAGATTTGTGGGTTGCTTACTGTCATA
CCAGAAATGGAGAGATGAGTGTTTATTCTGAGTGTGGGTGAAGAGGACACGAATATAATGGG

doi:10.6342/NTU201901184

GCCCCATTACAAAAGCAAAGAAAATGAAGTTAAATTTGTCCATTGCGTAGGGTGATAGACAAA
GGTACGTACGTTTTTAGACTCTAAAAAACTGTTTTAATCACCAGCTTTTTTTTATTTTTATTTTC
ACCATGTATACATGCAGAGGTCACACACCCTACAAAAACCTCACGGCTCTTCTTTGGAAAACC
TAACTACACACACTCACACACCAAACTAAGAGAAATTCACACTTCATAACCCTCTCTCTCTCT
CTCTTTCTCTCTCTGCACATATAGAGATACCATCAAACCCTAGCTACCCTTCTTTTTATTAGTAC
CTTTTTATGCTTTCAAGATTTTGTTTTCTCGATCATGGATTAATTAATGGTACCGTTGAACCAAA
ATGAATACAATACTAAGAAATACAATACGAAATGGTAGTAGTAATAATTAATAGTGTTGGTAGT
AGTAATGAAGAATTAATCATTATTTTGAGGGAGTGGCTGATTTTATCTGATGTTGCTGAAGTGG
AGGAACTGTAGCATAACTGTAGATTACATTTTGAATTGACAATAAATTTTTGTACTGCGCTAAA
AGTGAAGAAGATAAAGTTAAACTAGGTAGTTTTTTTTATTATTATTATCACCAATTTAATACCCT
ATTCAGTGCATCTGAACAAATTTTATTTGGAGATTAAAGAAAGGGTACAATACTTTTACTCCTG
AAAACCCAAAATTTTTCCCAATTCATCATATCTTCGTCCTCCATTTTTCACCTACACGCTAGCCT
TCCAGTCTTTCTCAGGCAAAGTCATTTTCTTTGGTGTAATATAAAGCAAAGACAAGAAAAATT
TGCATATAACTATATATATACACACACATTTATCATCAATAATAAATAAGTGATGCTAGAGTTATG
ATCTCTTGAGGAAAAAAAAAAGATAAAAACCTTAGTTCTCATTCTGGAGAAACCTTCAAACC
AGCTCTCACAGGTTGATTGCATAAACAATAAATATGGTTAAAAAATTCAAGAACTTAAGGGTT
TCTTTCCTTCTTTTTTTCTTTTTTATGTAAGAAATTAATTAGGGTTTATTAACCCTTCTTCCCCTC
CCCTCTCCCAAAAAAGAAAA

>*SsCYC*_B

TTCCTGCAAGAACGTATAGGCAATCTATTAACACACATTCTAGACTCATAAGATCCATTAAGAG
GAAAATTACTGTACCAAGTATTTCAATTTCTTTCAATCTTACACTTGAATTTTTTTATTTTTTTGA
AGAAAAAAAAAAAATTGTATTTACTGGAAAAGTGAAAGTAAATTAGGGTTCATTTATTCACACA
CAAATTTAATTTTCTTTCACCGAATCTGTATAAATGTATTATTCACTTACCATCACCGGAAACCA
TGTCTTTTTTCTTCAACTGAAGTAACCCTGTTTTTTGTTTTTTTTTTTTTGTTTTTCCTTTCCACC
AACTGATTGAAGCTGATCTGATCTGGTGTTCTTATAATAGAAGAGATGAAATTAAAGCAAACC
TTCAACAAACAGTTCAACGAGAACTCAAAAACCCAGTCTTTAAATTGTATTTAAAAATAATTA
CCCACCAAAAAAAAAAAAAAAAAAAAAGAATGTATCTGAAAATTACTATTCATACATAGGGCA
AGGAACCCACTCTTCAAGAATTAGGGTTTTCAGAAGTTTTCTTTAAATCAGGACGAGCACTAA
ATGTTTACAAACCCAAATTCTTCACTCTTAATTCACACTTCAAAATGGTTCAAACACAGAAAC
AAAAAAAAAAAAAGGGTTTCAAAAAAATTACACAAACATTCAAACTAAAAGCCAAAAAGAAA
AAGGAAAAATAAAAAAAAAAACTGTAAATTCTTTTAATAAATTGTTATTATTTTAACAGCTGCAA
AGAAGGAAATTAAGAAAAAGCTGTGGACCAAAAGAGAAGAGATTTGTGGGTTGCTTACTGTC
ATACCAGAAATGGAGAGATGAGTGTTTATTCTGAGTGTGGGTGAAGAGGACACGAATATAATG
GGGCCCCATTACAAAAGCAAAGAAAATGAAGTTAAATTTGTCCATTGCGTAGGGTGATAGACA
AAGGTACGTACGTTTTTAGACTCTAAAAAACTGTTTTAATCACCAGCTTTTTTTTATTTTTATTT
TCACCATGTATACATGCAGAGGTCACACACCCTACAAAAACCTCACGGCTCTTCTTTGGAAAA

doi:10.6342/NTU201901184

CCTAACTACACACACTCACACACCAAACTAAGAGAAATTCACACTTCATAACCCTCTCTCTCT
CTCTCTTTCTCTCTCTGCACATATAGAGATACCATCAAACCCTAGCTACCCTTCTTTTTATTAGT
ACCTTTTTATGCTTTCAAGATTTTGTTTTCTCGATCATGGATTAATTAATGGTACCGTTGAACCA
AAATGAATACAATACTAAGAAATACAATACGAAATGGTAGTAGTAATAATTAATAGTGTTGGTA
GTAGTAATGAAGAATTAATCATTATTTTGAGGGAGTGGCTGATTTTATCTGATGTTGCTGAAGT
GGAGGAACTGTAGCATAACTGTAGATTACATTTTGAATTGACAATAAATTTTTGTACTGCGCTA
AAAGTGAAGAAGATAAAGTTAAACTAGGTAGTTTTTTTTATTATTATTATCACCAATTTAATACC
CTATTCAGTGCATCTGAACAAATTTTATTTGGAGATTAAAGAAAGGGTACAATACTTTTACTCC
TGAAAACCCAAAATTTTTCCCAATTCATCATATCTTCGTCCTCCATTTTTCACCTACACGCTAGC
CTTCCAGTCTTTCTCAGGCAAAGTCATTTTCTTTGGTGTAATATAAAGCAAAGACAAGAAAAA
TTTGCATATAACTATATATATACACACACATTTATCATCAATAATAAATAAGTGATGCTAGAGTTA
TGATCTCTTGAGGAAAAAAAAAAAGATAAAAACCTTAGTTCTCATTCTGGAGAAACCTTCAAA
CCAGCTCTCACAGGTTGATTGCATAAACAATAAATATGGTTAAAAAATTCAAGAACTTAAGGG
TTTCTTTCCTTCTTTTTTTCTTTTTTATGTAAGAAATTAATTAGGGTTTATTAACCCTTCTTCCCC
TCCCCTCTCCCAAAAAAGAAAA

>SsCIB2

ATTTCCACCACAAAGCTGGGAAGCATGTCTACAATATGTTACAAATCTATTCTTGAAATTTACA
TTTTTATTTATTTTTTATTTAAAAAAAAAGAAAAAGAAAAGAAAAACAAAGGGAAGTAAATTG
TTATTTTCATTACTATGGGAATCCGACCAAAACCCGATTTAATCGGTGAGCTTGGAATTTAATG
CAATGATTGACTTACCCAAAACACGTTAACGATTTTTTCCCTCCTCCCCAGAGAAAAATTTATA
TTTAAAGTCGCATTTAATTTATTAACATTTGAGCCAAACGAAAAATAAAATAAAAAAATTGAAA
ATTAACTGTCTATAATAAATTATCTTTTTATTTTATTATTGTTTGACTTTAACCTTAATGGGCAATT
TTCAATTTGACCCAATCATTGAGGAATATAAAAATTGGAAAATAGATCTTACAATTTTAATTTAA
TGGGCTTTGGCCTATGATAATAATGTGAAATTCGAGGCCCAAAAATTAAGGTAAGTGTAAATTA
AATATCAAAGAATAAAAATATAATATAAACCCTAGATTTAGAAATAGGAATTAGGTAAGCATATC
ATCATTTATGAGATCATTAGCCTTATCTTTTTCCTCAAAATAGTATTTTTATAGTTTTGTTACGTA
CACTAATTAAAGTTTTTAGTATTTTGTTTAAGAAGAAATTTAAGGAAATTTTTTTTTTATTTTTTT
TGAATTTTACAGTCACGTAATATTCGAATTTGATATTTTTATATATAGTATTTTTTTTTTTTTTTG
TATTTGTATAAGAAAAAATCAAGAACTTGGTGTTAAAAGCGTGCTAGCACACACGTATCGTTT
GATAGGAACTTATATTCCATTATTGATTTCTTTTACTATTGTGTTTGTGCAATTGACAATATATTTA
AGTGAAAAATTATCATTTTTGTTCGTTTGACAAACATTTCATTCGGAATTCATGCAAAGTGTTT
TATTTCAAGTTTAAAATAAAAAATATTTACTCCTTTTATTGAAAAAGGCCTGAAAGAAAAAATT
GGAGACAGAAGCAATCTTTGTGAAAGAGACTCTTTTGTACACTTTCTCTTAAAATAATTTAAA
AGTGTCAATAATCAAAATTTAAGCGTGCAATTGTAATTAATATATAAGATTAAAAAGGAAATGT
GTAAAATCATCCACTAAATGACAAAAAAATTGCATATAATTCAAATATTTTAGATTTGAGAACT
CAAATTAATTTTCATTGACGTATATTTTTAATTTTTTTAAATATATATATATATGTATGTATATAATTT

doi:10.6342/NTU201901184

GTTATTTCAAGAAAGGTTGGACATTTTGGTGGTTTGACGTCTAAGGCCCGGCATTTAACCCTTT
CCCCCCACCGCTTTTAACGGCTACAAAAATTATTAAGAAAATTAAATTATAATATATGCTTTATT
TTTACAGTATGGTATAGAAAAGGTCAAAGAAAACTCTTTTCCTTTTCTGCTATAAAACTTCATC
ACTACCCTTCTTTTTTCCTTCCCCCTTTGCTTTCTTTTCTGATTTTCTTTTTTTCCCTTCAATTTC
CCTTTGCACTCTGCACTTTCTTTTTTTGCTTTTCTTTTTTCAGGTAATTTTTCTTTTCCTTTTCTT
GCTCCAACACACATTATTCCCCCTTCTTGCAAACCTTTCCTTTCATCACATTTTAGTTCATTGTG
TTGTTTTACCCTCCTATCTCGATTTCATCTGTTCTTTACAATTCTGTCTGTTTGTATGTATGAAGA
AGATTATTGGAGTTATTCTTGAATAAACTTTATCTGGGCATACATTTTCTTTCCGGAATTTAAGC
TTTTTTTGCAGTTTTTTTTTTTTTGTTTTTTGGAATCTTGATG

>SsNGAL1

CAGAGATGGTGTGGTCACAGGGAATCAGTTCCTAATATATAAAAAAATTAAGTACCCTCATGC
AAAGGGCAACTACTCTTTAATCAAAGCTTCAATTTTCTTCCATTTAGCCCACACCGATAGGCAT
AGCCCCACATACCAACTTTTGGCAACAATTTCTAAATAATAATAATTCTTTTTTTTCTTCTTCTA
ATTAAATTAGTACGTATAATTTTAAATGTATGTTTAGAAAATAAAGTTGAAAAGCCAGCAAAAT
TTGAATATTTAAAGACACCATTTCAATAGTCTGTTACGTAATAATTTTTTATTTTTTAAAATAAAA
AATATAATGACACAGTTGTTTCGCAATTATTTTTTAAAAAATAAAAAATTTACAAAACAAACTT
AAAAATAACTTATTTTATAACAAAACTCAAAATTTGTTTTTAAATTTTTATGTTACCCTATTATTC
AAGAATATATCTAAACAAAAATCATCAATTTTTCAAAATAAAATTTGATAAATAATTGTTTAATT
ACACAACTACATAGATACATAACTTTGCATTCTAACTTTTTGAAAAATTTAATAAAAGATTAATT
AAAAACCATTTTTAATTATTGTATTTTACATATTAAGCAGAAATACATTTAAACATACCCTTAATC
CACGGTATAACTACTGAAAATAAATAAACTATAGCTATTTTTACGTATGGAGTAAAAATTAATTA
ATATCATGTACTTTGCCACATTAATTAATTTGAACTTATTTTGCATTACACATATTGGGAAATTTA
AATTTGAAGTATTTAAGTAATAGTATTTAAAAAAAGTAGAGTTAGAGACCATAGACTAAAAATT
AAGAATTAATTGTGAATTAATAAGCAGCCTTAGTGATTTGTTGGGAAGGGCATCACCATTTGAT
CTGTTGGCCAGCACAGTAACATGGCAAGTCTTTTCTTAAATACTGAAATTAGGATATTCTATTTT
CACAATCATTTTAAATGTGTTATTTCTTGATGAATATTTAATTTTTTATTCTGTTTAAATATTGTGC
ACAAAAATTAATGTTTTTGGTATACTTAAAATTAAAATCTTTGTTGTGTAAACTGAACATTTTAC
AATAAACAATTATTTTTATTTTTTTTTAATTTTCAAAGAGAGAGAGTACTGTATATATATAGATTT
ATAAATTAAAGTGTGATATTTGGATTAATGAAATTATTACCTGTTGTGCCAAATTCAGTAGTTGA
CATTATTAGAAGTTTTTATACTGGAGATTGGATAGCTGAGAGCAGTACCTTTTAAAACCTAATT
TCCGTGTAGCCATGAAAAAGTTACAGTGGGGATATGAATATTGCACTGAATATAATGATTGATT
TTTTCCAATCACGCAAAACCTTAAAAGCAAAGTAAAACCACTCTACTCTACTTGTCCTTTCTTT
TATAATTTCTCTCTCCCAGAAAAAAAAACTCCAAACTCCTACTTATATATACTTACACTTTTACC
TTTATGTCTATCTTTCCTTCACTCTTTCTTGAATTTGCAACTCCAATTTTCTGCATCATATCAATT
CCTTCCAGGCCCCCCACTTGTTCCTCTAACTCACTTTAATTCTCTTTTCAATTCTTTTTCTTTTC
TTGCTGTTTTCATCAGACTAGTACTATAACACACATCATTACCAATATACAATACACACACTACA

doi:10.6342/NTU201901184

CGCACTGAG

>SsERF1

GAATCACTGGGATAACTCAGCCATCTGCAGGGGATGTGGGCCAGACATCCTCAGGCGACTGT
AAAGTCGTATCATATATCTGTCATGATAACAATATCATATGGATCCATCTGTCTCCTCAACTGAT
GCATCAGCCATCCAAGTATCACTATACGACTCTAGGTCTTACGGGCATTTCAAAACATATTTTC
ATATCTCTTTTAAATCCTGTTATCATAATAATCAATATTTGAAACCATATAAGCACGTAACTCAAG
GAAACCGACCAATCGATTCCTTAATGCCATAAACAGTAAATTAAGCCGATTTGGCTTACAATGA
GATGCCTAAAATTTTCCTATTAGCATATAATAGCTGTAACATAATTAAACAACAATCAATTACAT
TTCGTCATATATATCAATTTTCAGAATCCCCTCAACAGGGCCCTAATCTCATTCAATATATATATA
ATACCTAACTTATACCACGTAACAATTTCTCCAATATCAATCTCAAATAAGATATTTCATTCCAA
ATATTAATTCATAAATTTCCAGAATTCACGTAGAAAGACACAAACCCCATATTAATTTCAAATAT
AAGAATTTTCAAATCATAATCTTCAGATATCACAGAGAAAAGCACAAACCCTATTAATTTTCAG
GAGTTTTGCCCTTATCAGATCTATGGCTTGTCCCGGGCGTATTTTCACCTATCAAAACAATAATT
AAACATTGTATTAATATAATAACAAAGAAAATAATAAATCTGAAATTTTAGGGTTTAGCCAAAT
CTTACCAAAACTCAAATCTATTTGTACTAGTGTGTTCCTCTTATCACGTAGGCTACGATTCTTAA
ATTATTTTCGAATTTAATCGGTTGGAATTGCTAGTGTAAAGCCTGAAAAAATTGGTAAGGGTTT
TGAAAGGAAAGTTAAAAGCTTACTGTCTCGAACTCGAAGGAGAGAGAAAAGAAATGAAGTT
TCCCCCTTTTCTTTCCATTATTTATATAATTAGTATTATTATATTCCTTAATCTCACTTAATTGACTT
AATGAAACTAAAGAAAACAAATTGTTTTATTTTACAATCCAAATCTAAAACTCCACTACCGCC
ATATATATATTATTTTATTTATTATTTTAATATTTCTTTTTTCTTTCTTAGTGGATGTTACAATACCCT
TTGTGATAATGTGTTATTACCATCTAATTGTGTCATTTACAATTGTTTTTGTTATTAATTTAAGAA
CTTATCAATAGTCTATATACAAATATAAAATTTATGGTATGAAAATTCATATTTAAATATTATTTTA
AAACTATTATAAAGTCTTGTGCAAAATTTTCAAGTAAAAAAAGCTAATCACAATTCACAAACC
AGTTTTATGATTTGGTGACAAATGAATTTATTTTACTGAAGAGTTAAATTGTTATTAATGACGCA
TGAATGACGTAAAAAAAAAGTGATGGAAGCATCCATTATCCACCAAAATCAACTTGCCTCATTA
GTCACGTGTTTCTTTGGATGATTCCAATCCACTTGTTCCATCCACTTGCTCTTTATTTAGATGAA
GAACTTTCATGGAAAAATCAGCTCAGGTTGATTATTATCAATTCAAAATAATTTAATATTTTAAA
ATCTTTTATGGTGCATGCACCCTTTTCTTACATCTTTTATCTTATATAATCCCCCTCCTAACTTCTA
TCCCAAAAATTCAGAAAGTTCGATCAATTCATTTTTCTCTCTATAGGTCTTACAAATTCAAGAA
AATCCTAGATTTTGGAAAGCCCTTTAATTTTTTTGATCCGAAGCTTCTGCAATTTCCAAAAGAC
AAAAAAAAGAAAAAAAAAAGCAAGTATTTTTGTTGAATCCTTTTGGCTTAAATTGATCAAAG
ACTGTCACGGTTTTAAGAAAATCCAAGATTTTTGGAAAGAAGCCCCTTTTACTTCAATTGAT
AATTCTTC

doi:10.6342/NTU201901184