國立臺灣大學生命科學院生態學與演化生物學研究所

碩士論文

Institute of Ecology and Evolutionary Biology

College of Life Science

National Taiwan University

Master Thesis

大岩桐 *SsCYC* 在圓葉菸草的表現及性狀觀察與其組織培養再生系統之優化

Ectopic expression of *SsCYC* in *Nicotiana benthamiana* and optimizing regeneration system of *Sinningia speciosa*

郭聞喜

Wen-Hsi Kuo

指導教授：王俊能 博士

指導教授：洪傳揚 博士

Advisor: Chun-Neng Wang, Ph.D.

Advisor: Chwan-Yang Hong, Ph.D.

中華民國 103 年 12 月

December 2014

致謝

能夠完成這本論文，首先要感謝　王俊能老師和　洪傳揚老師這兩年多來的指導與叮嚀，讓學生聞喜在碩士班兩年半的時間中，無論是學業上或是生活態度上都受益良多，不斷突破自我向上提升。同時也要感謝　陳仁治老師及　蔡新聲老師擔任學生的口試委員並在實驗設計、未來規劃及論文的用詞上細心的指正。

剛進到實驗室的時候，感謝講話很快的徐卉汝學姊在實驗和態度上多方的提攜，「不要害別人實驗做不出來」，我一直時時提醒自己。50 嵐紅茶拿鐵 3 分糖去冰的郭立園學長帶剛進實驗室的我抽取 DNA，並且引領我踏進了不熟識的南機場夜市。十項全能的許皓鈞學長在大岩桐植物的認識與載體製作與親緣關係樹上的幫助和教導。坐在我座位後方的辛冠霆學長，總是貼心的關心我的實驗進況與吃飯時間。只有幾個月緣分的曾佐豪學長，學長的論文我一讀再讀讀了好多好多遍，好懷念學長還在時那實驗室會心一笑的氣氛。一直都很直接坦率的小孟學姐，原位雜交後的「所以你是嗎？」，總是那麼令人回味，還有頂樓的烤肉 party，讓我有機會深入認識這個實驗室，口試前學姊的關心，是那麼的窩心。

充滿稚氣微笑的運佑同窗，不管是實驗上的爭辯還是月下的豪腸，論文的堆疊、小房間的滷味似乎荏苒一瞬。坐在我位置旁邊很開朗很樂觀的何承妞，這一年多來的討論與切磋，教學相長再適合不過，還有很厲害的柳丁、雞湯和豆乾及熱情的招呼，讓我每天精神飽滿。搬到窗戶邊有點害羞的培安哥哥，貼心的觀照與幫忙。還有坐在我左邊的伊雯姊姊，回憶似乎太多排山倒海而來：那一顆有翅膀的白柚、布丁還有冰淇淋和數不清的實驗室夜晚，謝謝妳給了我美好的回憶，希望妳開心、有自信，我也不會讓妳失望。小小福控肉飯控的吳和為學弟，謝謝你在基因轉殖上多方的嘗試和合作，為這本論文貢獻良多。還有實驗室新進的學弟妹：育翎、馨儀、小馬、柏宏和國聖，雖然短短幾個月的相識，你們都是我碩士班階段不可缺少的一份子。謝謝小飆仔育翎在我實驗上多方的幫助、料理上的切磋，與貼心的問候及犀利的言詞，讓我深深體會我該趕快長大。謝謝小馬的笑容，實驗之餘的閒聊與二次元世界的分享。

另外，也要感謝洪老師家的趙雲洋學長在我對組織培養懵懂無知的時候，多方指導和拉拔。還有寬宇、喜歡綠色的科科、俊宗學長、188 學長在農桿菌培養和轉殖技術上的分享。

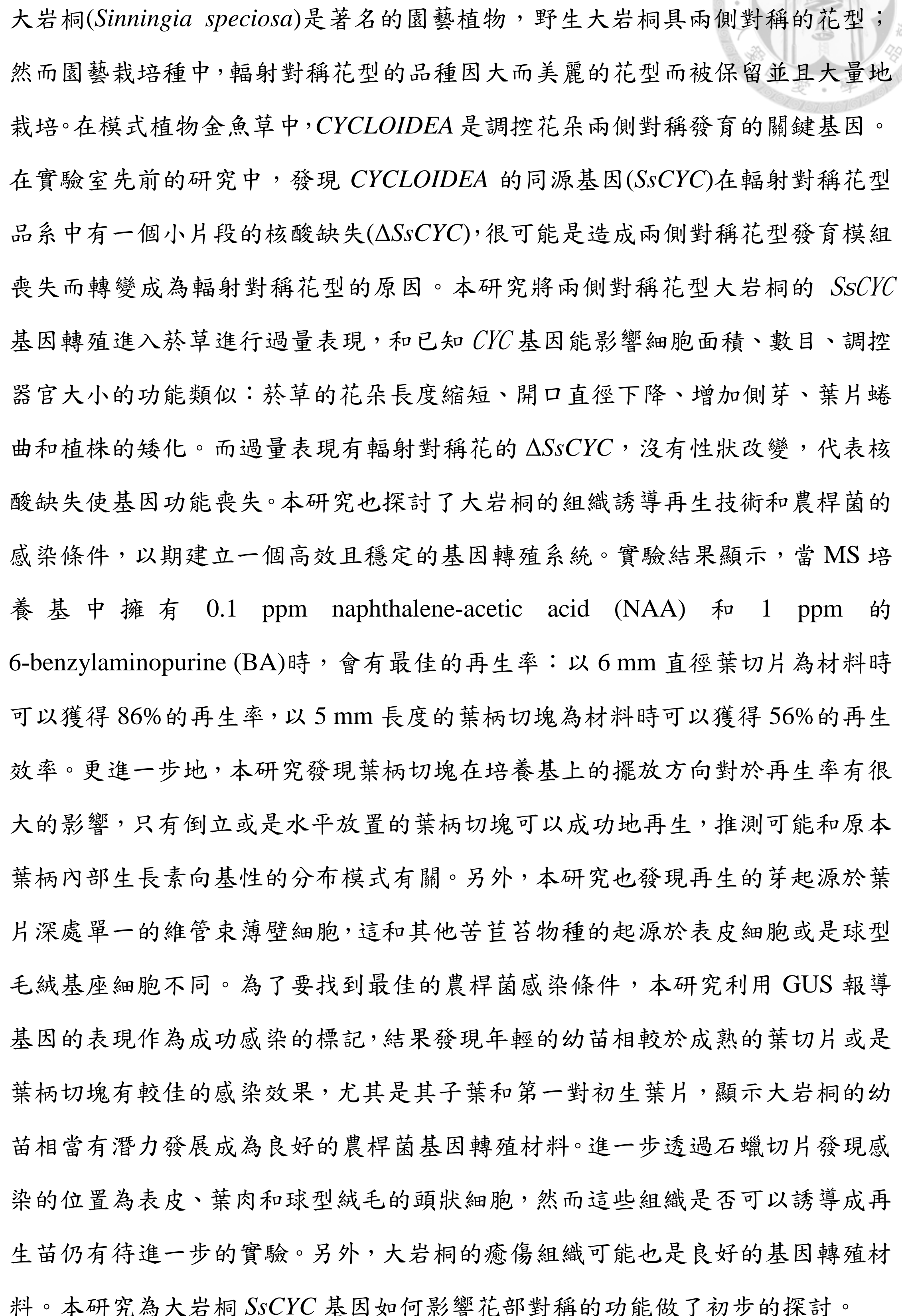
同學之外，碩士班的兩年多也結識了很多朋友：老王實驗室的阿華，謝謝你們家的大型冰箱、水果，還有玩耍與吃喝跟鼓勵。常來串門帶我去 Costco 的王蟲子、12 樓常會遇到的玉米、渼晨還有去過許多地方玩耍的林家琪、越長越大的豆花。宿舍的中文系學弟，晚上回家時的問候，讓疲累的一天有個溫馨的結束。以前的老屁股同學們：115 前成員小昆，總是擔心我寂寞，讓我的 FB 不斷充實。有獨到見解又好聊的謝秉，玩耍和吃飯 murmur 實驗，是個愉快的概念。還有後來就搬回高雄的黑兔于恬，陪我一起認識了許多台北市的夜市和夜景。好同學、好室友、好鄰居的棒子同學，翻翻地瓜、投投飲料、帶著笑容偶爾太鹹的土虱阿姨、校園書坊的臭豆腐、OK 的冰棒，為著論文、排山倒海的報告不就這麼過了？

認真的覺得要感謝的人太多了。為了要完成這本論文，讓我了解到自己能力的渺小，讓我了解到虛心學習和合作的重要，能夠在別人的教導或是成果之下做實驗是一件非常幸福的事情。在做人處事、人際關係方面，回首以前的不成熟，因為你們和我的交集，或開心、或難過、或生氣，這兩年多的成長是這麼來的。

最後，我還要感謝我的家人，我的爸爸、媽媽還有小妹，這 20 幾年來的支持。有你們，我才能順利完成碩士班的學業。謝謝。

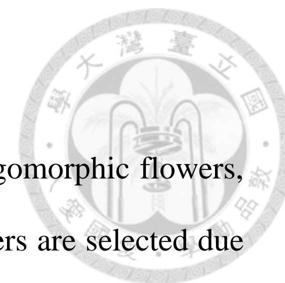
中華民國 103 年 12 月 郭閒喜

中文摘要



大岩桐(*Sinningia speciosa*)是著名的園藝植物，野生大岩桐具兩側對稱的花型；然而園藝栽培種中，輻射對稱花型的品種因大而美麗的花型而被保留並且大量地栽培。在模式植物金魚草中，*CYCLOIDEA* 是調控花朵兩側對稱發育的關鍵基因。在實驗室先前的研究中，發現 *CYCLOIDEA* 的同源基因(*SsCYC*)在輻射對稱花型品系中有一個小片段的核酸缺失(Δ*SsCYC*)，很可能是造成兩側對稱花型發育模組喪失而轉變成為輻射對稱花型的原因。本研究將兩側對稱花型大岩桐的 *SsCYC* 基因轉殖進入菸草進行過量表現，和已知 *CYC* 基因能影響細胞面積、數目、調控器官大小的功能類似：菸草的花朵長度縮短、開口直徑下降、增加側芽、葉片蜷曲和植株的矮化。而過量表現有輻射對稱花的 Δ*SsCYC*，沒有性狀改變，代表核酸缺失使基因功能喪失。本研究也探討了大岩桐的組織誘導再生技術和農桿菌的感染條件，以期建立一個高效且穩定的基因轉殖系統。實驗結果顯示，當 MS 培養基中擁有 0.1 ppm naphthalene-acetic acid (NAA) 和 1 ppm 的 6-benzylaminopurine (BA)時，會有最佳的再生率：以 6 mm 直徑葉切片為材料時可以獲得 86%的再生率，以 5 mm 長度的葉柄切塊為材料時可以獲得 56%的再生效率。更進一步地，本研究發現葉柄切塊在培養基上的擺放方向對於再生率有很大的影響，只有倒立或是水平放置的葉柄切塊可以成功地再生，推測可能和原本葉柄內部生長素向基性的分布模式有關。另外，本研究也發現再生的芽起源於葉片深處單一的維管束薄壁細胞，這和其他苦苣苔物種的起源於表皮細胞或是球型毛絨基座細胞不同。為了要找到最佳的農桿菌感染條件，本研究利用 GUS 報導基因的表現作為成功感染的標記，結果發現年輕的幼苗相較於成熟的葉切片或是葉柄切塊有較佳的感染效果，尤其是其子葉和第一對初生葉片，顯示大岩桐的幼苗相當有潛力發展成為良好的農桿菌基因轉殖材料。進一步透過石蠟切片發現感染的位置為表皮、葉肉和球型絨毛的頭狀細胞，然而這些組織是否可以誘導成再生苗仍有待進一步的實驗。另外，大岩桐的癒傷組織可能也是良好的基因轉殖材料。本研究為大岩桐 *SsCYC* 基因如何影響花部對稱的功能做了初步的探討。

Abstract



The native varieties of *Sinningia speciosa* (Gesneriaceae) bear zygomorphic flowers, but in horticultural varieties, large size showy actinomorphic flowers are selected due to human's preference. *CYCLOIDEA* has been demonstrated to have a major genetic control in zygomorphy in *Antirrhinum*. In actinomorphic varieties, we found a small fragment deletion in its *CYCLOIDEA* homologue (Δ*SsCYC*), which might indicate that the reversal to actinomorphy is a *SsCYC* loss of function mutant. I introduced *CYC* homologues from both zygomorphic cultivar (*SsCYC*) and actinomorphic cultivar (Δ*SsCYC*) into *Nicotiana benthamiana*, a closely related species to *Sinningia speciosa*, to verify whether Δ*SsCYC* has any effect on floral phenotype. I found that ectopic expression of *SsCYC* causes shorter longitudinal length of flowers, smaller floral opening diameters, induction of axillary shoots, curled leaves and dwarfism, agreed with *CYC*'s putative effects on cell proliferation or expansion. However, no visible phenotypic change could be observed in Δ*SsCYC* overexpression lines. I also optimized the genetic transformation system in *Sinningia speciosa*, focusing on tissue regeneration and *Agrobacterium* infection conditions. The MS medium supplied with 0.1 ppm naphthalene-acetic acid (NAA) and 1 ppm 6-benzylaminopurine (BA) was the best for shoot regeneration in both leaf and petiole explants. Eighty six percent and 56% regeneration rates were obtained from 6 mm diameter leaf explants and 5 mm petiole explants, respectively. Moreover, the orientation of petiole explants must be up-side down or horizontal to induce the shoot regeneration, which might relate to the endogenous basipetal distribution of auxin inside the petiole vascular tissue. It was found that the regenerative shoots of explants initiated from a single vascular parenchyma cell deep inside the regenerated tissue. This is different from other reported cases in Gesneriaceae species, in which their regenerative shoots usually

originated from an epidermis cell or a glandular trichome basal cell. To explore whether *Agrobacterium* infection can enter regenerative tissue, transient transformation using GUS reporters was applied. In contrast to mature leaf or petiole explants, I found that young seedlings, especially those in cotyledonary stage or with the first pair of primary leaves have much higher success of transformation. This opens an opportunity that young seedlings are potential material for transformation. By paraffin sections, the positive signals of transformation were seen in epidermis, mesophyll and glandular trichome head cells but not inside the regenerative shoots. Alternatively, the induced callus tissue might be a better transformation material. This study provides a preliminary study on the functions of *SsCYC* genes and guidelines for further optimization of transformation system of *Sinningia speciosa*.

Content

致謝 I

中文摘要 III

Abstract IV

Content VI

Index of Tables and Figures X

Abbreviations XIII

Introduction 1

Floral symmetry and its role in angiosperms diversification 1

Genetic mechanism of floral symmetry 1

Evolutionary history and conserved dorsal expression patterns of *CYC*-like genes among eudicots and monocots 2

Functional divergence between CYC's homologues 4

Derived actinomorphy from zygomorphic ancestors 6

Sinningia speciosa as a good model to study the genetics of floral symmetry 8

Functional analysis of SsCYC and ΔSsCYC in tobacco 12

Tissue culture and genetic transformation system of Sinningia speciosa 12

Aim of this study 15

Materials and Methods 16

Plant material and growth conditions 16

Sequence and phylogenetic analyses of the *SsCYC* gene 16

Southern blotting 17

Isolation of the upstream sequences of the *SsCYC* and Δ*SsCYC* genes 18

Promoter analysis of the *SsCYC* and Δ*SsCYC* genes 22
Vector construction for the functional analysis of the *SsCYC* and Δ*SsCYC* genes 22
Transformation of *Nicotiana benthamiana* leaf disk 24
Preparation of aseptic tobacco seedling 24
Leaf disk infection 25
Selection of transformed shoots 26
Clearing method for measuring the cell size..27
Genomic DNA extraction 27
Total RNA extraction and reverse transcription...29
Polymerase chain reaction 30
Preparation of aseptic seedlings for tissue culture...31
Tissue culture of *Sinningia speciosa*...32
Histological analysis by paraffin sections..33
Scanning electron microscope (SEM) 33
Transient transformation using GUS reporter gene 34
Statistical analysis..37
Results...38
Conserved domains of the *SsCYC* gene 38
Reconstruction the phylogenetic tree of *SsCYC* gene...45
Confirming the single copy of *SsCYC* gene in *Sinningia speciosa* 47
Isolation the upstream and coding sequences of *SsCYC* and Δ*SsCYC* 49
Comparison of upstream sequence between SsCYC and ΔSsCYC............49
Comparison of coding sequence between SsCYC and ΔSsCYC.................50
Construction the overexpressed *SsCYC* and Δ*SsCYC* transgenic lines................54

Abnormal phenotypes among *SsCYC* transgenic plants 61

Floral morphology measurement 66

Association between phenotypic abnormalities and mRNA expression levels 70

Floral tube cell shape and size measurements 74

Tissue culture conditions for shoot induction in *Sinningia speciosa* 78

Effects of petiole explant orientation on shoot regeneration 83

Callus induction from embryo 86

Histological observation of shoot regeneration 88

Scanning electron microscope observation..92

Testing the optimal explants source for *Agrobacterium*-mediated transformation 94

Discussion 98

SsCYC is a homologue to *AmCYC* by sharing domains and phylogeny 98

Δ*SsCYC* appears to lose its function due to a small fragment deletion 99

Δ*SsCYC* gene is selected by human for its derived actinomorphy 100

SsCYC represses the primary growth but promote the development of axillary buds 101

The downstream pathway of SsCYC is different from AmCYC...............103

SsCYC might cause male and female sterility in the plants with severe phenotypes 105

The results of functional analyses coincide with the genetic approach 106

Shoot regeneration has two hormone dependent pathways 107

Orientation and size of petiole explants have effects on the regeneration rate..109

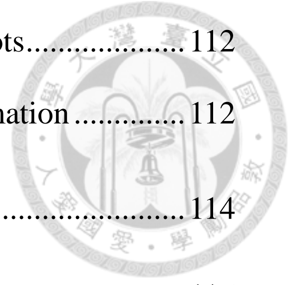
Shoots regeneration of *Sinningia speciosa* through direct embryogenesis........110

Meristem identities are unstable in the newly regenerative shoots....112

Young seedlings or callus might be a good source for transformation....112

Conclusion and Future Prospects....114

References....116



Index of Tables and Figures

Table 1 Genes used in construction of phylogenetic tree (Figure 2) 41

Table 2 Predicted trans-acting elements with known functions at the upstream regions of *SsCYC* and Δ*SsCYC*. 52

Figure 1 Floral morphologies of *Sinningia speciosa* 10

Figure 2 Conserved domains and phylogeny of the *SsCYC* gene 44

Figure 3 Phylogeny of the *SsCYC* gene 46

Figure 4 Southern blotting of detecting the copy number of *SsCYC* gene in Avanti (horticulture) and Espirito Santo (wild type) cultivars 48

Figure 5 Illustration of *SsCYC* and Δ*SsCYC* genes and their putative promoters 53

Figure 6 T-DNA insertion in T_0 transgenic populations 57

Figure 7 mRNA expression levels of *SsCYC* or Δ*SsCYC* in T_0 transgenic lines 59

Figure 8 Mature T_0 transgenic plant morphology 60

Figure 9 Different degrees of abnormality among the *SsCYC* transgenic plants 62

Figure 10 Promoting the growth of axillary buds in *SsCYC* transgenic plants 63

Figure 11 Leaf morphology of T_0 transgenic plants 65

Figure 12 Flower morphology of T_0 transgenic plants at fully elongated stage 67

Figure 13 Flower morphology measurement of *SsCYC*, Δ*SsCYC* transgenic tobacco plants. 69

Figure 14 Relative *SsCYC* mRNA expression level among T_0 population 71

Figure 15 Flower morphology according to the relative mRNA expression level in *SsCYC* transgenic plants 73

Figure 16 Cell shape and size of floral tube measurement of *SsCYC* and Δ*SsCYC* transgenic tobacco plants 77

Figure 17 Effect of cytokinin and auxin concentration and explant sources on regeneration rate and shoots per responsive explant..80

Figure 18 Effect of cytokinin and auxin concentration and explant sources on the morphology of shoot regeneration 82

Figure 19 Effects of petiole explant orientations on shoot regeneration rate and shoots per responsive explant...84

Figure 20 Effects of petiole explant orientations on the morphology of regeneration ... .85

Figure 21 Callus inductions from young seedlings...87

Figure 22 Process of shoot regeneration from leaf explants by paraffin section and dissecting microscope 91

Figure 23 Micro-scale observation of shoot regeneration by scanning electronic microscope 93

Figure 24 Tissue competence for *Agrobacterium* infection by GUS transient assay..97

Supplementary Table 1 Arbitrary Degenerative Primer 124

Supplementary Table 2 Putative transcription factor binding sites of *SsCYC*/Δ*SsCYC* 125

Supplementary Table 3 Primer list in this study 128

Supplementary Table 4 Flower morphology measurement of SsCYC and ΔSsCYC transgenic tobacco plants 130

Supplementary Table 5 Flower morphology according to the relative mRNA expression level in *SsCYC* transgenic plants 131

Supplementary Figure 1 Destination Vector pK2GW7,0..132

Supplementary Figure 2 Amino acid alignment of CYC-like gene for reconstruction the phylogeny..133

Supplementary Figure 3 Promoter regions of *SsCYC* and Δ*SsCYC* genes 135

Supplementary Figure 4 Nucleotide sequence of *SsCYC* and Δ*SsCYC* genes..138

Supplementary Figure 5 Amino acid sequence of *SsCYC* and Δ*SsCYC* genes..139

Supplementary Figure 6 Equal variance and normality tests of the floral length of three transgenic populations 140

Supplementary Figure 7 Equal variance and normality tests of the floral length of different mRNA expression levels among *p35S*::*SsCYC*::*c-Myc* transgenic population 141

Supplementary Figure 8 Preliminary result of Southern blotting of T_0 transgenic plants..142

Abbreviations

AD primer	Arbitrary degenerate primer
BA	6-benzyladenine purine
cDNA	Complimentary DNA
CTAB	Hexadecyl trimethyl-ammonium bromide
CYC	CYCLOIDEA
DICH	DICHOTOMA
DIV	DIVARICATA
DNA	Deoxyribonucleic acid
GUS	Beta-glucuronidase
GS primer	Gene-specific primer
MS medium	Murashige and Skoog medium
NAA	Naphthalene acetic acid
RAD	RADIALIS
RNA	Ribonucleic acid
SsCYC	Homologous gene of *CYC* in native zygomorphic S*inningia speciosa*
Δ*SsCYC*	Homologous gene of *CYC* in horticultural actinomorphic *Sinningia speciosa* with deletion
SsCYC transgenic plant	Tobacco T_0 transgenic plant with *p35S*::*SsCYC*::c-*Myc*
Δ*SsCYC* transgenic plant	Tobacco T_0 transgenic plant with *p35S*::c-*Myc*:: Δ*SsCYC* or *p35S*:: Δ*SsCYC*::c-*Myc*
RT	Reverse transcription
TAIL PCR	Thermo asymmetry interlaced PCR
TCP	**T**EOSINTE BRANCHED 1, **C**YCLOIDEA and **P**ROLIFERATING CELL FACTOR 1
TE	Tris-EDTA
RFLP	Restriction fragment length polymorphism

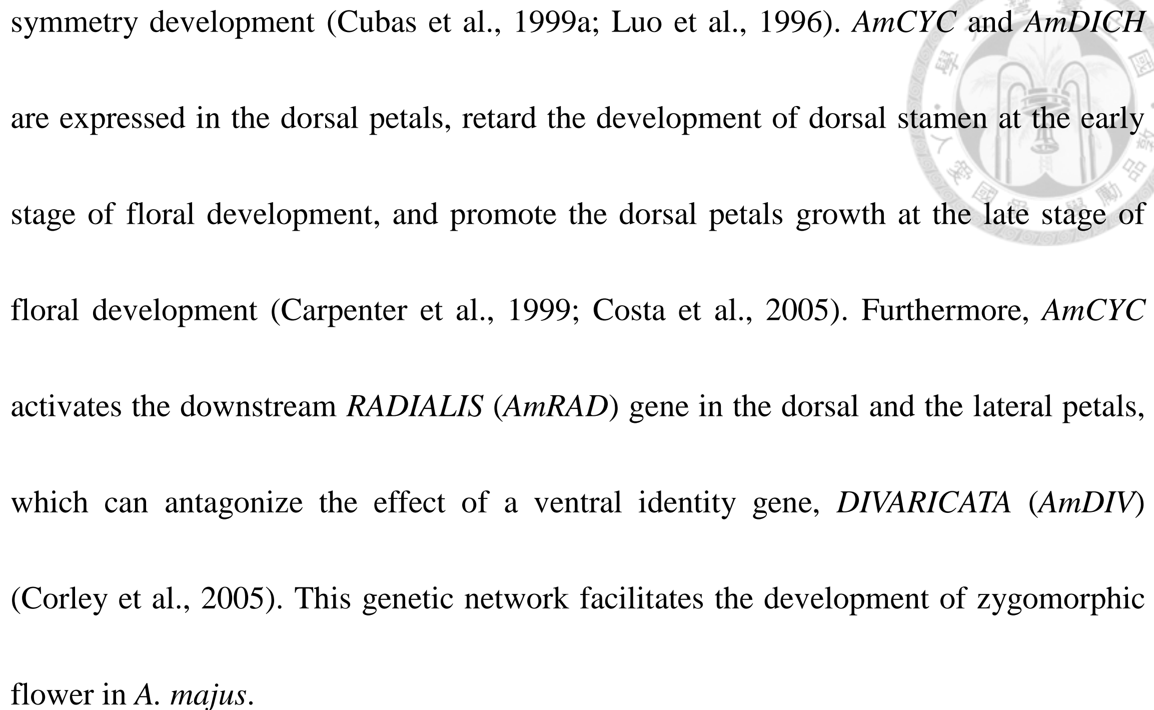
Introduction

Floral symmetry and its role in angiosperms diversification

The evolution of zygomorphy (dorsoventral asymmetry, monosymmetry) is one of the most important events in angiosperms diversification. Zygomorphic flower is thought to have first evolved at about 85 million years ago, coincided with the first radiation of angiosperms and considered playing a key role in angiosperm diversification (Crepet, 2008; Crepet and Niklas, 2009; Crepet et al., 2004). In addition, by comparing species numbers from zygomorphic and actinomorphic sister groups, more species evolved in the zygomorphic groups than that of the actinomorphic ones. This supports the hypothesis that the zygomorphic clade might have higher speciation rate than their sister actinomorphic groups (Sargent, 2004). Generally, zygomorphic flowers, usually bearing elaborated corolla shapes and asymmetric positions of stamens and carpels, are considered having higher the pollination efficiency and specificity, thus increase the speciation rate (Ushimaru et al., 2009; Vamosi and Vamosi, 2010).

Genetic mechanism of floral symmetry

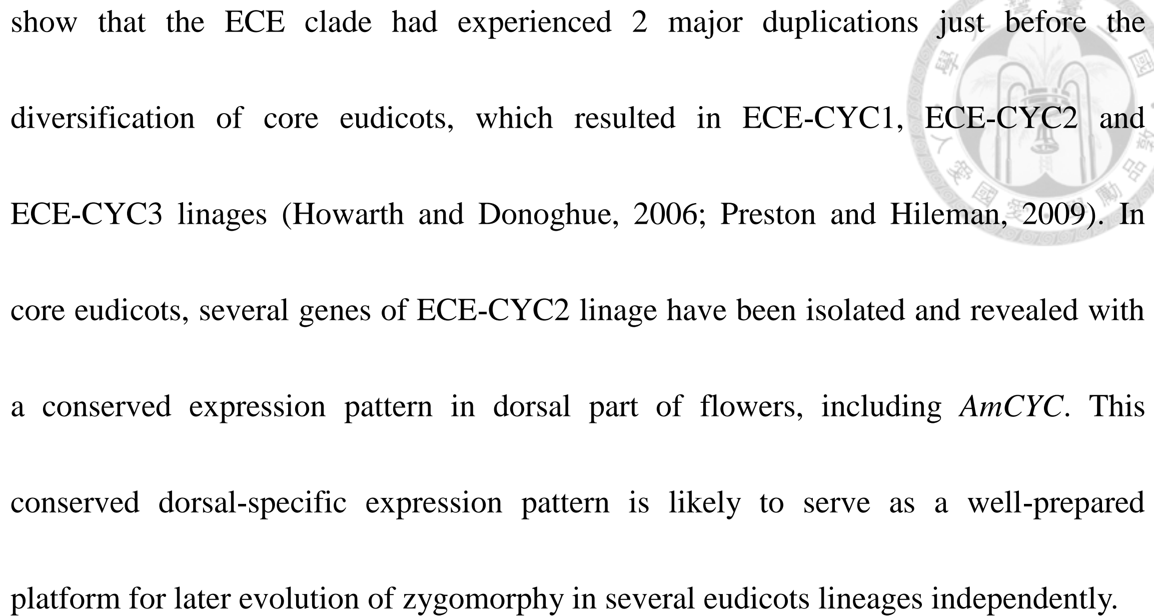
The genetic mechanism of this extraordinary evolution was first revealed in the model plant *Antirrhinum majus* (Plantagainaceae). Two TCP family genes, the *CYCLOIDEA* (*AmCYC*) gene and its paralog *DICHOTOMA* (*AmDICH*) gene play key roles in floral



symmetry development (Cubas et al., 1999a; Luo et al., 1996). *AmCYC* and *AmDICH* are expressed in the dorsal petals, retard the development of dorsal stamen at the early stage of floral development, and promote the dorsal petals growth at the late stage of floral development (Carpenter et al., 1999; Costa et al., 2005). Furthermore, *AmCYC* activates the downstream *RADIALIS* (*AmRAD*) gene in the dorsal and the lateral petals, which can antagonize the effect of a ventral identity gene, *DIVARICATA* (*AmDIV*) (Corley et al., 2005). This genetic network facilitates the development of zygomorphic flower in *A. majus*.

Evolutionary history and conserved dorsal expression patterns of *CYC*-like genes among eudicots and monocots

Phylogenetic studies of *AmCYC* reveal that this genes belong to TCP gene family (**T**EOSINTE BRANCHED 1, **C**YCLOIDEA and **P**ROLIFERATING CELL FACTOR 1), which is characterized by an unique non-canonical basic-Helix-Loop-Helix domain (TCP domain) (Cubas et al., 1999a). Generally, TCP genes are transcription factors that regulate patterns of cell division and proliferation (Martin-Trillo and Cubas, 2010). This gene family can be divided into 2 subfamilies, PCF (PROLIFERATING FACTOR 1) and CYC/TB1 (CYCLOIDEA/TEOSINTE BRANCHED 1) subfamilies. Within the CYC/TB1 subfamily, *AmCYC* gene belongs to the ECE clade, which is characterized by a glutamate-cysteine-glutamate motif between TCP and R domains. Recent phylogenies



show that the ECE clade had experienced 2 major duplications just before the diversification of core eudicots, which resulted in ECE-CYC1, ECE-CYC2 and ECE-CYC3 linages (Howarth and Donoghue, 2006; Preston and Hileman, 2009). In core eudicots, several genes of ECE-CYC2 linage have been isolated and revealed with a conserved expression pattern in dorsal part of flowers, including *AmCYC*. This conserved dorsal-specific expression pattern is likely to serve as a well-prepared platform for later evolution of zygomorphy in several eudicots lineages independently.

After the first discovery of *AmCYC* gene, the homologous genes in several plant lineages have been found to participate in the evolution of zygomorphy independently (Damerval and Manuel, 2003; Preston and Hileman, 2009). For example, in the *Helianthus annuus* (Asteraceae, Asterids), the expression of *HaCYC*s gene (ECE-CYC2) are detected in the zygomorphic ray florets but not in the actinomorphic disk florets, suggesting that the *HaCYC*s genes participate in the development of capitulum inflorescences (Chapman et al., 2012). In the *Lotus japonicas* (Leguminosae, Rosids), the *LjCYC*s gene (ECE-CYC2) not only participate in the establishment of the dorsal identity of flowers but also regulate the inner asymmetry development of petals (Feng et al., 2006; Xu et al., 2013). Even outside the eudicots, in *Oryza sativa* (Poaceae, Monocots), the *RETARDED PALE1* (*REP1*), a *CYC*'s homologue not belonging to ECE-CYC2, can promote the growth of dorsal palea (homologous organ of bract or

petal) in florets, which relates to the development of zygomorphy (Yuan et al., 2009). The previous works extend our knowledge that *CYC*-like genes have a conserved or similar dorsal expression patterns in the development of zygomorphic flowers among the eudicots and monocots. However, the phenotypic effects of *CYC*-like genes on floral development seem to be quite diverse among different species and most of them still remain unclear.

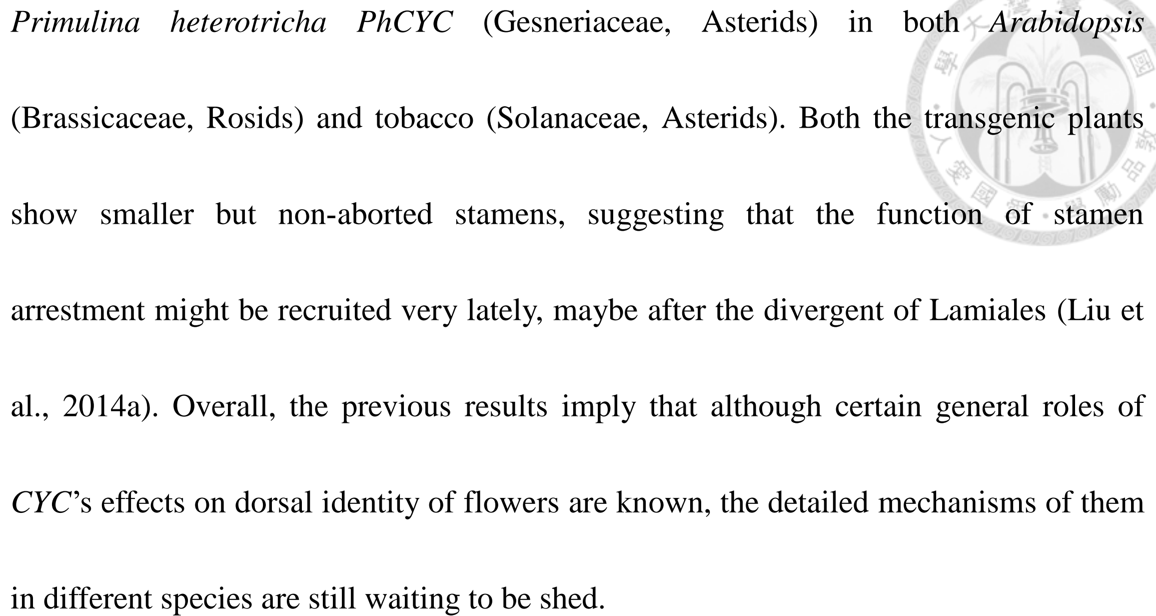
Functional divergence between CYC's homologues

The development of zygomorphic flower is attributed to the unequal growth of the dorsal and ventral petals. How *CYC* homologues regulate such unequal growth is revealed by the functional analyses. In *Antirrhinum majus* (Plantagainaceae), dorsal petals are longer than the ventral petal. Overexpression of the *AmCYC* in *Arabidopsis* increases the petal size by increasing the cell size. These indicated that *AmCYC* can promote the growth of the dorsal petals by cell elongation rather than cell proliferation (Costa et al., 2005).

However, *CYC*'s role on cell growth is not conserved among the other zygomorphic flower species. As *CYC*'s homologue in *Primulina heterotricha* (Gesneriaceae), the dorsal petals are smaller than the ventral petal. Overexpression of *PhCYC1C* in *Arabidopsis* decreases the petal size by decreasing the cell area, indicating that *PhCYC1C* can repress the growth of the dorsal petals by retarding the cell elongation

(Liu et al., 2014a; Yang et al., 2012), which is opposite to *Antirrhinum majus*. In another case, *Iberis amara* (Brassicaceae), a close relative of *Arabidopsis,* has dorsal petals smaller than the ventral petal, and overexpression of its *CYC* homologue, *IaTCP1*, in *Arabidopsis* decreases the petal size by decreasing the number of cells rather than reducing the cell size (Busch and Zachgo, 2007). Based on above, these functional analyses prove that the unequal growth of the dorsal and ventral petals can be resulted from different developmental mechanisms.

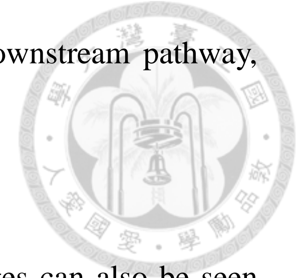
On the other hand, the expression of *AmCYC* also correlates the abortion of dorsal stamen in *Antirrhinum majus*, suggesting that *CYC*'s homologues might repress or arrest the development of stamens (Carpenter et al., 1999). This hypothesis is also supported by the functional analysis of *Gerbera* hybrid (Asteraceae) *GhCYC2*. Overexpression of *GhCYC2* gene retards the stamens and inhibits the pollen production similar to the morphology of ray florets (Broholm et al., 2008). Nevertheless, the abortion or retardation of the dorsal stamen is not found in the functional analyses of *CYC*'s homologues in *Lotus japonica* (Leguminosae) (Feng et al., 2006), *Pisum sativum* (Leguminosae) (Wang et al., 2008) and *Iberis amara* (Brassicaceae) (Busch and Zachgo, 2007). This inconsistency indicates that the later recruited functions on stamen development might be different from two major linages of core eudicots: the asterids and rosids. This hypothesis was further revised by the study of overexpression of the



Primulina heterotricha PhCYC (Gesneriaceae, Asterids) in both *Arabidopsis* (Brassicaceae, Rosids) and tobacco (Solanaceae, Asterids). Both the transgenic plants show smaller but non-aborted stamens, suggesting that the function of stamen arrestment might be recruited very lately, maybe after the divergent of Lamiales (Liu et al., 2014a). Overall, the previous results imply that although certain general roles of *CYC*'s effects on dorsal identity of flowers are known, the detailed mechanisms of them in different species are still waiting to be shed.

Derived actinomorphy from zygomorphic ancestors

Although it seems that the zygomorphic flowers take many advantages over the actinomorphic flowers in pollination syndromes, derived actinomorphy (reversal evolution) from a zygomorphic ancestor has been reported in many studies. Some of them are preferred by natural selection, for instance, the preference of wind pollination; others are occasional mutant or are selected by human. The genetic mechanisms of the derived actinomorphy are even more complex than the primary evolution of zygomorphic flowers. For example, in *Cadia purpurea* (Leguminosae), a rare actinomorphic species in Legume family, *LegCYC1B* extends its expression to the whole corolla, which is different from its zygomorphic relative, *Lupinus nanus* (Leguminosae) whose *LegCYC1B* is dorsal expressed (Citerne et al., 2006). In *Plantago*, an actinomorphic sister group to *Antirrhinum*, the *PlCYC1A* has lost its dorsal



expression pattern during the evolution and cannot turn on the downstream pathway, contributing to the zygomorphy development (Preston et al., 2011).

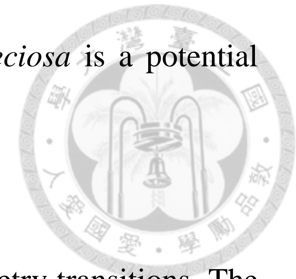
In addition to the interspecific floral symmetry reversal, the changes can also be seen intraspecifically. The most well-known case is *Linaria vulgaris*, which drew the attention of Linnaeus more than 250 years ago. The common *Linaria vulgaris* has a classic zygomorphic flower with 2 elongated dorsal petals and a spurred ventral petal. However, within the natural populations, the actinomorphic mutants with all ventral-like petals with spurs are occasionally recorded. Through the methylation sensitive RFLP analysis shows that the *Lvcyc* gene in the actinomorphic individuals is heavily methylated and silenced compare to that in the common zygomorphic *Linaria vulgaris*, suggesting that DNA methylation might also be involved in the reversion of floral zygomorphy (Cubas et al., 1999b).

Still other cases of derived actinomorphy are found in horticultural plants, in which the actinomorphic phenotypes are selected by human due to their preferences and commercial values. For example, in *Saintpaulia* sp. and *Sinningia speciosa*, the actinomorphic cultivars can only be seen in horticulture but not in the wild (Citerne and Cronk, 1999; Citerne et al., 2000). How human selection affected the very recent genetic changes of floral symmetry reversion in these important horticultural species remains largely unknown. Detailed study of these mechanisms might provide insights

into how human selection strongly affected *CYC*-like gene regulations on floral symmetry reversal among these cultivars.

***Sinningia speciosa* as a good model to study the genetics of floral symmetry**

Sinningia speciosa is a widely cultivated ornamental plant with many flower traits available, such as color, shape, number of whorls and symmetry patterns (actinomorphy or zygomorphy). These traits have been domestically selected and preserved within the last 200 years since its first collection on 1815. The rapid changes of these flower traits reminisce of the saltatory evolution of diverse floral traits in angiosperm during Turonian stage (in late Cretaceous), including the first fossil record of zygomorphic flower (Crepet, 2008). Thus, understanding the genetic regulations behind these floral traits development in *Sinningia speciosa*, particularly the rapid reversal of floral symmetry, may give us cues on the evolution history of angiosperms diversification. In addition, these cultivars with contrast floral symmetry patterns belong to the same species, thus the crossing between zygomorphic and actinomorphic cultivars can generate fertile offspring. This allows us to explore whether simple or complicate genetic control is involved with this transition. Moreover, the genome size of *Sinningia speciosa* is small, only about twice the size of *Arabidopsis* (Zaitlin and Pierce, 2010). This is suitable for comparative transcriptomes or genome analysis, which are the

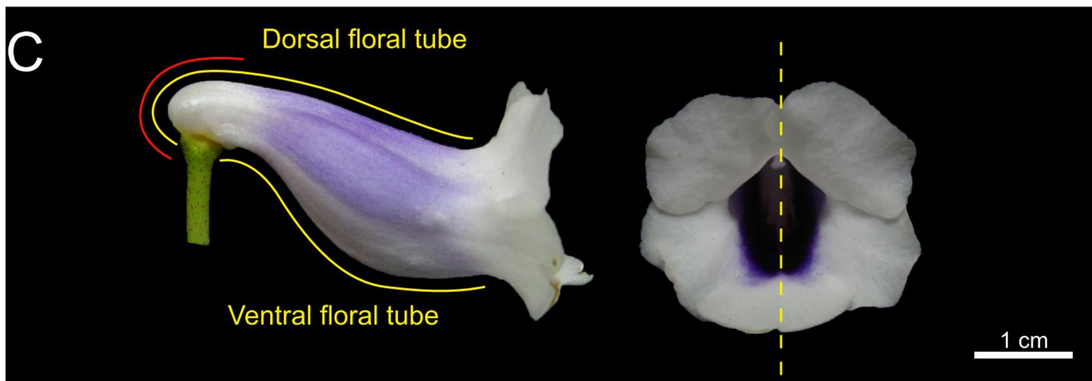


ongoing projects in our lab. All these facts suggest *Sinningia speciosa* is a potential good model species for studying the floral traits evolution.

In this study, we focused on finding *CYC*'s role on the floral symmetry transitions. The native cultivar of *Sinningia speciosa* has nodding, zygomorphic flowers with white, pink or purple color, which takes the advantage of bee pollination (Fig. 1 A & C) (Zaitlin, 2012). After only the 30 years in horticulture, the actinomorphic flower cultivars with erect pedicles were selected by human due to their showy large size flowers (Fig. 1 B & D) (Citerne and Cronk, 1999). This saltatory change implies that human might select on some of the key genes responsible for floral symmetry development. Our previous data has shown that there is a small fragment deletion in the *CYC*'s homologous gene (Δ*SsCYC*) in actinomorphic cultivar. In addition, this deletion is conserved among all available actinomorphic cultivars, implying that all present actinomorphic cultivars might originate from this single deletion event. On the other hand, in native zygomorphic cultivar, the dorsal specific expression of *SsCYC* gene (the *CYC*'s homologous gene) has been revealed by real time PCR (unpublished data in our lab), supporting the hypothesis that the reversal into actonomorphy is resulted from the loss function of *CYC*'s homologue. However, lacking the data of functional analysis, whether the deletion in the Δ*SsCYC* is the main factor leading to the floral symmetry reversion still remains ambiguous.

A
8 cm

B
8 cm


C
Dorsal floral tube
Ventral floral tube
1 cm

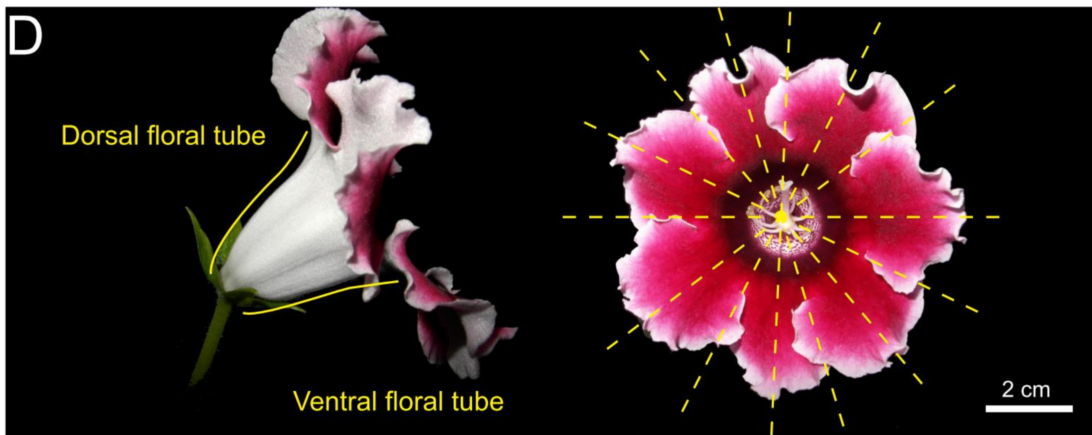
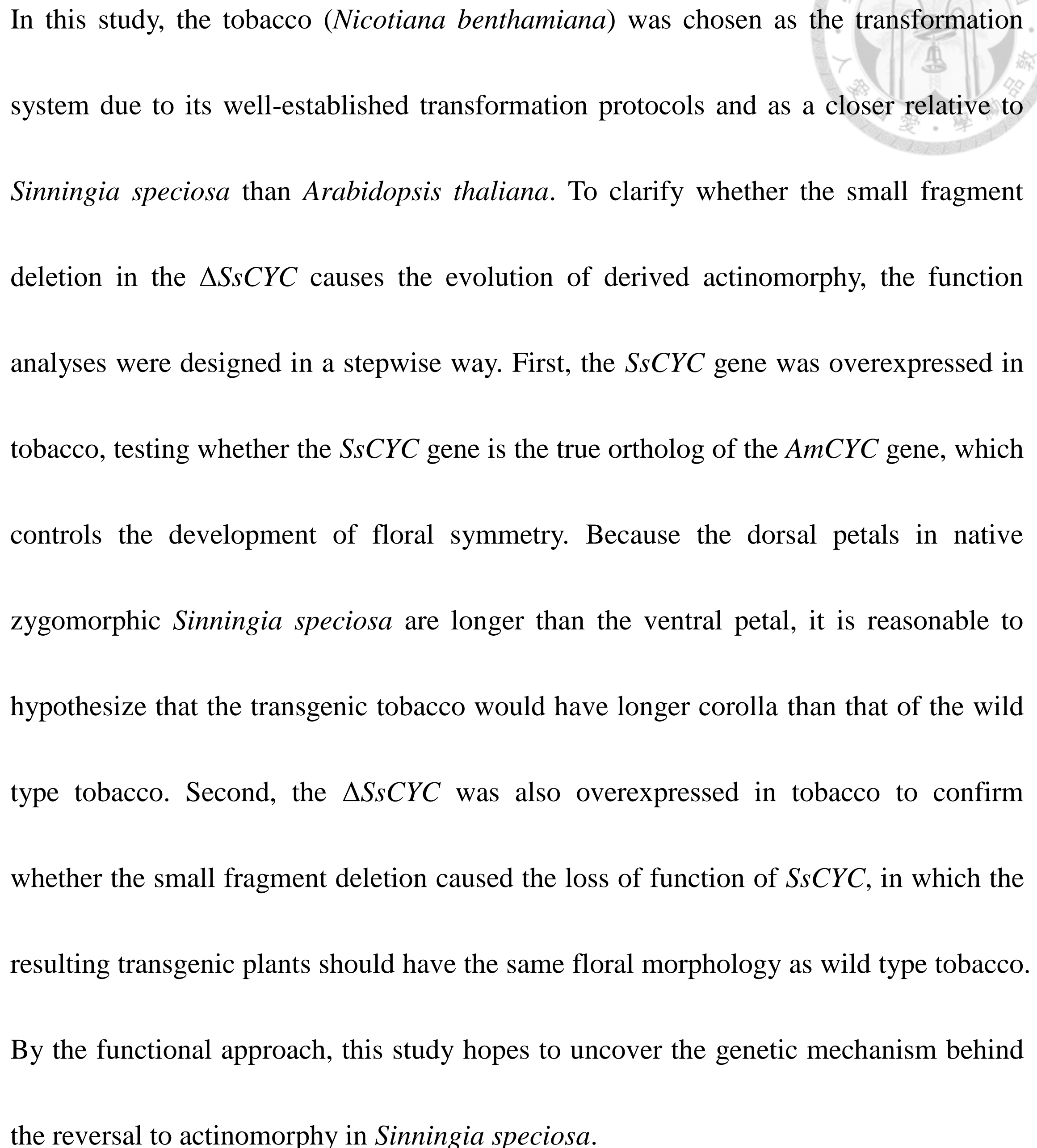

D
Dorsal floral tube
Ventral floral tube
2 cm

Figure 1 Floral morphologies of *Sinningia speciosa*

A, Native zygomorphic cultivar 'Carangola'. **B**, horticultural actinomorphic cultivar 'Avanti'. **C**, floral morphology of native zygomorphic cultivar 'Carangola' (from Hsu, Hao-Chun, personal communication). **D**, floral morphology of horticultural actinomorphic cultivar 'Avanti'. Left, side views. Right, front views. Dashed lines, the symmetry planes. Yellow solid line, the floral tube. Red solid line, the nectar spur.

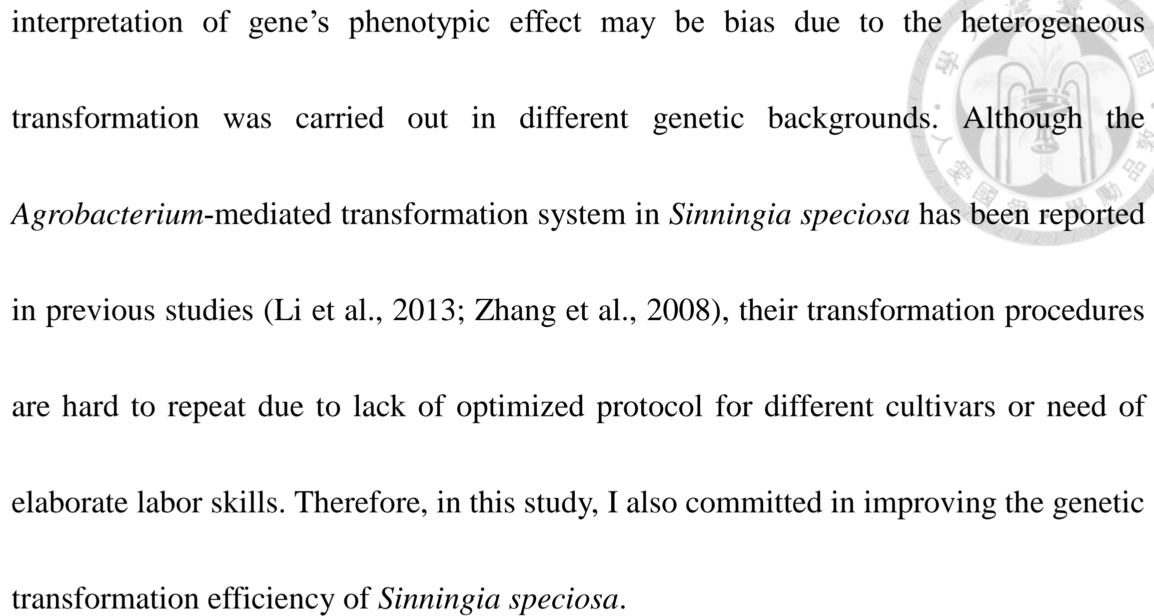
Functional analysis of SsCYC and ΔSsCYC in tobacco



In this study, the tobacco (*Nicotiana benthamiana*) was chosen as the transformation system due to its well-established transformation protocols and as a closer relative to *Sinningia speciosa* than *Arabidopsis thaliana*. To clarify whether the small fragment deletion in the Δ*SsCYC* causes the evolution of derived actinomorphy, the function analyses were designed in a stepwise way. First, the *SsCYC* gene was overexpressed in tobacco, testing whether the *SsCYC* gene is the true ortholog of the *AmCYC* gene, which controls the development of floral symmetry. Because the dorsal petals in native zygomorphic *Sinningia speciosa* are longer than the ventral petal, it is reasonable to hypothesize that the transgenic tobacco would have longer corolla than that of the wild type tobacco. Second, the Δ*SsCYC* was also overexpressed in tobacco to confirm whether the small fragment deletion caused the loss of function of *SsCYC*, in which the resulting transgenic plants should have the same floral morphology as wild type tobacco. By the functional approach, this study hopes to uncover the genetic mechanism behind the reversal to actinomorphy in *Sinningia speciosa*.

Tissue culture and genetic transformation system of Sinningia speciosa

For non-model organisms, functional analyses by genetic transformation are often conducted in the closely related model organisms due to the limited genetic transformation system or reliable protocols within the same species. However,



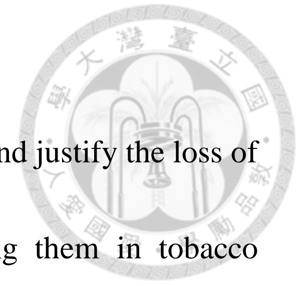
interpretation of gene's phenotypic effect may be bias due to the heterogeneous transformation was carried out in different genetic backgrounds. Although the *Agrobacterium*-mediated transformation system in *Sinningia speciosa* has been reported in previous studies (Li et al., 2013; Zhang et al., 2008), their transformation procedures are hard to repeat due to lack of optimized protocol for different cultivars or need of elaborate labor skills. Therefore, in this study, I also committed in improving the genetic transformation efficiency of *Sinningia speciosa*.

A detailed test of tissue culture conditions of shoot regeneration is a crucial prerequisite for developing a stable genetic transformation system because a stable transgenic line must regenerate from a single or few successful transformed cells. Various parameters may have impacts on the efficiency of plant tissue culture, for example: the source of explant, the orientation how explants attach to the medium, the composition of nutrient medium, the exogenous plant growth regulators and other abiotic or biotic factors. In recent years, tissue culture technique of *Sinningia speciosa* has been extensively studied (Chae et al., 2012; Nhut et al., 2006; Pang et al., 2006; Scaramuzzi et al., 1999; Wuttisit and Kanchanapoom, 1996; Xu et al., 2009). Unfortunately, it is hard to compare the culture conditions among these studies due to inconsistency of culture period and different explant size. Moreover, previous studies have not addressed the effect of explant orientation on the regeneration efficiency. Nor have these indicated the explant

type most suitable for shoot regeneration, on which the successfulness of transformation protocol relies and whether regenerated shoots can be infected by certain *Agrobacterium* strains.

On the other hand, different plant tissues might have different competencies for *Agrobacterium* infection. For maximizing the transformation efficiency, choosing the adequate tissue sources is also very important. Therefore, this study not only examined the best tissue culture conditions, including the hormone concentration and combination, the explant orientations and the developmental process of the regenerative shoots, but also compared the infective competencies of different tissues by the *Agrobacterium* *GUS* reporter assay.

Aim of this study



The aim of this study is to conduct the function analysis of *SsCYC* and justify the loss of function of a small fragment deleted Δ*SsCYC* by overexpressing them in tobacco (*Nicotiana benthamiana*). In addition, in order to conduct functional analyses within the same homogenous species, *Sinningia speciosa*, this study also aims to explore the best tissue culture conditions and *Agrobacterium*-mediated genetic transformation system of *Sinningia speciosa*.

Materials and Methods

Plant material and growth conditions

The native zygomorphic cultivar ‘Carangola’ was kindly provided by Dr. Cecilia Koo Botanic Conservation Center (*Sinningia speciosa* ‘Carangola’, accession number: K015265). The horticultural actinomorphic cultivar ‘Avanti’ (Gloxinia Avanti, blue and white) was purchased from a local seed distributer, Taiwan Horticultural Co., Ltd. Plants were grown in 4 inch diameter pots, placed in an incubator at 25℃, 16 h daylight and a relative humidity of 80%. The potting medium was a mixture of peat (Potgrond H, klasmann Deilmann Gmbh, Germany), vermiculite and perlite in 2: 1: 1 ratio.

Sequence and phylogenetic analyses of the *SsCYC* gene

In order to compare the conserved domains between the SsCYC and other CYC/TB1 ECE clade members, the positions of TCP domain, ECE domain and the R domain were located by the studies of Cubas et al. (1999a) and Howarth and Donoghue (2005), and the secondary structures of TCP domain were annotated followed by Cubas et al. (1999a).

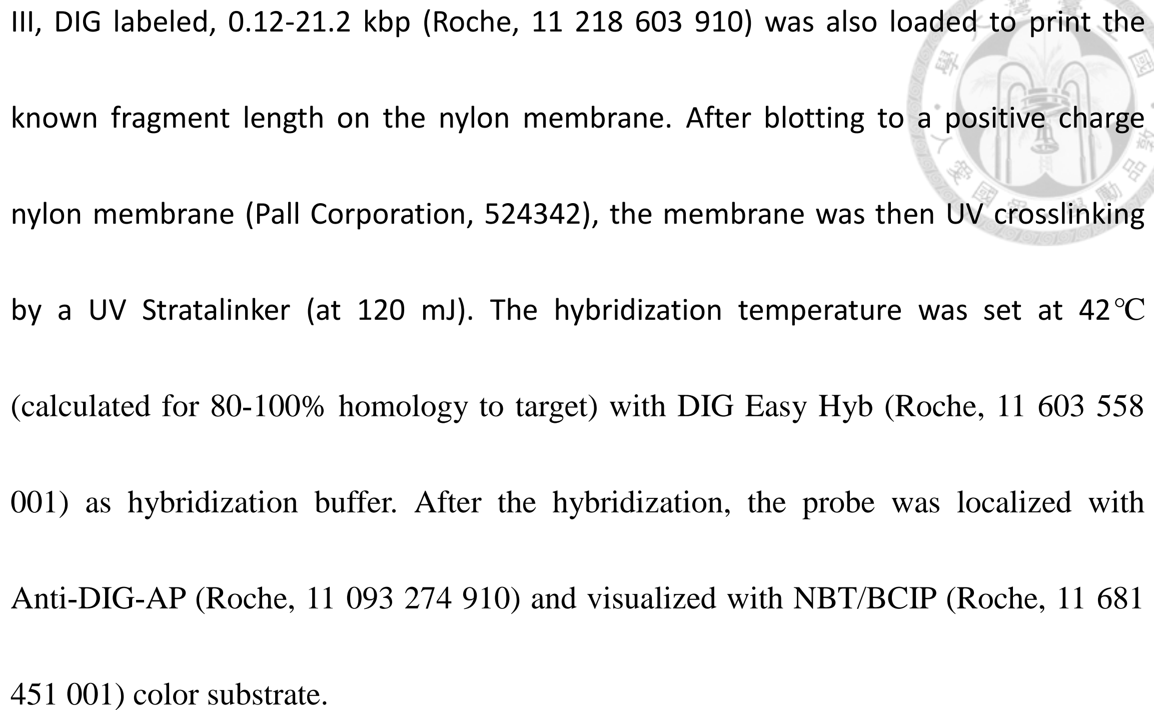
To ascertain the phylogenetic position of *SsCYC* gene within the ECE-CYC2 clade, the Maximum likelihood (ML) and Bayesian inference (BI) tree was reconstructed based on the whole coding nucleotide sequences, which were aligned by MAFFT version 7

software (http://mafft.cbrc.jp/alignment/software/)(Supplementary figure 2). For ML tree, the aligned matrix without partition was analyzed by GARI 2.0 software with TrN+I+γ model, which was prior selected by jmodeltest2 software (https://code.google.com/p/jmodeltest2/). ML nodal support was calculated by analyzing 500 bootstrap replicates in GARLI software. MrBayes software was used for BI tree reconstruction. 1,000,000 generations were employed and sampled every 500 generations with Running Markov chain. The diagnostic frequency is 5000.

Southern blotting

In order to confirm the single copy of CYC-like gene in *Sinningia speciose*, a non-radiative DIG based Southern blotting technique was applied in this study. All the procedures were followed by Roche's DIG Application Manual for Filter Hybridization. The probe was designed before the TCP domain and ending after the R domain, which was synthesized by PCR DIG Probe Synthesis Kit (Roche, 1 1 636 090 910) with 1: 6 ratio of DIG-dUTP: dTTP. 10 pg of plasmid DNA, harboring the wild type *SsCYC* gene, was used as the template in a standard 50 μL reaction. The PCR product was used directly as the probe for hybridizing the genomic DNA from both wild type (Espirito Santo) and horticulture (Avanti) cultivars.

For each lane, about 20 μg genomic DNA was digested with selected restriction enzyme and loaded directly into the gel. At the same time, 5 μL DNA Molecular Weight Marker



III, DIG labeled, 0.12-21.2 kbp (Roche, 11 218 603 910) was also loaded to print the known fragment length on the nylon membrane. After blotting to a positive charge nylon membrane (Pall Corporation, 524342), the membrane was then UV crosslinking by a UV Stratalinker (at 120 mJ). The hybridization temperature was set at 42℃ (calculated for 80-100% homology to target) with DIG Easy Hyb (Roche, 11 603 558 001) as hybridization buffer. After the hybridization, the probe was localized with Anti-DIG-AP (Roche, 11 093 274 910) and visualized with NBT/BCIP (Roche, 11 681 451 001) color substrate.

Isolation of the upstream sequences of the *SsCYC* and Δ*SsCYC* genes

Thermo asymmetry interlaced PCR (TAIL PCR) was applied to isolate the upstream sequence of the *SsCYC* and Δ*SsCYC* genes. TAIL PCR is a fast and efficient method to amplify unknown sequences adjacent to the known sequence. This method uses both gene-specific primers (GS primer), which usually have high melting temperatures, and arbitrary degenerate primers (AD primer), which usually have low melting temperatures, to amplify the sequence from the known end and the unknown end, respectively. Alternative cycles of high and low annealing temperatures and nested program enhance the amplified specificity and suppress unspecific products. The protocol in this study was modified from Liu et al. (1995) and Liu and Whittier (1995). The AD primer sequence was adopted from Singer and Burke (2003) (Supplementary table 1). The

recipe and program for the first thermal cycle of TAIL PCR listed:

Single Reaction for Primary TAIL PCR

Reagents/stock solution	Final conc.	Adding volume
TaKaRa Ex Taq (5 units/μL)	0.5 U	0.1 μL
10× Ex Taq buffer	1× (2.0 mM $MgCl_2$)	2 μL
dNTP mixture (2.5 mM each)	0.2 mM	1.6 μL
4× AD primer stock		5 μL
10 μM GS1 primer stock	0.2 μM	0.4 μL
DNA Template (20 ng/μL)	20-60 ng	1, 2 or3 μL
Water		Add to 20 μL

Thermal Cycle for Primary TAIL PCR

Step	Temperature	Time
1	94℃	3 min
2	94℃	30 s
3	62℃	1 min
4	72℃	2.5 min
5	Go to step2 and repeat 4 times	
6	94℃	30 s
7	25℃ (25% ramping)	3 min
8	72℃ (32% ramping)	2.5 min
9	98℃	10 s
10	68℃	1 min
11	72℃	2.5 min
12	98℃	10 s
13	68℃	1 min
14	72℃	2.5 min
15	98℃	10 s
16	44℃	1 min
17	72℃	2.5 min

18	Go to step 9 and repeat 14 times	
19	72℃	5 min
20	4℃	hold

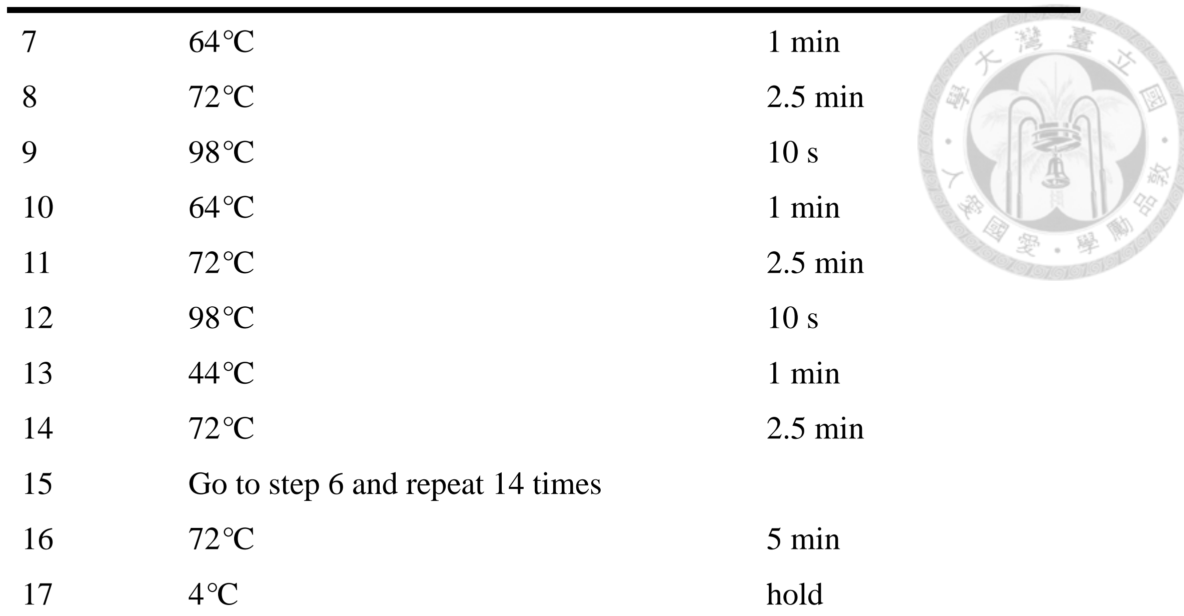
After the primary run, 4 μL of product was run on gel electrophoresis, a multi ladder pattern with partial smear should be seen on the gel. The remaining product was diluted to 1: 50 with sterile distilled water and continued the secondary thermal cycle of TAIL PCR:

Single Reaction for Secondary/Tertiary TAIL-PCR

Reagents/stock solution	Final conc.	Adding volume
TaKaRa Ex Taq (5 units/μL)	0.5 U	0.1 μL
10× Ex Taq buffer	1×	2 μL
dNTP mixture (2.5 mM each)	0.2 mM	1.6 μL
4× AD primer stock		5 μL
10 μM GS2 primer stock	0.2 μM	0.4 μL
1:50 diluted 1st reaction product	1:1000	1 μL
Water		9.9 μL

Thermal Cycle for Secondary TAIL PCR

Step	Temperature	Time
1	94℃	3 min
2	98℃	10 s
3	64℃	1 min
4	72℃	2.5 min
5	Go to step2 and repeat 4 times	
6	98℃	10 s



7	64℃	1 min
8	72℃	2.5 min
9	98℃	10 s
10	64℃	1 min
11	72℃	2.5 min
12	98℃	10 s
13	44℃	1 min
14	72℃	2.5 min
15	Go to step 6 and repeat 14 times	
16	72℃	5 min
17	4℃	hold

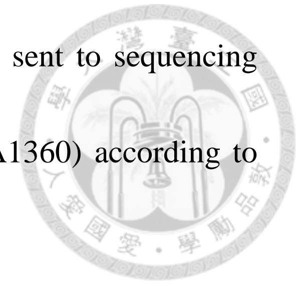
After the secondary run, 4 μL of product was run on gel electrophoresis, the ladders should be less but brighter than the primary run. The remaining product was diluted to 1: 50 with sterile distilled water and continued the secondary thermal cycle of TAIL PCR: (the recipe of the tertiary run was the same as the secondary run)

Thermal Cycle for Tertiary TAIL PCR

Step	Temperature	Time
1	94℃	3 min
2	98℃	10 s
3	44℃	1 min
4	72℃	2.5 min
5	Go to step 2 and repeat 20times	
6	72℃	5 min
7	4℃	hold

After the tertiary run, 4 μL of product was run on gel electrophoresis, there should be

only 1 to few bands remained. The largest band was isolated and sent to sequencing service directly or cloned into pGEM® -T easy vector (Promega, A1360) according to the manufacturer's instructions.

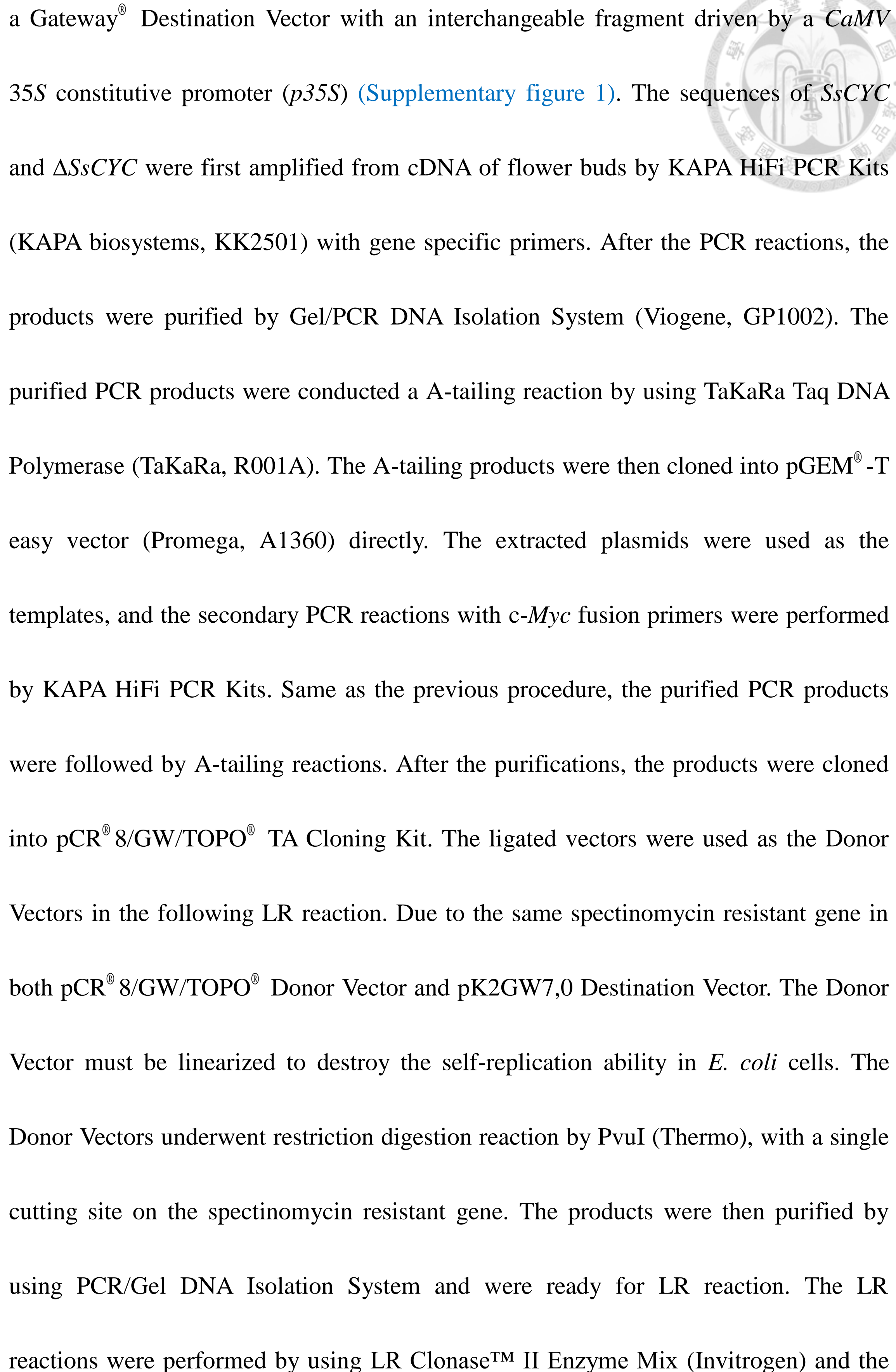


Promoter analysis of the *SsCYC* and Δ*SsCYC* genes

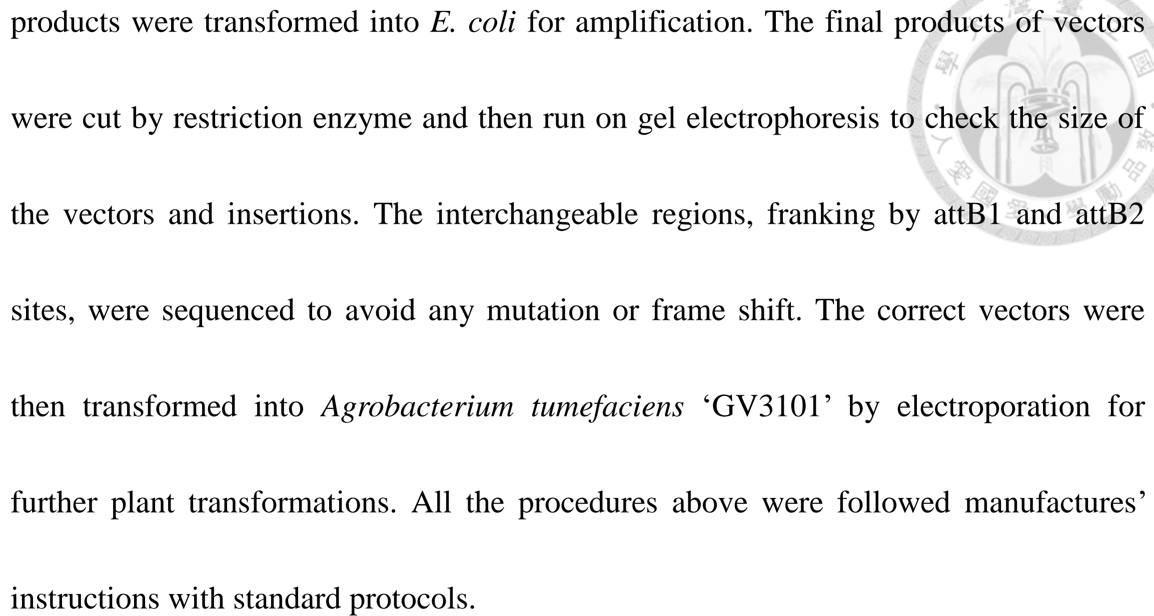
To compare the *cis*-regulatory elements at the upstream sequenc of the *SsCYC* and Δ*SsCYC* genes, the upstream sequences obtained from TAIL PCR were submitted to an online TSSP (Plants Pol II promoter region and start of transcription) tool [using RegSite Plant DB (Softberry Inc., Last Update: May 03, 2014); http://linux1.softberry.com/berry.phtml?topic=tssp&group=programs&subgroup=promoter]. The predicted transcription start site with the highest score was adopted in this study. The putative transcription factor binding sites were retrieved from TSSP result as accession numbers. The list of accession numbers is available at http://www.softberry.com/berry.phtml?topic=regsitelist.

Vector construction for the functional analysis of the *SsCYC* and Δ*SsCYC* genes

All the vector constructions in this study were applied with Gateway® Recombination Cloning Technology. In order to examine the function of SsCYC and ΔSsCYC, three different vectors were constructed: *p35S*::*SsCYC*::c-*Myc*, *p35S*::Δ*SsCYC*::c-*Myc* and *p35S*::c-*Myc*::Δ*SsCYC*. The backbone vectors for all these 3 constructs were pK2GW7,0,



a Gateway® Destination Vector with an interchangeable fragment driven by a *CaMV* 35*S* constitutive promoter (*p35S*) (Supplementary figure 1). The sequences of *SsCYC* and Δ*SsCYC* were first amplified from cDNA of flower buds by KAPA HiFi PCR Kits (KAPA biosystems, KK2501) with gene specific primers. After the PCR reactions, the products were purified by Gel/PCR DNA Isolation System (Viogene, GP1002). The purified PCR products were conducted a A-tailing reaction by using TaKaRa Taq DNA Polymerase (TaKaRa, R001A). The A-tailing products were then cloned into pGEM®-T easy vector (Promega, A1360) directly. The extracted plasmids were used as the templates, and the secondary PCR reactions with c-*Myc* fusion primers were performed by KAPA HiFi PCR Kits. Same as the previous procedure, the purified PCR products were followed by A-tailing reactions. After the purifications, the products were cloned into pCR® 8/GW/TOPO® TA Cloning Kit. The ligated vectors were used as the Donor Vectors in the following LR reaction. Due to the same spectinomycin resistant gene in both pCR® 8/GW/TOPO® Donor Vector and pK2GW7,0 Destination Vector. The Donor Vector must be linearized to destroy the self-replication ability in *E. coli* cells. The Donor Vectors underwent restriction digestion reaction by PvuI (Thermo), with a single cutting site on the spectinomycin resistant gene. The products were then purified by using PCR/Gel DNA Isolation System and were ready for LR reaction. The LR reactions were performed by using LR Clonase™ II Enzyme Mix (Invitrogen) and the



products were transformed into *E. coli* for amplification. The final products of vectors were cut by restriction enzyme and then run on gel electrophoresis to check the size of the vectors and insertions. The interchangeable regions, franking by attB1 and attB2 sites, were sequenced to avoid any mutation or frame shift. The correct vectors were then transformed into *Agrobacterium tumefaciens* ‘GV3101’ by electroporation for further plant transformations. All the procedures above were followed manufactures’ instructions with standard protocols.

Transformation of *Nicotiana benthamiana* leaf disk

To exam whether *SsCYC* and Δ*SsCYC* have the similar function as *Antirrhinum majus AmCYC* and other homologues, functional analysis has been done in *Nicotiana benthamiana*, a Asterids relative to *Sinningia speciosa*. Stable transgenic lines were constructed by *Agrobacterium* mediated transformation. Transformation protocol was modified from Horsch et al. (1985). The protocol can be divided into 3 parts: 1) Preparation of aseptic tobacco seedling, 2) leaf disk infection and 3) selection of transformed shoots.

Preparation of aseptic tobacco seedling

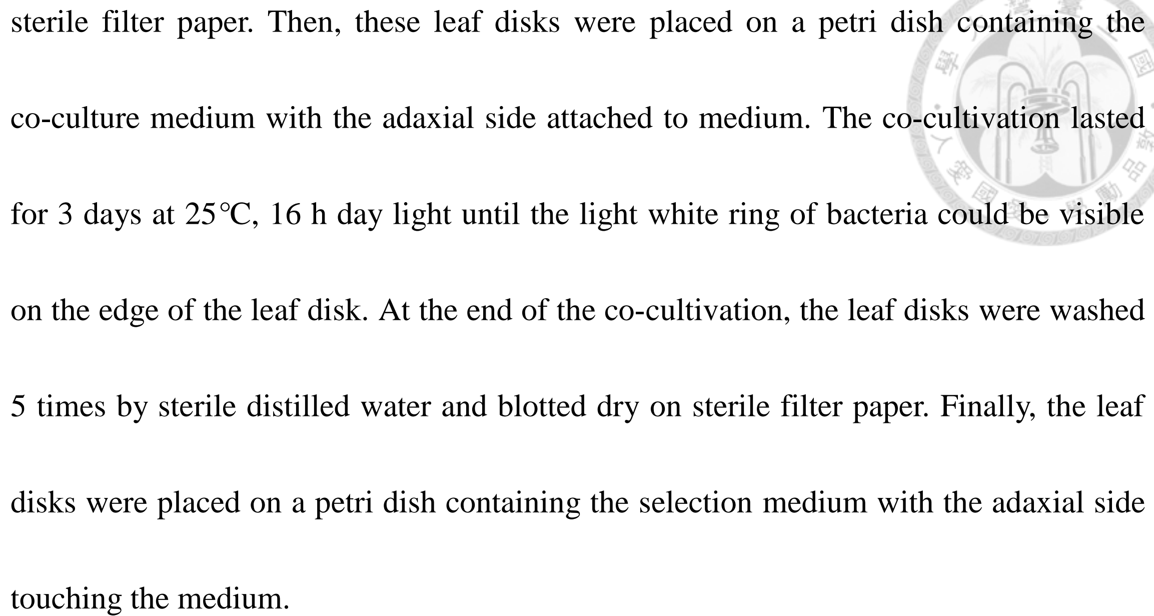
Seeds of *Nicotiana benthamiana* were sterilized by immersion in 70% ethanol for 30 seconds and in 1% sodium hypochlorite for 15 minutes. The seeds were then washed 5 times by sterile distilled water. After sterilization, the seeds were sowed on 1× MS

medium (Murashige and Skoog, 1962), solidified with 3 g/L Phytagel^{TM} (P8169, Sigma) in a 9 cm diameter, 2 cm height petri dish. After 1 week of culture, the young seedlings were transferred to glass cans (9 cm diameter and 11 cm height), which contained the same medium. Through 5-6 weeks of growth, the aseptic seedlings became ready for leaf disk infection.

Leaf disk infection

Two days before the infection, the *Agrobacterium* harboring the target vector was precultured in 3 mL of YEP medium (1 L, 10 g yeast extract, 10 g Bactopeptone. 5 g NaCl) with 50 ppm spectinomycin and 10 ppm gentamycin. The preculture was incubated overnight at 28°C in an orbital shaker set at 200 rpm. On the next day, 1 mL of the preculture was inoculated a 50 mL *Agrobacterium* culture with the same medium and antibiotics. After the OD_{600} reached 0.8-1.0, the culture was ready for use. The culture was transferred to a 50 mL centrifuge tube and was centrifuged at 6000× g and 4℃ for 8 minutes. After the centrifuge, the supernatant was discarded and equal volume of YEP medium without antibiotics was added to resuspend the *Agrobacterium*.

The young leaves of tobacco seedlings were cut into 5 mm × 5 mm disks with 100 explants per construct. For inoculation, all the disks were transferred into the *Agrobacterium* culture for about 1 minute with shaking. After the inoculation, the leaf disks were recovered with forceps and excess liquid was eliminated by blotting on



sterile filter paper. Then, these leaf disks were placed on a petri dish containing the co-culture medium with the adaxial side attached to medium. The co-cultivation lasted for 3 days at 25℃, 16 h day light until the light white ring of bacteria could be visible on the edge of the leaf disk. At the end of the co-cultivation, the leaf disks were washed 5 times by sterile distilled water and blotted dry on sterile filter paper. Finally, the leaf disks were placed on a petri dish containing the selection medium with the adaxial side touching the medium.

Selection of transformed shoots

After one week of selection, the leaf disks were sub-cultured on the fresh selection medium to maintain the selection pressure and effect of antibiotics. Through 2-3 weeks of selection, the putative transformed shoots appeared at the edge of the leaf disks. After the shoots had reached 0.5 cm in length, these were cut at the base carefully and transfer to a 9 cm diameter, 11 cm height glass cans, which contained root induction medium. Due to the presence of antibiotics, only the successfully transformed shoots could root while others became bleached and showed abnormal growth. After about 3-4 weeks of culture, the rooted shoots were ready to transfer into pots for further growth. For acclimation, pots were covered with plastic bags to reduce the evaporation for 1 week. After that, the plastic bags were removed.

Each transgenic plant was assigned a unique ID, for example: A1-1a. The first

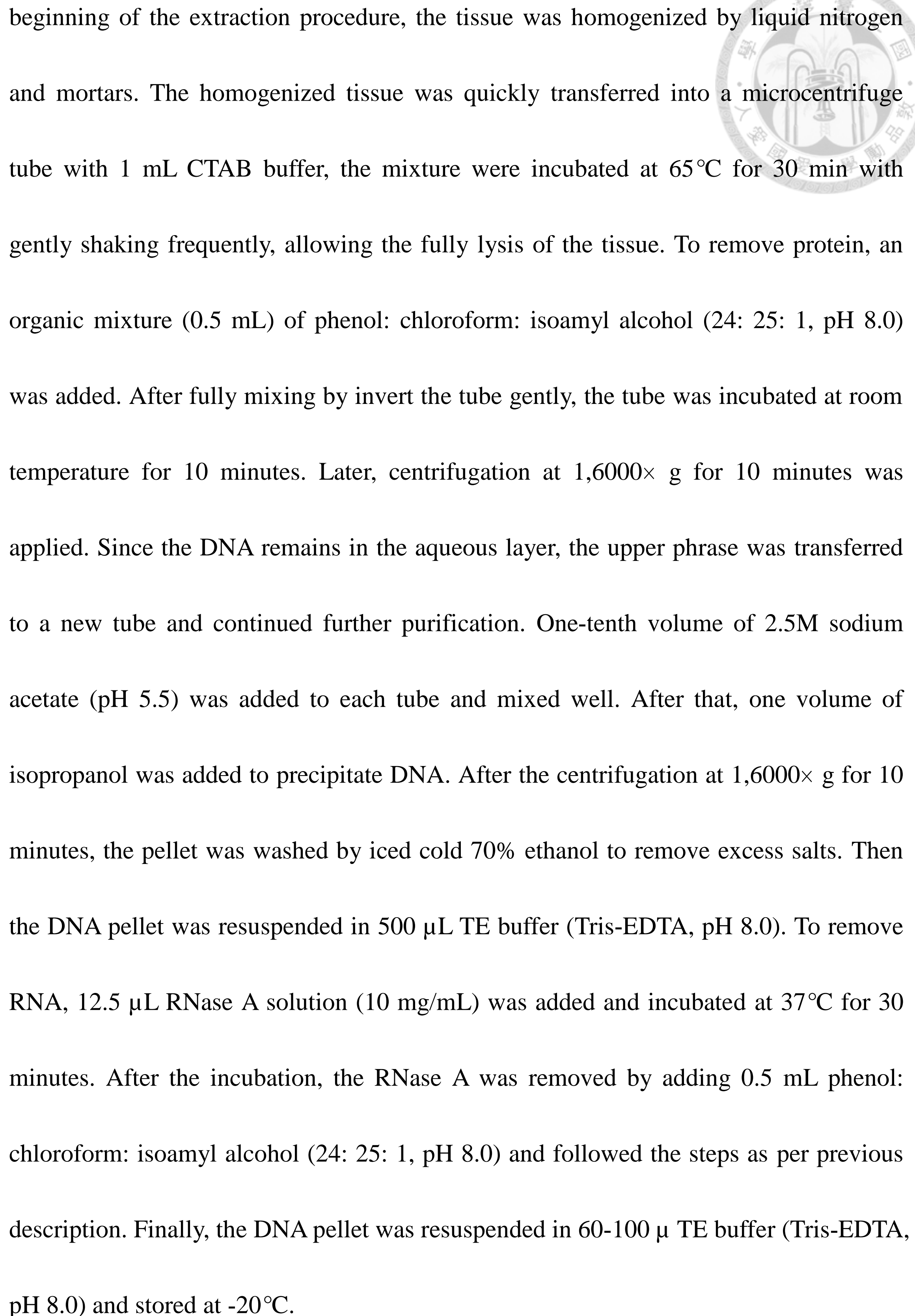
uppercase letter presents the type of vectors being transformed. With 'A' stands for p35S::Δ*SsCYC*::c-*Myc*; 'B' stands for p35S::c-*Myc*::Δ*SsCYC*; 'C' stands for p35S::*SsCYC*::c-*Myc*. The number next to the letter shows the ID of petri dish during the selection. The number after the '-' indicates the ID inside a petri dish. In some cases, the explant was cut into 2 or more pieces after the inoculation, the pieces from the same explant were donate as 'a', 'b', 'c' … at the end. Only one regenerative shoot was taken from the explant with a unique ID to assure the independent transformation event.

Clearing method for measuring the cell size

To measure the floral tube cell size of the transgenic plants, clearing method was applied in this study. The dissected floral tubes were put into a 2.0 mL microcentrifuge tubes with 1 M NaOH and incubated overnight at 30℃. After washing 5 times with distilled water to remove the NaOH, the samples were stained with 0.5% safranin-O for 1 hour. Then, de-staining was followed by washing with the distilled water again. The probable stained samples were mounted with 100% glycerol directly and pictured by the light microscope (BX51, Olympus, Japan). The pictures were later analyzed by image J software to calculate the cell sizes.

Genomic DNA extraction

In order to isolate plant DNA, CTAB (Hexadecyl trimethyl-ammonium bromide) method described by Rogers and Bendich (1985) was applied with modification. At the



beginning of the extraction procedure, the tissue was homogenized by liquid nitrogen and mortars. The homogenized tissue was quickly transferred into a microcentrifuge tube with 1 mL CTAB buffer, the mixture were incubated at 65℃ for 30 min with gently shaking frequently, allowing the fully lysis of the tissue. To remove protein, an organic mixture (0.5 mL) of phenol: chloroform: isoamyl alcohol (24: 25: 1, pH 8.0) was added. After fully mixing by invert the tube gently, the tube was incubated at room temperature for 10 minutes. Later, centrifugation at 1,6000× g for 10 minutes was applied. Since the DNA remains in the aqueous layer, the upper phrase was transferred to a new tube and continued further purification. One-tenth volume of 2.5M sodium acetate (pH 5.5) was added to each tube and mixed well. After that, one volume of isopropanol was added to precipitate DNA. After the centrifugation at 1,6000× g for 10 minutes, the pellet was washed by iced cold 70% ethanol to remove excess salts. Then the DNA pellet was resuspended in 500 μL TE buffer (Tris-EDTA, pH 8.0). To remove RNA, 12.5 μL RNase A solution (10 mg/mL) was added and incubated at 37℃ for 30 minutes. After the incubation, the RNase A was removed by adding 0.5 mL phenol: chloroform: isoamyl alcohol (24: 25: 1, pH 8.0) and followed the steps as per previous description. Finally, the DNA pellet was resuspended in 60-100 μ TE buffer (Tris-EDTA, pH 8.0) and stored at -20℃.

CTAB buffer (100 mL)

Regent/Stock solution	Adding weight/volume
CTAB (Hexadecyl trimethyl-ammonium bromide)	2.0 g
1 M Tris, pH 8.0	10.0 mL
0.5 M EDTA, pH 8.0 (Ethylenediaminetetraacetic acid)	4.0 mL
5 M NaCl	28.0 mL
Water	56.0 mL

Adjust to pH8.0 with NaOH and store at room temperature

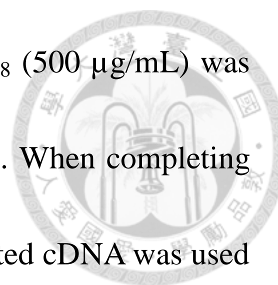
Add the following items (per 1 mL CTAB buffer) and incubate on 65℃ dry bath for 30 minutes before use. In this step, PVPP might saturate with water and be ready to use.

Regent/Stock solution	Adding weight/volume
β-Mercaptoethanol	5 μl
Polyvinylpolypyrrolidone (PVPP)	20 mg

Total RNA extraction and reverse transcription

In order to detect the expression levels of transgenes in the transgenic plant, the total RNA of young leaves were isolated by TRIzol® Reagent. The protocol of RNA extraction was adapted from the manufacture's instructions with a modification. Additional phrase separation by phenol: chloroform (5:1, pH 4.5) was applied before the precipitation step to further eliminate DNA and protein contaminations.

Total RNA of the transgenic plants was reverse transcribed into complementary DNA (cDNA) by using M-MLV (Moloney Murine Leukemia Virus) Reverse Transcriptase (RT) (Invitrogen™, 28025-013) and oligo $(dT)_{18}$ primer. In a 20 μL reaction, 2 μg of



total RNA was added at the beginning as the template. Oligo $(dT)_{18}$ (500 μg/mL) was used as the primer to reverse transcribe the mRNA from the 3' end. When completing the reaction, 80 μL water was added to dilute the cDNA and the diluted cDNA was used directly as a template for amplification in PCR.

Polymerase chain reaction

The general PCR and RT PCR carried out in this were performed by Applied Biosystem 2720 and 2700 thermo cycle. For RT PCR, the *Actin* gene, a housekeeping gene, was chosen as the positive control. The primers were designed based on the sequence downloaded from NCBI (AY594294). The general recipe and program are listed as below:

Single Reaction for general PCR or RT PCR

Regent/Stock solution	Adding weight/volume
Taq DNA Polymerase 2× Master Mix (Ampliqon)	10 μL
Forward primer (10 μM)	1 μL
Reverse primer (10 μM)	1 μL
Sterile distilled water	As required
Template (DNA or cDNA)	As required
Total volume	20 μL

Thermal Cycle for general PCR or RT PCR

Step	Temperature	Time
1	94°C	5 min
2	94°C	30 s

3	Annealing temperature	30 s
4	72℃	1 min/kb
5	Go to step 2 and repeat multiple times depending on the descriptions	
6	72℃	10 min
7	4℃	hold

Preparation of aseptic seedlings for tissue culture

To prepare aseptic plant source for tissue culture, *Sinningia speciosa* 'Avanti' seeds were surface sterilized by 70% ethanol for 30 s then by 1% sodium hypochlorite solution for 10 min. After sterilization, seeds were rinsed with sterilized water for at least 5 times thoroughly to remove sodium hypochlorite solution. The seeds were cultured on 1/2 MS medium (Murashige and Skoog, 1962), solidified with 3 g/L Phytagel™ (P8169, Sigma) in a 9 cm diameter, 2 cm height petri dish. After 7 days, the seeds began to germinate. The small seedlings were left in the petri dish for about a month until these grew big enough to be transplanted. Then the seedlings were transplanted to 9 cm diameter, 11 cm height glass cans, which contained the same medium and were sealed with surgery tape. After another 1 to 2 months, the aseptic seedlings of *Sinningia speciosa* became ready for use as the source of the explants. All the aseptic plant materials were kept in an incubator at 27℃ and 50% relative humidity and placed under cool white fluorescent light at 40-60 μmol m^{-2} s^{-1} with a photoperiod of 16 h light and 8 h dark.

Tissue culture of *Sinningia speciosa*

To optimize the best conditions for shoot regeneration of *Sinningia speciosa*, different combinations of plant hormones were examined. 1× MS medium supplemented with naphthalene acetic acid (NAA; 0.1, 0.2 ppm) and 6-benzyladenine purine (BA; 0, 1, 2, 3 ppm) were autoclaved (121℃, 20 min) prior pouring into to the petri dish (50 mL medium in a 9 cm diameter, 2 cm height plate). Adventitious shoot formation was induced in leaves and petioles of 1-2 month-old aseptically grown seedlings. To eliminate the size effect of the explants which may have an effect on the response, uniformed leaf explants were obtained by using a hole puncher (6 mm diameter). Petioles were cut into 5 mm or 2.5 mm pieces. To investigate the effect of orientation of explants on to the medium, different orientations were designed for petiole explants. As shown in Fig. 18, there were 3 orientations: proximal side vertically attached to the medium, distal side vertically attaching to the medium and horizontally attached to the medium. There were 12 leaf explants or 9 petiole explants in each petri dish with 4 replicates in each treatment. To reduce the effect of incubation period, the observations were recorded from 30 days to 65 days after culture initiation. In each record, 2 data were carefully documented: regeneration rate (number of responsive explants/number of total explants) and shoots per responsive explant (total number of shoots/number of responsive explants). All cultures were incubated under the same physical conditions as

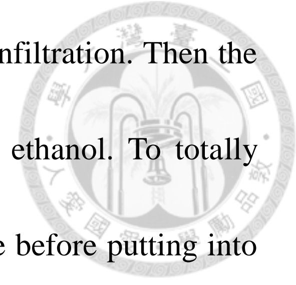
described above.

Histological analysis by paraffin sections

To investigate the early stages of shoot regeneration, histological observation was done by paraffin section. Leaf explants cultured on medium containing 2 ppm BA and 0.2 ppm NAA were sampled every 7 days from 0 day to 42 days after culture. They were fixed for 6 h in FAA solution (50% ethanol: formalin: acetic acid = 90:5:5; v/v/v) or 0.1 M phosphate buffer (pH 7.0) with 4% paraformaldehyde and 2.5% glutaldehyde. During the fixation, the samples were kept in a vacuum oven (about 40 cmHg) to help infiltration. Then the samples were dehydrated via t-butanol alcohol/ethanol series and finally into pure t-butanol alcohol and were embedded into paraffin block. 10 μm thick continuous sections were cut by a rotary microtome. The sections were stained with safranin-O overnight and then counter-stained with fast green. After mounting with DPX, sections were observed under a light microscope (BX51, Olympus, Japan).

Scanning electron microscope (SEM)

In order to clarify the early growth of newly regenerative shoots, scanning electron microscopy (SEM) analysis was done with a FEI Inspect S50 microscope. Leaf explants cultured on the medium, same as for histological observation, were sampled every 7 days from 0 day to 42 days after culture. They were fixed for 6 h in 0.1 M phosphate buffer (pH 7.0) with 4% paraformaldehyde and 2.5% glutaldehyde. During the fixation,



the samples were kept in a vacuum oven (about 40 cmHg) to help infiltration. Then the samples were dehydrated via ethanol series and finally into pure ethanol. To totally remove the liquid inside the tissue, ethanol was replaced by acetone before putting into a critical point dryer. Inside the dryer, the acetone was substituted with CO_2, and then evaporated. After coating with gold particles to strength the tissue, the samples were ready for SEM observation.

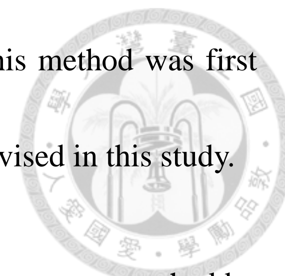
Transient transformation using GUS reporter gene

In order to find the tissue which was most susceptible to *Agrobacterium* infection, the transformation efficiency was tested by GUS transient transformation in various tissues of *Sinningia speciosa*. Considering the low transformation competence of *Sinningia speciosa*, a sophisticated *Agrobacterium* culture and a pre-induction procedure were applied here.

The *Agrobacterium* stain 'GV3101' and 'EHA105' harboring pCambia2301 vectors were cultured in 3 mL YEP medium with 50 ppm Kanamycin (for 'GV3101', additional 10 ppm Gentamycin was added) overnight at 28℃ and shaking at 200 rpm. Next day, 100 μL of culture was added into 50 mL AB medium with 20 μM acetosyringone and 0.5% glucose and 50 ppm Kanamycin (for 'GV3101', additional 10 ppm Gentamycin was added). The culture was incubated for 12-24 h at 28℃ with shaking at 200 rpm until the OD_{600} reached 0.8-1.0. The cultures were then transferred into 50 mL

centrifuge tube and centrifuged at 6000× g for 8 minutes at 4℃. After discard the supernatant, the *Agrobacterium* cell was resuspended in 50 mL sterile 1×MS medium with 200 μM acetosyringone and 0.5% glucose. At this stage, the *Agrobacterium* became ready for co-cultivation.

The leaf explants (5 mm^2), petiole explants (5 mm long segment) and 14-day-old seedlings *in toto* were prepared from aseptic seedlings of *Sinningia speciosa* 'Avanti'. For inoculation of leaf explants and petiole explants, the explants were soaked in the *Agrobacterium* medium for 15-20 minutes. After the inoculation, the explants were blotted dry on the sterile filter paper and transferred to co-culture medium. For leaf explants, the adaxial sides were touching the medium; for petiole explants, the explants were placed horizontally on the medium. On the other hand, for the 14-day-old seedlings *in toto*, in order to infiltrate through the epidermis layer, an external pressure was applied here to facilitate the infiltration. At the beginning, about 10 mL of *Agrobacterium* culture was poured into a 20 mL sampling bottle. Then, about 20 seedlings were carefully transferred into the bottle by forceps. After the transfer, the bottle containing the seedlings was put into a 50 mL syringe tube. Holding the tube firmly by one hand on a rubber stopper, and pushing the pestle forward by the other hand till the pestle contacting the upper rim of the sampling bottle. This procedure was repeated at least 5 times for full infiltration. The infiltrated seedlings were blotted dry



on the sterile filter paper and transferred to co-culture medium. This method was first discovered by Wu, Ho-Wei (personal communication) and further revised in this study.

After 3-4 days of co-culture in dark at 25℃, the explants and seedlings were washed by 0.1 M phosphate buffer (pH 7.0) and stained with GUS staining solution overnight in dark at 37℃. The explants and seedlings with positive signals were fixed and de-coloration in FAA (50% ethanol: formalin: acetic acid = 18: 1: 1) solution for 6 hours. After the fixation, the samples were photographed directly under the light microscope or continued histological analysis by paraffin sections.

AB medium 1 L

Regent/Stock solution	Adding weight/volume
20× Buffer A	50 mL
20× Buffer B	50 mL
Glucose	5 g

20× Buffer A 1 L

Regent/Stock solution	Adding weight/volume
K_2HPO_4	60 g
NaH_2PO_4	20 g

20× Buffer B 1 L

Regent/Stock solution	Adding weight/volume
NH_4Cl	20 g

$MgSO_4 \cdot 7H_2O$	6 g
KCl	3 g
$CaCl_2$	0.2 g
$FeSO_4 \cdot 7H_2O$	0.05 g

Statistical analysis

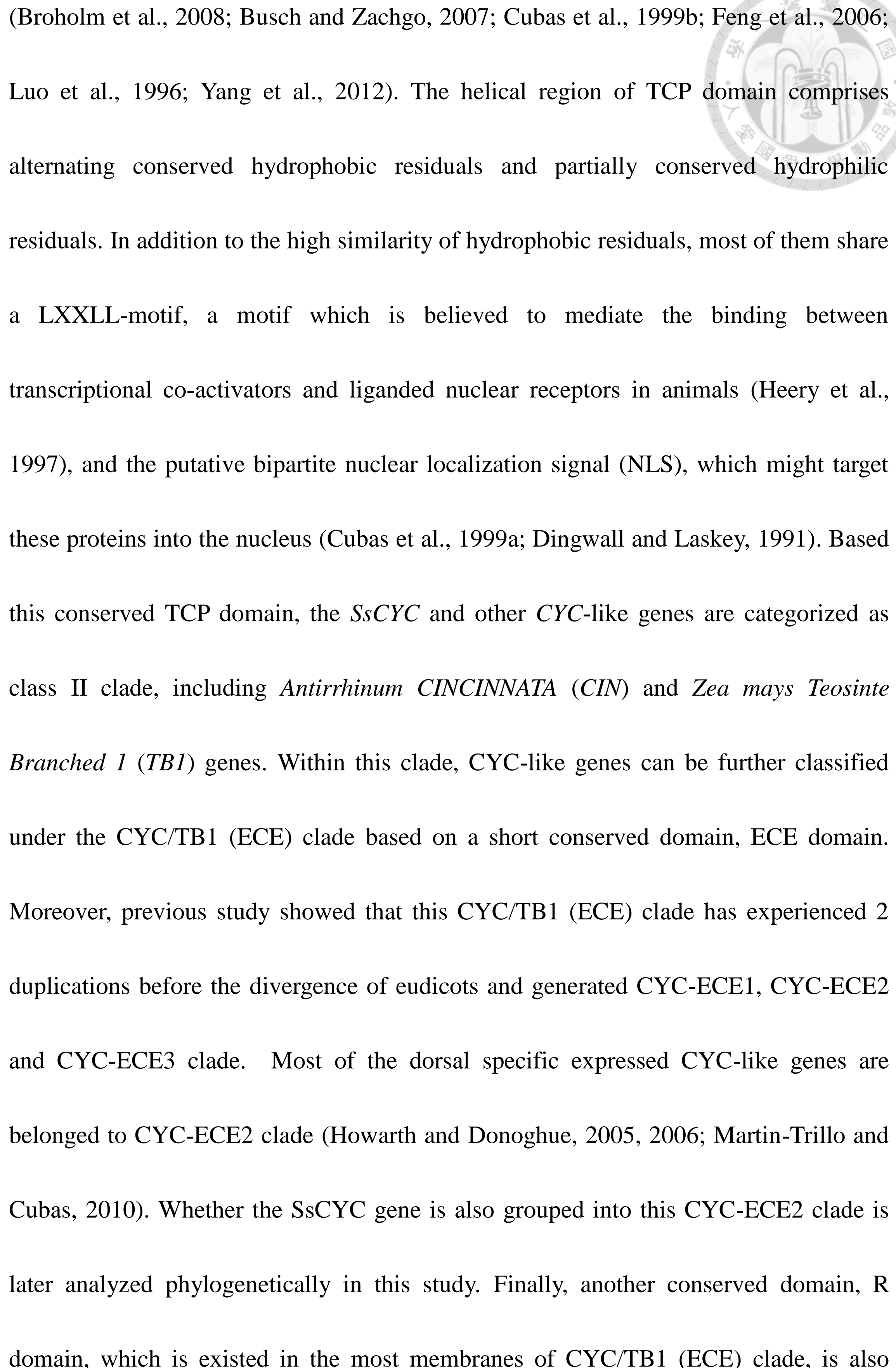
All the statistical analysis in this study was performed with R software (R Core Team, 2013). Kruskal-Wallis Rank Sum Test was used to analysis the effect of different transgenic treatment or different mRNA expression level on flower morphology (total length, tube length and lobe length), since the data did not meet the assumptions of analysis of variance (ANOVA) (Supplementary figure 6 & 7). If found significance, multiple comparison test after Kruskal-Wallis (Giraudoux, 2013) was performed to test the differences between each group. Analysis of variance was used to analysis the effect of different transgenic treatment on the cell size of floral tube. If found significance, the data were grouped by Turkey multiple comparison.

Results

Conserved domains of the *SsCYC* gene

The *CYCLOIDEA* homologous gene in *Sinningia speciosa* (*SsCYC*) was first isolated by Citerne et al. (2000) with conserved-designed primer and low stringent PCR condition. Since *CYCLOIDEA* (*CYC*) gene and its homologues in various lineages have been shown playing a key role in floral symmetry development, the *SsCYC* gene is believed to be the human selection target for floral symmetry reversion in horticultural cultivars. In order to obtain the full coding sequence of *SsCYC* gene, the TAIL PCR and 3'-RACE techniques were applied to extend the 5' and 3' unknown regions (Hsu, Hao-Chun, unpublished). In this study, the PCR primers were designed based on this known sequence and the full coding sequence of *SsCYC* gene (wild type, zygomorphic) and Δ*SsCYC* (horticultural, actinomorphic) were successfully isolated.

Proteins sharing conserved domains are usually in the same gene family and have the similar functions. Just like other CYC-like genes, the *SsCYC* encoded protein contains three conserved domains: TCP domain (Fig. 2A), ECE domain (Fig. 2B) and R domain (Fig. 2C). The TCP domain is predicted to form a non-canonical basic-Helix-Loop-Helix, which can bind to target DNA sequence as a transcription factor. The TCP domain shows high similarity between SsCYC and the other CYC homologues which have been functionally studied (red arrow head in Fig. 2A)



(Broholm et al., 2008; Busch and Zachgo, 2007; Cubas et al., 1999b; Feng et al., 2006; Luo et al., 1996; Yang et al., 2012). The helical region of TCP domain comprises alternating conserved hydrophobic residuals and partially conserved hydrophilic residuals. In addition to the high similarity of hydrophobic residuals, most of them share a LXXLL-motif, a motif which is believed to mediate the binding between transcriptional co-activators and liganded nuclear receptors in animals (Heery et al., 1997), and the putative bipartite nuclear localization signal (NLS), which might target these proteins into the nucleus (Cubas et al., 1999a; Dingwall and Laskey, 1991). Based this conserved TCP domain, the *SsCYC* and other *CYC*-like genes are categorized as class II clade, including *Antirrhinum CINCINNATA* (*CIN*) and *Zea mays Teosinte Branched 1* (*TB1*) genes. Within this clade, CYC-like genes can be further classified under the CYC/TB1 (ECE) clade based on a short conserved domain, ECE domain. Moreover, previous study showed that this CYC/TB1 (ECE) clade has experienced 2 duplications before the divergence of eudicots and generated CYC-ECE1, CYC-ECE2 and CYC-ECE3 clade. Most of the dorsal specific expressed CYC-like genes are belonged to CYC-ECE2 clade (Howarth and Donoghue, 2005, 2006; Martin-Trillo and Cubas, 2010). Whether the SsCYC gene is also grouped into this CYC-ECE2 clade is later analyzed phylogenetically in this study. Finally, another conserved domain, R domain, which is existed in the most membranes of CYC/TB1 (ECE) clade, is also

found within SsCYC protein. The R domain is abundant of polar residuals (arginine, lysine and glutamic acid) and is predicted to form a hydrophilic alpha-helix, which may mediate protein-protein interactions (Lupas et al., 1991).

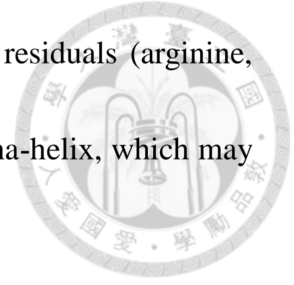


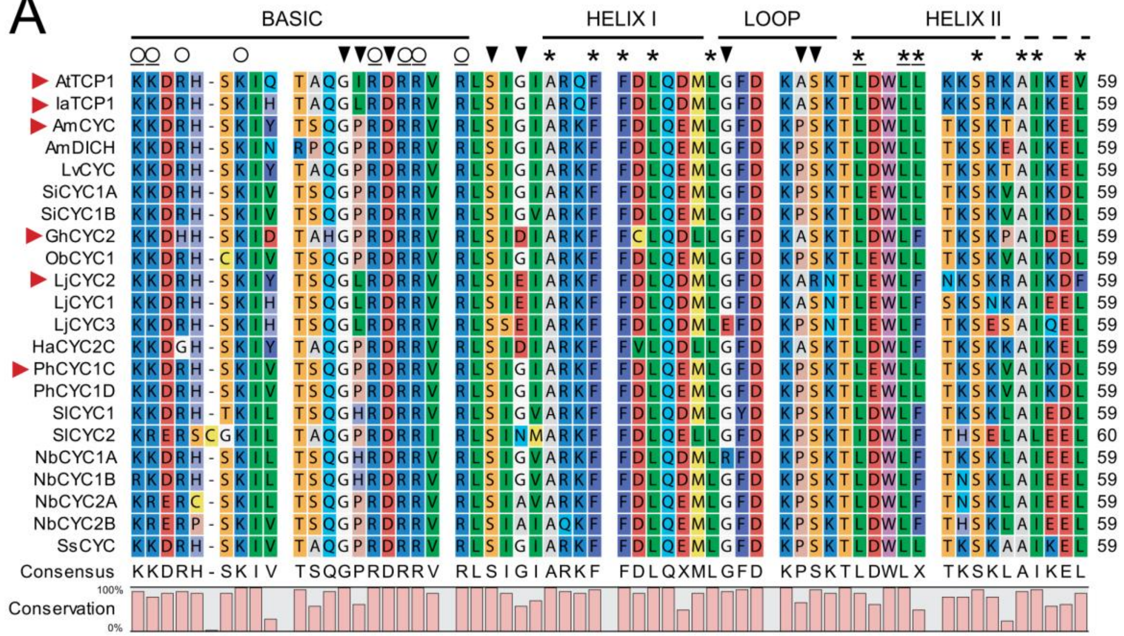
Table 1

Genes used in construction of phylogenetic tree (Figure 2)

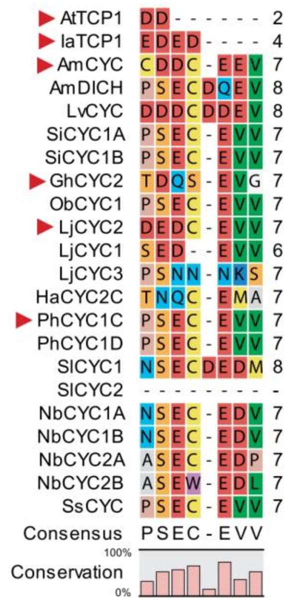
Gene	Accession number(s)	Description	Reference
AmCYC	Y16313	*Antirrhinum majus* mRNA for cycloidea protein	Luo et al. (1996)
AmDICH	AF199465	*Antirrhinum majus* flower asymmetry protein DICHOTOMA mRNA, complete cds	(Carpenter et al., 1999)
AtTCP1	NM_001160982	*Arabidopsis thaliana* transcription factor *TCP1* mRNA, complete cds	(Theologis et al., 2000)
AtTCP12	NM_105554	*Arabidopsis thaliana* transcription factor *TCP12* mRNA, complete cds	(Theologis et al., 2000)
AtTCP18	NM_001125184	*Arabidopsis thaliana* transcription factor *TCP18* mRNA, complete cds	(Salanoubat et al., 2000)
GhCYC2	EU429303	*Gerbera* hybrid cultivar *CYCLOIDEA*-like 2 (*CYC2*) mRNA, complete cds	(Broholm et al., 2008)
HaCYC2c	EU088370	*Helianthus annuus* cycloidea-like 2c protein gene, complete cds	(Chapman et al., 2008)
IaTCP1	EU145779	*Iberis amara TCP1* (*TCP1*) mRNA, complete cds	(Busch and Zachgo, 2007)
LjCYC1	DQ202475	*Lotus corniculatus* var. *japonicus* CYC1 mRNA, complete cds	(Feng et al., 2006)
LjCYC2	DQ202476	*Lotus corniculatus* var. *japonicus CYC2* mRNA, complete cds	(Feng et al., 2006)
LjCYC3	DQ202477	*Lotus corniculatus* var. *japonicus CYC3* mRNA, complete cds	(Feng et al., 2006)
LvCYC	AF161252	*Linaria vulgaris* cycloidea-like protein gene, complete cds	(Cubas et al., 1999b)
NbCYC1B	Niben.v0.4.2.Scf8735 complement (24209..25249)	*Nicotiana benthamiana* putative *CYC* homologue, blast from Sol database	This study
NbCYC1A	Niben.v0.4.2.Scf54682 (1728..2711)	*Nicotiana benthamiana* putative *CYC* homologue, blast from Sol database	This study
NbCYC2A[¶]	Niben.v0.3.Scf25092789 (6721..8064)	*Nicotiana benthamiana* putative *CYC* homologue, blast from Sol database	This study
NbCYC2B[¶]	Niben.v0.3.Scf25291400 (7204..8540)	*Nicotiana benthamiana* putative *CYC* homologue, blast from Sol database	This study
ObCYC1	FJ710517	*Oreocharis benthamii* transcription factor *CYC1* (*CYC1*) mRNA, complete cds	(Song et al., 2009)

PhCYC1C	JX020500	*Primulina heterotricha* TCP transcription factor *CYC1C* (*CYC1C*) gene, complete cds	(Liu et al., 2014a)
PhCYC1D	JX020501	*Primulina heterotricha* TCP transcription factor *CYC1D* (*CYC1D*) gene, complete cds	(Liu et al., 2014a)
SiCYC1A	EF127811	*Saintpaulia velutina* CYCLOIDEA protein (*CYC1A*) gene, CYC1A-zygomorphic WT allele, partial cds	Wang, 2007 (unpublished)
SiCYC1B	EF127812	*Saintpaulia velutina* CYCLOIDEA protein (*CYC1B*) gene, CYC1B-zygomorphic WT allele, partial cds	Wang, 2007 (unpublished)
SlCYC1	HM921068	*Solanum lycopersicum CYCLOIDEA1* gene, partial sequence	(Martin-Trillo et al., 2011)
SlCYC2	HM921069	*Solanum lycopersicum CYCLOIDEA2* gene, partial sequence	(Martin-Trillo et al., 2011)
SsCYC	This study	*Sinningia speciosa CYCLOIDEA* mRNA, complete cds	This study

¶ Putative intron was removed by ORF prediction in CLC software



B



C

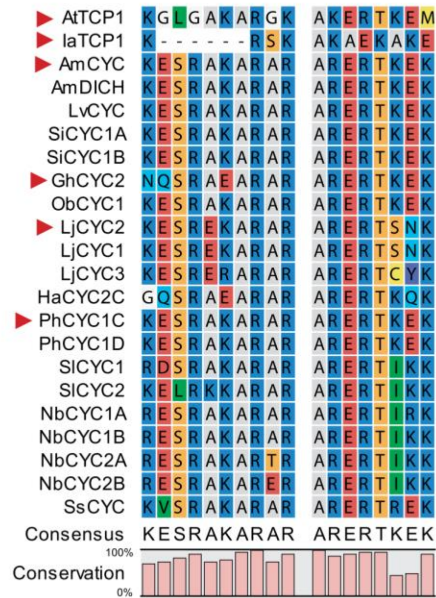
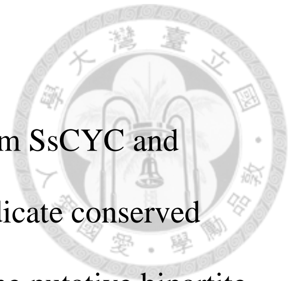


Figure 2 Conserved domains and phylogeny of the *SsCYC* gene



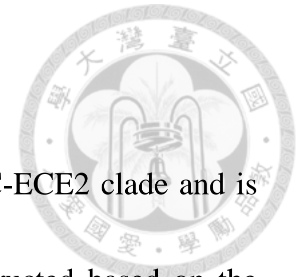
A, TCP domain. Alignment of the predicted amino acid sequence from SsCYC and other members of the CYC/TB1 ECE clade. Circles along the top indicate conserved basic residues; underlined circles indicate residuals forming part of the putative bipartite NLS; asterisks indicate conserved hydrophobic residues in the helices; underline asterisks indicated the LXXLL motif; black arrowheads point to residues (glycine or proline) that disrupt α-helix formation. The variable predicted length of helix II is indicated by a dash line.

B, ECE domain. Containing a glutamic acid-cysteine-glutamic acid motif.

C, R domain. Alignment of the predicted amino acid sequence from SsCYC and other members of the CYC/TB1 subfamily. Note that SsCYC has very similar sequence to AmCYC and other CYC homologues which are reported to having relation with floral symmetry development.

Red arrow head, the genes which have been functional studied.

Reconstruction the phylogenetic tree of *SsCYC* gene



In order to know whether the *SsCYC* gene is belonged to the CYC-ECE2 clade and is close to other CYC-like genes, the phylogenetic tree was reconstructed based on the alignment of the full coding nucleotides (Supplementary Fig 1). The tree included the well characterized dorsal specific ECE-CYC2 homologues and *Arabidopsis TCP12* (ECE-CYC3), *TCP18* (ECE-CYC1) as the out group (Fig. 3 A). Within the ECE-CYC2 lineage, the *SsCYC* gene (highlight in yellow background in Fig 3 A) formed a monophyletic clade (branch support value BS=100, BP=100) with the homologues of the other Gesneriaceae species. Then, this monophyletic clade is the sister group to the clade of *AmCYC*, *AmDICH* and *LvCYC* (Plantagainaceae) and subtended by other *CYC* homologues in Solanaceae, including those of *Solanum lycopersicum* and *Nicotiana benthamiana*. After combining the sequences of Asteraceae, the CYC homologues form a large monophyletic clade, corresponding to Asterids (branch support value BS=62, BP=100). This large Asterids group is then subtended by Brassicaceae (Rosids) and Leguminosae (Rosids). All the results indicate that the *SsCYC* gene is a member of floral dorsally expressed ECE-CYC2 clade and is likely to participate in zygomorphy development.

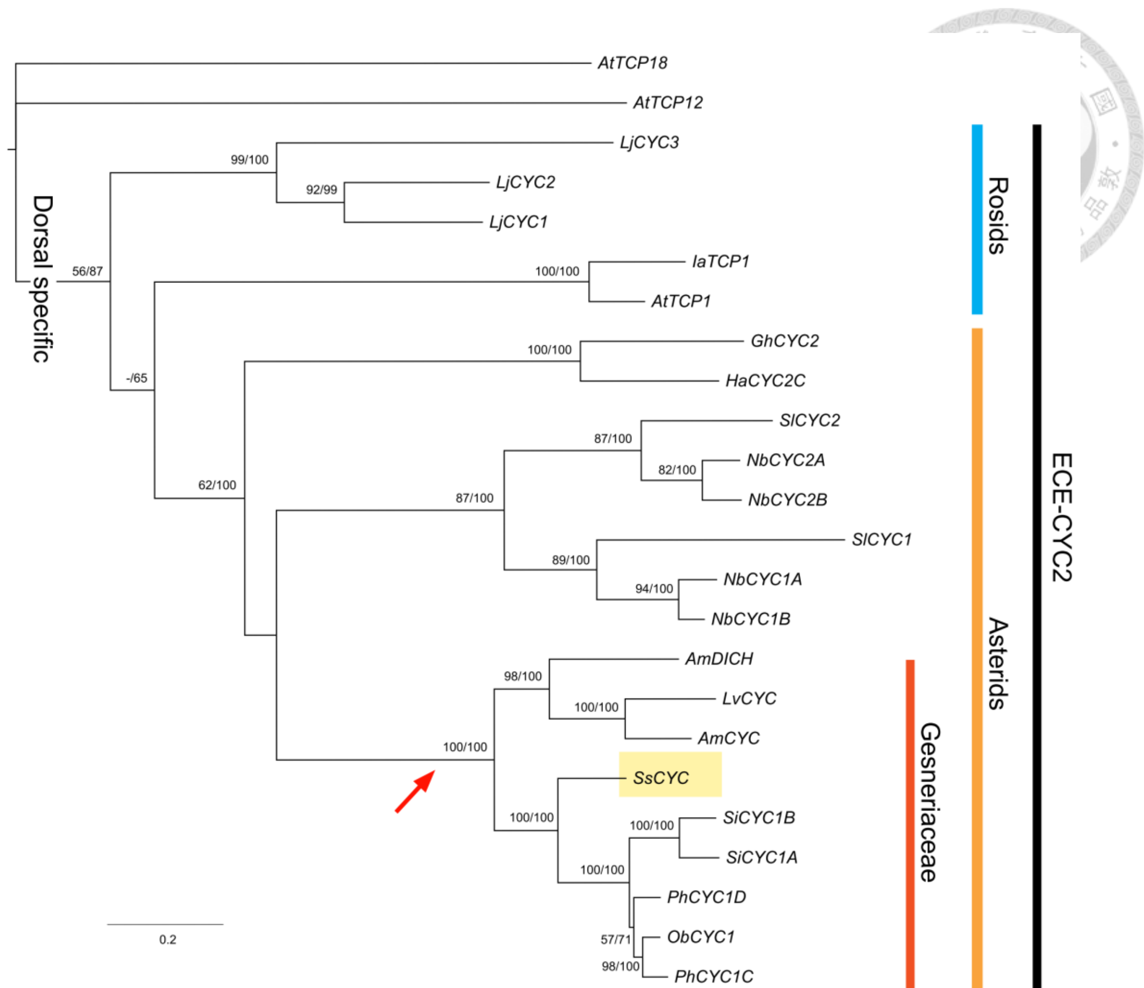
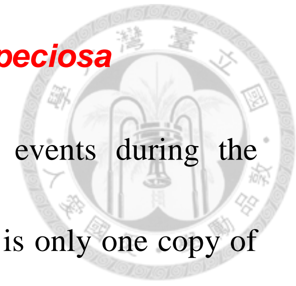


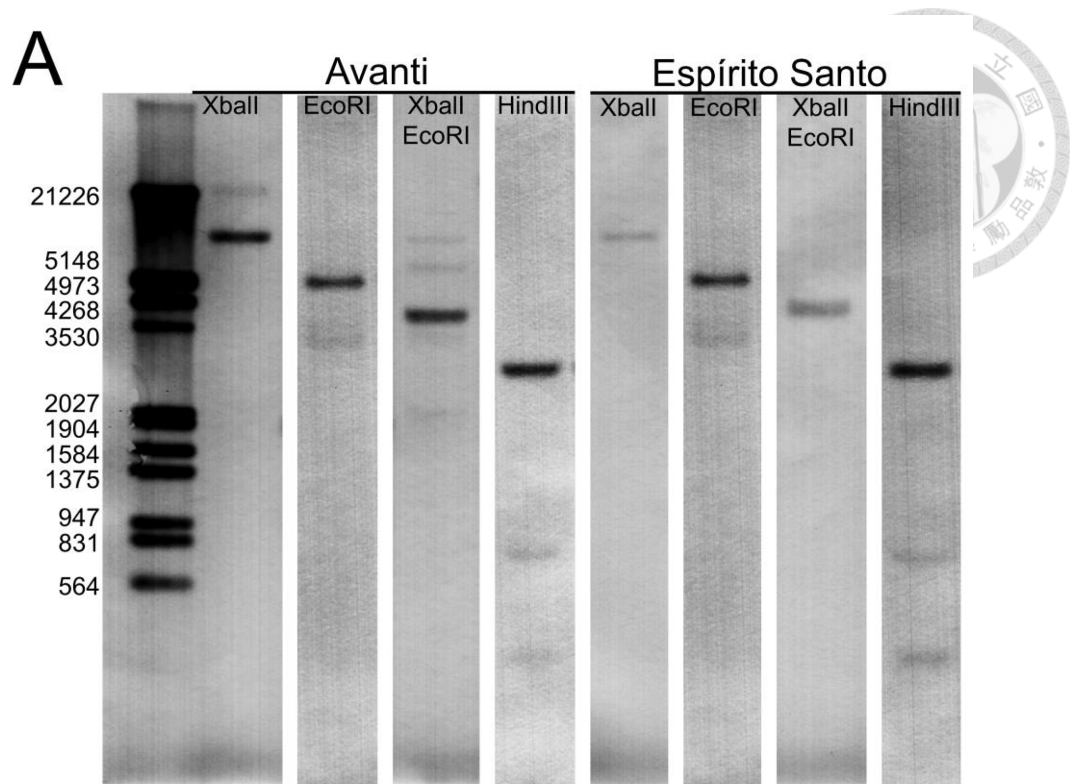
Figure 3 Phylogeny of the *SsCYC* gene

Maximum Likelihood tree and Bayesian tree showing the relation between *SsCYC* and other dorsal specific ECE-CYC2 homologues (Preston and Hileman, 2009). The percentage of bootstrap samples in which particular clades are monophyletic is indicated for those clades with support of 50% or more. (Bootstrap value of Maximum Likelihood method/ Bootstrap value of Bayesian inference method). Red arrow, the Gesneriaceae monophyletic clade.



Confirming the single copy of *SsCYC* gene in *Sinningia speciosa*

CYC-like genes have experienced several duplication and lost events during the angiosperms evolution. Although previous study showed that there is only one copy of CYC-like gene in the new world Gesneriaceae lineage, including *Sinningia speciosa*, there might be a misleading due to the inefficient PCR procedures (Citerne et al., 2000). To avoid gene redundancy effects, confirming the copy number of *SsCYC* gene by Southern blotting is necessary before the functional analysis. In this study, a 664 bp wild type specific probe was used to hybridize both wild type and horticultural cultivars' genomic DNA. The hybridization temperature (42℃) was calculated by the function provided in Roche's manual for 80-100% homology. As shown in Fig. 4, only one major band can be seen in the each lane of Avanti (horticultural, actinomorphic) and Espirito Santo (wild type, zygomorphic) cultivars, indicating that there is only one copy of CYC-like gene in *Sinningia speciose*. Moreover, the enzyme-digested fragment lengths are almost identical among the Avanti and Espirito Santo cultivars, implying that the genomic structures of these cultivars are very similar. Maybe this similarity is duo to the very recent domestication of the horticultural cultivars since 1815 AD (Citerne and Cronk, 1999). Although there are few weak bands other the major band, the contrasted weakness of these bands indicates the low homology of these sequences. The probe might hybridize to the other members of TCP gene family rather than the genes in the CYC-ECE2 clade.



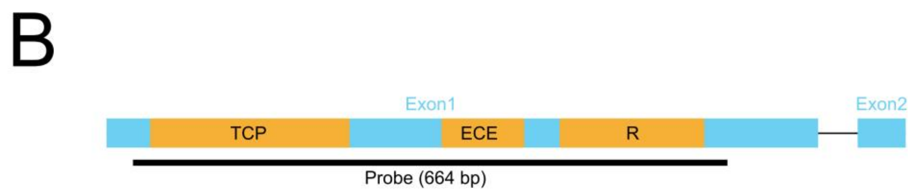


Figure 4 Southern blotting of detecting the copy number of *SsCYC* gene in Avanti (horticulture) and Espirito Santo (wild type) cultivars

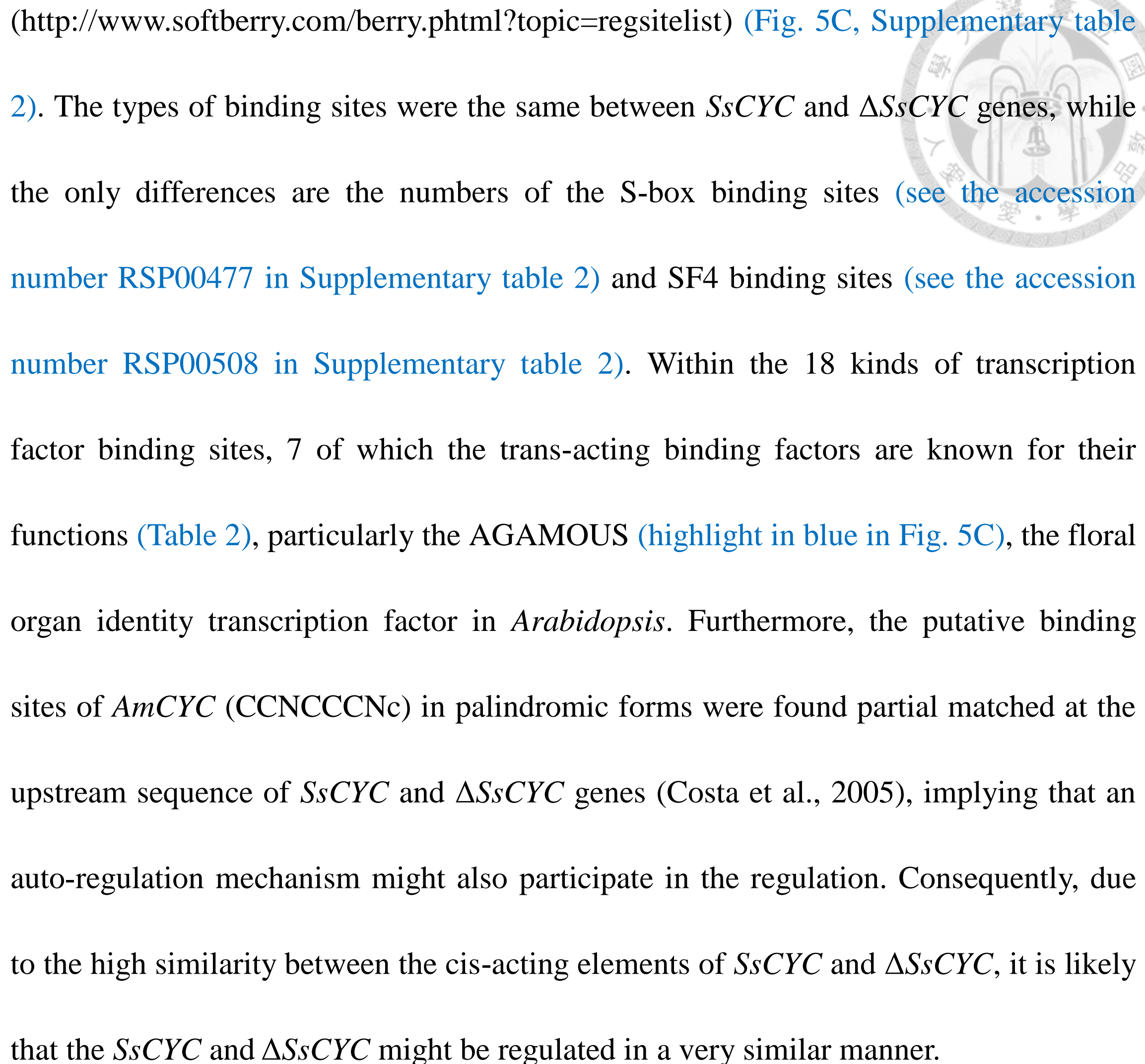
A, the result of southern blotting. The cultivars and restriction enzyme types are showed at the top of each lane. The most left lane is the marker and the size (bp) is annotated beside. This image is a combination of 2 membrane. Lanes for XbalI and XbalI/EcoRI are on the same membrane while lanes for EcoRI and HindIII are on the same membrane. B, the probe design. The exon 1 and exon 2 are showed as blue boxes. The three conserved domains are showed as orange box.

Isolation the upstream and coding sequences of *SsCYC* and Δ*SsCYC*

How a gene controls the phenotype can be categorized as regulatory level and function level. At the regulatory level, a gene can be either activated or repressed by cis- or trans-regulation in different spatial and temporal environment through the developmental program; at functional level, for AmCYC, a transcription factor, it can regulate its downstream genes through DNA binding or interact with other protein. To clarify whether there are sequence differences between ΔSsCYC from the actinomorphic cultivar 'Avanti' and the *SsCYC* from native zygomorphic cultivar 'Carangola', coding sequence and their upstream regulatory regions (1668 bp and 1656 bp, respectively) were isolated by standard PCR and TAIL PCR. Sequence alignments done by CLC Sequence Viewer 7 (http://www.clcbio.com) revealed 96.23% sequence similarity between the *SsCYC* and Δ*SsCYC* upstream regions and 98.36% sequence similarity between the *SsCYC* and Δ*SsCYC* coding regions. These sequence difference regions were marked with purple color in supplementary figure 3.

Comparison of upstream sequence between SsCYC and ΔSsCYC

To predict the promoters of *SsCYC* and Δ*SsCYC* genes, their upstream sequences were submitted to an online Transcription Start Site Prediction (TSSP) tool (http://linux1.softberry.com/berry.phtml?topic=tssp&group=programs&subgroup=promoter). The most likely transcription start sites (TSS) were predicted at the 919 and 921 bp upstream the start codons respectively (Fig. 5). In addition, 18 kinds of transcription factor binding sites were found in RegSite Database



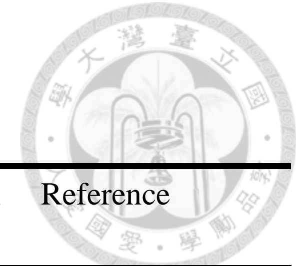
(http://www.softberry.com/berry.phtml?topic=regsitelist) (Fig. 5C, Supplementary table 2). The types of binding sites were the same between *SsCYC* and Δ*SsCYC* genes, while the only differences are the numbers of the S-box binding sites (see the accession number RSP00477 in Supplementary table 2) and SF4 binding sites (see the accession number RSP00508 in Supplementary table 2). Within the 18 kinds of transcription factor binding sites, 7 of which the trans-acting binding factors are known for their functions (Table 2), particularly the AGAMOUS (highlight in blue in Fig. 5C), the floral organ identity transcription factor in *Arabidopsis*. Furthermore, the putative binding sites of *AmCYC* (CCNCCCNc) in palindromic forms were found partial matched at the upstream sequence of *SsCYC* and Δ*SsCYC* genes (Costa et al., 2005), implying that an auto-regulation mechanism might also participate in the regulation. Consequently, due to the high similarity between the cis-acting elements of *SsCYC* and Δ*SsCYC*, it is likely that the *SsCYC* and Δ*SsCYC* might be regulated in a very similar manner.

Comparison of coding sequence between SsCYC and ΔSsCYC

Based on the alignment of *SsCYC* and Δ*SsCYC*, a small fragment deletion was found in Δ*SsCYC* (Fig. 5). Except this, the sequences between them are almost the same. In Δ*SsCYC*, a 10 bp deletion is located at 49-58 bp relative to *SsCYC* from the start codon (Fig. 5B). In addition, this deletion caused the frame shift and probably changes the encoded amino acid sequence. Furthermore, in Δ*SsCYC*, a premature stop codon is formed only 135 bp downstream from the start codon.

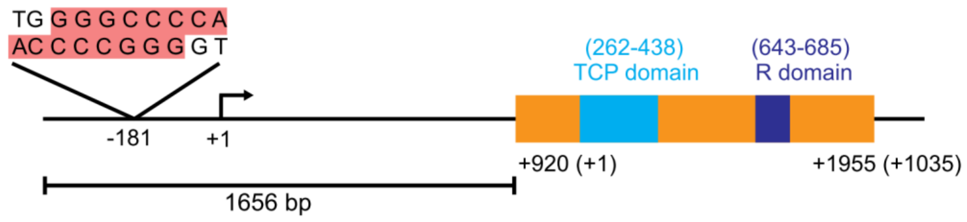
Table 2

Predicted trans-acting elements with known functions at the upstream regions of *SsCYC* and Δ*SsCYC*

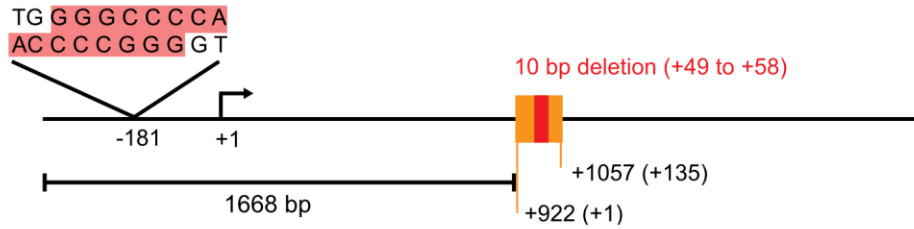


Binding factor	Organism	Known functions	Count in *SsCYC*	Count in Δ*SsCYC*	Reference
TGA1a	Tobacco (*Nicotiana tabacum*)	Stimulate the transcription of class II promoter	1	1	Yamazaki et al. (1990)
AGAMOUS	Arabidopsis (*Arabidopsis thaliana*)	Floral organ identity gene	1	1	Bowman et al. (1991)
Dof1	Maize (*Zea mays*)	Involving in carbon metabolism	4	4	Yanagisawa (2000)
ABI3	Arabidopsis (*Arabidopsis thaliana*)	key regulator of gene expression during embryo maturation	1	1	Nambara et al. (1994)
SEF3	Soybean (*Glycine max*)	A positive regulator of transcription of the beta-conglycinin, alpha' subunit gene in seed development	1	1	Allen et al. (1989)
SEF4	Soybean (*Glycine max*)	A positive regulator of transcription of the beta-conglycinin, alpha' subunit gene in seed development	12	13	Allen et al. (1989)
RITA-1	Rice (*Oryza sativa*)	Seed development	2	2	Izawa et al. (1994)

A



B



C

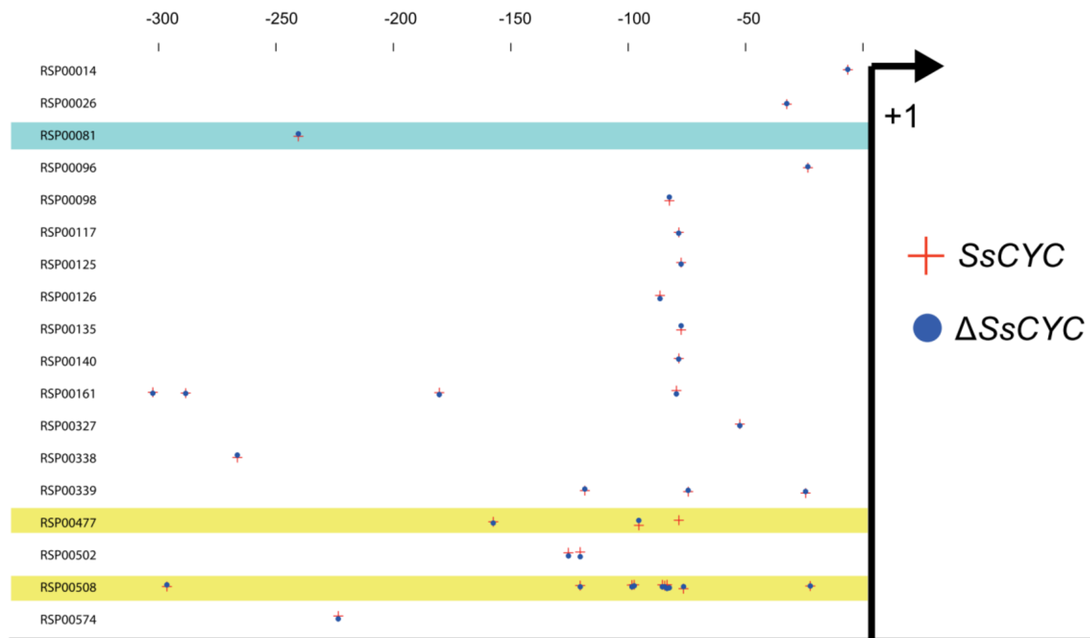
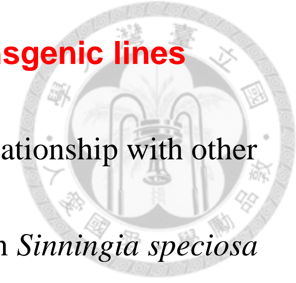


Figure 5 Illustration of *SsCYC* and Δ*SsCYC* genes and their putative promoters.

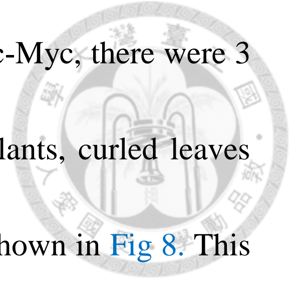
A, *SsCYC* gene, the homologue of *CYCLOIDEA* from zygomorphic *Sinningia speciosa*. **B**, Δ*SsCYC* gene, the homologue of *CYCLOIDEA* from actinomorphic *Sinningia speciosa*. +**1**, putative transcription start site (TSS). The pink boxes show the putative self-binding sites at the 181 bp upstream of TSS. Note that Δ*SsCYC* has a putative premature stop codon before the TCP and the R domains. **C**, the predicted transcription bind sites on *SsCYC* and Δ*SsCYC*. Red cross, on the sequence of *SsCYC*. Blue dot, on the sequence of Δ*SsCYC*. The names of the transcription factor binding sites are shown as the Accession Number (see in Supplementary table 2). The difference between *SsCYC* and Δ*SsCYC* are highlight with yellow background. The AGAMOUS binding sites are highlight with blue background.



Construction the overexpressed *SsCYC* and Δ*SsCYC* transgenic lines

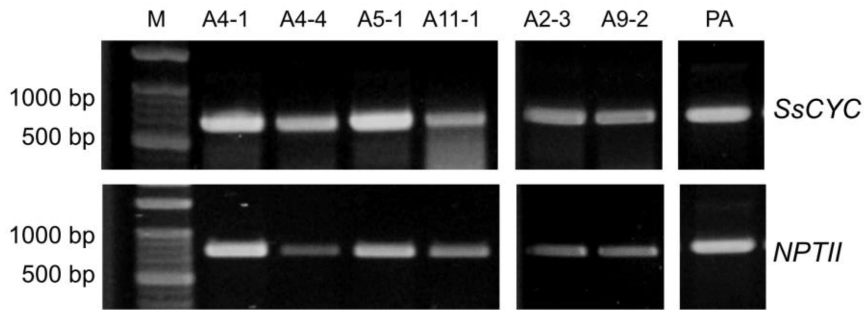
In spite of sharing the conserved domains and close phylogenetic relationship with other *CYC*-like genes, whether the *SsCYC* is the true orthologue of *CYC* in *Sinningia speciosa* still requires confirmation from functional studies. In addition to this, whether the small fragment deletion in Δ*SsCYC* is the key reason for floral symmetry reversion also needs functional proving. In this study, functional analyses by transformation both gene copies into *Nicotiana benthamiana* were done. Through *Agrobacterium* mediated leaf disk transformation, several independent T_0 transgenic lines were obtained. The putative transgenic plants were verified by checking the insertion of the *NPTII* and *SsCYC*/Δ*SsCYC* genes (Fig. 6). In addition, the mRNA expression levels of the *SsCYC*/Δ*SsCYC* gene were also examined by RT PCR with the *Actin* gene as the background control. As can be seen in Fig. 7, there were different levels of mRNA expression among the same T_0 population.

For the construct *p35S*::Δ*SsYC*::c-*Myc*, 6 independent transgenic lines were obtained from 104 explants, with the transformation rate = 5.7%. For the construct *p35S*::c-*Myc*::Δ*SsCYC*, 4 independent transgenic lines were obtained from 76 explants, with the transformation rate = 5.3%. For the construct *p35S*::*SsCYC*::c-*Myc*, 13 independent transgenic lines were obtained, with the transformation rate = 16%. There was no visible abnormal phenotype seen in the T_0 populations of *p35S*::Δ*SsYC*::c-*Myc*

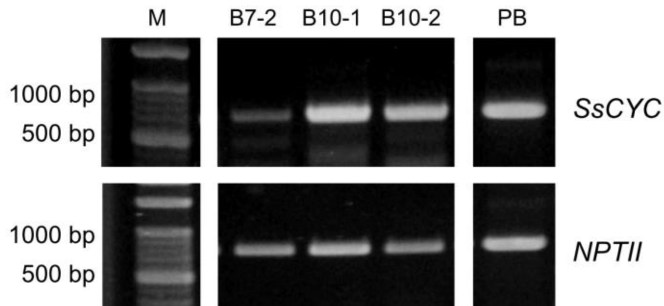


and *p35S*::c-*Myc*::Δ*SsCYC* constructs. However, for *p35S*::SsCYC::c-Myc, there were 3 out of 13 showing abnormal phenotypes. For example, dwarfish plants, curled leaves and small flowers. The pictures of the 3 transgenic populations are shown in Fig 8. This implies that 1) Δ*SsCYC* gene might have lost its function due to the small sequence deletion and frame shift. 2) *SsCYC* could change the growth patterns of tobacco plants. The detailed descriptions of the abnormalities in *p35S*::*SsCYC*::c-*Myc* transgenic population are presented in the following sections.

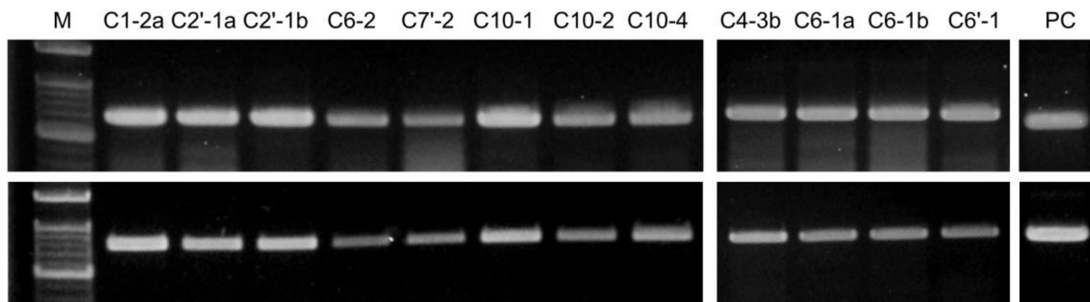
A



B



C



D

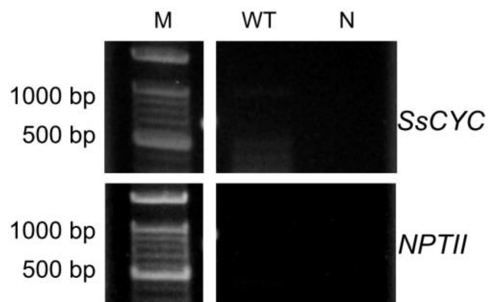
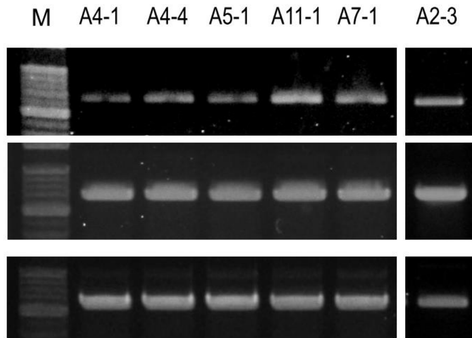
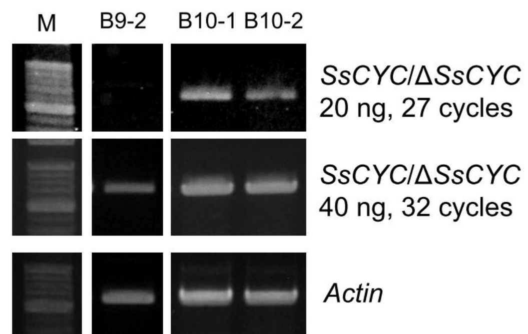
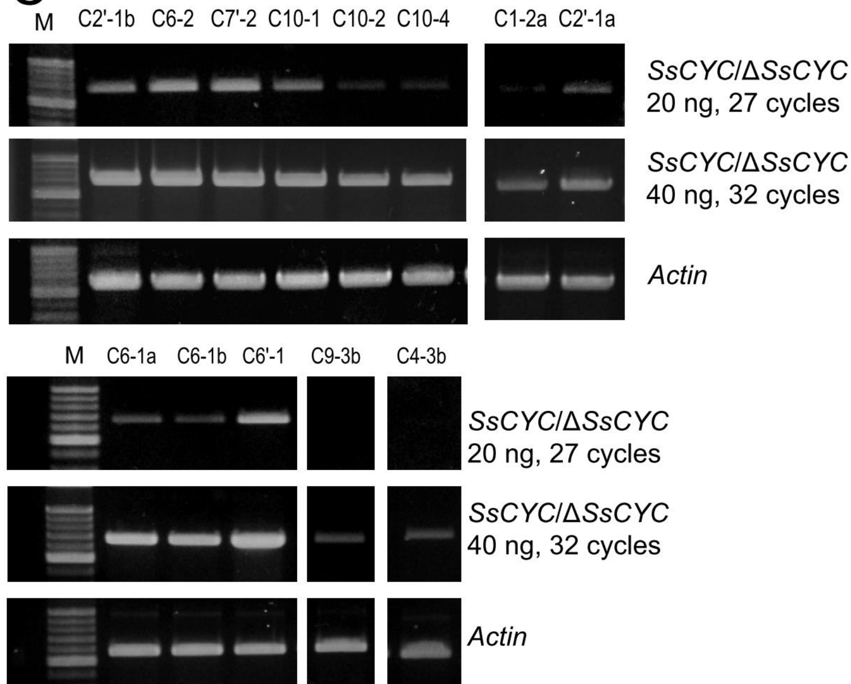


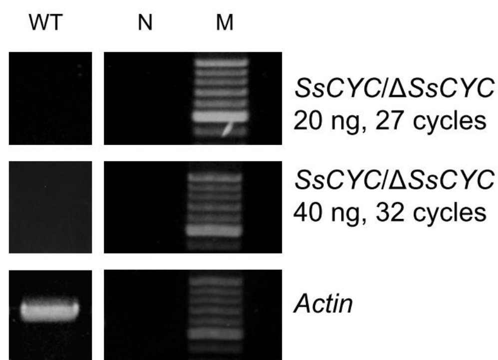
Figure 6 T-DNA insertion in T_0 transgenic populations.

A**,** transgenic lines of *p35S*::Δ*SsCYC*::*c-Myc*; **B**, transgenic lines of *p35S*::*c-Myc*::*ΔSsCYC*; **C**, transgenic lines of *p35S*::*SsCYC*::*c-Myc*. PA, positive control of "A" transgenic lines. PB, positive control of "B" transgenic lines. PC, positive control of "C" transgenic lines. WT, wild type. N, no template control. M, marker.


A
M
A4-1
A4-4
A5-1
A11-1
A7-1
A2-3


B
M
B9-2
B10-1
B10-2
SsCYC/ΔSsCYC
20 ng, 27 cycles
SsCYC/ΔSsCYC
40 ng, 32 cycles
Actin


C
M
C2'-1b
C6-2
C7'-2
C10-1
C10-2
C10-4
C1-2a
C2'-1a
SsCYC/ΔSsCYC
20 ng, 27 cycles
SsCYC/ΔSsCYC
40 ng, 32 cycles
Actin
M
C6-1a
C6-1b
C6'-1
C9-3b
C4-3b
SsCYC/ΔSsCYC
20 ng, 27 cycles
SsCYC/ΔSsCYC
40 ng, 32 cycles
Actin


D
WT
N
M
SsCYC/ΔSsCYC
20 ng, 27 cycles
SsCYC/ΔSsCYC
40 ng, 32 cycles
Actin

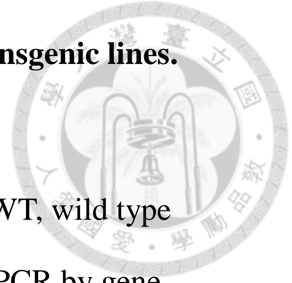


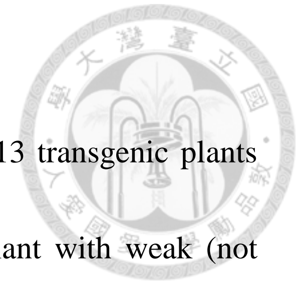
Figure 7 mRNA expression levels of *SsCYC* or *ΔSsCYC* in T_0 transgenic lines.

A, transgenic lines of *p35S*::Δ*SsCYC*::c-*Myc*; **B**, transgenic lines of *p35S*::c-*Myc*::Δ*SsCYC*; **C**, transgenic lines of *p35S*::*SsCYC*::c-*Myc*. WT, wild type *Nicotiana benthamiana*. N, no template control. M, marker. *SsCYC*, PCR by gene specific primers of *SsCYC* or Δ*SsCYC*. The gene names, amount of RNA template and PCR cycles are listed at the right of each row.

Figure 8 Mature T_0 transgenic plant morphology.

Left, transgenic plant of *p35S*::Δ*SsCYC*::c-*Myc*. **Middle**, transgenic plant of *p35S*::c-*Myc*::Δ*SsCYC*. **Right**, transgenic plant of *p35S*::*SsCYC*::c-*Myc*. Note that transgenic plant of p35*S*::*SsCYC*::c-*Myc* is shorter than the other two transgenic plant. Scale bar, 10 cm.

Abnormal phenotypes among *SsCYC* transgenic plants



Within the *SsCYC* T_0 population (*p35S*::*SsCYC*::c-*Myc*), 3 out of 13 transgenic plants had severe phenotypes. As can be seen in Fig. 9, a transgenic plant with weak (not visible) phenotypes (Fig. 9A) and 3 transgenic plants with different degrees of abnormalities (Fig. 9B, C & D). Generally, the plant height declined with the increase of abnormalities. In addition, comparison with wild type and Δ*SsCYC* transgenic populations (including both *p35S*::Δ*SsYC*::c-*Myc* and *p*35*S*::c-*Myc*:: Δ*SsCYC*), *SsCYC* transgenic plants had more developed axillary shoots (Fig. 10), indicating that ectopic *SsCYC* expression might suppress the primary growth but promote the secondary growth of axillary shoots. For detailed leaf morphology, as shown in Fig. 11, the *SsCYC* transgenic plants had smaller juvenile leaves (Fig. 11A) and adult leaves (Fig. 11B) than wild type and Δ*SsCYC* transgenic plants. Moreover, the leaves of *SsCYC* transgenic plants became thicker and curled toward the abaxial sites, implying ectopic expression of the *SsCYC* might also have participated in the leaf development.

Figure 9 Different degrees of abnormality among the *SsCYC* transgenic plants. **A**, weak phenotype, no visible difference can be seen between wild type. **B**, medium phenotype. **C**, strong phenotype. **D**, very strong phenotype. Note that the phenotypes of induction of axillary shoots, curled leaves and dwarfism can be seen in B-D. Scale bar, 10 cm.

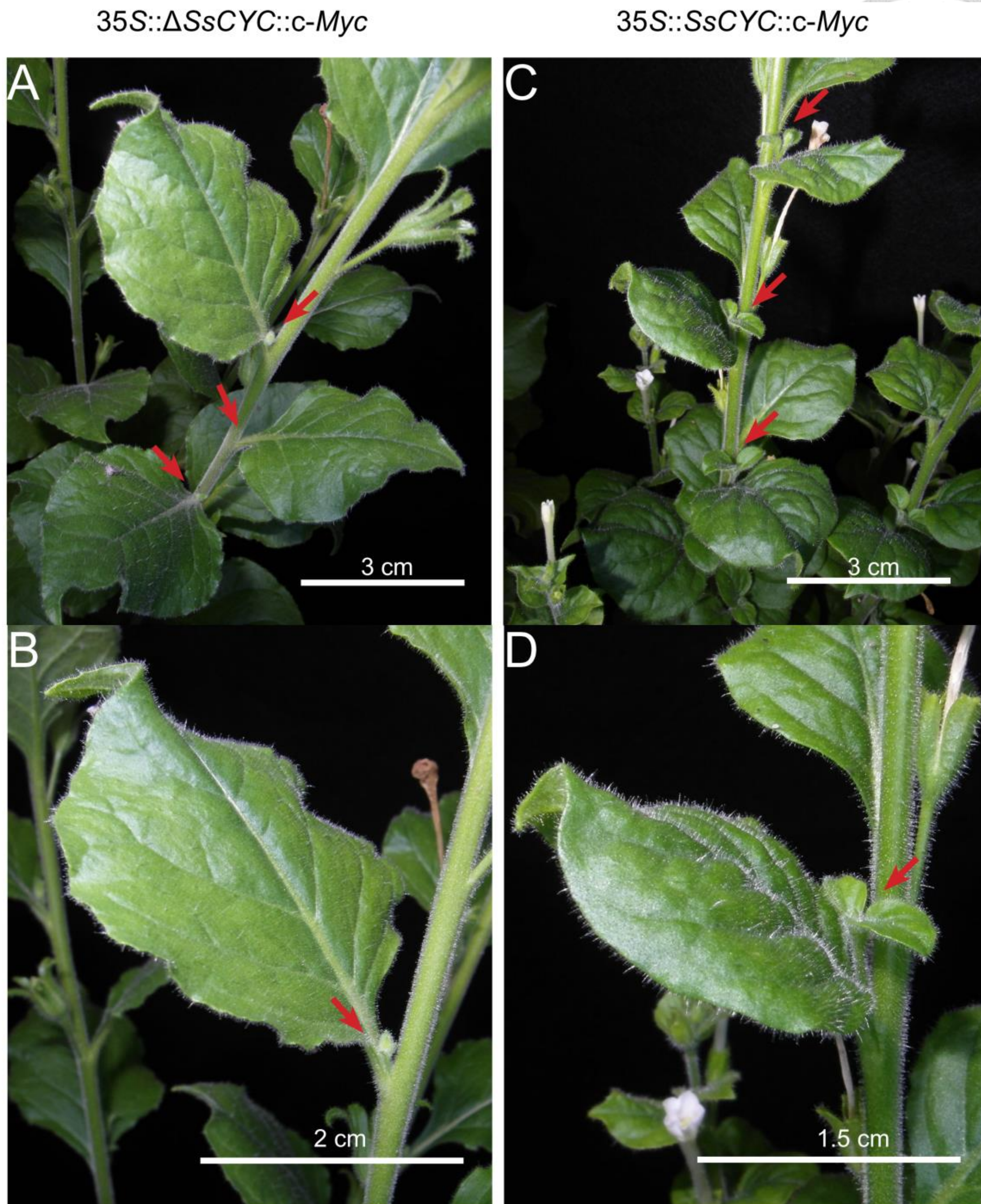
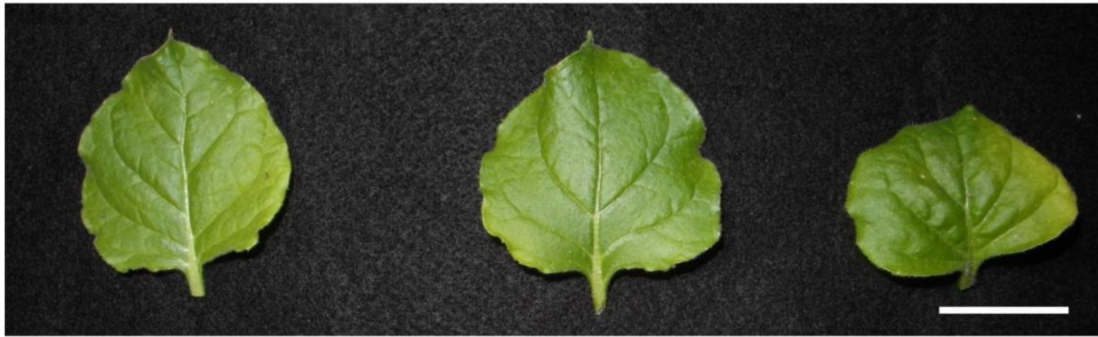


Figure 10 Promoting the growth of axillary buds in *SsCYC* transgenic plants.

A-B, Δ*SsCYC* transgenic plant with no phenotype. **C-D,** *SsCYC* transgenic plant with sever phenotype. Arrow, the position of axillary shoot. Note that C and D have the developed axillary shoots while A and B do not.

A

35*S*::Δ*SsCYC*::c-*Myc*　　35*S*::c-*Myc*::Δ*SsCYC*　　35*S*::*SsCYC*::c-*Myc*

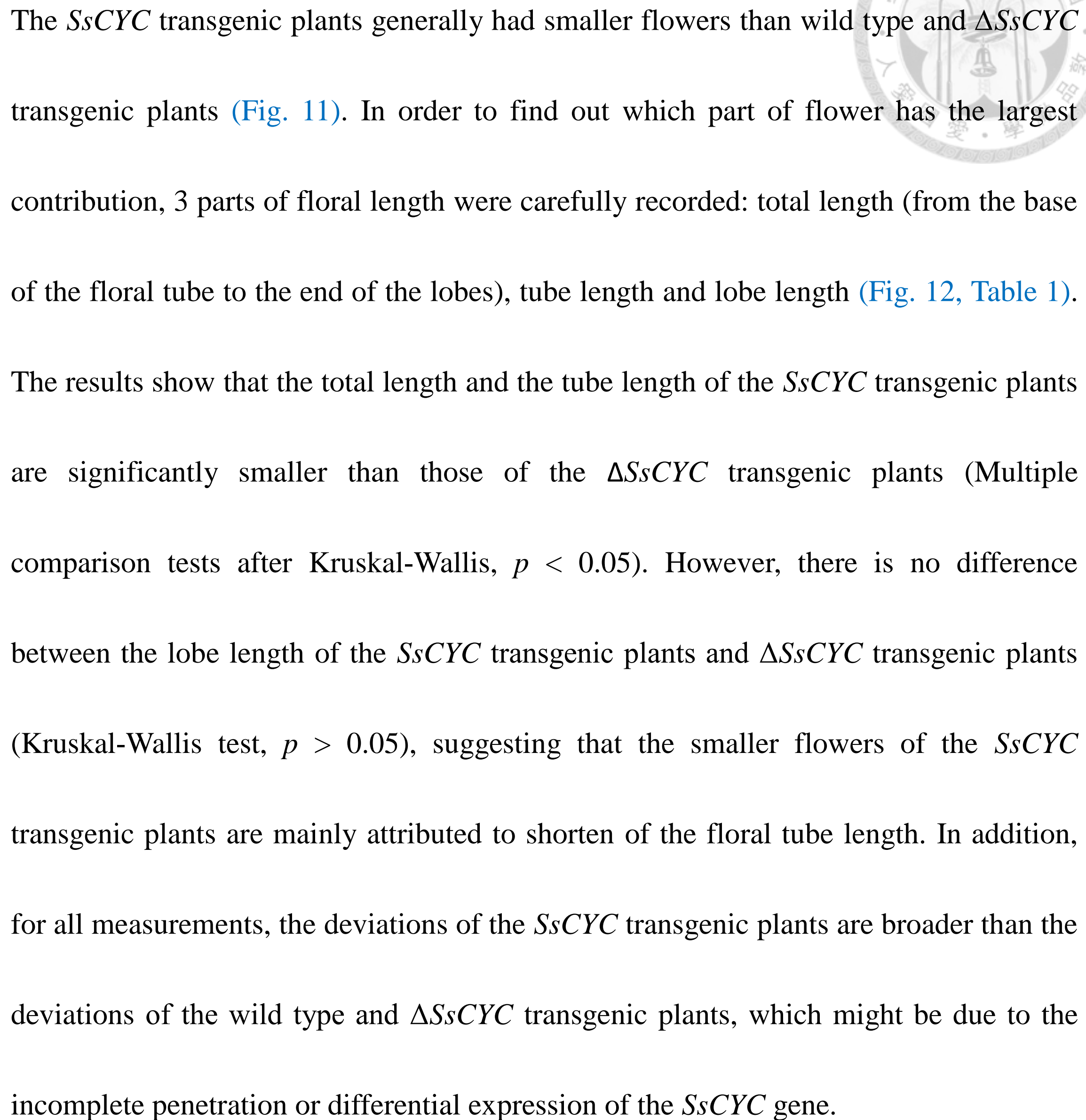


B

35*S*::Δ*SsCYC*::c-*Myc*　　35*S*::c-*Myc*::Δ*SsCYC*　　35*S*::*SsCYC*::c-*Myc*

Figure 11 Leaf morphology of T_0 transgenic plants.

A, Juvenile leaves. **B**, adult leaves. **Left**, ectopic expression of Δ*SsCYC* with Myc tag fusion at the C-terminal end. **Middle**, ectopic expression of Δ*SsCYC* with Myc tag fusion at the N-terminal end. **Right**, ectopic expression of *SsCYC* with Myc tag fusion at the C-terminal end. Note the smaller size and curling toward abaxial side. The upper row, adaxial side. The lower row, abaxial side. Scale bar, 3 cm.



Floral morphology measurement

The *SsCYC* transgenic plants generally had smaller flowers than wild type and Δ*SsCYC* transgenic plants (Fig. 11). In order to find out which part of flower has the largest contribution, 3 parts of floral length were carefully recorded: total length (from the base of the floral tube to the end of the lobes), tube length and lobe length (Fig. 12, Table 1). The results show that the total length and the tube length of the *SsCYC* transgenic plants are significantly smaller than those of the Δ*SsCYC* transgenic plants (Multiple comparison tests after Kruskal-Wallis, $p < 0.05$). However, there is no difference between the lobe length of the *SsCYC* transgenic plants and Δ*SsCYC* transgenic plants (Kruskal-Wallis test, $p > 0.05$), suggesting that the smaller flowers of the *SsCYC* transgenic plants are mainly attributed to shorten of the floral tube length. In addition, for all measurements, the deviations of the *SsCYC* transgenic plants are broader than the deviations of the wild type and Δ*SsCYC* transgenic plants, which might be due to the incomplete penetration or differential expression of the *SsCYC* gene.

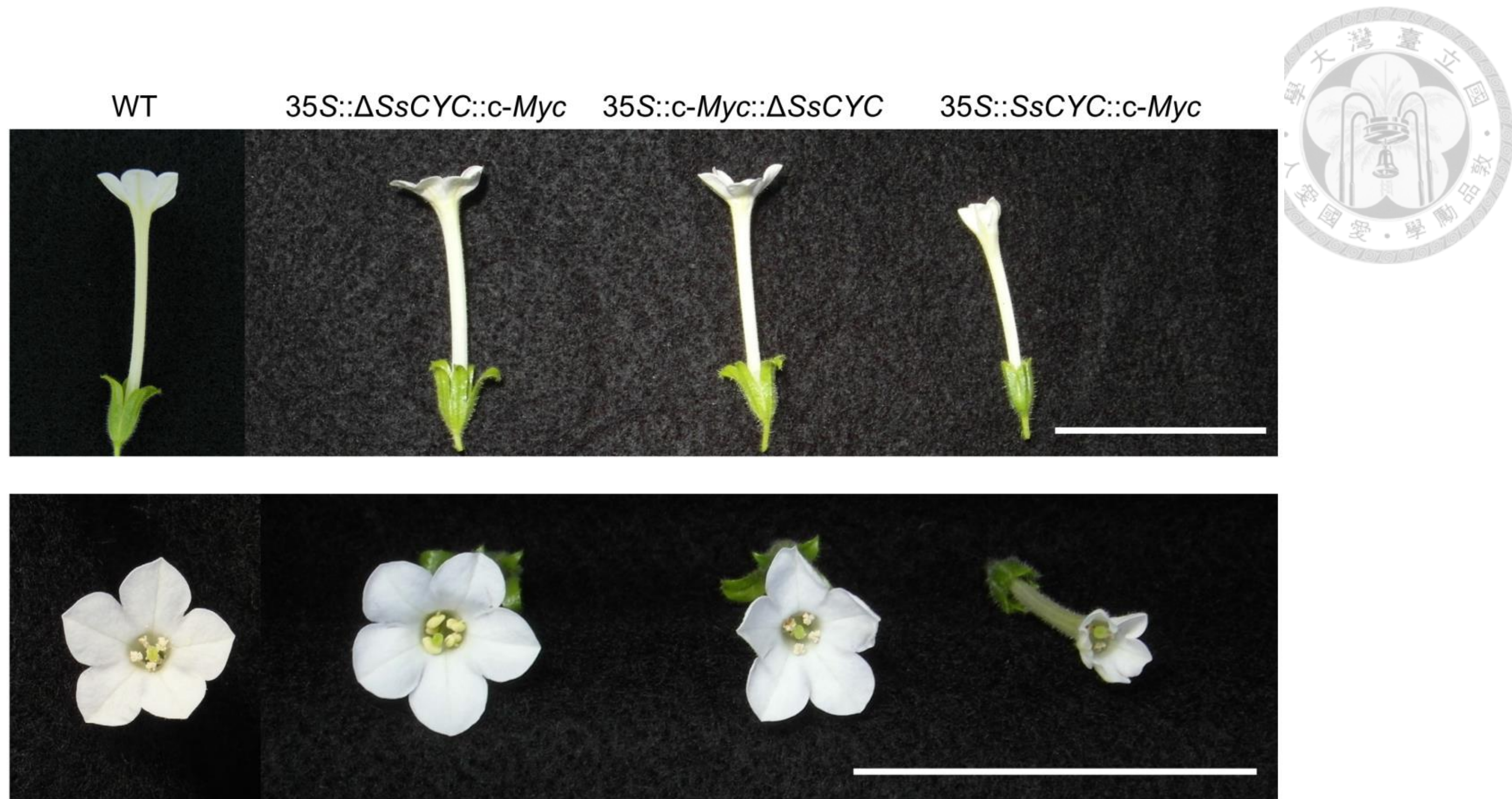
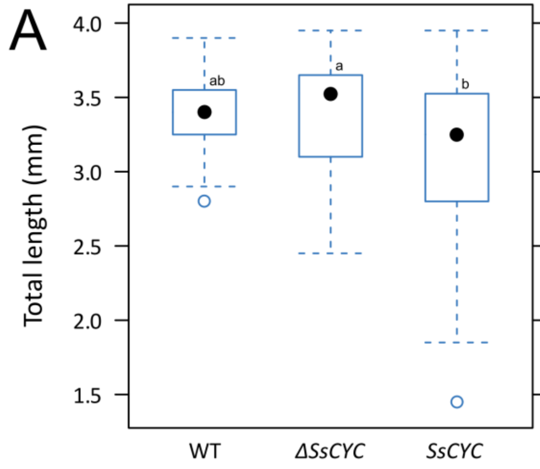
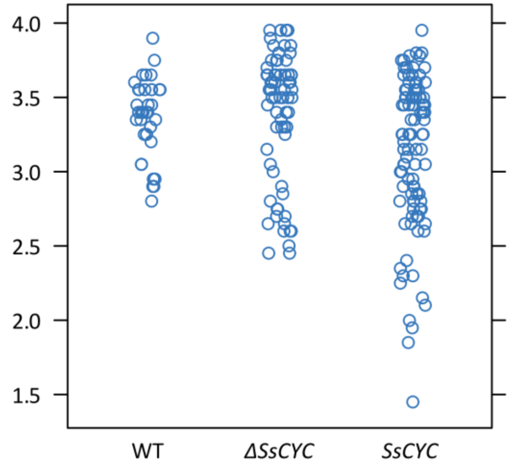
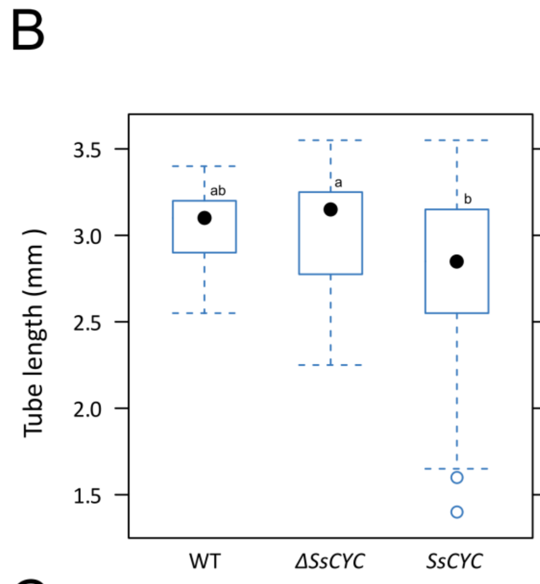
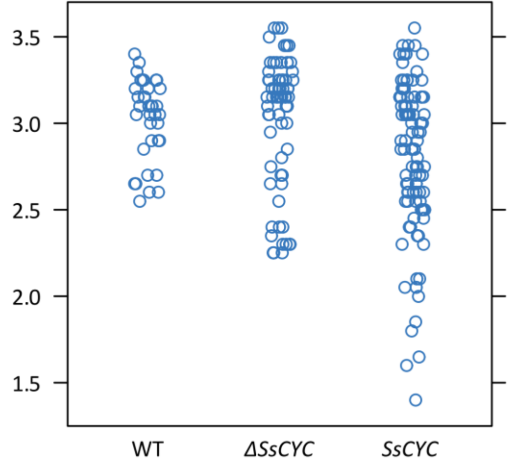


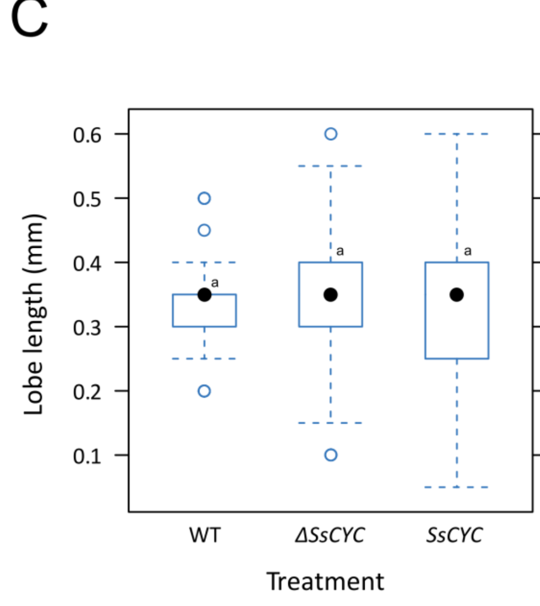
Figure 12 Flower morphology of T_0 transgenic plants at fully elongated stage.
Left, ectopic expression of Δ*SsCYC* with c-*Myc* tag fusion at the C-terminal end. **Middle**, ectopic expression of Δ*SsCYC* with *c-Myc* tag fusion at the N-terminal end. **Right**, ectopic expression of *SsCYC* with *c-Myc* tag fusion at the C-terminal end. Note the smaller size and curling toward adaxial side. All the flowers were pictured at the fully open stage (stage 15-16). The upper row, the side view. The lower row, the front view. Scale bar, 3 cm.


A
Total length (mm)
4.0
3.5
3.0
2.5
2.0
1.5
ab
a
b
WT
ΔSsCYC
SsCYC


4.0
3.5
3.0
2.5
2.0
1.5
WT
ΔSsCYC
SsCYC


B
Tube length (mm)
3.5
3.0
2.5
2.0
1.5
ab
a
b
WT
ΔSsCYC
SsCYC


3.5
3.0
2.5
2.0
1.5
WT
ΔSsCYC
SsCYC


C
Lobe length (mm)
0.6
0.5
0.4
0.3
0.2
0.1
a
a
a
WT
ΔSsCYC
SsCYC
Treatment

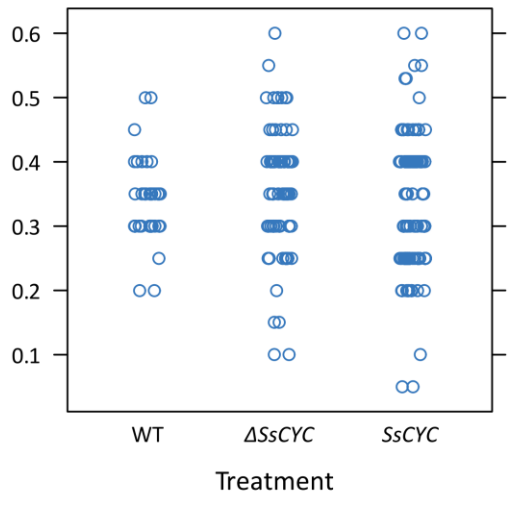
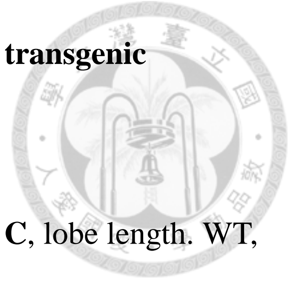

0.6
0.5
0.4
0.3
0.2
0.1
WT
ΔSsCYC
SsCYC
Treatment

Figure 13 Flower morphology measurement of *SsCYC*, Δ*SsCYC* transgenic tobacco plants.



A, total length (tube length + lobe length) of flowers. **B**, tube length. **C**, lobe length. WT, wild type. Δ*SsCYC*, ectopic expression of *CYC* homologue from actinomorphic cultivar. *SsCYC*, ectopic expression of *CYC* homologue from zygomorphic cultivar. **Left column**, boxplot with black dot as medium, box as the first and the third quartiles, dash line as the minimum and the maximum values within the fences (the first quartile – 1.5X IQR and the third quartile + 1.5X IQR), empty dot as the outliner (the values outside the fences). Different letters on the boxes indicate that the mediums are significantly different (Multiple comparison test after Kruskal-Wallis , $p < 0.05$). **Right column**, dot plot reveals the distribution of the raw data.

Association between phenotypic abnormalities and mRNA expression levels

To test whether various degree of phenotypic abnormality among the *SsCYC* transgenic plants has any correlation with its *SsCYC* mRNA overexpression levels, the expression level of the *SsCYC* mRNA among these plants were compared (Fig. 14). The mRNA expression levels were ranked as 4 orders (Fig. 14) and associated with the floral total length, tube length and lobe length (Fig. 15).

For total length (from the base of floral tube to the end of floral lobe) (Fig. 15A) and tube length (Fig. 15B), there is no difference between the level 0 (wild type control), level 1, level 2 and level 3 transplants (Kruskal-Wallis test, $p > 0.05$). However, the total length and tube length of the level 4 treatment (highest *SsCYC* expression) is significantly smaller than the other four treatments (Multiple comparison tests after Kruskal-Wallis, $p < 0.05$). For lobe length (Fig. 15C), there is no difference between the level 0, level 1 and level 3 treatments (Kruskal-Wallis test, $p > 0.05$). In addition, there is no difference between the level 0, level 2 and level 3 treatments (Kruskal-Wallis test, $p > 0.05$). However, the lobe length of the level 4 treatment is significantly smaller than the other four treatments (Multiple comparison tests after Kruskal-Wallis, $p < 0.05$).

The results indicate that the mRNA expression levels played an important role in the floral morphology, which might explain the large deviation of the measurements of the *SsCYC* transgenic plants (Fig. 13). Only the highest expression level (rank 4 in Fig. 15) *SsCYC* gene can penetrate its effects to phenotypes.

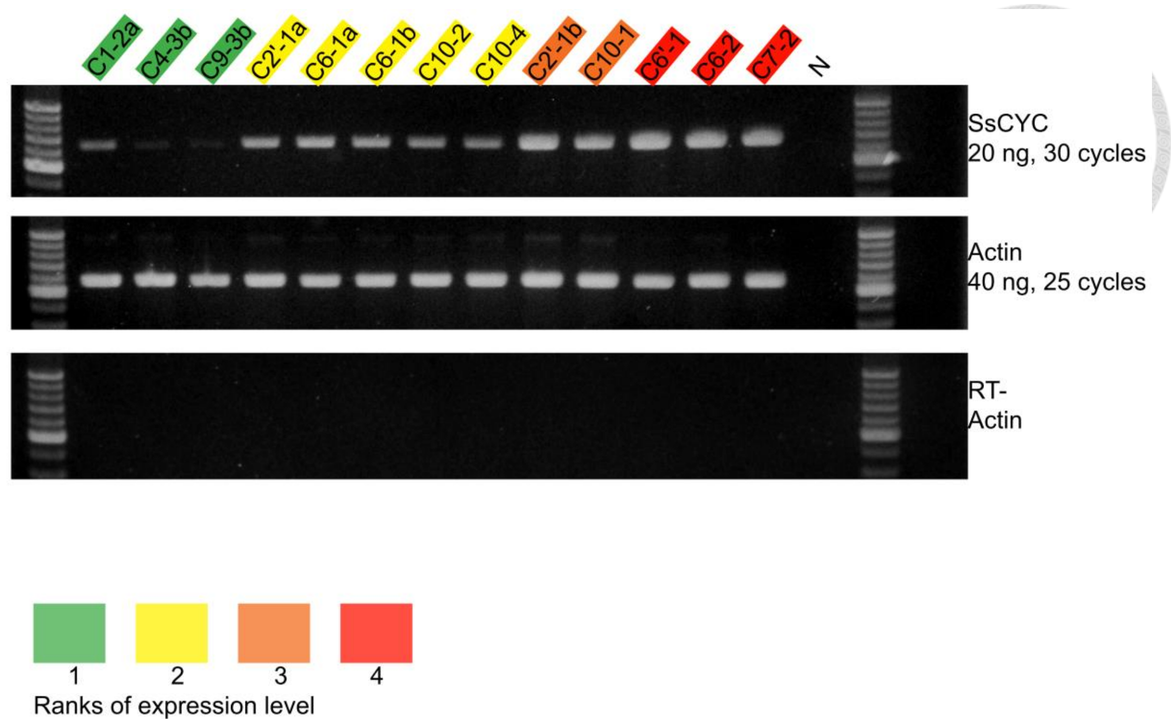
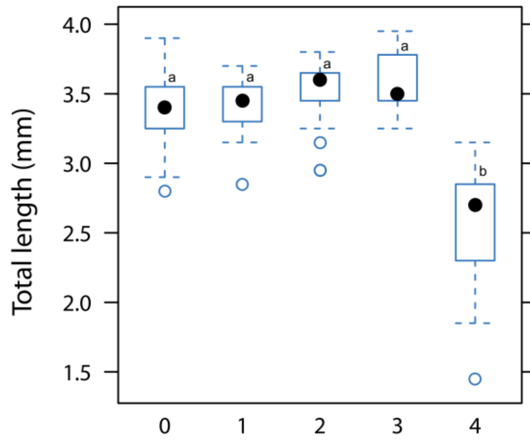
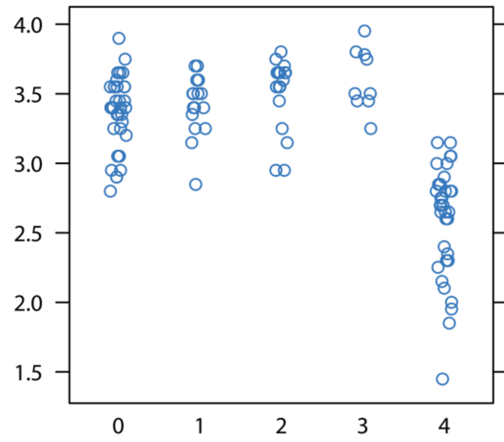


Figure 14 Relative *SsCYC* mRNA expression level among T_0 population.

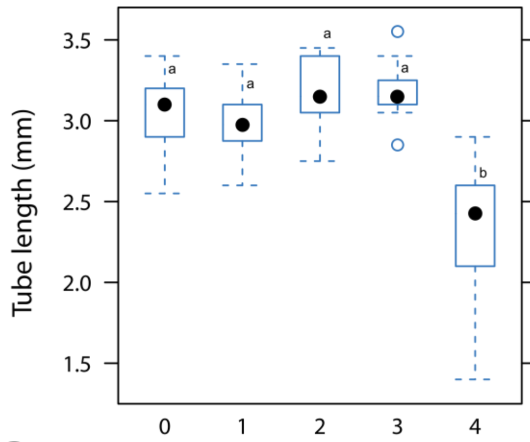
The first row, the RT-PCR result of *SsCYC*, 20 ng of total RNA was reverse transcribed into cDNA by oligo $d(T)_{18}$ primer and then followed by 30 cycles of PCR amplification. **The second row**, the RT-PCR result of *Actin*, the housekeeping gene, 40 ng of total RNA was reverse transcribed into cDNA by oligo $d(T)_{18}$ primer and then followed by 25 cycles of PCR amplification. **The third row**, the RT minus control of Actin, 40 ng of total RNA was directly used PCR amplification, 25 cycles. The relative *SsCYC* mRNA expression levels are ranked as 1-4: **1 (green)** stands for the lowest expression level; **2 (yellow)** stands for medium-low expression; **3 (orange)** stands for medium-high expression level; **4 (red)** stands for the highest expression level. **N**, no template control. **M**, marker, the major bands stand for 500 bp and 1000 bp.

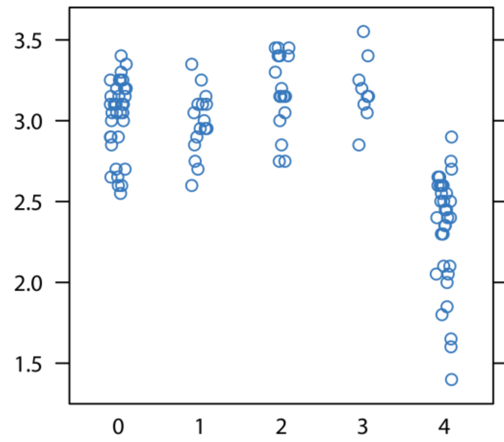
A



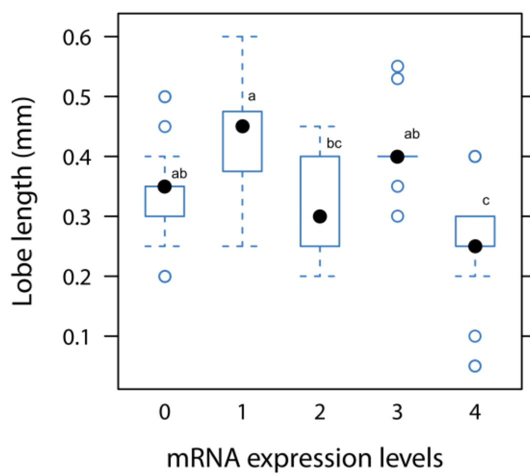


B





C



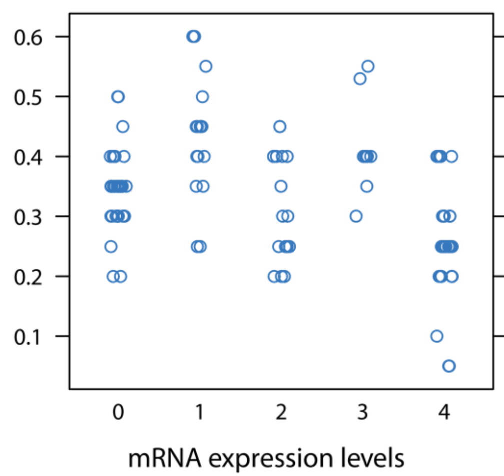
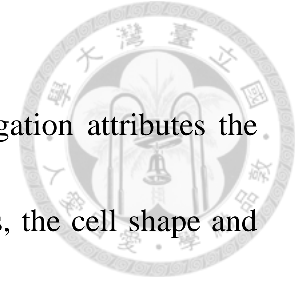


Figure 15 Flower morphology according to the relative mRNA expression level in *SsCYC* transgenic plants

A, total length (tube length + lobe length) of flowers. **B**, tube length. **C**, lobe length. The relative mRNA expression levels are classed as rank: **0**, there is no expression detected (wild type). 1-4, the expression levels are grouped by order: **1**, the lowest expression level. **4**, the highest expression level. **Left column**, boxplot with black dot as medium, box as the first and the third quartiles, dash line as the minimum and the maximum values within the fences (the first quartile − 1.5X IQR and the third quartile + 1.5X IQR), empty dot as the outliner (the values outside the fences). Different letters on the boxes indicate that the mediums are significantly different (Multiple comparison test after Kruskal-Wallis , $p < 0.05$). **Right column**, dot plot reveals the distribution of the raw data.

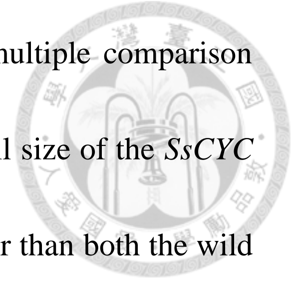


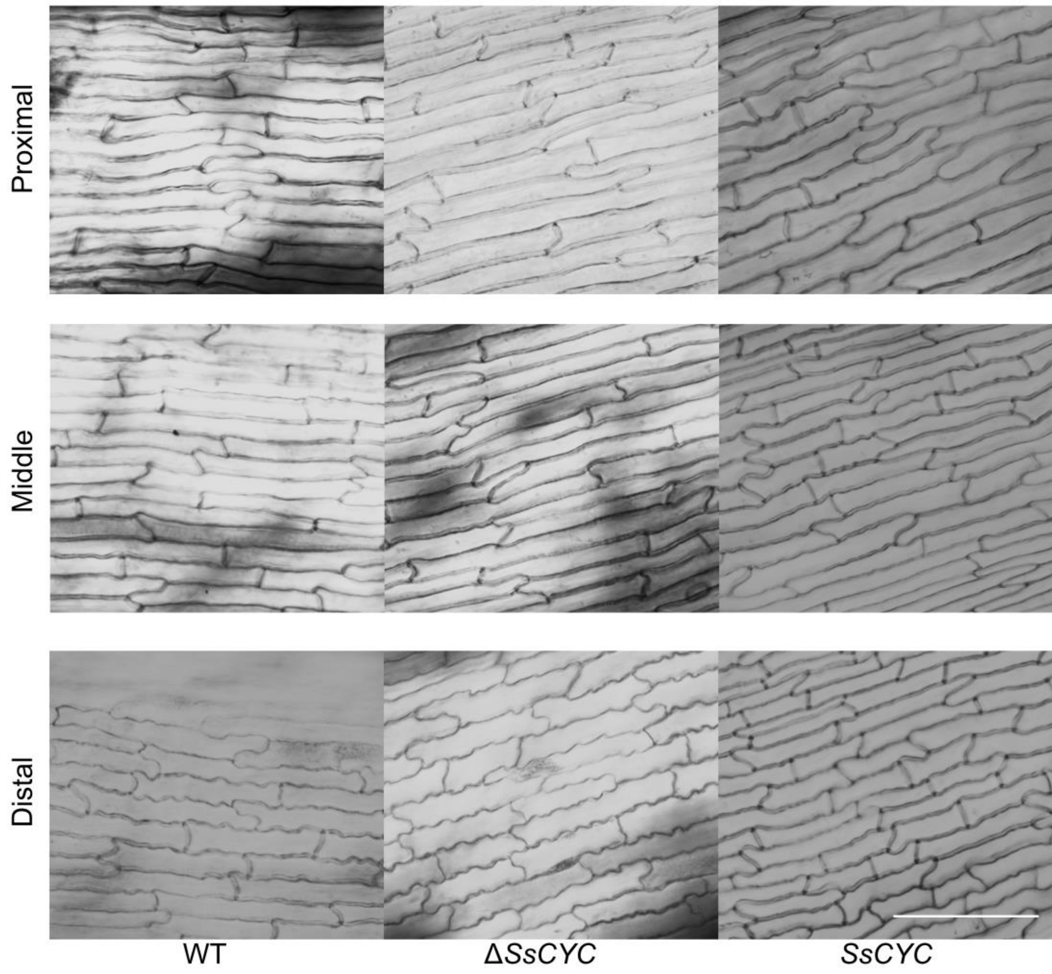
Floral tube cell shape and size measurements

In order to further examine whether the cell division or cell elongation attributes the shorten total length and tube length of the *SsCYC* transgenic plants, the cell shape and cell area of the floral tubes were measured by clearing method. As seen in Fig. 16A, for wild type and Δ*SsCYC* transgenic plants, the cell shapes were rectangular with smooth edges at the proximal and middle sites while the cell shape was rectangular with serrated edge at the distal site. However, for *SsCYC* transgenic plants, the serrate of edges became weak but can be seen in the proximal, middle and distal sites. This shows that ectopic expression *SsCYC* could modify the cell shape in both spatial identity and intensity.

For statistical analysis of cell area in a floral tube, 20 cells of 3 sites were measured: the proximal site (the site closest to the pedicle), the distal site (the site closest to the floral opening) and the middle site (the site at the middle of proximal and distal sites). For the proximal site (Fig. 16B, left), the cell size of Δ*SsCYC* transgenic plants (1981.26 ± 105.04 μm^2, n = 20) were significantly higher than the wild type (1411.69 ± 78.34 μm^2, n = 20) and the SsCYC transgenic plants (1432.05 ± 71.20 μm^2, n = 20) (Tukey multiple comparison after ANOVA, $p < 0.05$). For the middle site (Fig. 16B, middle), the cell size of the wild type (1483.13 ± 78.60 μm^2, n = 20) the ΔSsCYC transgenic plants (1889.79 ± 92.95 μm^2, n = 20) and the SsCYC transgenic plant (969.14 ± 49.19

μm^2, n = 20) are significantly different from each other (Tukey multiple comparison after ANOVA, $p < 0.05$). For the distal site (Fig. 16B, right), the cell size of the *SsCYC* transgenic plant (721.03 ± 37.75 μm^2, n = 20) is significantly smaller than both the wild type (1524.65 ± 59.49 μm^2, n = 20) and the ΔSsCYC transgenic plant (1505.07 ± 61.32 μm^2, n = 20) (Tukey multiple comparison after ANOVA, $p < 0.05$).




A
Proximal
Middle
Distal
WT
ΔSsCYC
SsCYC

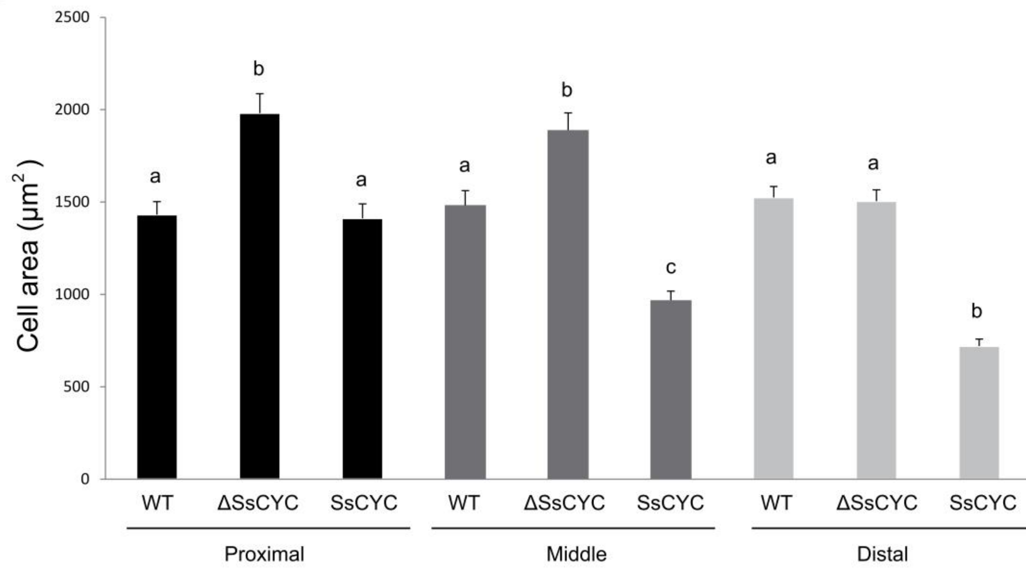
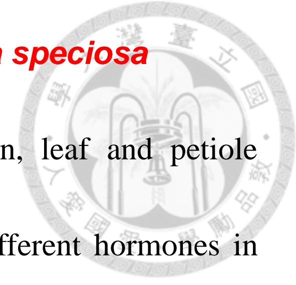

B
2500
2000
1500
1000
500
0
Cell area (μm2)
a
b
a
a
b
c
a
a
b
WT
ΔSsCYC
SsCYC
WT
ΔSsCYC
SsCYC
WT
ΔSsCYC
SsCYC
Proximal
Middle
Distal

Figure 16 Cell shape and size of floral tube measurement of *SsCYC* and Δ*SsCYC* transgenic tobacco plants

A, cell morphology under light microscope. Proximal, the proximal site of the floral tube which is close the pedicel. Middle, the middle site of the floral tube. Distal, the distal site of the floral tube which is close to the floral opening. **B**, cell area of the proximal, middle and distal sites. Different letters on the bars indicate that the mean values are significantly different (Tukey multiple comparisons after one-way ANOVA test, $p < 0.05$). The data is presented as mean ± stand error. Each treatment has 20 cells from a single T_0 plant. Scale bar, 100 μm.

Tissue culture conditions for shoot induction in *Sinningia speciosa*

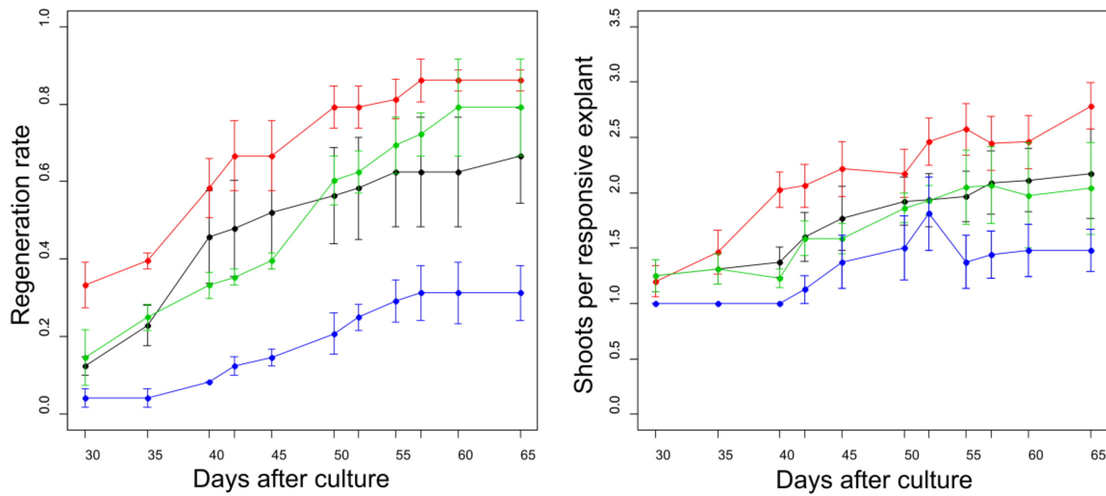


To find the best tissue culture conditions for shoot regeneration, leaf and petiole explants were inoculated on the medium with combinations of different hormones in different concentrations. For all treatments, after 7 days of culture, the leaf explants expended their size about 2-4 times, started to curl toward abaxial side and the edge of explants became swollen, especially at the site of secondary veins. However, cutting edge of petiole explants were only slightly swollen. After 14 days of culture, in the treatment with 0.1 ppm NAA only, several adventitious roots formed at the edge of both leaf and petiole explants. However, in the treatment with both NAA and BA (0.1 ppm NAA, 1 ppm BA; 0.1 ppm NAA, 2 ppm BA; 0.1 ppm NAA, 3 ppm BA), no root formation could be seen, suggesting that the exogenous BA might inhibit the re-differentiation of roots. After 30 days of culture, visible shoots appeared on the explants among all treatments, and the regeneration rate and shoots per responsive explant were carefully recorded till 65 days by which the shoot regeneration had reached a stationary phrase (Fig. 17). The best hormone combination for shoot regeneration was 0.1 ppm NAA together with 1.0 ppm BA in both leaf and petiole explants. For leaf explants, this treatment had almost 90% regeneration rate with 2.5-3 shoots per responsive explant at the end of culture (Fig. 17A); for petiole explants, up to 50% regeneration rate with 1.5-2 shoots per responsive explant was obtained (Fig. 17B).

On the other hand, the treatment with 0.1 ppm NAA, 0 ppm BA and the treatment with 0.1 ppm NAA, 2 ppm BA had similar performances in shoot regeneration, which are 10-30% lower than the regeneration rate of the best condition, while the treatment with 0.1 ppm NAA, 3 BA had the lowest shoot regeneration response.

However, for leaf explants, though not the best shoot regeneration response, the treatment with 0.1 ppm NAA, 0 ppm BA had the largest well-development shoots among other treatments (Fig. 18A, B, C, and D). The effect of this treatment was quite different from petiole explants, which had almost no regenerative shoots at the end of the culture (Fig. 18E, F, G, and H). Furthermore, as can be seen in Fig. 18E, the regenerative roots formed exclusively at the proximal site (closer to the main stem than distal), which indicates that some endogenous hormone gradient existed to direct root regeneration sites.

A. Shoot induction from leaf



B. Shoot induction from petiole

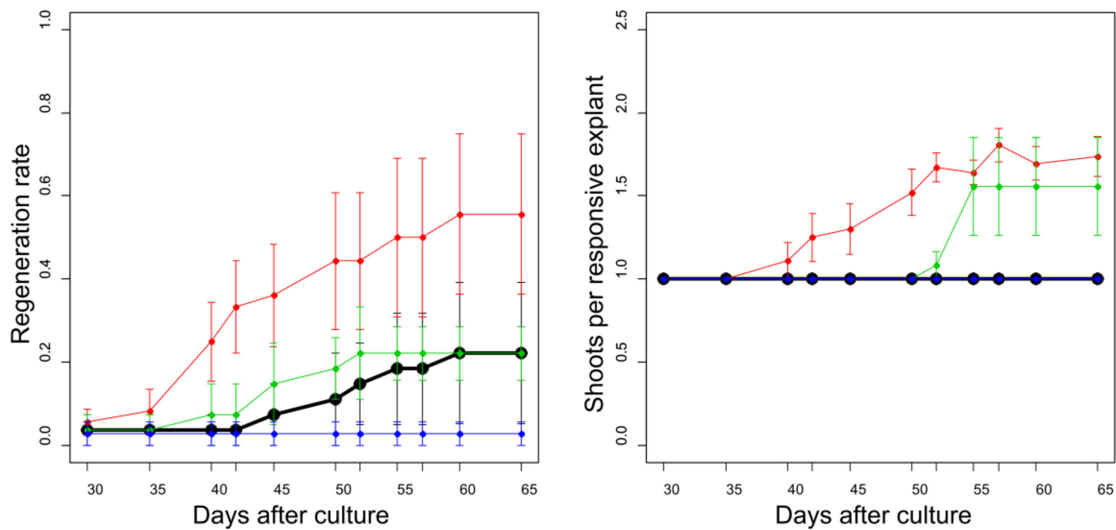
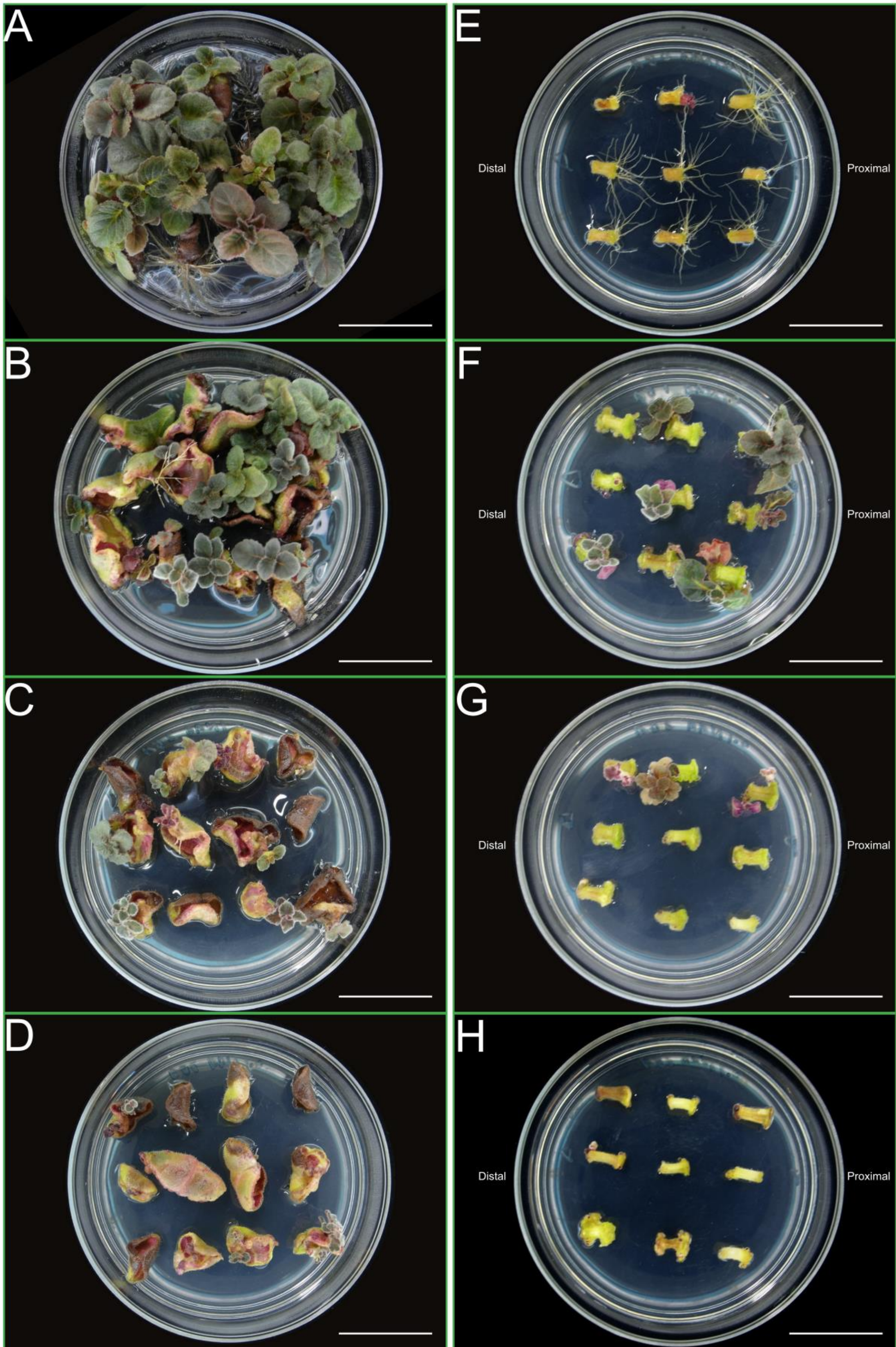


Figure 17 Effect of cytokinin and auxin concentration and explant sources on regeneration rate and shoots per responsive explant.

A, shoot induction from leaf explants. **B**, shoot induction from petiole explants. Regeneration rate is calculated as number of responsive explants (the shooting explants) divided by total explants. Shoots per responsive explant is defined as total number of regenerative shoots divided by total explants. **Black**, 0.1 ppm NAA and 0 ppm BA. **Red**, 0.1 ppm NAA and 1 ppm BA. **Green**, 0.1 ppm NAA and 2 ppm BA. **Blue**, 0.1 ppm NAA and 3 ppm BA. The data are presented as mean ± stand error. Each treatment has 4 replicates and each replicate has 12 leaf explants or 9 petiole explants.


A
B
C
D
E
Distal
Proximal
F
Distal
Proximal
G
Distal
Proximal
H
Distal
Proximal

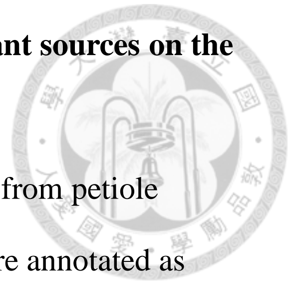
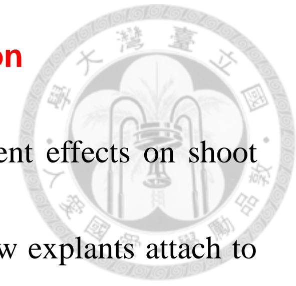


Figure 18 Effect of cytokinin and auxin concentration and explant sources on the morphology of shoot regeneration.

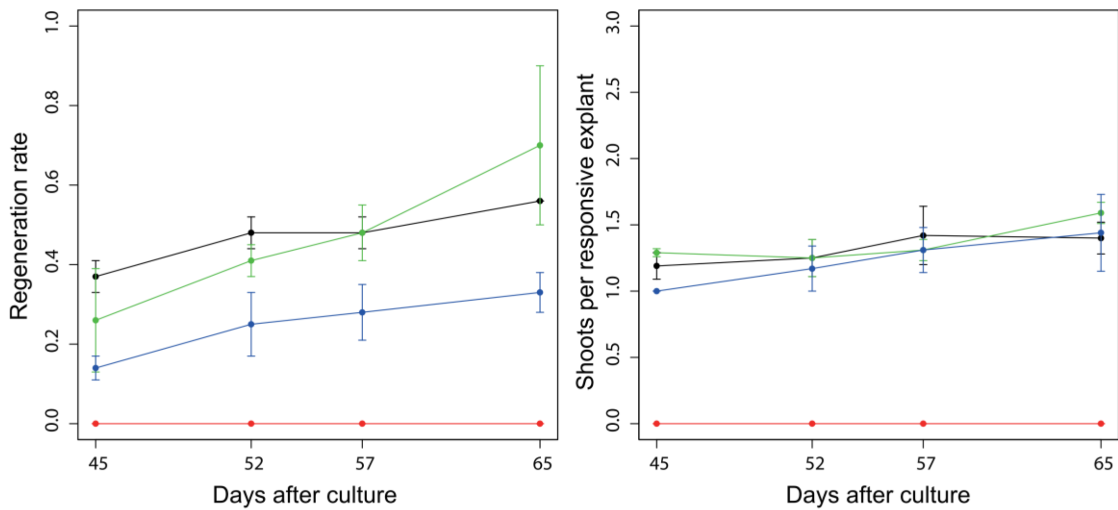
A-D, shoot regeneration from leaf explants. **E-H**, shoot regeneration from petiole explants in horizontal orientation. The direction of petiole explants are annotated as distal and proximal next to the petri dish. **A & E**, 0.1 ppm NAA. **B &F**, 1 ppm BA and 0.1 ppm NAA. **C & G**, 2 ppm BA and 0.1 ppm NAA. **D & H**, 3 ppm BA and 0.1 ppm NAA. All the pictures were taken after 65 days of culture. Scale bar, 3 cm.

Effects of petiole explant orientation on shoot regeneration



In order to confirm whether there are endogenous hormone gradient effects on shoot regeneration of petiole explants, different orientations by which how explants attach to the culture medium was tested. Fig. 19A shows that only the insertion of the distal part (hormone from the medium will be supplied basipetally to petiole) in the medium (black, picture in Fig. 20A) and horizontal (green, picture in Fig. 20B) placement allowed explants to induce new shoots. While the petiole is inserted via proximal site in the medium (now hormone from the medium will be supplied to petiole acropetally, which is against its original hormone movement direction from leave to stem) (red, picture in Fig. 20C), the explants did not induce any shoots. This might be a result of polar transportation within the vascular bundles inside the petiole explants. We further cut the petiole explants into half-length (2.5 mm) and placed the distal site into the medium (blue, picture in Fig. 20D). Surprisingly, we found that the shoots per responsive explant remained approximately the same as the original one (Fig. 19A, right) but the regeneration rate became only half (Fig. 19A, left). This probably indicates that certain size of petiole explants (at least > 2.5mm) was necessary for tissue shoot regeneration.

A



B

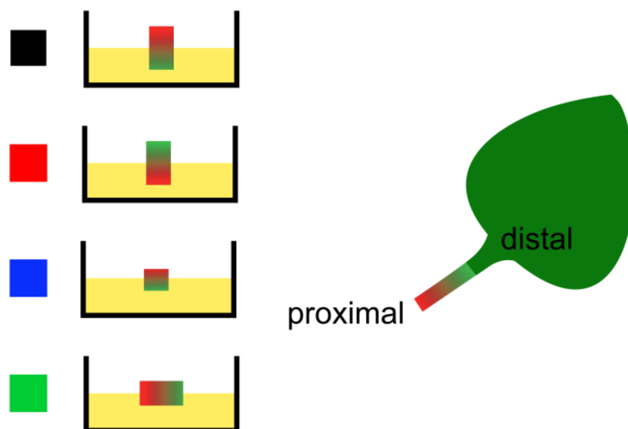


Figure 19 Effects of petiole explant orientations on shoot regeneration rate and shoots per responsive explant.

A, the regeneration rate and shoots per responsive explant from 45 to 65 days after culture. **B**, an illustration showing the corresponding orientations to the colors. **Black**, 5 mm long petiole with the distal end in the medium. **Red**, 5 mm long petiole with the proximal end in the medium. **Blue**, 2.5 mm petiole with the distal end in the medium. **Green**, 5 mm long petiole with the both ends in the medium (horizontal position). The data are presented as mean ± stand error. All treatments have 3 replicates and each replicate has 9 petiole explants (except for 2.5 mm petiole with the distal end in the medium, which has 12 petiole explants).

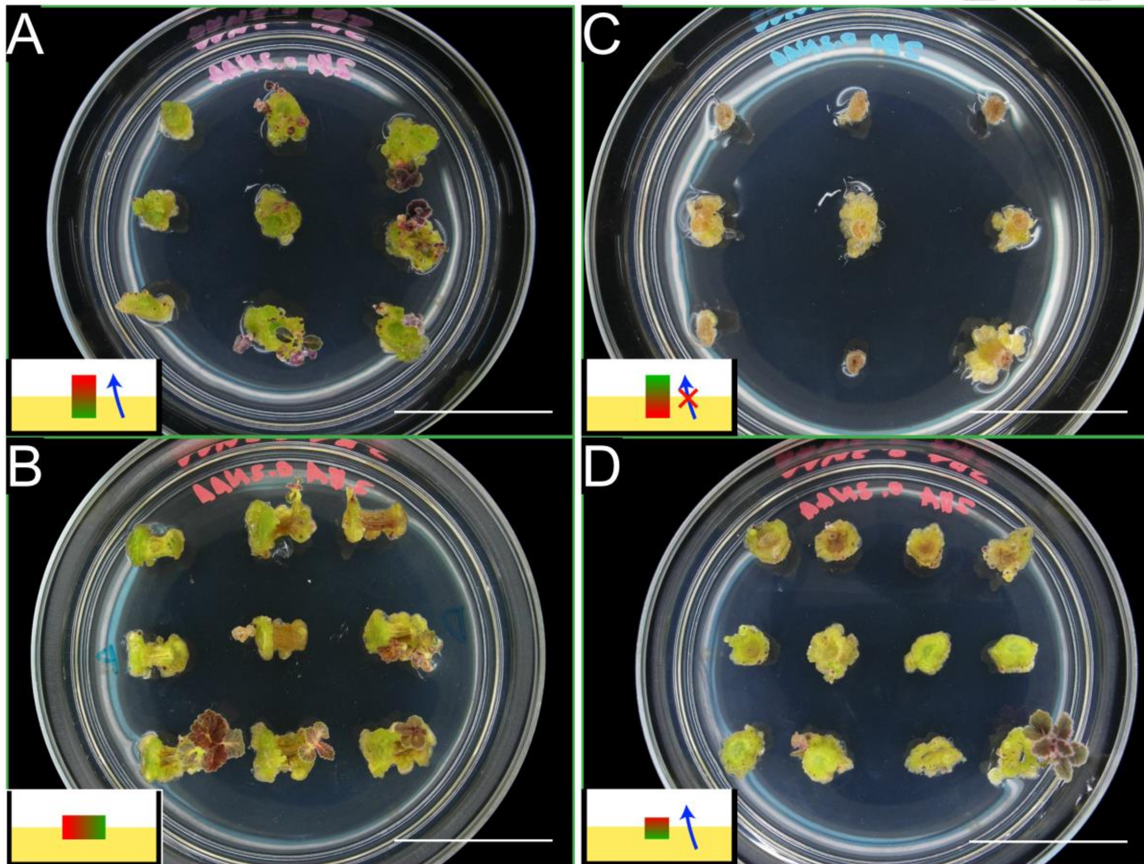
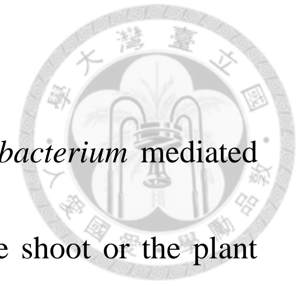


Figure 20 Effects of petiole explant orientations on the morphology of regeneration.

A, 5 mm long petiole with the distal end in the medium. **B**, 5 mm long petiole with the both ends in the medium (horizontal position). **C**, 5 mm long petiole with the proximal end in the medium. **D**, 2.5 mm petiole with the distal end in the medium. **E**, an illustration showing the relative proximal and distal position on a petiole. The blue arrow indicates the native auxin distribution inside a petiole.

Callus induction from embryo



Callus is a cluster of unorganized cells and is often used in *Agrobacterium* mediated genetic transformation. When the plant tissue is difficult to induce shoot or the plant tissue is hardly being infected, callus is a good substitute. After the inoculation, the successfully transformed cells could be selected and grow into a whole plant. In this study, the preliminary result shows that callus of *Sinningia speciosa* could be induced by culturing the seed on 1× MS medium supported with 1 ppm 2, 4-D (Fig. 21D, E, F). However, compared with *Nicotiana benthamiana* (Fig. 21A, B, C), the induced callus of *Sinningia speciosa* was much smaller, suggesting that the source of explant and the hormone conditions needs to be further optimization.

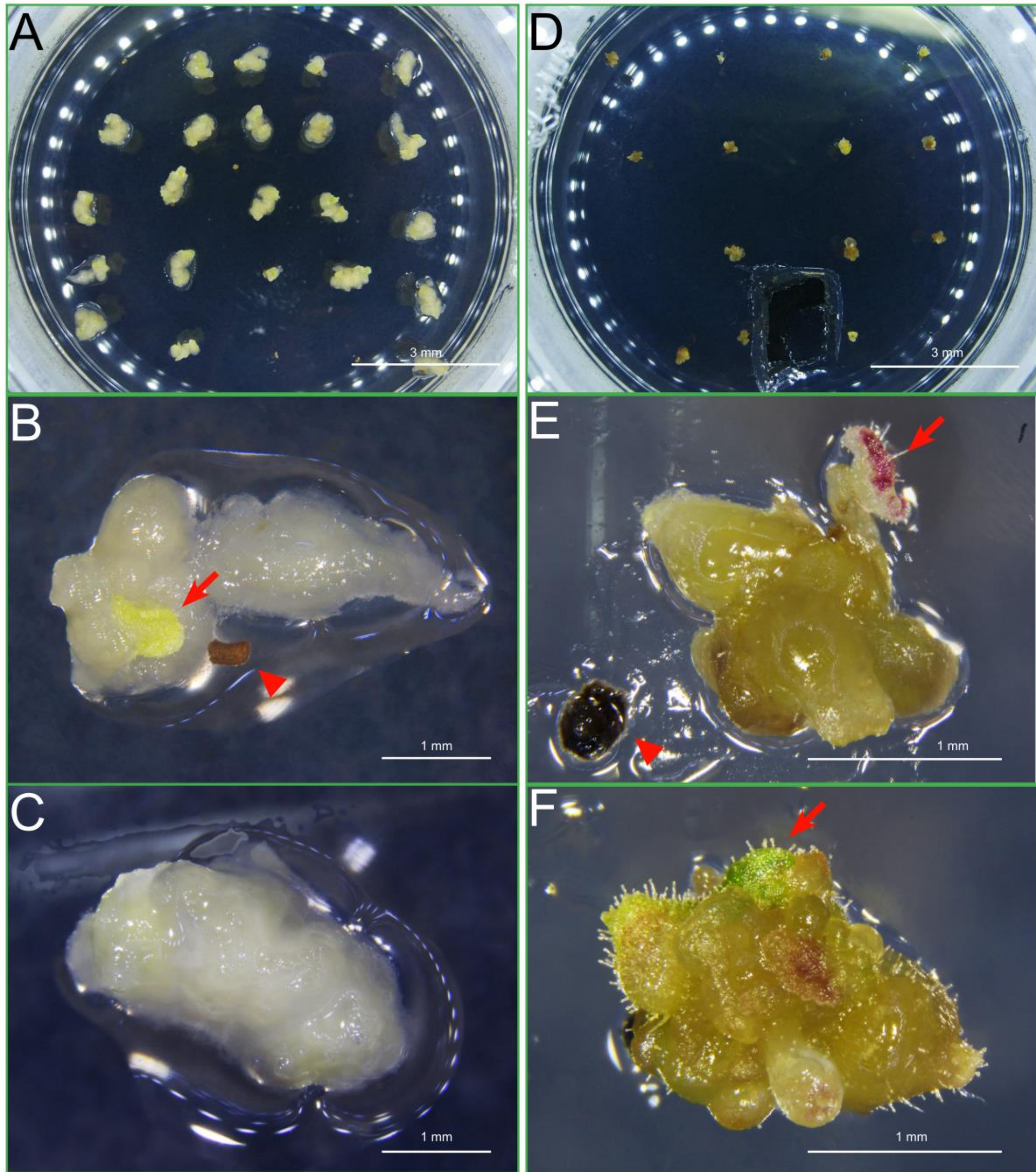
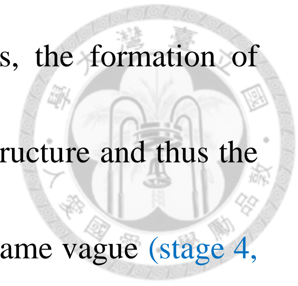


Figure 21 Callus inductions from young seedlings.

A-C, 60 days induction from the seeds of *Nicotiana benthamiana*. **A**, the whole petri dish. **B-C**, detail morphology of the two different induced calli. **D-F**, 60 days induction from the seeds of *Sinningia speciosa*. **D**, the whole petri dish. **E-F**, detail morphology of the two different induced calli. Arrow, the remaining cotyledon. Arrow head, the seed coat.

Histological observation of shoot regeneration

Understanding the developmental process of shoot regeneration is crucial for the efficient explant selection and the genetic transformation. Fig. 22 provides a detailed histological study which shows various shoot regeneration stages. Before the culture initiation, the leaf of *Sinningia speciosa* composed of 2 layers of palisade mesophyll tissue (the red line in Fig. 22A) (the lower one was not fully differentiated), 4-5 layers of sponge mesophyll tissue and a single layer of epidermis tissue at the each side of the leaf (Fig. 22A). After 14 days of culture, the upper epidermis and sponge mesophyll tissue gave rise to several layers of callus-like cells (the red line in Fig. 22B). In addition, especially near the cutting edge of the explants, the vascular tissue had developed into several loosely organized vascular structures (Fig. 22B). Moreover, about 8 meristematic tissues (red arrow, Fig. 22B), characterized with small size and dense cytoplasm, can be seen around the re-differentiated vascular tissues. This meristematic tissue originated from a single phloem parenchyma cell, which underwent continuous cell divisions without cell expansion (stage 1, Fig. 22C). Then, accompanied by repeated cell divisions and expansions, formed a globular shape structure (stage 2, Fig. 22D). Note that this structure was still connected to the vascular tissue within the explant. The globular shape structure continued to differentiate into a spindle shape structure in which the bipolarity was starting to establish, and the meristematic cells



began to form at both polar ends (stage 3, Fig. 22E). After this, the formation of vascular bundle initiated from the outer most layer of the spindle structure and thus the boundary between the newly regenerative shoot and the explant became vague (stage 4, Fig. 22F). Finally, the regenerative shoot protruded out from the explant. As can be seen, the structure of newly regenerative leaves had well differentiated palisade and sponge mesophyll tissues (stage 5, the red arrows in Fig. 22G). By morphological observation, the newly regenerative shoot grew out by breaking the dome-shell structure located at the edge of the leaf explant (Fig. 22H). This study provides the first insight into the detailed developmental process of *Sinningia speciosa*.

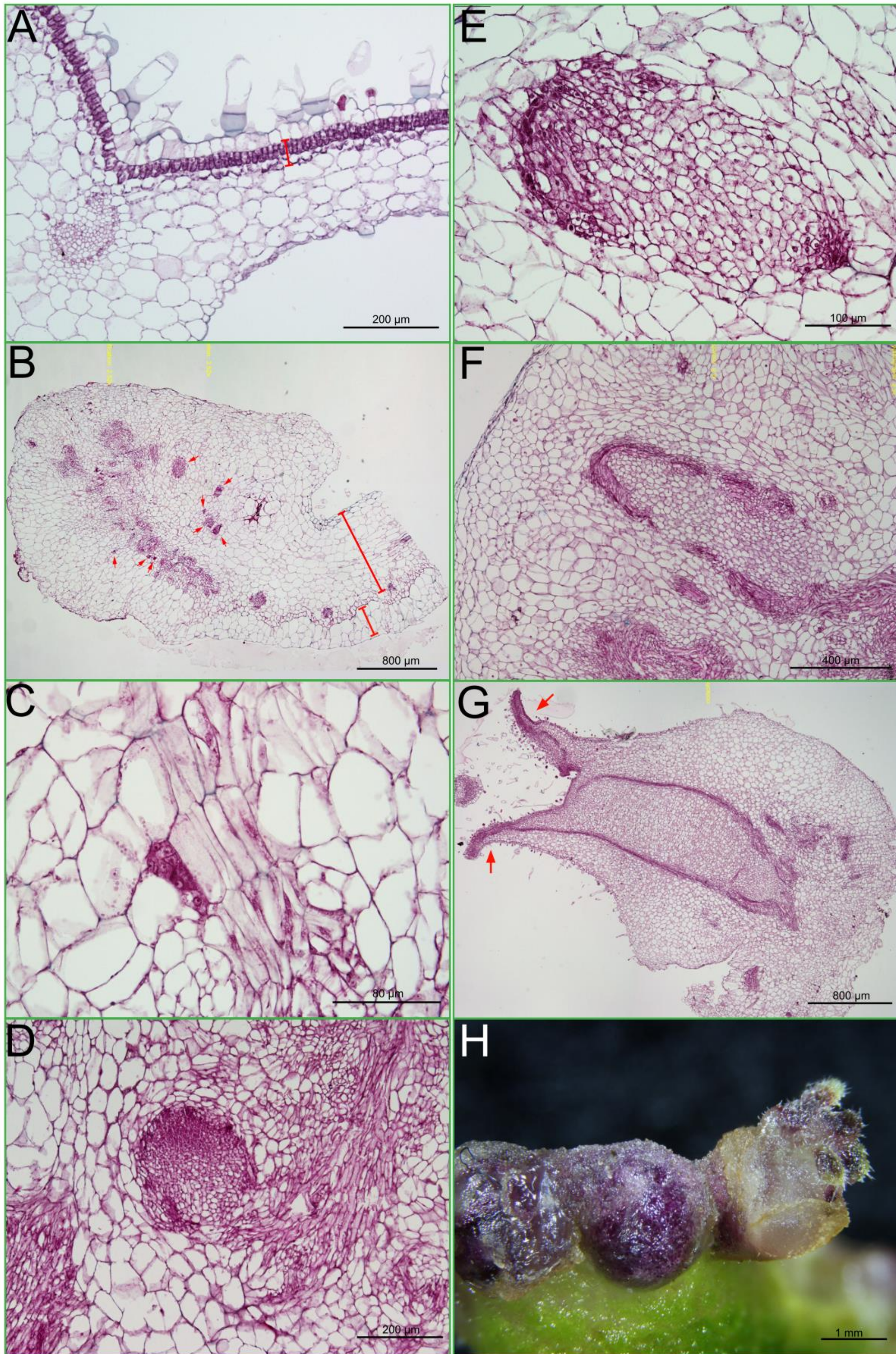
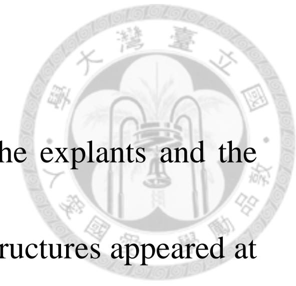

A
200 µm
B
800 µm
C
80 µm
D
200 µm
E
100 µm
F
400 µm
G
800 µm
H
1 mm

Figure 22 Process of shoot regeneration from leaf explants by paraffin section and dissecting microscope.

A-G, histological observation of the process of shoot regeneration. **A**, leaf before culture initiation. **B**, callus formation near the cutting site with several meristematic-like tissues (arrow) scatting around the vascular bundles. **C**, stage 1, continuous cell divisions without cell elongation in a phloem parenchyma cell, characterized by large nuclei and dense cytoplasm. **D**, stage 2, cell division along with cell elongation, resulting a globular embryo-like structure. The cells remain undifferentiated in this stage. **E**, stage 3, formation of the bipolar structure. The cells begin to differentiate. **F**, stage 4, initiation of vascular tissue. The boundary of newly developed embryo-like structure becomes vague in this stage. **G**, stage 5, protruding of the newly developed leaf and the formation of apical meristem. **H**, the morphology of the newly developed shoot (by dissecting microscope). Note the shoot protruding out of dome structure by breaking the epidermis layer of the explant.

Scanning electron microscope observation



Scanning electron microscope shows the micro morphologies of the explants and the newly regenerative shoots. After 20 days of culture, several dome structures appeared at the edge of the leaf explant, especially near the secondary vein (Fig. 23A & B). This could be a result of continuous internal cell divisions and enlargement. Around the 30 days of culture, the dome structures started to break. Thus, the newly regenerative shoots could come out (Fig. 23C). It is worth to note that the meristems in the newly regenerative shoots were able to further develop into multiple adventitious shoots (Fig. 23D). On the other hand, there were also shoots which regenerated from the edge of the leaf explants without forming the dome structure (Fig 23E). On the newly regenerative shoots, several stomata and trichomes were observed (Fig 23F).

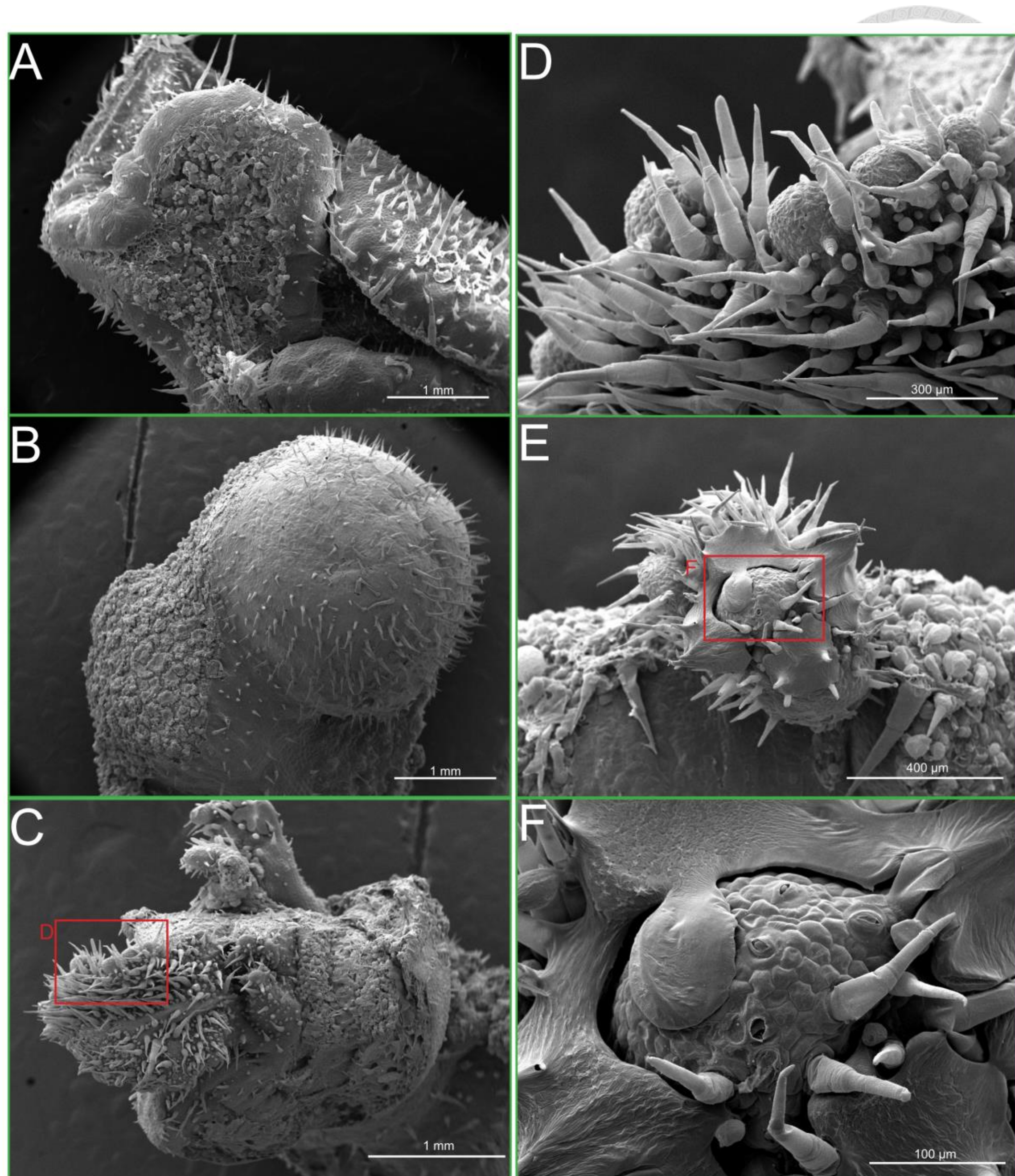
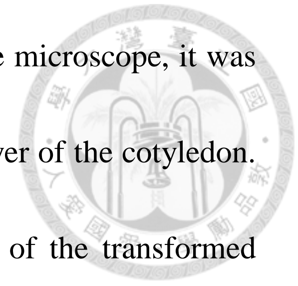


Figure 23 Micro-scale observation of shoot regeneration by scanning electronic microscope.

A, leaf a week after culture initiation. Note the callus formation extensively at the site of secondary vein. **B-F**, leaf 5 week after culture initiation. **B**, hemi-spherical tumor-like structure rising from the site next to the secondary vein. **C**, regenerative shoots appearing by breaking the surface of hemi-spherical tumor-like structure. **D**, enlargement of “C”. Note several primordium-like structures appearing on the edge of the newly formed leaf. **E**, regenerative shoots appearing at the cutting site of the explant. Note the remaining epidermis layer of the explant. **F**, enlargement of “E”. Note the stomata on the newly formed leaf.

Testing the optimal explants source for *Agrobacterium*-mediated transformation

To generate a reliable transformation system, the high competence of tissue for *Agrobacterium* infection is extremely important. In this study, GUS reporter system was used to examine the transformation efficiency of different *Sinningia speciosa* tissues. Young leaves, petioles and 14-day-old seedlings with fully opened cotyledons were inoculated with either *Agrobacterium* strain 'GV3101::pMP90' or 'EHA105' harboring pCambia2301 vector. After 4 days of co-culture, the GUS activity was detected by a chromogenic method. Successfully transformed tissues could produce GUS protein, which changed the transparent substrate, like x-gluc, into blue a precipitate. The positive signals could only be seen in the 14-day-old seedlings but not in both young leaf and petiole explants. For inoculation with *Agrobacterium* strain 'GV3101::pMP90', 14 out of 28 seedlings showed GUS activities. For inoculation with *Agrobacterium* strain 'EHA105', 24 out of 28 seedlings showed GUS activities, showing that EHA105 might have higher virulence for *Sinningia speciosa* compared to strain GV3101. Microscopic observation (Fig. 24A & B) showed that the successfully transformed cells distributed in basal parts of both the cotyledon and the first pair of primary leaves, especially around the wounding sites. Perhaps the wound facilitated the *Agrobacterium* infection in the tissues. However, the tissues without wounds could also be successfully



transformed (Fig. 24C & D). By adjusting the focusing plane of the microscope, it was found that the transformed tissues were located at the mesophyll layer of the cotyledon. Furthermore, histological sections provide more precise locations of the transformed cells. Fig. 24E, F & G show that the positive signals of transformation were located in the mesophyll cells and the glandular trichome head cell. Fig. 24H shows that the signals were also located in the epidermis cells. By understanding the types of tissues which are susceptible of being transformed and the most suitable *Agrobacterium* strains can provide a great help in improving the genetic transformation system of *Sinningia speciosa* in the future.

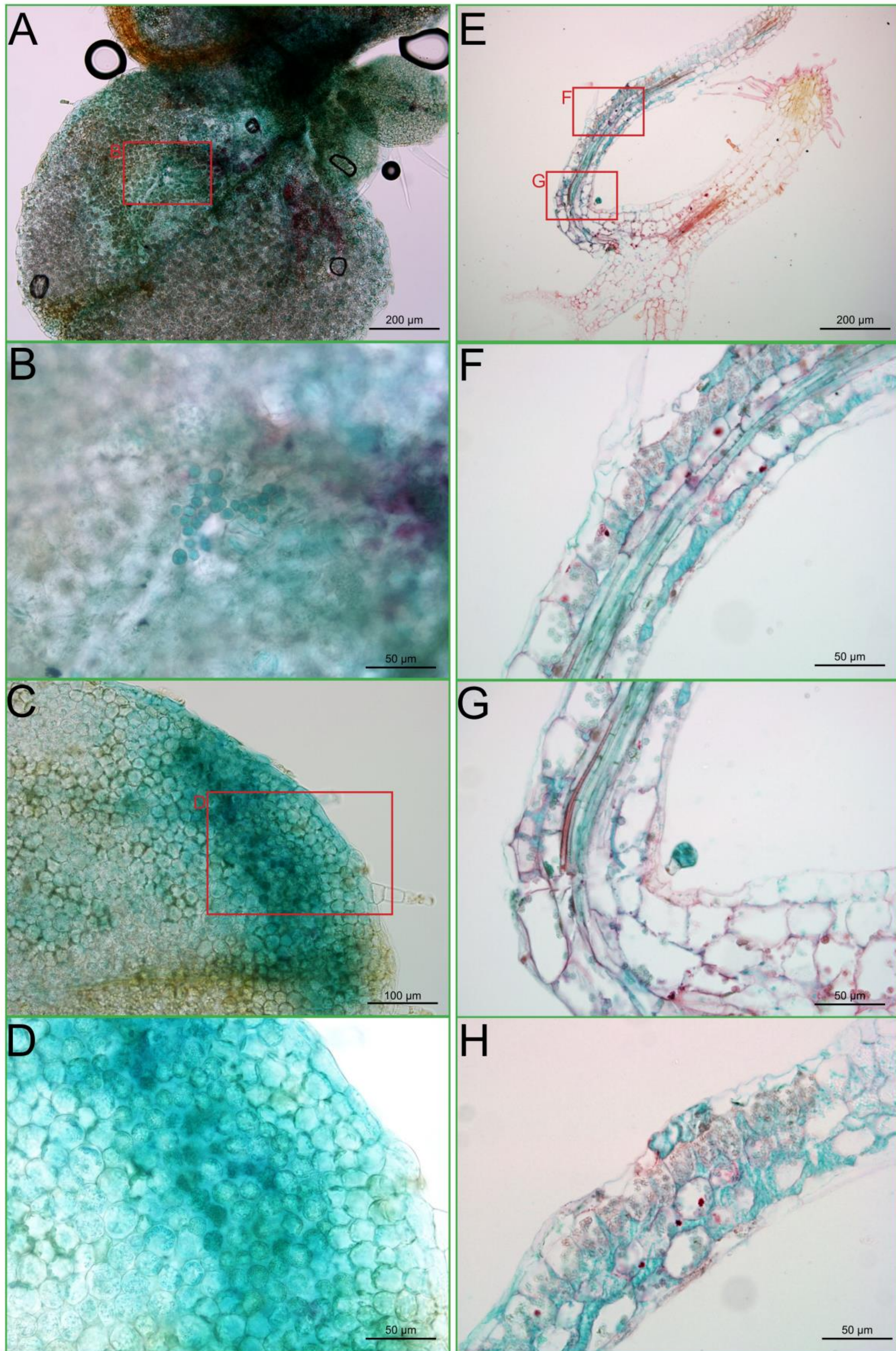
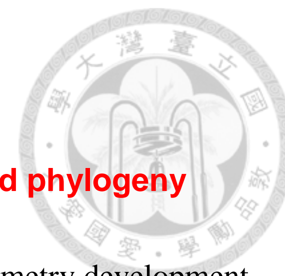

A
B
200 µm
E
F
G
200 µm
B
50 µm
F
50 µm
C
D
100 µm
G
50 µm
D
50 µm
H
50 µm

Figure 24 Tissue competence for *Agrobacterium* infection by GUS transient assay. **A-D,** whole mount of transformed seedlings. **E-H**, paraffin section of transformed seedlings. **A**, blue signals on a 14 day-old seedling. **B**, the enlargement of "A". Note that there are several transformed cells around a wound, which is made by forceps during inoculation. **C**, blue signals near the edge of the cotyledon. **D**, enlargement of "D". Note the successfully transformed mesophyll cells. **E**, blue signals on a cotyledon. **F-G**, the enlargement of "E". Note that the blue signals distribute in epidermis, mesophyll, vascular tissues and a glandular trichome head cell. **H**, the cotyledon with multiple transformed regions. This work was cooperated with Wu, Ho-Wei.

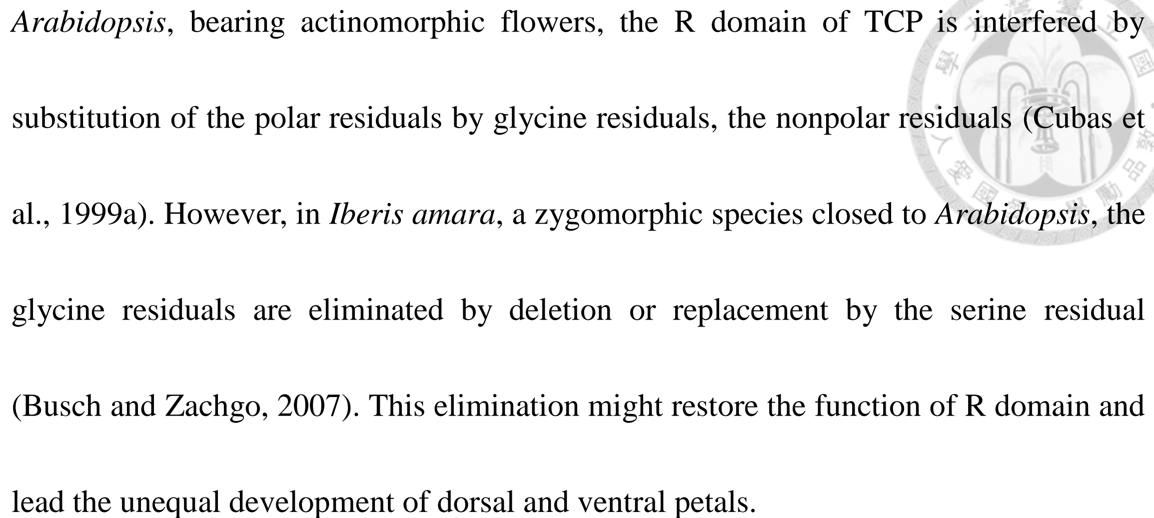
Discussion



SsCYC is a homologue to *AmCYC* by sharing domains and phylogeny

AmCYC, also known as *CYCLOIDEA*, plays a key role in floral symmetry development. After the first isolation of this gene, its orthologs have been found in several plant lineages and some of them have been proved functionally. In this study, we aimed to confirm the homology between *CYCLOIDEA* and putative *SsCYC* genes, which was previously isolated by Citerne et al. (2000). In the wild type cultivar 'Carangola', a sequence of 1038 bp with an open reading frame was isolated. The encoded *SsCYC* protein shares conserved domains and motifs with other *CYC*-like genes, indicating that *SsCYC* protein might target to nuclear and participate in regulation of transcription. In addition, by phylogenetic analysis of the whole coding nucleotides, the phylogenetic relationship of *SsCYC* gene is grouped with other ECE-CYC2 clade genes with high branch support values. Moreover, Southern blotting result indicates that there is only one copy of *CYC*-like gene in *Sinningia speciosa*, consisting the previous hypothesis that single copy of *CYC*-like gene in new world Gesneriaceae species (Citerne et al., 2000). These results support that *SsCYC* gene is the homologue to *AmCYC* gene and probably the only ortholog in *Sinningia speciosa*.

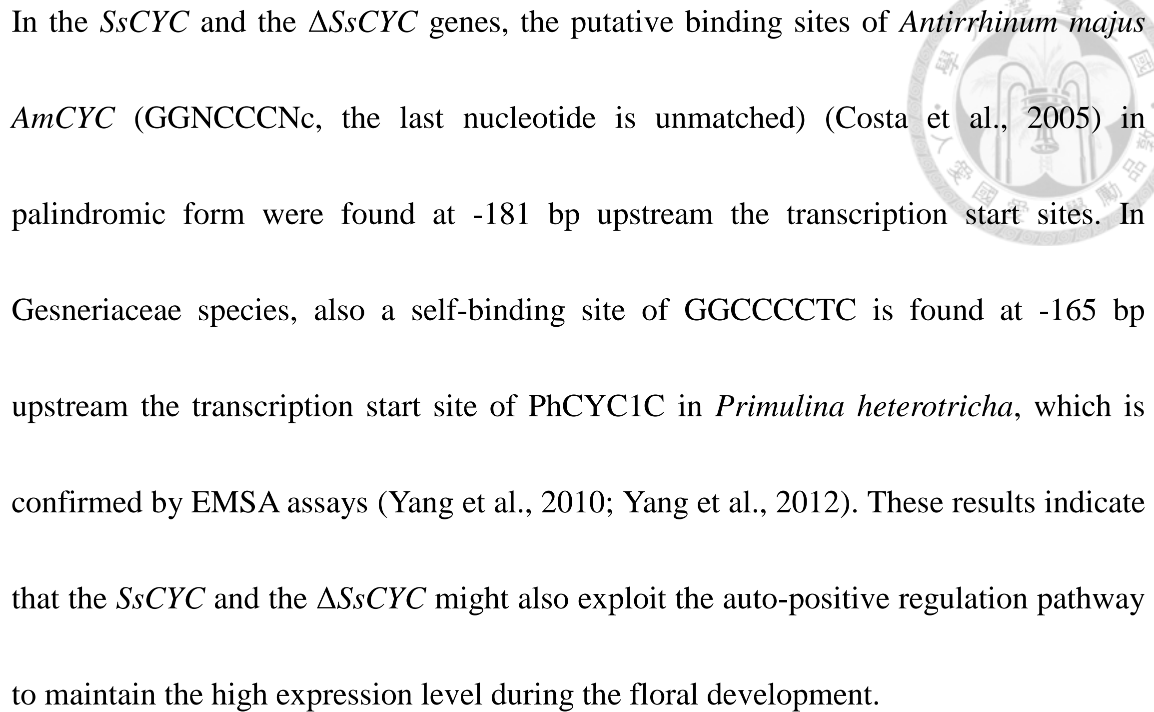
On the other hand, by comparing the R domain of IaTCP1 and *Arabidopsis* TCP1, R domain seems having a role in regulation of floral symmetry development. In



Arabidopsis, bearing actinomorphic flowers, the R domain of TCP is interfered by substitution of the polar residuals by glycine residuals, the nonpolar residuals (Cubas et al., 1999a). However, in *Iberis amara*, a zygomorphic species closed to *Arabidopsis*, the glycine residuals are eliminated by deletion or replacement by the serine residual (Busch and Zachgo, 2007). This elimination might restore the function of R domain and lead the unequal development of dorsal and ventral petals.

Δ*SsCYC* appears to lose its function due to a small fragment deletion

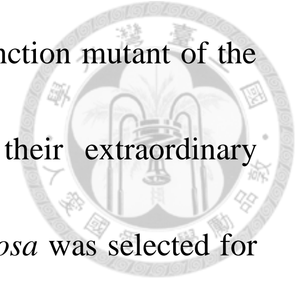
In this study, candidate gene approach was used to examine whether the differences between *SsCYC* and Δ*SsCYC* genes are the main reasons for floral symmetry reversion. At the gene regulation level, the upstream sequences about 1.7 kb were sent to online TSSP tool and compared with RedSite Database. All types of predicted transcription factor binding sites are shared among the *SsCYC* and the *ΔSsCYC*, suggesting that the *SsCYC* and the Δ*SsCYC* have no major difference of sequences in regulatory regions. Interestingly, among the putative transcription factor binding sites, one of those is targeted by the AGAMOUS transcription factor. The *AGAMOUS* gene belongs to the C class floral organ identity genes, which participates in the development of stamens and carpels in *Arabidopsis* (Bowman et al., 1991). This result suggests that the *SsCYC* and the Δ*SsCYC* genes might be regulated by the genes determining the floral organ identities although further studies are needed.



In the *SsCYC* and the Δ*SsCYC* genes, the putative binding sites of *Antirrhinum majus AmCYC* (GGNCCCNc, the last nucleotide is unmatched) (Costa et al., 2005) in palindromic form were found at -181 bp upstream the transcription start sites. In Gesneriaceae species, also a self-binding site of GGCCCCTC is found at -165 bp upstream the transcription start site of PhCYC1C in *Primulina heterotricha*, which is confirmed by EMSA assays (Yang et al., 2010; Yang et al., 2012). These results indicate that the *SsCYC* and the Δ*SsCYC* might also exploit the auto-positive regulation pathway to maintain the high expression level during the floral development.

Δ*SsCYC* gene is selected by human for its derived actinomorphy

To examine whether the loss of function of the Δ*SsCYC* leads to the floral symmetry reversion in horticultural actinomorphic cultivars, *ΔSsCYC* was overexpressed in Tobacco in this study. The results show that there is no visible phenotypic difference between the wild type and Δ*SsCYC* transgenic plants, while phenotypes like dwarfish plant, curled leaves, promotion of the growth of axillary buds, small flower and sterility were evident in the transgenic plants. *SsCYC* resulting phenotypes were also reported in the transgenic plants of other *CYC* homologues overexpression (Broholm et al., 2008; Busch and Zachgo, 2007; Costa et al., 2005; Liu et al., 2014a; Yang et al., 2012), suggesting that the *SsCYC* is very likely to play a key role in the regulation of floral symmetry development as a result of human selection. Similar to the famous sunflowers



in Van Gogh's painting, the gain of function mutant and loss of function mutant of the *CYC* homologue were selected by human unconsciously for their extraordinary inflorescence traits (Chapman et al., 2012). Peloric *Sinningia speciosa* was selected for large actinomorphic and showy flowers.

SsCYC represses the primary growth but promote the development of axillary buds

In the *SsCYC* transgenic plants with abnormal phenotypes, the plants showed dwarfish phenotypes, in which the plants were shorter than the wild type and Δ*SsCYC* transgenic plants. In addition, the *SsCYC* transgenic plants had retarded growth of the leaves, which were smaller and curled toward the abaxial sides. However, the *SsCYC* transgenic plants had more developed axillary buds than the wild and the Δ*SsCYC* transgenic plants, showing that the *SsCYC* could reduce the apical dominance and promote the growth of axillary buds. These effects are similar to the cases of overexpression the *Antirrhinum majus AmCYC*, *Iberis amara IaTCP1* and *Primulina heterotricha PhCYC1C* in Arabidopsis plants, showing dwarfish plants, numerous secondary shoots and retarded growth of leaves (Busch and Zachgo, 2007; Costa et al., 2005; Liu et al., 2014b; Yang et al., 2012). This similarity suggests the functional conservation between the *SsCYC* and other homologues. However, *TCP1* gene enhanced expression by activation tag in *Arabidopsis* showed elongated leaf phenotype (Choe et al., 2001),

which is different from the effect of the *SsCYC*, suggesting that there are some functional differences between the *SsCYC* and the *TCP1*. These functional differences coincide with the substitution of polar residuals into non polar ones of the R domain within the TCP1 protein, which are important for protein-protein interaction (Cubas et al., 1999a).

Furthermore, TCP1 has been reported as a positive regulator of *DWARF4* (*DEF4*) gene, encoding a key enzyme in brassinosteroid biosynthesis of which the *def4* mutation causes severe dwarfish plants (Guo et al., 2010). Overexpression of the *DEF4* in the *Arabidopsis* causes higher inflorescences, additional secondary shoots and more seed production compared to wild type (Choe et al., 2001). Although the *SsCYC* transgenic plants show dwarfish phenotype which is opposite to the case of *DEF4* overexpression, *SsCYC* transgenic plants do have the similar phenotype of additional secondary shoots. This indicates that the abnormal phenotypes of the *SsCYC* transgenic plants are perhaps related to changes of endogenous brassinosteroid level and distribution.

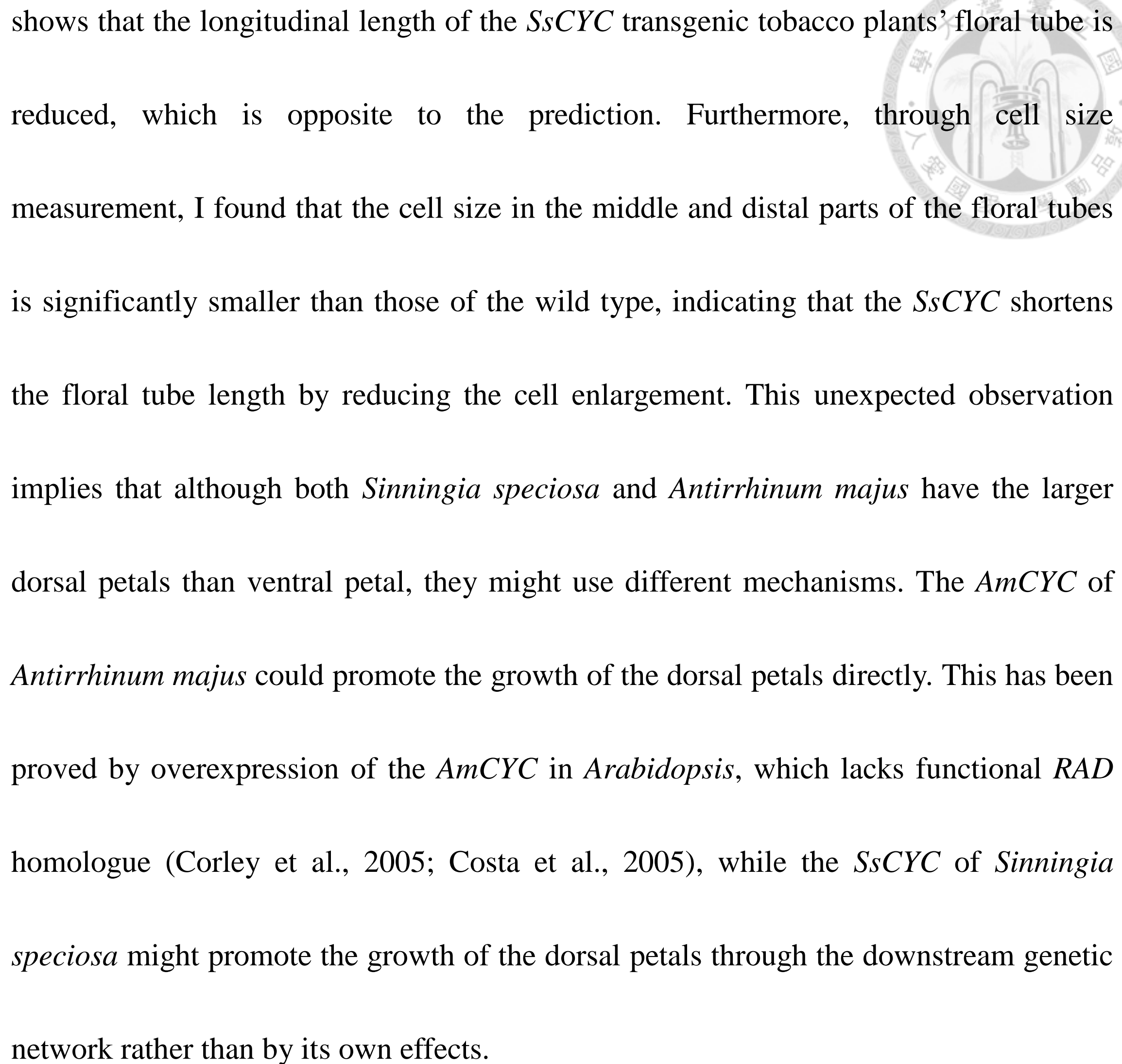
On the other hand, for promoting the growth of the secondary shoots, *SsCYC* might have an opposite effect to the *TB1* gene, the *CYC* homologue in *Zea mays*. In *tb1* mutant, numerous secondary shoots with the male flowers on the top can be seen (Doebley et al., 1997). Rather, this might be a competitive antagonistic effect of that SsCYC proteins compete against the endogenous *TB1*-like gene in *Nicotiana benthamiana* for the

dimerization partners or downstream trans-regulated targets.

Finally, the phenotype of curled leaves toward the abaxial side is not yet reported in the previous functional studies of ECE-CYC2 genes, suggesting that *SsCYC* might have recruited a new function leading the unequal growth of the *Nicotiana benthamiana* leaves. This is reminiscent of the curled leaf in *cin* mutant (Crawford et al., 2004). The CIN gene belongs to the sister clade of ECE genes. Alternatively, it is might be a side effect of the species in transformation system, which is not *Arabidopsis*.

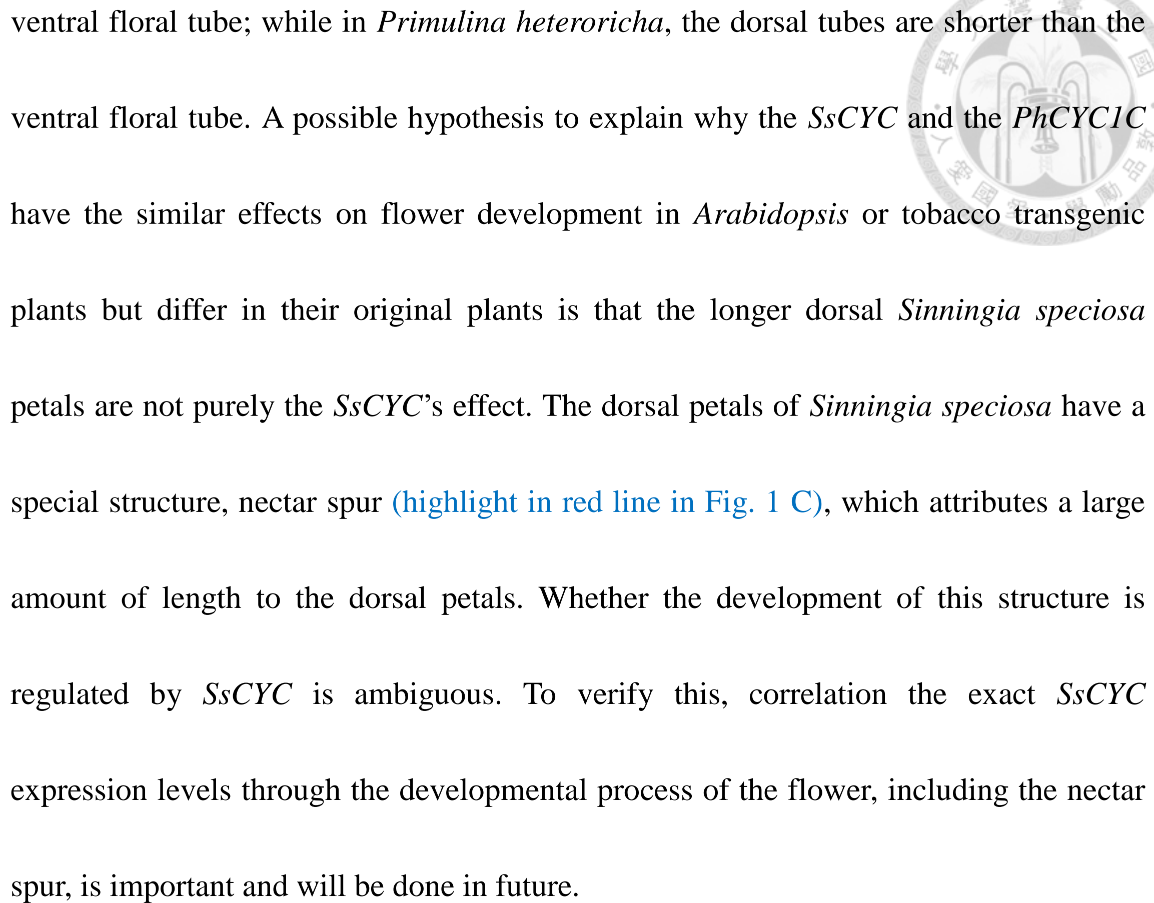
The downstream pathway of SsCYC is different from AmCYC

The native zygomorphic cultivar of *Sinningia speciosa* has the dorsal floral tube longer than the ventral floral tube, thus leading to a nodding position of the floral opening which can unify the incoming direction of the pollinators and enhance the pollination efficiency. By measuring the cell length of the dorsal and the ventral petals' epidermis cells in the native zygomorphic cultivar 'Carangola' shows that the cell length of the dorsal petal is about 1.5 times longer than that of the ventral petal (Hsu, Hao-Chun, unpublished data), suggesting that the cell enlargement rather than the cell proliferation is the main reason of zygomorphy development in *Sinningia speciosa*. Therefore, it is reasonable to predict that *SsCYC* transgenic plants should have the longer floral tubes, caused by the larger cell size, which is similar to the case of overexpression of the *Antirrhinum majus AmCYC* in *Arabidopsis* (Costa et al., 2005). However, this study



shows that the longitudinal length of the *SsCYC* transgenic tobacco plants' floral tube is reduced, which is opposite to the prediction. Furthermore, through cell size measurement, I found that the cell size in the middle and distal parts of the floral tubes is significantly smaller than those of the wild type, indicating that the *SsCYC* shortens the floral tube length by reducing the cell enlargement. This unexpected observation implies that although both *Sinningia speciosa* and *Antirrhinum majus* have the larger dorsal petals than ventral petal, they might use different mechanisms. The *AmCYC* of *Antirrhinum majus* could promote the growth of the dorsal petals directly. This has been proved by overexpression of the *AmCYC* in *Arabidopsis*, which lacks functional *RAD* homologue (Corley et al., 2005; Costa et al., 2005), while the *SsCYC* of *Sinningia speciosa* might promote the growth of the dorsal petals through the downstream genetic network rather than by its own effects.

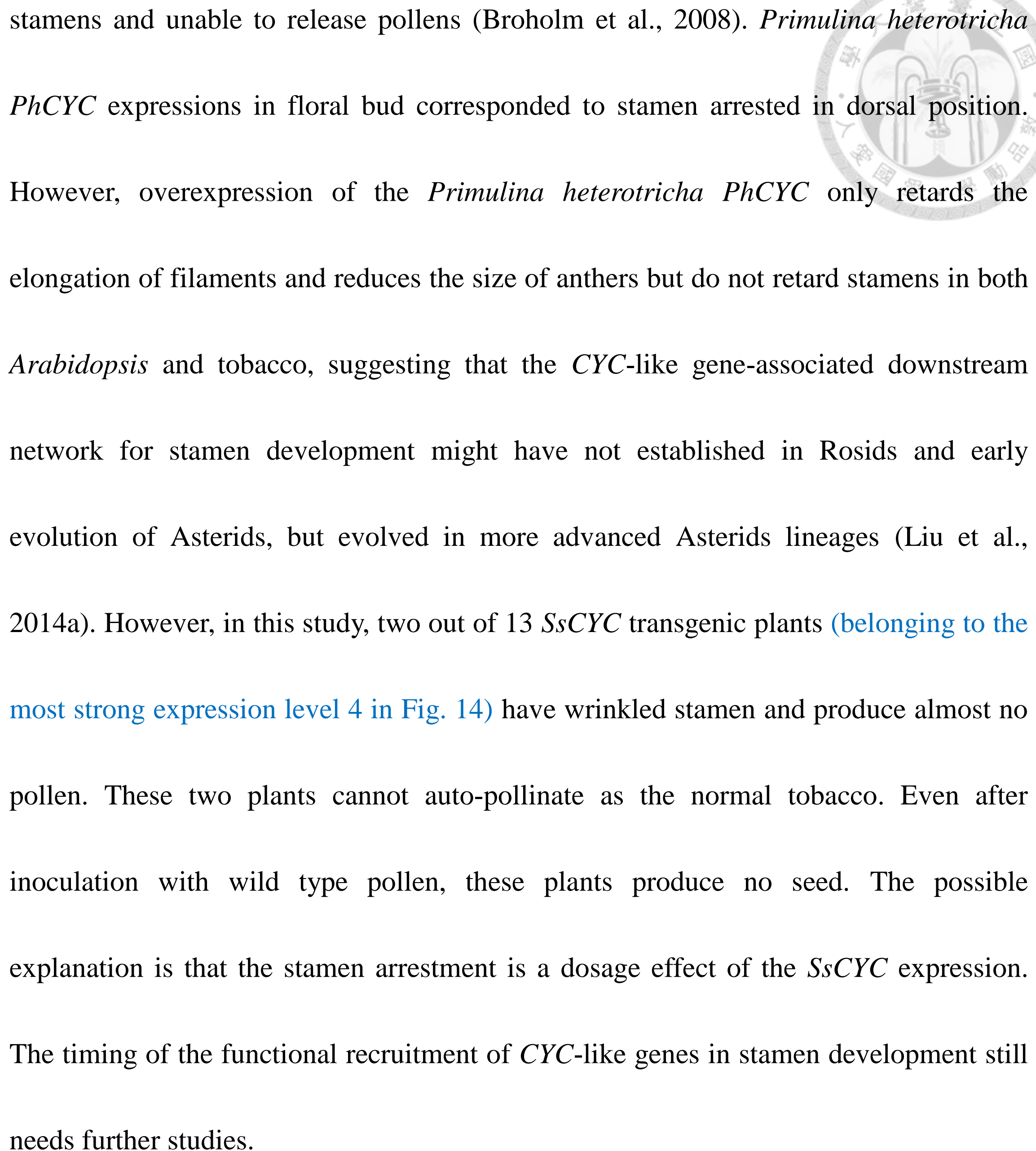
This difference can be further elucidated by comparing the functional analysis of the *PhCYC1C* belonging to *Primulina heterotricha* (Gesneriaceae), a closer relative species to *Sinningia speciosa* than *Antirrhinum majus*. Overexpression of the *PhCYC1C* in *Arabidopsis* and tobacco also reduces the petal area by reducing the cell enlargement (Liu et al., 2014a; Yang et al., 2012). However, this functional conservation raises a contradiction to explain the floral morphology between *Sinningia speciosa* and *Primulina heterotricha*. In *Sinningia speciosa*, the dorsal floral tubes are longer than the



ventral floral tube; while in *Primulina heteroricha*, the dorsal tubes are shorter than the ventral floral tube. A possible hypothesis to explain why the *SsCYC* and the *PhCYC1C* have the similar effects on flower development in *Arabidopsis* or tobacco transgenic plants but differ in their original plants is that the longer dorsal *Sinningia speciosa* petals are not purely the *SsCYC*'s effect. The dorsal petals of *Sinningia speciosa* have a special structure, nectar spur (highlight in red line in Fig. 1 C), which attributes a large amount of length to the dorsal petals. Whether the development of this structure is regulated by *SsCYC* is ambiguous. To verify this, correlation the exact *SsCYC* expression levels through the developmental process of the flower, including the nectar spur, is important and will be done in future.

SsCYC might cause male and female sterility in the plants with severe phenotypes

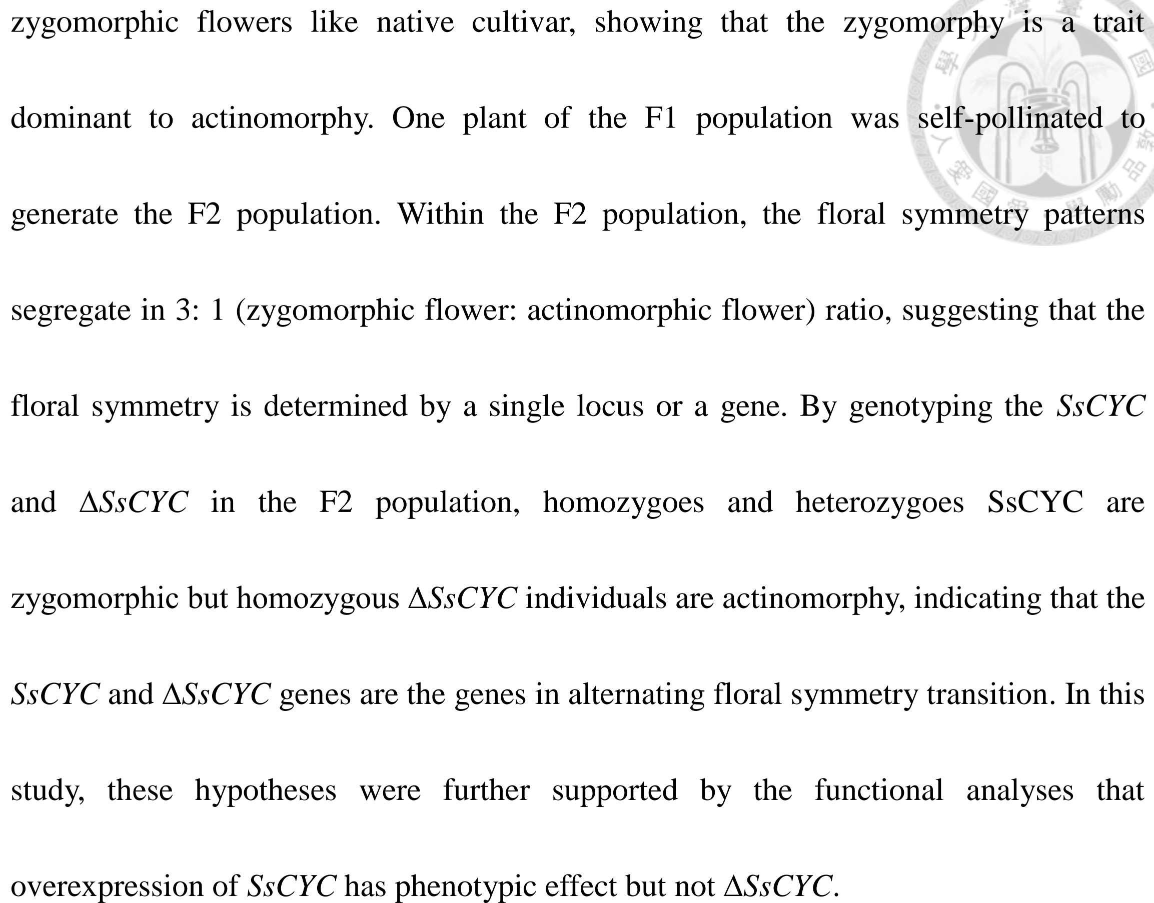
In the native zygomorphic cultivar of *Sinningia speciosa*, the dorsal specific expression of the *SsCYC* gene is concordance to the abortion of the dorsal stamen, which is similar to the dorsal stamen abortion in *Antirrhinum majus* (Carpenter et al., 1999). In addition, in the horticultural actinomorphic cultivar of *Sinningia speciosa*, all five or six stamens are fertile, corresponded to the loss of function of the Δ*SsCYC*, reinforcing the relation of *SsCYC* and the stamen arrestment. This relation can also be seen in *Gerbera* hybrid in which that overexpression of the *GhCYC2* in *Gerbera* hybrid causes the brownish



stamens and unable to release pollens (Broholm et al., 2008). *Primulina heterotricha PhCYC* expressions in floral bud corresponded to stamen arrested in dorsal position. However, overexpression of the *Primulina heterotricha PhCYC* only retards the elongation of filaments and reduces the size of anthers but do not retard stamens in both *Arabidopsis* and tobacco, suggesting that the *CYC*-like gene-associated downstream network for stamen development might have not established in Rosids and early evolution of Asterids, but evolved in more advanced Asterids lineages (Liu et al., 2014a). However, in this study, two out of 13 *SsCYC* transgenic plants (belonging to the most strong expression level 4 in Fig. 14) have wrinkled stamen and produce almost no pollen. These two plants cannot auto-pollinate as the normal tobacco. Even after inoculation with wild type pollen, these plants produce no seed. The possible explanation is that the stamen arrestment is a dosage effect of the *SsCYC* expression. The timing of the functional recruitment of *CYC*-like genes in stamen development still needs further studies.

The results of functional analyses coincide with the genetic approach

To understand the genetic mechanism behind the development of zygomorphy, in which the different growth rate of dorsal and ventral petals are critical, the intercross line of the native zygomorphic cultivar with the horticultural actinomorphic cultivar has been done (Hsu, Hao-Chun, unpublished data). In the F1 population, all the descendants bore

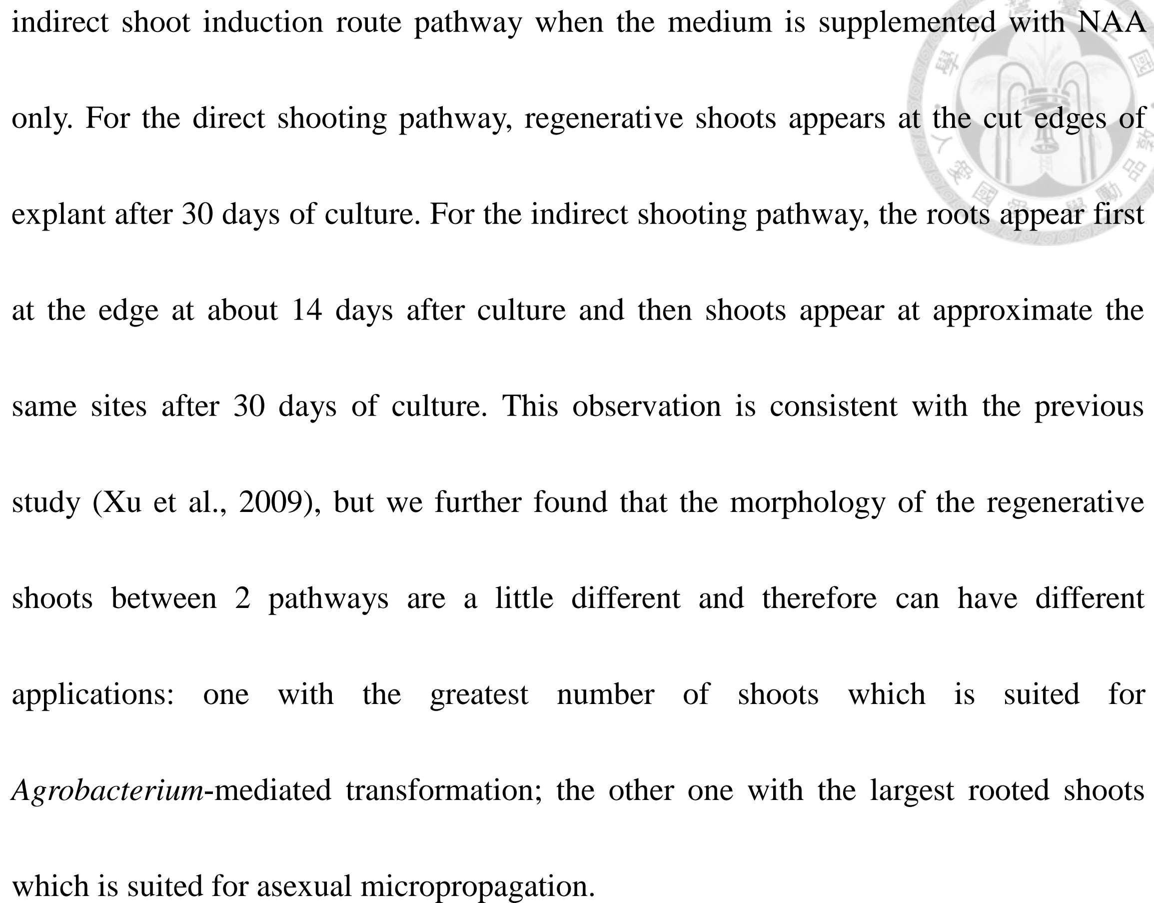


zygomorphic flowers like native cultivar, showing that the zygomorphy is a trait dominant to actinomorphy. One plant of the F1 population was self-pollinated to generate the F2 population. Within the F2 population, the floral symmetry patterns segregate in 3: 1 (zygomorphic flower: actinomorphic flower) ratio, suggesting that the floral symmetry is determined by a single locus or a gene. By genotyping the *SsCYC* and Δ*SsCYC* in the F2 population, homozygoes and heterozygoes SsCYC are zygomorphic but homozygous Δ*SsCYC* individuals are actinomorphy, indicating that the *SsCYC* and Δ*SsCYC* genes are the genes in alternating floral symmetry transition. In this study, these hypotheses were further supported by the functional analyses that overexpression of *SsCYC* has phenotypic effect but not Δ*SsCYC*.

Shoot regeneration has two hormone dependent pathways

Tissue culture technique and *in vitro* shoot induction conditions are important for either asexually manipulation of horticultural traits or establishment of a genetic transformation system. This study aims to provide a highly efficient shoot regeneration protocol with clearly defined explant size and duration after culture initiation. Thus, the results of study can be comparable with other studies and can be replicated by other researchers.

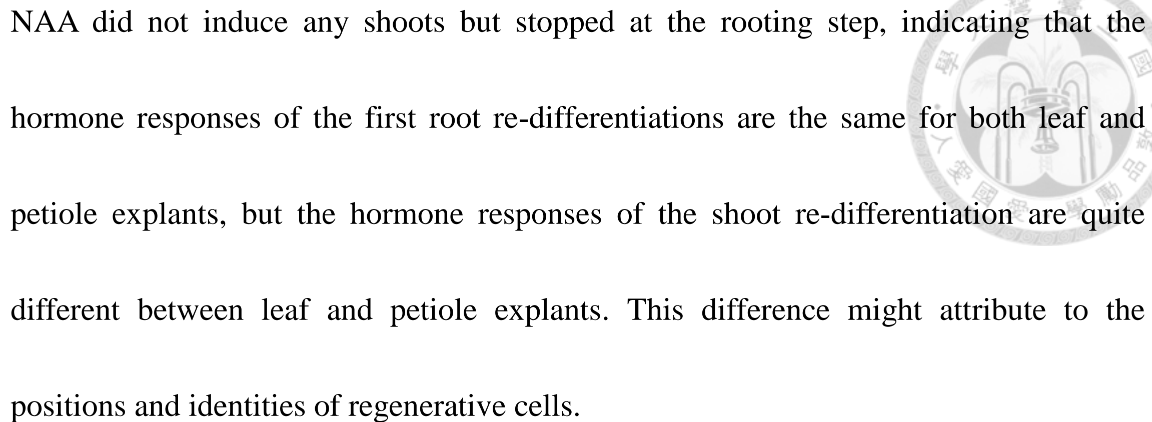
We find that there are 2 different routes of hormone induced shoot regeneration: direct shoot induction route when the medium is supplemented with both NAA and BA and



indirect shoot induction route pathway when the medium is supplemented with NAA only. For the direct shooting pathway, regenerative shoots appears at the cut edges of explant after 30 days of culture. For the indirect shooting pathway, the roots appear first at the edge at about 14 days after culture and then shoots appear at approximate the same sites after 30 days of culture. This observation is consistent with the previous study (Xu et al., 2009), but we further found that the morphology of the regenerative shoots between 2 pathways are a little different and therefore can have different applications: one with the greatest number of shoots which is suited for *Agrobacterium*-mediated transformation; the other one with the largest rooted shoots which is suited for asexual micropropagation.

For both leaf and petiole explants, we find that the medium supported with 0.1 ppm NAA and 1 ppm BA had the best regeneration rate which with the highest number of shoots from each explant. This condition can be applied to the shoot induction step of *Agrobacterium*-mediated transformation of *Sinningia speciosa* in which large numbers of independent transgenic lines are desired.

However, for leaf explants, the largest shoots with normal morphology were obtained in the treatment of 0.1 ppm NAA and 0 ppm BA. Although the treatment does not have the highest regeneration rate, this condition can be applied in maintenance of special horticultural strains; while for petiole explants, medium supported with only 0.1 ppm

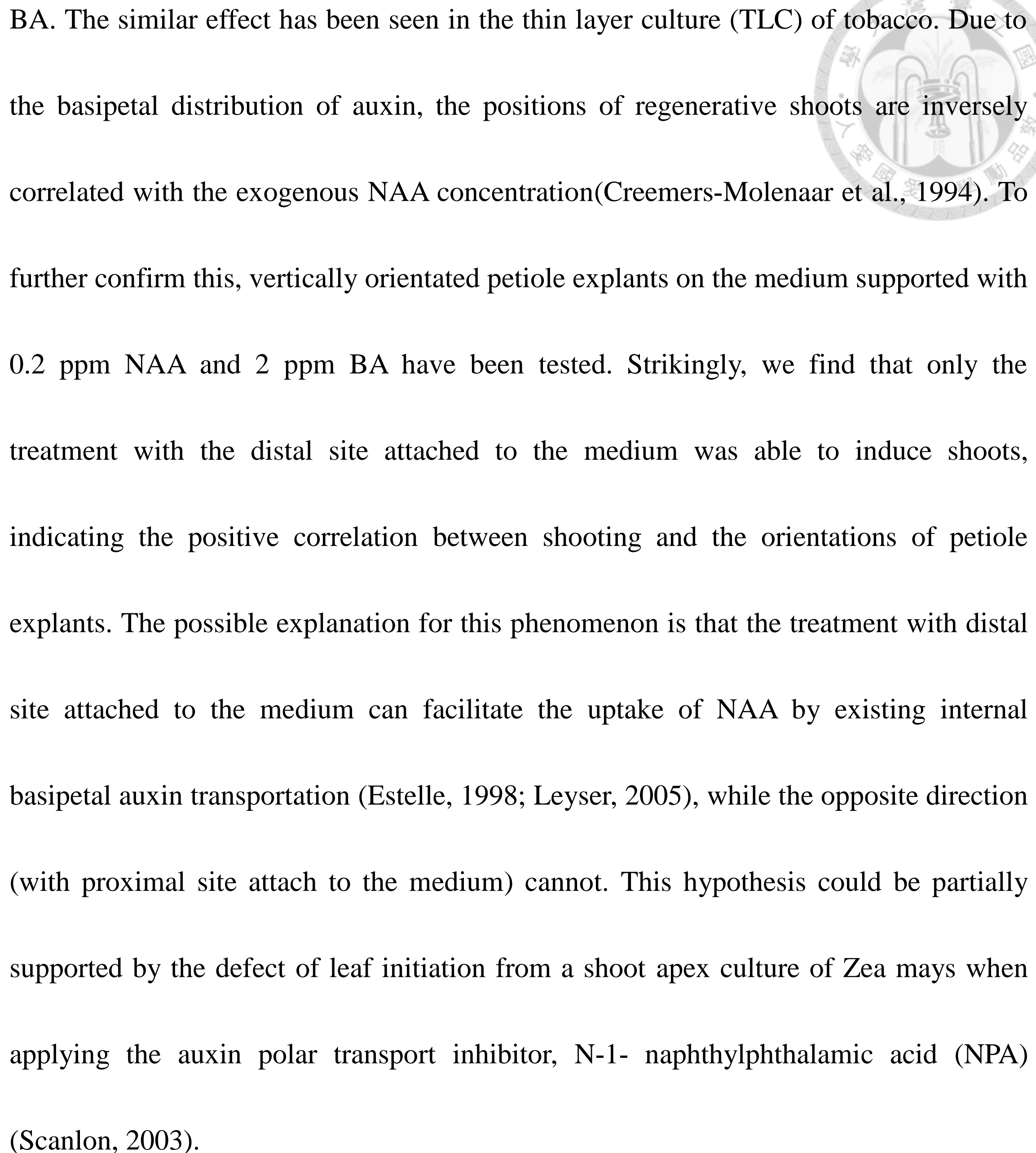


NAA did not induce any shoots but stopped at the rooting step, indicating that the hormone responses of the first root re-differentiations are the same for both leaf and petiole explants, but the hormone responses of the shoot re-differentiation are quite different between leaf and petiole explants. This difference might attribute to the positions and identities of regenerative cells.

Moreover, we find that the medium supported with 0.2 ppm NAA and 2 ppm BA has the regeneration rate similar to the medium with 0.1 ppm NAA and 1 ppm BA and 0.2 ppm NAA and 3 ppm BA and the medium with 0.2 ppm NAA and 4 ppm NAA. The results indicate, overall, that the ratio of auxin and cytokinin rather than the actual dosage is the main factor that controls the shoot regeneration.

Orientation and size of petiole explants have effects on the regeneration rate

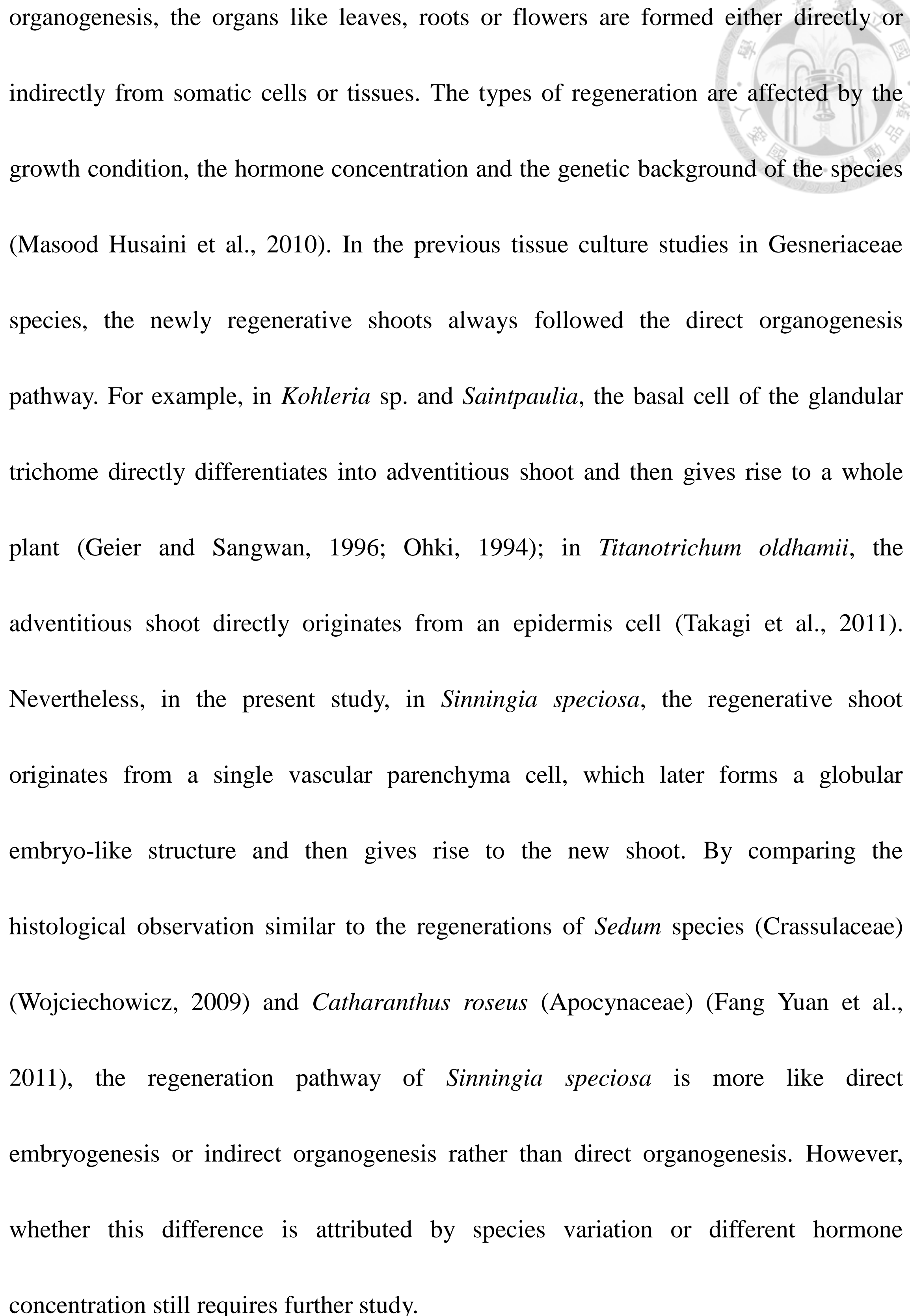
How explant orientation affects the shoot regeneration is rarely reported in previous tissue culture studies. Interestingly, in *Sinningia speciosa*, the shoot regeneration of petiole has a strong positional effect. In the treatment with 0.1 ppm NAA, the regenerative roots grew exclusively at the proximal ends, while in the treatment with 0.1 ppm NAA and 1 ppm BA, the regenerative shoots grew randomly at the both ends, showing no polar specificity. This reveals that the positional effect is mainly attributed to the internal NAA asymmetric distribution (basipetal polar transportation) rather than



BA. The similar effect has been seen in the thin layer culture (TLC) of tobacco. Due to the basipetal distribution of auxin, the positions of regenerative shoots are inversely correlated with the exogenous NAA concentration(Creemers-Molenaar et al., 1994). To further confirm this, vertically orientated petiole explants on the medium supported with 0.2 ppm NAA and 2 ppm BA have been tested. Strikingly, we find that only the treatment with the distal site attached to the medium was able to induce shoots, indicating the positive correlation between shooting and the orientations of petiole explants. The possible explanation for this phenomenon is that the treatment with distal site attached to the medium can facilitate the uptake of NAA by existing internal basipetal auxin transportation (Estelle, 1998; Leyser, 2005), while the opposite direction (with proximal site attach to the medium) cannot. This hypothesis could be partially supported by the defect of leaf initiation from a shoot apex culture of Zea mays when applying the auxin polar transport inhibitor, N-1- naphthylphthalamic acid (NPA) (Scanlon, 2003).

Shoots regeneration of *Sinningia speciosa* through direct embryogenesis

Shoot regeneration from tissue culture can be grouped into 2 pathways: embryogenesis and organogenesis, although it is sometimes hard to judge them mutually exclusive. In embryogenesis, an embryo-like structure is developed from somatic cells or tissues, and then gives rise to a whole plant, similar to the development of a zygotic embryo. In



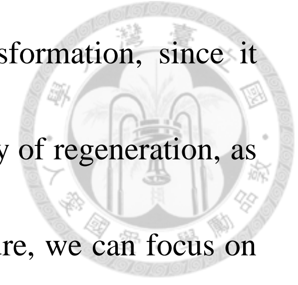
organogenesis, the organs like leaves, roots or flowers are formed either directly or indirectly from somatic cells or tissues. The types of regeneration are affected by the growth condition, the hormone concentration and the genetic background of the species (Masood Husaini et al., 2010). In the previous tissue culture studies in Gesneriaceae species, the newly regenerative shoots always followed the direct organogenesis pathway. For example, in *Kohleria* sp. and *Saintpaulia*, the basal cell of the glandular trichome directly differentiates into adventitious shoot and then gives rise to a whole plant (Geier and Sangwan, 1996; Ohki, 1994); in *Titanotrichum oldhamii*, the adventitious shoot directly originates from an epidermis cell (Takagi et al., 2011). Nevertheless, in the present study, in *Sinningia speciosa*, the regenerative shoot originates from a single vascular parenchyma cell, which later forms a globular embryo-like structure and then gives rise to the new shoot. By comparing the histological observation similar to the regenerations of *Sedum* species (Crassulaceae) (Wojciechowicz, 2009) and *Catharanthus roseus* (Apocynaceae) (Fang Yuan et al., 2011), the regeneration pathway of *Sinningia speciosa* is more like direct embryogenesis or indirect organogenesis rather than direct organogenesis. However, whether this difference is attributed by species variation or different hormone concentration still requires further study.

Meristem identities are unstable in the newly regenerative shoots

In angiosperm, the shoot apical meristems are strictly restricted to the specific sites by the regulation of plant hormone, including the auxin and the cytokinin (Su et al., 2011). However, in this study, the newly regenerative shoots usually appear randomly as clusters at the edge of the explants, indicating the loose regulation of meristem identities. In addition, several ectopic primordia can be seen at the edge of the newly regenerative leaf. These primordia might later give rise to another shoot or the serrated structures around the leaf, resulting in a cluster of multiple shoots or leaf-like structures. This instability might be attributed to the application of exogenous hormones in the medium, which can influence the regulation of meristem.

Young seedlings or callus might be a good source for transformation

To establish a reliable and highly efficient *Agrobacterium* mediated transformation system, a detailed study of tissue regeneration, enhanced tissue competence of *Agrobacterium* infection are important issues. In this study, explants like leaf, petiole and 14-day-old seedling were tested for their competence. The result shows that the visible positive transformation signals could only be seen in the cotyledons and the first pair of primary leaves, suggesting that the seedlings might be a good substitute for the leaf explants, which have been used in previous studies (Li et al., 2013; Zhang et al., 2008).



In addition, callus is also an excellent material for genetic transformation, since it usually has high competence for *Agrobacterium* infection and ability of regeneration, as in the case of rice genetic transformation (Hiei et al., 1994). In future, we can focus on establishing the protocol of the seedling or callus-based *Agrobacterium*-mediated transformation system and conduct the study of functional analysis of floral developmental genes.

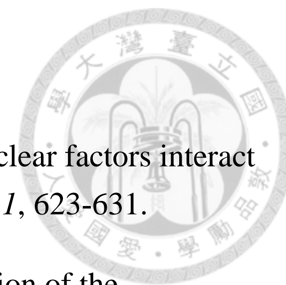
Conclusion and Future Prospects

In this study, the *CYCLOIDEA* homologous gene in *Sinningia speciosa*, *SsCYC*, is identified by shared conserved domains and phylogeny. In addition, single copy of *SsCYC* gene is confirmed by Southern blotting. This implies that *SsCYC* might be the only key gene, participating in the regulation of floral symmetry development. This is further confirmed by the functional analysis, in which ectopic expression of the *SsCYC* gene in tobacco caused small floral size, small and curled leaf, promotion of the growth of axillary buds and dwarfish plant. These *SsCYC* phenotypic effects are more or less consistent with the functional analyses of the *CYCLOIDEA* and other ECE-CYC2 genes. Moreover, to deduce the genetic control of reversal to actinomorphy in horticultural cultivar, the Δ*SsCYC* gene was also transformed into tobacco. The transgenic plants had no visible phenotypic changes, indicating that the Δ*SsCYC* might have lost its function due to a small fragment deletion in the coding sequence, disrupting the reading frame. For tissue culture conditions, I found that the MS medium with 0.1 ppm NAA and 1 ppm BA gave rise to the highest regeneration rate. The regenerative shoots originated from a single vascular parenchyma cell, which is more like direct embryogenesis rather than direct organogenesis as in the other Gesneriaceae species. Finally, the competence of different tissues for *Agrobacterium* infection was tested by GUS reporter system. The result shows that the positive transformed signals could only be seen in the young

seedlings rather than the mature leaf or petiole explants, suggesting that the young seedlings might be a good material for genetic transformation. However, the paraffin sections show that the positive signals were mainly located in epidermis, mesophyll and glandular trichome head cells, which may not grow into new shoots. An unorganized callus tissue may be a potential alternative.

The future work of this project may include: 1) Check the insertion copies of the T-DNA in the T_0 transgenic tobacco plants by Southern blotting (partial of this work has been done in **Supplementary Figure 8**). 2) Establish the transgenic tobacco of empty vector control. 3) Grow the T_1 and T_2 generation of the transgenic tobacco plants. 4) Record the detailed traits phenotype of the T_1 and T_2 populations, including the plant height, leaf number, floral size, cell shape and size and other important quantitative traits. 5) Establish the transgenic lines of *Sinningia speciosa* using regeneration tissues from the young seedlings or callus.

References



Allen, R.D., Bernier, F., Lessard, P.A., and Beachy, R.N. (1989). Nuclear factors interact with a soybean beta-conglycinin enhancer. The Plant Cell Online *1*, 623-631.

Bowman, J.L., Drews, G.N., and Meyerowitz, E.M. (1991). Expression of the Arabidopsis floral homeotic gene *AGAMOUS* is restricted to specific cell types late in flower development. The Plant Cell Online *3*, 749-758.

Broholm, S.K., Tahtiharju, S., Laitinen, R.A., Albert, V.A., Teeri, T.H., and Elomaa, P. (2008). A TCP domain transcription factor controls flower type specification along the radial axis of the *Gerbera* (Asteraceae) inflorescence. Proceedings of the National Academy of Sciences of the United States of America *105*, 9117-9122.

Busch, A., and Zachgo, S. (2007). Control of corolla monosymmetry in the Brassicaceae *Iberis amara*. Proceedings of the National Academy of Sciences of the United States of America *104*, 16714-16719.

Carpenter, R., Copsey, L., Vincent, C., Clark, J., and Coen, E. (1999). Control of Organ Asymmetry in Flowers of *Antirrhinum*. Cell *99*, 367-376.

Chae, S.C., Kim, H.H., and Park, S.U. (2012). Ethylene inhibitors enhance shoot organogenesis of gloxinia (*Sinningia speciosa*). TheScientificWorldJournal *2012*, 859381.

Chapman, M.A., Leebens-Mack, J.H., and Burke, J.M. (2008). Positive selection and expression divergence following gene duplication in the sunflower *CYCLOIDEA* gene family. Mol Biol Evol *25*, 1260-1273.

Chapman, M.A., Tang, S.X., Draeger, D., Nambeesan, S., Shaffer, H., Barb, J.G., Knapp, S.J., and Burke, J.M. (2012). Genetic Analysis of Floral Symmetry in Van Gogh's Sunflowers Reveals Independent Recruitment of *CYCLOIDEA* Genes in the Asteraceae. PLoS Genet *8*.

Choe, S., Fujioka, S., Noguchi, T., Takatsuto, S., Yoshida, S., and Feldmann, K.A. (2001). Overexpression of *DWARF4* in the brassinosteroid biosynthetic pathway results in increased vegetative growth and seed yield in *Arabidopsis*. The Plant Journal *26*, 573-582.

Citerne, H., and Cronk, Q.C.B. (1999). The origin of the Peloric *Sinningia*. The New Plantsman.

Citerne, H.L., Moller, M., and Cronk, Q.C.B. (2000). Diversity of *cycloidea*-like genes in Gesneriaceae in relation to floral symmetry. Ann Bot-London *86*, 167-176.

Citerne, H.L., Pennington, R.T., and Cronk, Q.C. (2006). An apparent reversal in floral symmetry in the legume *Cadia* is a homeotic transformation. Proceedings of the National Academy of Sciences of the United States of America *103*, 12017-12020.

Corley, S.B., Carpenter, R., Copsey, L., and Coen, E. (2005). Floral asymmetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum*. Proceedings of the National Academy of Sciences of the United States of America *102*, 5068-5073.

Costa, M.M., Fox, S., Hanna, A.I., Baxter, C., and Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. Development *132*, 5093-5101.

Crawford, B.C., Nath, U., Carpenter, R., and Coen, E.S. (2004). *CINCINNATA* controls both cell differentiation and growth in petal lobes and leaves of *Antirrhinum*. Plant Physiol *135*, 244-253.

Creemers-Molenaar, J., Hakkert, J., Van Staveren, M., and Gilissen, L. (1994). Histology of the morphogenic response in thin cell layer explants from vegetative tobacco plants. Ann Bot-London *73*, 547-555.

Crepet, W.L. (2008). The Fossil Record of Angiosperms: Requiem or Renaissance? . Annals of the Missouri Botanical Garden *95*, 3-33.

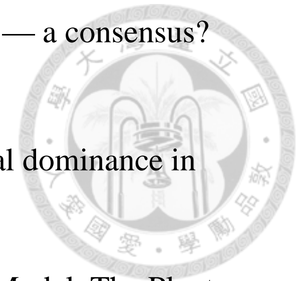
Crepet, W.L., and Niklas, K.J. (2009). Darwin's second 'abominable mystery': Why are there so many angiosperm species? American journal of botany *96*, 366-381.

Crepet, W.L., Nixon, K.C., and Gandolfo, M.A. (2004). Fossil evidence and phylogeny: the age of major angiosperm clades based on mesofossil and macrofossil evidence from Cretaceous deposits. American journal of botany *91*, 1666-1682.

Cubas, P., Lauter, N., Doebley, J., and Coen, E. (1999a). The TCP domain: a motif found in proteins regulating plant growth and development. Plant J *18*, 215-222.

Cubas, P., Vincent, C., and Coen, E. (1999b). An epigenetic mutation responsible for natural variation in floral symmetry. Nature *401*, 157-161.

Damerval, C., and Manuel, M. (2003). Independent evolution of *Cycloidea*-like sequences in several angiosperm taxa. C R Palevol *2*, 241-250.



Dingwall, C., and Laskey, R.A. (1991). Nuclear targeting sequences — a consensus? Trends Biochem Sci *16*, 478-481.

Doebley, J., Stec, A., and Hubbard, L. (1997). The evolution of apical dominance in maize. Nature *386*, 485-488.

Estelle, M. (1998). Polar Auxin Transport: New Support for an Old Model. The Plant Cell Online *10*, 1775-1778.

Fang Yuan, Quan Wang, Qifang Pan, Guofeng Wang, Jingya Zhao, Yuesheng Tian, and Kexuan Tang (2011). An efficient somatic embryogenesis based plant regeneration from the hypocotyl of *Catharanthus roseus*. Afr J Biotechnol *10*.

Feng, X., Zhao, Z., Tian, Z., Xu, S., Luo, Y., Cai, Z., Wang, Y., Yang, J., Wang, Z., Weng, L., *et al.* (2006). Control of petal shape and floral zygomorphy in *Lotus japonicus*. Proceedings of the National Academy of Sciences of the United States of America *103*, 4970-4975.

Geier, T., and Sangwan, R.S. (1996). Histology and chimeral segregation reveal cell-specific differences in the competence for shoot regeneration and *Agrobacterium*-mediated transformation in *Kohleria* internode explants. Plant Cell Rep *15*, 386-390.

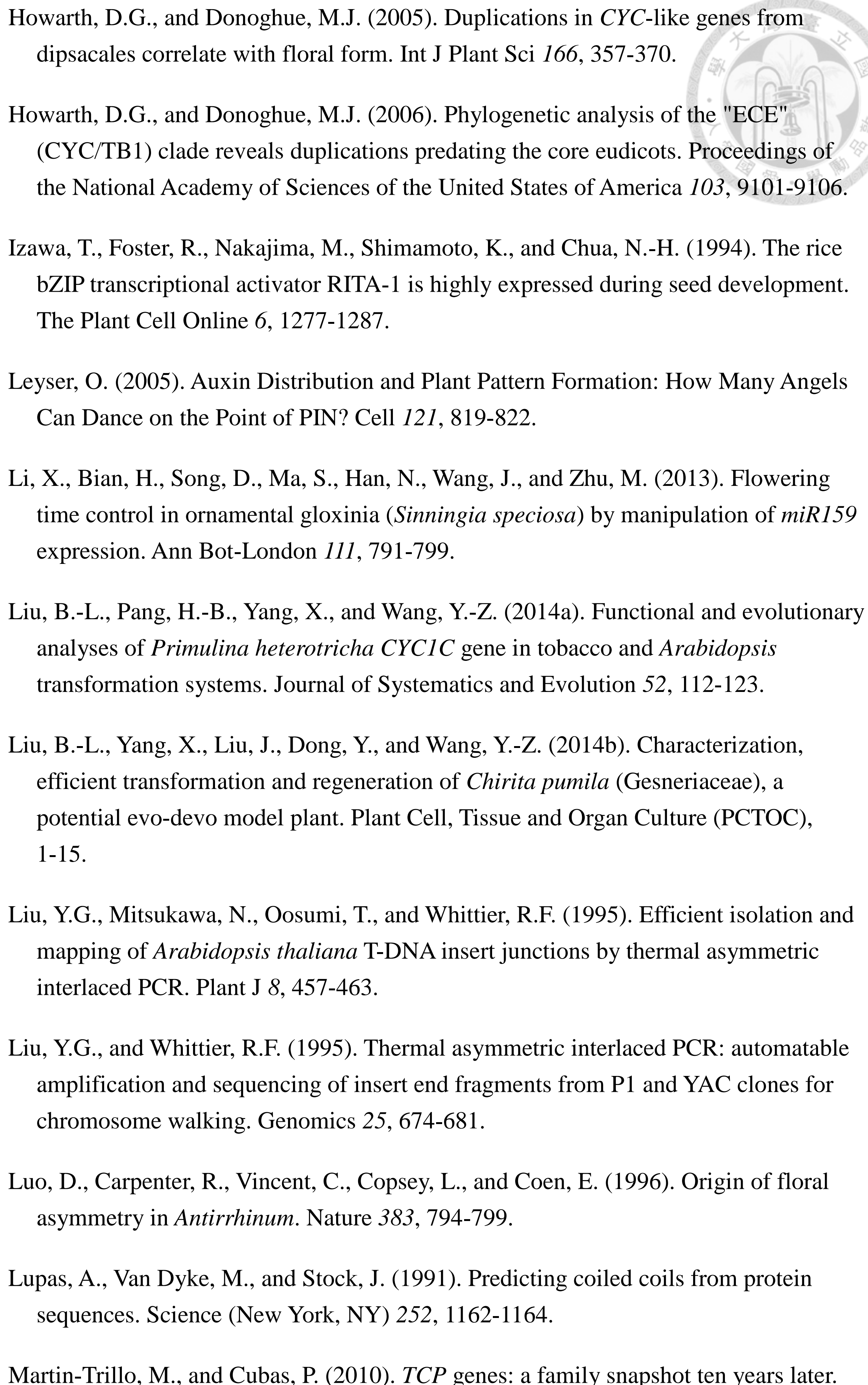
Giraudoux, P. (2013). pgirmess: Data analysis in ecology. R package version 1.5.8. http://CRAN.R-project.org/package=pgirmess.

Guo, Z., Fujioka, S., Blancaflor, E.B., Miao, S., Gou, X., and Li, J. (2010). *TCP1* modulates brassinosteroid biosynthesis by regulating the expression of the key biosynthetic gene *DWARF4* in *Arabidopsis thaliana*. Plant Cell *22*, 1161-1173.

Heery, D.M., Kalkhoven, E., Hoare, S., and Parker, M.G. (1997). A signature motif in transcriptional co-activators mediates binding to nuclear receptors. Nature *387*, 733-736.

Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. (1994). Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J *6*, 271-282.

Horsch, R., Fry, J., Hoffmann, N., Eichholtz, D., Rogers, S.a., and Fraley, R. (1985). A simple and general method for transferring genes into plants. Science *227*, 1229-1231.



Howarth, D.G., and Donoghue, M.J. (2005). Duplications in *CYC*-like genes from dipsacales correlate with floral form. Int J Plant Sci *166*, 357-370.

Howarth, D.G., and Donoghue, M.J. (2006). Phylogenetic analysis of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots. Proceedings of the National Academy of Sciences of the United States of America *103*, 9101-9106.

Izawa, T., Foster, R., Nakajima, M., Shimamoto, K., and Chua, N.-H. (1994). The rice bZIP transcriptional activator RITA-1 is highly expressed during seed development. The Plant Cell Online *6*, 1277-1287.

Leyser, O. (2005). Auxin Distribution and Plant Pattern Formation: How Many Angels Can Dance on the Point of PIN? Cell *121*, 819-822.

Li, X., Bian, H., Song, D., Ma, S., Han, N., Wang, J., and Zhu, M. (2013). Flowering time control in ornamental gloxinia (*Sinningia speciosa*) by manipulation of *miR159* expression. Ann Bot-London *111*, 791-799.

Liu, B.-L., Pang, H.-B., Yang, X., and Wang, Y.-Z. (2014a). Functional and evolutionary analyses of *Primulina heterotricha CYC1C* gene in tobacco and *Arabidopsis* transformation systems. Journal of Systematics and Evolution *52*, 112-123.

Liu, B.-L., Yang, X., Liu, J., Dong, Y., and Wang, Y.-Z. (2014b). Characterization, efficient transformation and regeneration of *Chirita pumila* (Gesneriaceae), a potential evo-devo model plant. Plant Cell, Tissue and Organ Culture (PCTOC), 1-15.

Liu, Y.G., Mitsukawa, N., Oosumi, T., and Whittier, R.F. (1995). Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J *8*, 457-463.

Liu, Y.G., and Whittier, R.F. (1995). Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. Genomics *25*, 674-681.

Luo, D., Carpenter, R., Vincent, C., Copsey, L., and Coen, E. (1996). Origin of floral asymmetry in *Antirrhinum*. Nature *383*, 794-799.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. Science (New York, NY) *252*, 1162-1164.

Martin-Trillo, M., and Cubas, P. (2010). *TCP* genes: a family snapshot ten years later.

Trends in plant science *15*, 31-39.

Martin-Trillo, M., Grandio, E.G., Serra, F., Marcel, F., Rodriguez-Buey, M.L., Schmitz, G., Theres, K., Bendahmane, A., Dopazo, H., and Cubas, P. (2011). Role of tomato *BRANCHED1*-like genes in the control of shoot branching. Plant J *67*, 701-714.

Masood Husaini, A., Mercado, J.A., Teixeira da Silva, J.A., and Schaart, J.G. (2010). Review of Factors Affecting Organogenesis, Somatic Embryogenesis and *Agrobacterium tumefaciens*-Mediated Transformation of Strawberry.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant *15*, 473-497.

Nambara, E., Keith, K., McCourt, P., and Naito, S. (1994). Isolation of an internal deletion mutant of the *Arabidopsis thaliana ABI3* gene. Plant Cell Physiol *35*, 509-513.

Nhut, A., Nguyet, N.A., Phuc, H.T., Huy, N.P., Uyen, P.N., Vi, T.K., Hai, N.T., Binh, N., and Thien, N.Q. (2006). Primary designs of tube-shaped nylon film culture system in shoot regeneration of *Sinningia* spp. Leaf explants. Paper presented at: Proceedings of International Workshop on Biotechnology in Agriculture.

Ohki, S. (1994). Scanning electron microscopy of shoot differentiationin vitro from leaf explants of the African violet. Plant Cell, Tissue and Organ Culture *36*, 157-162.

Pang, J.-L., Wang, L.-L., Hu, J.-Q., Xiang, T.-H., and Liang, H.-M. (2006). Synergistic promotion of gibberellin and cytokinin on direct regeneration of floral buds from in vitro cultures of sepal segments in *Sinningia speciosa* hiern. In Vitro Cell Dev Biol Plant *42*, 450-454.

Preston, J.C., and Hileman, L.C. (2009). Developmental genetics of floral symmetry evolution. Trends in plant science *14*, 147-154.

Preston, J.C., Martinez, C.C., and Hileman, L.C. (2011). Gradual disintegration of the floral symmetry gene network is implicated in the evolution of a wind-pollination syndrome. Proceedings of the National Academy of Sciences of the United States of America *108*, 2343-2348.

R Core Team (2013). R: A Language and Environment for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rogers, S.O., and Bendich, A.J. (1985). Extraction of DNA from milligram amounts of

fresh, herbarium and mummified plant tissues. Plant molecular biology *5*, 69-76.

Salanoubat, M., Lemcke,K., Rieger,M., Ansorge,W., Unseld,M.,, Fartmann, B., Valle,G., Blocker,H., Perez-Alonso,M., Obermaier,B.,, Delseny, M., Boutry,M., Grivell,L.A., Mache,R., Puigdomenech,P., De, Simone, V., Choisne,N., Artiguenave,F., Robert,C., Brottier,P.,, Wincker, P., Cattolico,L., Weissenbach,J., Saurin,W., Quetier,F.,, Schafer, M., Muller-Auer,S., Gabel,C., Fuchs,M., Benes,V.,, Wurmbach, E., Drzonek,H., Erfle,H., Jordan,N., Bangert,S.,, Wiedelmann, R., Kranz,H., Voss,H., Holland,R., Brandt,P.,, Nyakatura, G., Vezzi,A., D'Angelo,M., Pallavicini,A., Toppo,S.,, Simionati, B., Conrad,A., Hornischer,K., Kauer,G., Lohnert,T.H.,, *et al.* (2000). Sequence and analysis of chromosome 3 of the plant Arabidopsis thaliana. Nature *408*, 820-823.

Sargent, R.D. (2004). Floral symmetry affects speciation rates in angiosperms. Proceedings Biological sciences / The Royal Society *271*, 603-608.

Scanlon, M.J. (2003). The Polar Auxin Transport Inhibitor N-1-Naphthylphthalamic Acid Disrupts Leaf Initiation, KNOX Protein Regulation, and Formation of Leaf Margins in Maize. Plant Physiol *133*, 597-605.

Scaramuzzi, F., Apollonio, G., and D’Emerico, S. (1999). Adventitious shoot regeneration from *Sinningia speciosa* leaf discs in vitro and stability of ploidy level in subcultures. In Vitro Cell Dev Biol Plant *35*, 217-221.

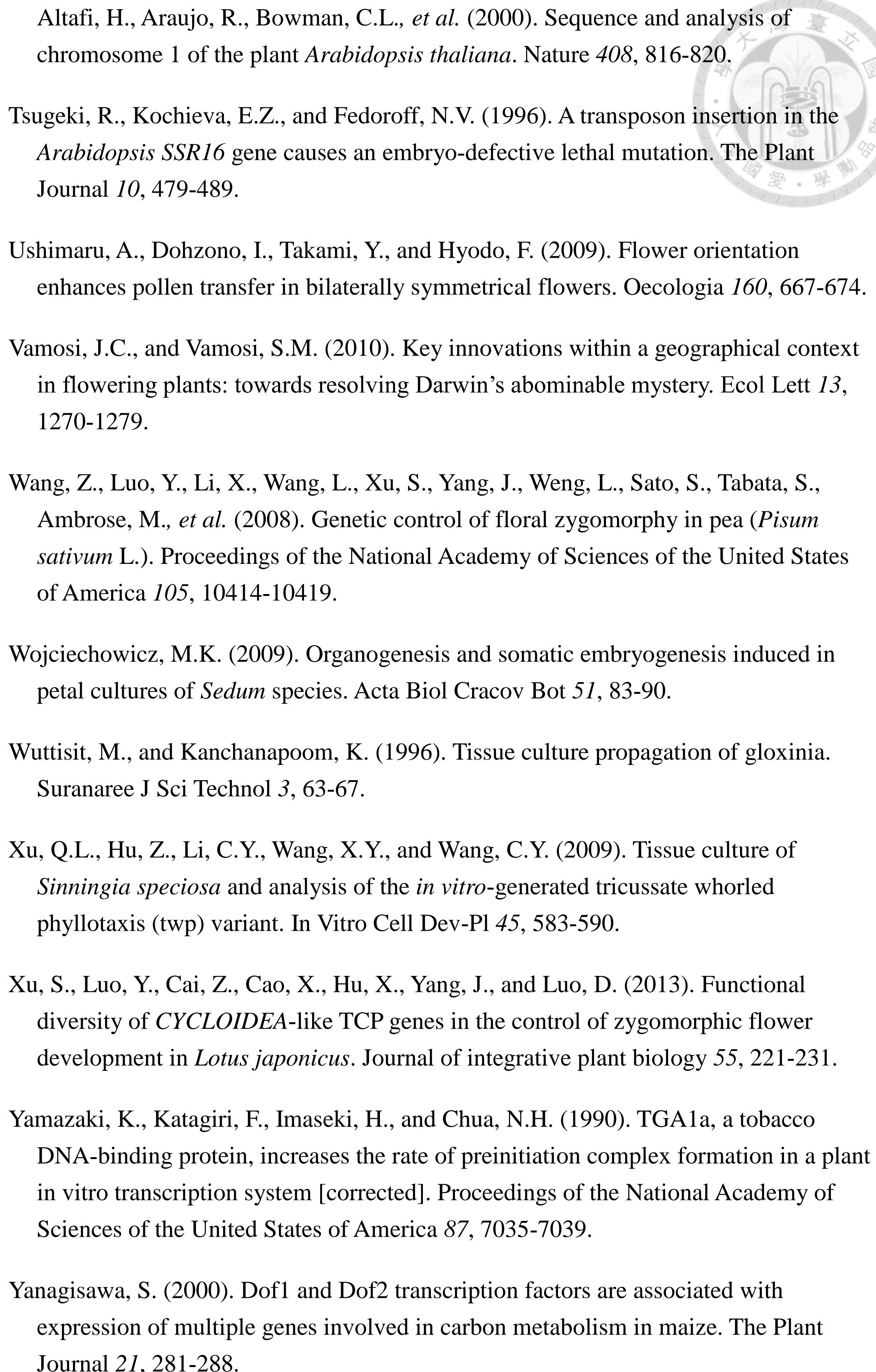
Singer, T., and Burke, E. (2003). High-Throughput TAIL-PCR as a Tool to Identify DNA Flanking Insertions. In Plant Functional Genomics, E. Grotewold, ed. (Humana Press), pp. 241-271.

Song, C.F., Lin, Q.B., Liang, R.H., and Wang, Y.Z. (2009). Expressions of ECE-CYC2 clade genes relating to abortion of both dorsal and ventral stamens in *Opithandra* (Gesneriaceae). Bmc Evol Biol *9*, 244.

Su, Y.H., Liu, Y.B., and Zhang, X.S. (2011). Auxin-cytokinin interaction regulates meristem development. Molecular plant *4*, 616-625.

Takagi, H., Sugawara, S., Saito, T., Tasaki, H., Yuanxue, L., Kaiyun, G., Han, D.-S., Godo, T., and Nakano, M. (2011). Plant regeneration via direct and indirect adventitious shoot formation and chromosome-doubled somaclonal variation in *Titanotrichum oldhamii* (Hemsl.) Solereder. Plant Biotechnol Rep *5*, 187-195.

Theologis, A., Ecker, J.R., Palm, C.J., Federspiel, N.A., Kaul, S., White, O., Alonso, J.,



Altafi, H., Araujo, R., Bowman, C.L., *et al.* (2000). Sequence and analysis of chromosome 1 of the plant *Arabidopsis thaliana*. Nature *408*, 816-820.

Tsugeki, R., Kochieva, E.Z., and Fedoroff, N.V. (1996). A transposon insertion in the *Arabidopsis SSR16* gene causes an embryo-defective lethal mutation. The Plant Journal *10*, 479-489.

Ushimaru, A., Dohzono, I., Takami, Y., and Hyodo, F. (2009). Flower orientation enhances pollen transfer in bilaterally symmetrical flowers. Oecologia *160*, 667-674.

Vamosi, J.C., and Vamosi, S.M. (2010). Key innovations within a geographical context in flowering plants: towards resolving Darwin's abominable mystery. Ecol Lett *13*, 1270-1279.

Wang, Z., Luo, Y., Li, X., Wang, L., Xu, S., Yang, J., Weng, L., Sato, S., Tabata, S., Ambrose, M., *et al.* (2008). Genetic control of floral zygomorphy in pea (*Pisum sativum* L.). Proceedings of the National Academy of Sciences of the United States of America *105*, 10414-10419.

Wojciechowicz, M.K. (2009). Organogenesis and somatic embryogenesis induced in petal cultures of *Sedum* species. Acta Biol Cracov Bot *51*, 83-90.

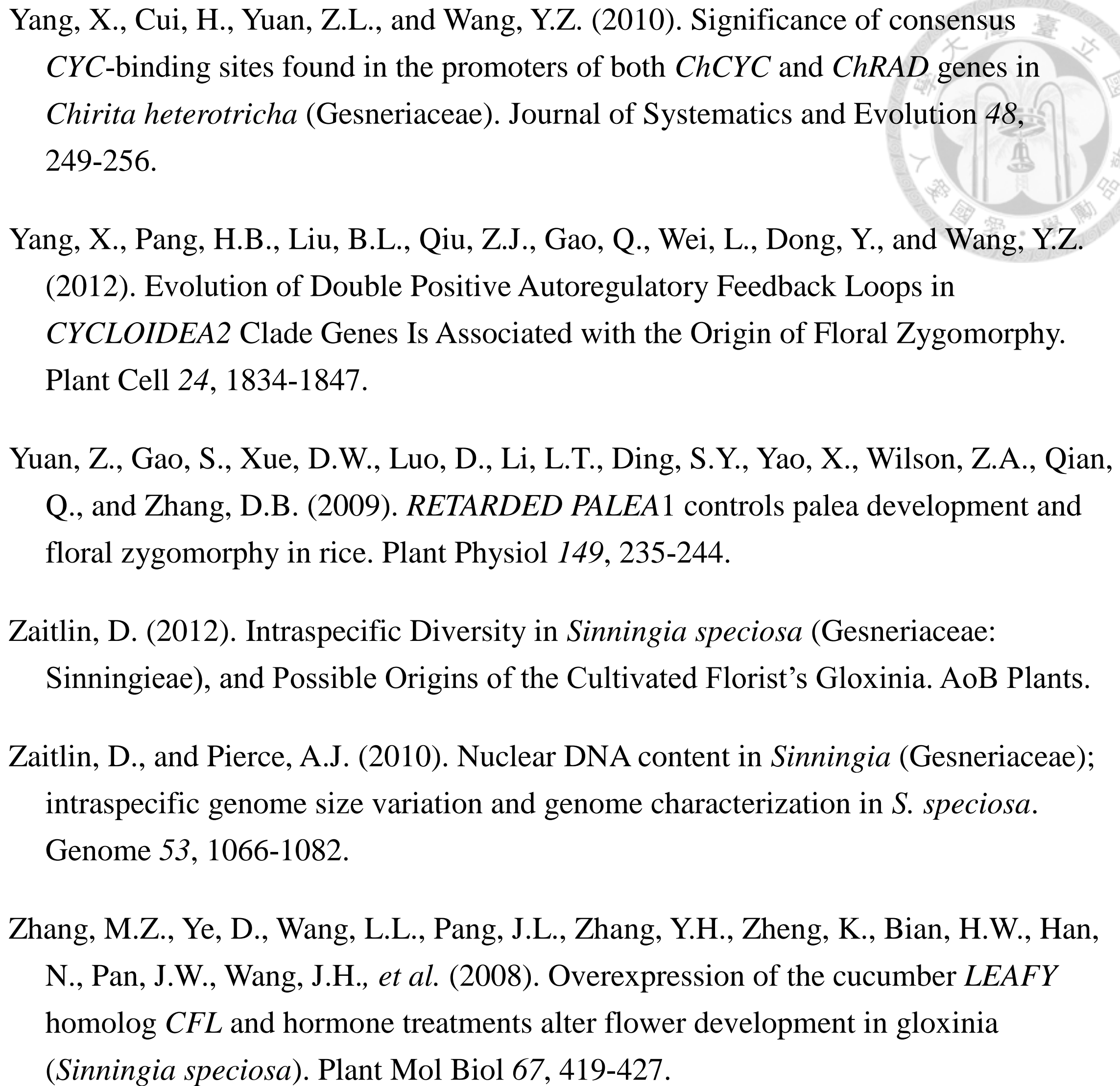
Wuttisit, M., and Kanchanapoom, K. (1996). Tissue culture propagation of gloxinia. Suranaree J Sci Technol *3*, 63-67.

Xu, Q.L., Hu, Z., Li, C.Y., Wang, X.Y., and Wang, C.Y. (2009). Tissue culture of *Sinningia speciosa* and analysis of the *in vitro*-generated tricussate whorled phyllotaxis (twp) variant. In Vitro Cell Dev-Pl *45*, 583-590.

Xu, S., Luo, Y., Cai, Z., Cao, X., Hu, X., Yang, J., and Luo, D. (2013). Functional diversity of *CYCLOIDEA*-like TCP genes in the control of zygomorphic flower development in *Lotus japonicus*. Journal of integrative plant biology *55*, 221-231.

Yamazaki, K., Katagiri, F., Imaseki, H., and Chua, N.H. (1990). TGA1a, a tobacco DNA-binding protein, increases the rate of preinitiation complex formation in a plant in vitro transcription system [corrected]. Proceedings of the National Academy of Sciences of the United States of America *87*, 7035-7039.

Yanagisawa, S. (2000). Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. The Plant Journal *21*, 281-288.



Yang, X., Cui, H., Yuan, Z.L., and Wang, Y.Z. (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*, 249-256.

Yang, X., Pang, H.B., Liu, B.L., Qiu, Z.J., 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. Plant Cell *24*, 1834-1847.

Yuan, Z., Gao, S., Xue, D.W., Luo, D., Li, L.T., Ding, S.Y., Yao, X., Wilson, Z.A., Qian, Q., and Zhang, D.B. (2009). *RETARDED PALEA*1 controls palea development and floral zygomorphy in rice. Plant Physiol *149*, 235-244.

Zaitlin, D. (2012). Intraspecific Diversity in *Sinningia speciosa* (Gesneriaceae: Sinningieae), and Possible Origins of the Cultivated Florist's Gloxinia. AoB Plants.

Zaitlin, D., and Pierce, A.J. (2010). Nuclear DNA content in *Sinningia* (Gesneriaceae); intraspecific genome size variation and genome characterization in *S. speciosa*. Genome *53*, 1066-1082.

Zhang, M.Z., Ye, D., Wang, L.L., Pang, J.L., Zhang, Y.H., Zheng, K., Bian, H.W., Han, N., Pan, J.W., Wang, J.H.*, et al.* (2008). Overexpression of the cucumber *LEAFY* homolog *CFL* and hormone treatments alter flower development in gloxinia (*Sinningia speciosa*). Plant Mol Biol *67*, 419-427.

Supplementary Table 1

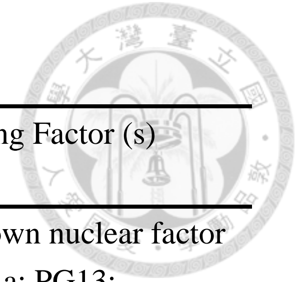
Arbitrary Degenerative Primer

Primer name	Primer sequence (5’-3’)	Length	Degeneracy	Average Tm	Average GC	Reference
AD1	NGTCGASWGANAWGAA	16 bp	128-fold	46.6 °C	43.8%	AD2 in Liu et al. (1995)
AD2	TGWGNAGSANCASAGA	16 bp	128-fold	49.2°C	50%	AD1 in Liu and Whittier (1995)
AD3	AGWGNAGWANCAWAGG	16 bp	128-fold	46.6°C	43.8%	AD2 in Liu and Whittier (1995)
AD4	STTGNTASTNCTNTGC	16 bp	256-fold	47.9°C	46.9%	AD5 in Tsugeki et al. (1996)
AD5	NTCGASTWTSGWGTT	15 bp	64-fold	43.3°C	43.3%	AD1 in Liu et al. (1995)
AD6	WGTGNAGWANCANAGA	16 bp	256-fold	45.3°C	40.6%	AD1 in Liu et al. (1995)

Primer name	4X stock concentration	Final concentration[b]
AD1	12 µM	3 µM
AD2	12 µM	3 µM
AD3	12 µM	3 µM
AD4	16µM	4 µM
AD5	8 µM	2 µM
AD6	16 µM	4 µM

[a]This table was adapted from Singer and Burke (2003).

[b]These concentrations of AD primers were found to yield good results empirically, which are non-linearly correlated with their degeneracy (Liu et al., 1995).



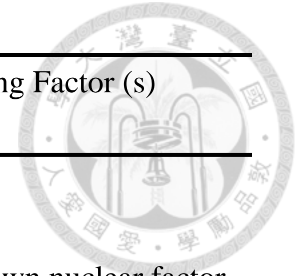
Supplementary Table 2 Putative transcription factor binding sites of SsCYC/ΔSsCYC

treat	Accession	occurrence	Organism/Species	Gene Name	Regulatory Element Name	Binding Factor (s)
SsCYC	RSP00014	1	Arabidopsis (*Arabidopsis thaliana*)	CHS	MRE-core	unknown nuclear factor
SsCYC	RSP00026	1	Tobacco (*Nicotiana tabacum*)	G13	-141 sequence	TGA1a; PG13;
SsCYC	RSP00081	1	Arabidopsis (*Arabidopsis thaliana*)	synthetic oligonucleotides	AGAMOUS BS	AGAMOUS
SsCYC	RSP00096	1	Maize (*Zea mays*)	GapC4	GT-box	tobacco nuclear factors
SsCYC	RSP00098	1	Pea (*Pisum sativum*)	rbcS-3.6	AT-1	AT-1
SsCYC	RSP00117	1	Tobacco (*Nicotiana plumbaginifolia*)	cab-E	AT-1 (2)	unknown nuclear factor
SsCYC	RSP00125	1	Tobacco (*Nicotiana plumbaginifolia*)	cab-E	AT-1 (3)	unknown nuclear factor
SsCYC	RSP00126	1	Tomato (*Lycopersicon esculentum*)	rbcS-3A	AT-1 (1)	unknown nuclear factor
SsCYC	RSP00135	1	Tobacco (*Nicotiana plumbaginifolia*)	cab-E	AT-1 (4)	unknown nuclear factor
SsCYC	RSP00140	1	Pea (*Pisum sativum*)	rbcS-3.6	AT-1 (2)	AT-1
SsCYC	RSP00161	4	Maize (*Zea mays*)	Synthetic oligonucleotids	Dof1 BSopt	Dof1
SsCYC	RSP00327	1	*Brassica napus*	napA	RY	ABI3
SsCYC	RSP00338	1	Soybean (*Glycine max*)	beta-conglicinin alpha'	SEF3 BS	SEF3
SsCYC	RSP00339	3	Soybean (*Glycine max*)	beta-conglicinin alpfa'	SEF4 BScons	SEF4
SsCYC	RSP00477	3	Barley (*Hordeum vulgare*), *Brassica napus*	rbcSF1	S-box	unknown nuclear factor
SsCYC	RSP00502	2	Rice (*Oryza sativa*)	rifa-7-P-glucuronidase transgene	A-box	Different bZIP factors, including RITA-1
SsCYC	RSP00508	9	Soybean (*Glycine max*)	beta-conglicinin alfa'	SEF4 BS	SEF4

Supplementary Table 2 (continued)

treat	Accession	occurrence	Organism/Species	Gene Name	Regulatory Element Name	Binding Factor (s)
SsCYC	RSP00574	1	Tomato (*Lycopersicon esculentum*)	rbcS1	AT-rich K	unknown nuclear factor
ΔSsCYC	RSP00014	1	Arabidopsis (*Arabidopsis thaliana*)	CHS	MRE-core	unknown nuclear factor
ΔSsCYC	RSP00026	1	Tobacco (*Nicotiana tabacum*)	G13	-141 sequence	TGA1a; PG13;
ΔSsCYC	RSP00081	1	Arabidopsis (*Arabidopsis thaliana*)	synthetic oligonucleotides	AGAMOUS BS	AGAMOUS
ΔSsCYC	RSP00096	1	Maize (*Zea mays*)	GapC4	GT-box	tobacco nuclear factors
ΔSsCYC	RSP00098	1	Pea(*Pisum sativum*)	rbcS-3.6	AT-1	AT-1
ΔSsCYC	RSP00117	1	Tobacco (*Nicotiana plumbaginifolia*)	cab-E	AT-1 (2)	unknown nuclear factor
ΔSsCYC	RSP00125	1	Tobacco (*Nicotiana plumbaginifolia*)	cab-E	AT-1 (3)	unknown nuclear factor
ΔSsCYC	RSP00126	1	Tomato (*Lycopersicon esculentum*)	rbcS-3A	AT-1 (1)	unknown nuclear factor
ΔSsCYC	RSP00135	1	Tobacco (*Nicotiana plumbaginifolia*)	cab-E	AT-1 (4)	unknown nuclear factor
ΔSsCYC	RSP00140	1	Pea (*Pisum sativum*)	rbcS-3.6	AT-1 (2)	AT-1
ΔSsCYC	RSP00161	4	Maize (*Zea mays*)	Synthetic oligonucleotids	Dof1 BSopt	Dof1
ΔSsCYC	RSP00327	1	*Brassica napus*	napA	RY	ABI3
ΔSsCYC	RSP00338	1	Soybean (*Glycine max*)	beta-conglicinin alpha'	SEF3 BS	SEF3
ΔSsCYC	RSP00339	3	Soybean (*Glycine max*)	beta-conglicinin alfa'	SEF4 BScons	SEF4
ΔSsCYC	RSP00477	2	Barley (*Hordeum vulgare*) *Brassica napus*	rbcSF1	S-box	unknown nuclear factor
ΔSsCYC	RSP00502	2	Rice (*Oryza sativa*)	rifa-7-P-glucuronidase transgene	A-box	Different bZIP factors, including RITA-1

Supplementary Table 2 (continued)

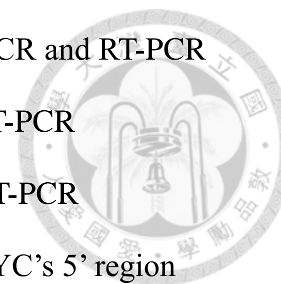


treat	Accession	occurrence	Organism/Species	Gene Name	Regulatory Element Name	Binding Factor (s)
ΔSsCYC	RSP00508	10	Soybean (*Glycine max*)	beta-conglicinin alfa'	SEF4 BS	SEF4
ΔSsCYC	RSP00574	1	Tomato (*Lycopersicon esculentum*)	rbcS1	AT-rich K	unknown nuclear factor

Supplementary Table 3

Primer list in this study

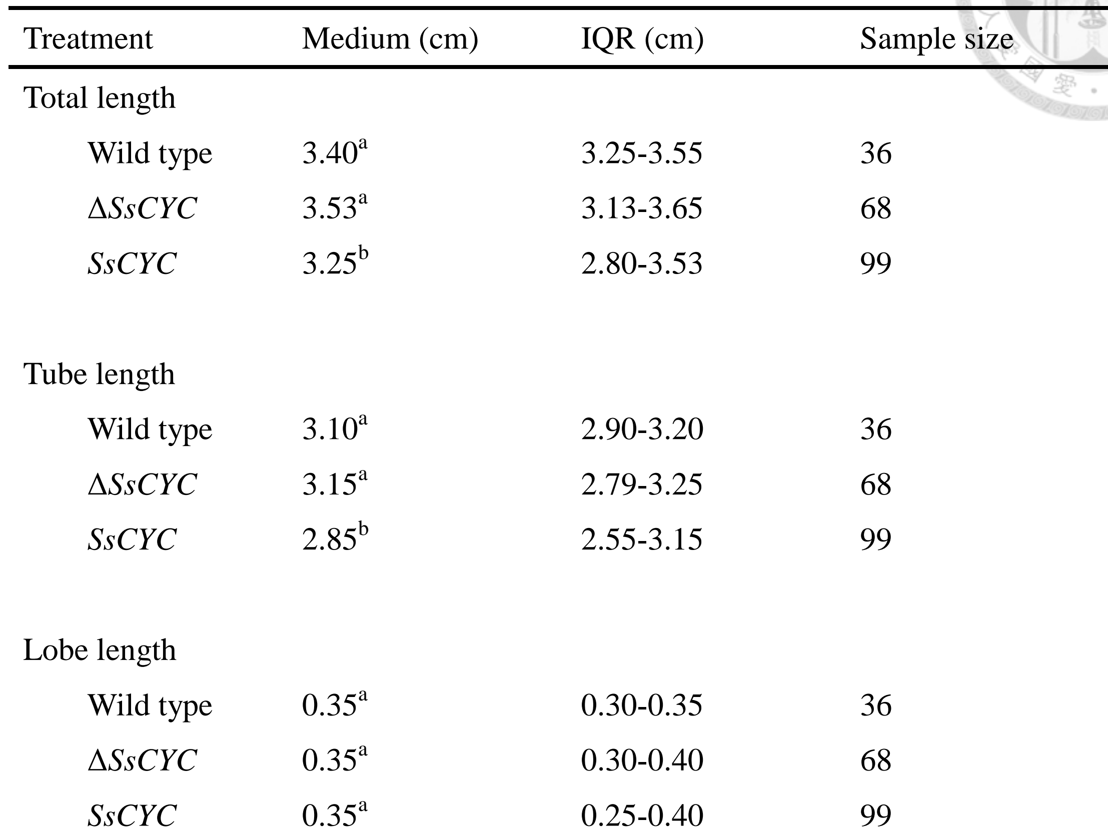
Primer name	Sequence (5'-3')	T_m	Note
NPTFo	TCA GAA GAA CTC GTC AAG AA	47.7	F primer for NPTII gene
NPTBaR	AAC AAG ATG GAT TGC ACG CA	49.7	R primer for NPTII gene
SsCYC_F	ATG TTT AGC AAG AGC ACA TAC CTG AA	54.8	F primer for SsCYC in vector construction
SsCYC_R	CTA TAT ATT TGA AGA GCT ATT CAT GAA CTT GTG	56.9	R primer for SsCYC in vector construction
SsCYC_C_cmyc_R	CTA GAG ATC TTC TTC AGA AAT AAG TTT CTG TTC TAT ATT TGA AGA GCT ATT CAT GAA CTT	66.0	R primer with c-Myc tag for SsCYC in vector construction
p35SS_682_F	TCA AAA GGA CAG TAG AAA AGG AAG G	54.4	F primer for checking the vector
SsCYC_172FLE_R	CCT CAA GAA ACG GGG CAT TC	53.8	R primer for checking the vector
SsCYC_790NNV_F	CTA ATA ATG TAC CTG CTG CAA CAA C	54.4	F primer for checking the vector
SsCYC_118_F	CAC CAT GAC ATG CTT TCC AGC	54.4	F primer for SsCYC/ΔSsCYC gene PCR and RT-PCR
SsCYC_760_R	AGT GGA GGG TAG GCT GAA TTT G	54.8	R primer for SsCYC gene PCR and RT-PCR



dSsCYC_760_R	AGT GGA TGG TAG GCT GAA TTT G	53.0	R primer for ΔSsCYC gene PCR and RT-PCR
NbActin_F	TGT GTT GGA CTC GGG TGA TG	53.8	F primer for NbActin gene RT-PCR
NbActin_R	AGA GCC TCC TAT CCA GAC ACT	54.4	R primer for NbActin gene RT-PCR
SsCYC_139_R	GGG CAT TCA CGG CTA AGT AAT	52.4	R primer for sequencing SsCYC's 5' region
SsCYC_5_622_F	TAT CTG ATG TTG CTG AAG TGG AG	53.5	F primer for sequencing SsCYC's 5' region
SsCYC_5_1500_F	ACC ATT CAT ACA TAG GGC AAG G	53.0	F primer for sequencing SsCYC's 5' region
SsCYC_tail1	AAG AAA ATG ACT TTG CCT GAG AAA GAC TGG AAG G	62.0	GS primer for 1st run TAIL PCR
SsCYC_tail2	TTG GGT TTT CAG GAG TAA AAG TAT TGT ACC CTT TCT	61.0	GS primer for 2nd run TAIL PCR
SsCYC_tail3	ACC TAG TTT AAC TTT ATC TTC TTC ACT TTT AGC GC	59.7	GS primer for 3rd run TAIL PCR

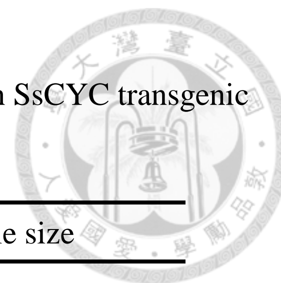
Supplementary Table 4

Flower morphology measurement of *SsCYC* and Δ*SsCYC* transgenic tobacco plants



Treatment	Medium (cm)	IQR (cm)	Sample size
Total length			
Wild type	3.40[a]	3.25-3.55	36
Δ*SsCYC*	3.53[a]	3.13-3.65	68
SsCYC	3.25[b]	2.80-3.53	99
Tube length			
Wild type	3.10[a]	2.90-3.20	36
Δ*SsCYC*	3.15[a]	2.79-3.25	68
SsCYC	2.85[b]	2.55-3.15	99
Lobe length			
Wild type	0.35[a]	0.30-0.35	36
Δ*SsCYC*	0.35[a]	0.30-0.40	68
SsCYC	0.35[a]	0.25-0.40	99

Different letters beside the medium indicate that the mediums are significantly different (Multiple comparison test after Kruskal-Wallis , $p < 0.05$). IQR, Interquartile range which is presented as 1st quartile to 3rd quartile.



Supplementary Table 5

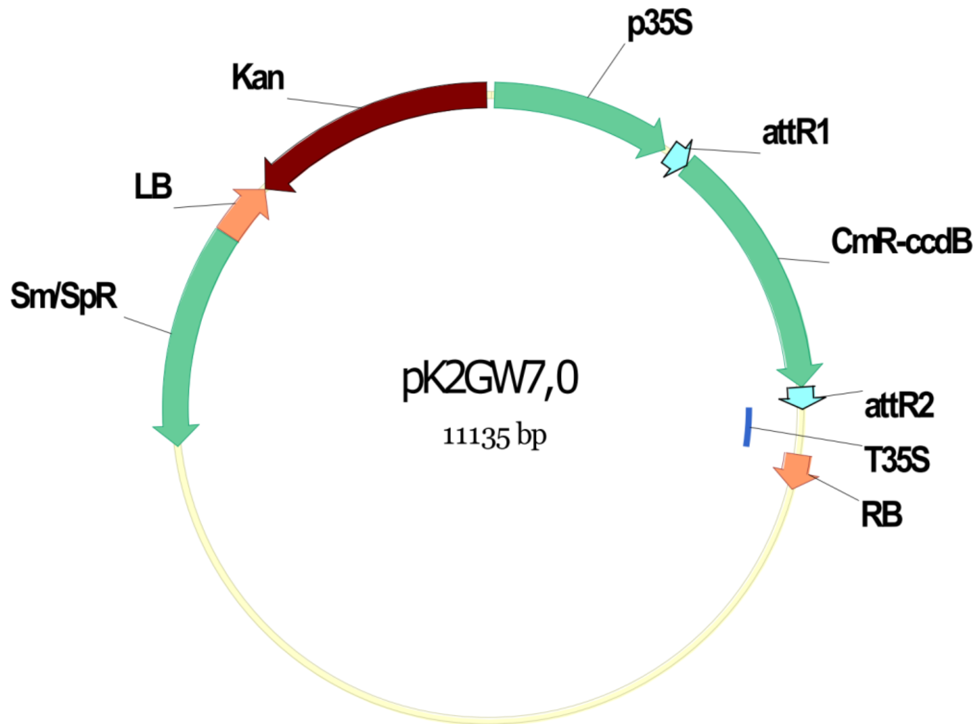
Flower morphology according to the relative mRNA expression level in SsCYC transgenic plants

Treatment	Medium (cm)	IQR (cm)	Sample size
Total length			
Level 0	3.40[a]	3.25-3.55	36
Level 1	3.45[a]	3.33-3.53	16
Level 2	3.60[a]	3.45-3.65	17
Level 3	3.50[a]	3.45-3.78	9
Level 4	2.70[b]	2.31-2.84	34
Tube length			
Level 0	3.10[a]	2.90-3.20	36
Level 1	2.98[a]	2.89-3.10	16
Level 2	3.15[a]	3.05-3.40	17
Level 3	3.15[a]	3.10-3.25	9
Level 4	2.43[b]	2.10-2.60	34
Lobe length			
Level 0	0.35[ab]	0.30-0.35	36
Level 1	0.45[a]	0.39-0.46	16
Level 2	0.30[bc]	0.25-0.40	17
Level 3	0.40[ab]	0.40-0.40	9
Level 4	0.25[c]	0.25-0.30	34

Different letters beside the medium indicate that the mediums are significantly different (Multiple comparison test after Kruskal-Wallis , $p < 0.05$). IQR, Interquartile range which is presented as 1st quartile to 3rd quartile.

Supplementary Figure 1

Destination Vector pK2GW7,0



Supplementary Figure 1 Destination Vector pK2GW7,0

The destination vector used in tobacco transformation. **Sm/SpR**, spectinomycin resistant gene for bacteria. **Kan** (NPTII), kanamycin resistant gene for plant. **P35S**, 35S constitutive promoter. **T35S**, 35S terminator. **CmR-ccdB**, chloramphenicol resistance gene and bacteria lethal gene. **attR1 & attR2**, gateway recombination sites for LR reaction. **LB**, left border. **RB**, right border.

Supplementary figure 2

```
                                          20                                      40                                     60                                     80
                                           |                                       |                                      |                                      |
AtTCP12  ---------- ---------- --MFPSLDTN ---GYDLFDP FIP------H QT-TMFPSF- ---------- ---ITHI--- 31
AtTCP18  MNN------- ---------- -NIFSTTTTI -------NDD YMLFPYNDHY SS-QPLLPFS ---P--SSSI NDILIHST-- 47
AtTCP1   ---------- ---------- --MSSSTNDY NDGNNNGVYP ---------- LS-LYLSSL- --------SG HQDIIHN--- 35
IaTCP1   ---------- ---------- --MSSSNNDY ---SSNGAYP ---------- FS-LYLSSL- --------SG QQD------- 28
AmCYC    ---------- ---------- --MFGKNTYL -------HLP -QVSSSLHSR AA-TSVVDL- --------NG NEIQLH---- 36
AmDICH   ---------- ---------- --MFNKKRYL -------QYP -QVPPSLNPR AS-TSVVSL- --------NG NEILLH---- 36
LvCYC    ---------- ---------- --MFGKNTYL -------HLP HQVSTSLNSR A----VVDL- --------NG NEIQLHL--- 35
SiCYC1A  ---------- ---------- --MFGKNSYL -------HPP -QVSRPLQTR GP-TSSVDLS --------DG AEILL----- 36
SiCYC1B  ---------- ---------- --MFGKNSYL -------HPP -QVSQPLQTR GP-TSSIDLI --------DE AEILLH---- 37
GhCYC2   ---------- ---------- --MFSTNPFQ ---------- ELNSSSHVFP PP-NSFLDH- --------EQ DDLFHH---- 34
ObCYC1   ---------- ---------- --MFGKSPYL -------HLP -QVSQSLQSR GA-TSAVDLI --------NG AEILLHD--- 38
LjCYC2   ---------- ---------- --MFPFSSS- ---------P YPSFPSSSS- ---SPFHPFS FLNPENVSAS NNTLVLDPLS 44
LjCYC1   ---------- ---------- --MFPNSSY- ---------P YPYFPSSSSS SP-SPYPPFN FLNPGNA-ST NNIFLHDPPP 46
LjCYC3   ---------- ---------- --MFSSTSYH -----DSINP FPSSFPSSSY PS-LPFL--- -IDP------ ---------- 32
HaCYC2C  MNLYLNFFVN YYCYCFFHFQ ITMFSSNPF- ---------P QIPSPIHVSH PF-NSFFDL- --------ER NDVYFN---- 56
PhCYC1C  ---------- ---------- --MFGKSSYL -------HPP -QVSQSLQSR GS-TSAIDIV --------NG DEILLHD--- 38
PhCYC1D  ---------- ---------- --MLSKSSYL -------HPR -QVSQSLESR GS-TSAVDLV --------NG AEILLHD--- 38
SlCYC1   ---------- ---------- --MFPFGNSS -----NGGNP ILH------- ---SSFL--- ---------N NQILLH---- 27
SlCYC2   ---------- ---------- --MFSASNSS ---TH--DNP LPHYISSSFH TS-SPFLGF- --------TG NQILLH---- 39
NbCYC1A  ---------- ---------- --MFPSGNSN -------GNP ILQCSIPSFH SSNSSFHGL- --------NG NQILLH---- 38
NbCYC1B  ---------- ---------- --MFPSGNSS -------GNP ILQFSIPSFH SSNSSFLGL- --------NG NQILLH---- 38
NbCYC2A  ---------- ---------- --MFPASNST -------GNP -PPHPSLSFH SS-SPFLGL- --------NG SQILLH---- 36
NbCYC2B  ---------- ---------- --MFPASNST -------GNP -PPNPSLSFH SS-SPFLGL- --------NG NQILLH---- 36
SsCYC    ---------- ---------- --MFGKNTYL -------HVP -QVSPSLQSR AS-TSLVDL- --------NG GEILLH---- 36

                                         100                                     120                                    140                                    160
                                           |                                       |                                      |                                      |
AtTCP12  -------QS PNSH--HHYS SP-SFPFSSD FL-------- ------ESFD ESFLINQFLL ---------- ---QQQDVAA 73
AtTCP18  -------SNT SNNHLDHHHQ FQQPSPFSHF EF-------A PDCALL---T SFHPENNGHD DNQTIPNDNH HPSLHFPLNN 110
AtTCP1   ---------- ---P--YNHQ LK----ASPG HM-------- ---------- ---VSAVP-- ---------- ----ESLIDY 59
IaTCP1   ---------- ---P--YNHQ LI----ASQG QM-------- ---------- ---VSAVP-- ---------- ----HSLVDY 52
AmCYC    ---------- ---------- DM----LSGH YL-------- ---------- --TTANAP-- ---------- -----VLEST 55
AmDICH   ---------- --------HH DV----ISGY YL-------- ---------- ---ASNPQ-- ---------- -----NLEPD 56
LvCYC    ---------- ---Q--HHHH DI----LSGH YL-------- ---------- ----TNGP-- ---------- -----VLEST 57
SiCYC1A  ---------- ---H--HHQD DV----LSGH YL-------- ---------- ---AENTP-- ---------- -----FLGVS 59
SiCYC1B  ---------- ---H--HHQD GV----LSGH YL-------- ---------- ---ATNTP-- ---------- -----FLGVS 60
GhCYC2   ---------- -------RSN HP---FVSGD SFFGAMADFK ---------- ---------- DSGGQ----- ----HQLFSG 65
ObCYC1   ---------- ---H--HQKQ DV----LSGH YL-------- ---------- ---ATNAS-- ---------- -----FLDAS 61
LjCYC2   VPYIPQTHH- HHHH--HHNN PP----ALPE TL-------- ---------T NLAVAGSP-- ---------- ---------- 78
LjCYC1   LPSVPYTNNT TNTH--HHHA PP----TCSE TF-------- ---------T NWALADA--- ---------- -----MLKQE 85
LjCYC3   ---------- ENIN--NFPQ DP----LVSH PFIHD-SNIN DAPIFTHDTT NLAVSVS--- ---------- ----HLLEQQ 78
HaCYC2C  ---------- --HH--EDNN NP---FVSGD SFLHSYSSFA PQPSTLPPVT DYIPSNAQEL DSQNQ----- ----LLDHEG 110
PhCYC1C  ---------- ---H--QQQQ DM----LSSH YL-------- ---------- ---ATNAP-- ---------- -----FIETS 61
PhCYC1D  ---------- ---H--HHHQ DM----LSDH YL-------- ---------- ---AENVS-- ---------- -----FLEVS 61
SlCYC1   ---------- -------QHD LP----THHH YL-------- ---------- ---AAANGHS IDSYA----- ---------- 50
SlCYC2   ---------- ------QYYQ NQ----FSSH YY-------- ---------- --LAKN---- ---------- ---------- 55
NbCYC1A  ---------- -------QHQ DQ----VSTH YL-------- ---------- ---AANNGYL IDNSSS---- ----AMMFAN 68
NbCYC1B  ---------- -------QHQ DQ----LSTH YL-------- ---------- ---AANNGHL IDNNSS---- -----VMFAN 67
NbCYC2A  ---------- -------HYQ TQ----LSSH HF-------- ---------- ---AKNM--- ---------- -----TLSAN 56
NbCYC2B  ---------- -------HYQ NQ----FSSH HF-------- ---------- ---AKNT--- ---------- -----TLSAN 56
SsCYC    ---------- ---H--HHHH DM----LSSH YL-------- ---------- ---AVNAP-- ---------- -----FLEAS 59

                                         180                                     200                                    220                                    240
                                           |                                       |                                      |                                      |
AtTCP12  NVVESPWKFC KKLELKKKN- -----EKCVD GSTSQEVQ-- WR---RTVKK RDRH-SKICT AQGPRDRRMR LSLQIARKFF 141
AtTCP18  TIVEQPTEP- ----SETINL I----EDSQR -ISTSQDP-- KMKKAKKPSR TDRH-SKIKT AKGTRDRRMR LSLDVAKELF 177
AtTCP1   MAFKSNN--- ----V--VN- ---------- -QQGFEFP-E VSKEIKKVVK KDRH-SKIQT AQGIRDRRVR LSIGIARQFF 116
IaTCP1   IKFNSNSST- ----C--VN- ---------- -QQGFRCPEE VPREIKKSVK KDRH-SKIHT AQGLRDRRVR LSIGIARQFF 112
AmCYC    ALFNNNNNF- NHDVVNGLN- --------RD PS---PTF-- PT---KQAVK KDRH-SKIYT SQGPRDRRVR LSIGIARKFF 116
AmDICH   ALFNGFHHD- ----VGGTN- --------GD PLVLANTL-- AK---KHTPK KDRH-SKINR PQGPRDRRVR LSIGIARKFF 116
LvCYC    ALFNNFNN-- QHVVSHGMNC H------DPD PANMAETF-- QT--KQSTVK KDRH-SKIYT AQGPRDRRVR LSIGIARKFF 124
SiCYC1A  TLYSQD---- ----LGVGS- --------ED PSSLTNTL-- SR---RQVVK KDRH-SKIVT SQGPRDRRVR LSIGIARKFF 116
SiCYC1B  TLYSQD---- ----ASEGT- --------ED PSALANAL-- SR---RQTVK KDRH-SKIVT SQGPRDRRVR LSIGVARKFF 117
GhCYC2   SGLEYNDEY- -------NN- ---------- -TLESGVS-- KMKNKKKISK KDHH-SKIDT AHGPRDRRVR LSIDIARKFF 122
ObCYC1   ALYDQD---- ----VGGCN- --------ED PCALANTF-- SR---KQTAK KDRH-CKIVT SQGPRDRRVR LSIGIARKFF 118
LjCYC2   ---------- -------DNC FAMPKPDPSG PHYGISCL-- LT---KRPAK KDRH-SKIYT SQGLRDRRVR LSIEIARKFF 135
LjCYC1   NLTGGPHHH- ----YNLSN- ---------- -FLTRKLP-- AA---KKPAK KDRH-SKIHT SQGLRDRRVR LSIEIARKFF 142
LjCYC3   DPLGGNNYA- ---SSIVPNF L----RASKP -AASASAA-- AA---MATTK KDRH-SKIHT SQGLRDRRVR LSSEIARKFF 143
HaCYC2C  SGLQYCDNY- -------SD- ---------- -LLESVVY-- PSKKKVAISK KDGH-SKIYT AQGPRDRRVR LSIDIARKFF 167
PhCYC1C  TLYNQD---- ----VGGSN- --------ED PSALASTF-- SI---KQMVK KDRH-SKIVT SQGPRDRRVR LSIGIARKFF 118
PhCYC1D  TLYNQD---- ----VGGSN- --------ED PSALANTF-- SR---NQTVK KDRH-SKIVT SQGPRDRRVR LSIGIARKFF 118
SlCYC1   ---------- -------TN- ---------- -----NVA-- INNKSKKQVK KDRH-TKILT SQGHRDRRVR LSIGVARKFF 94
SlCYC2   ---------- --------N- -----EDYCD -NSLRSFP-- MK---KKSKK RERSCGKILT AQGPRDRRIR LSINMARKFF 105
NbCYC1A  NVIHNR---- -------SN- ---------- -NQEIGCP-- LNKMMKKPVK KDRH-SKILT SQGHRDRRVR LSIGVARKFF 122
NbCYC1B  NVIHNK---- -------SN- ---------- -NQEISFP-- LNKMMKKPMR KDRH-SKILT SQGHRDRRVR LSIGVARKFF 121
NbCYC2A  NIINSPN--- -------YN- -----QKHRD -NSIRSFP-- IN---KKPKK RERC-SKILT SQGPRDRRVR LSIAVARKFF 113
NbCYC2B  NIINSPN--- -------YN- -----QNHSG -NSLRSFP-- IN---KKPKK RERP-SKIVT SQGPRDRRVR LSIAIAQKFF 113
SsCYC    SLYNQDA--- ----IVGLN- --------ED PSAMANTF-- PR---KQTVK KDRH-SKIVT AQGPRDRRVR LSIGIARKFF 117

                                         260                                     280                                    300                                    320
                                           |                                       |                                      |                                      |
AtTCP12  DLQDMLGFDK ASKTIEWLFS KSKTSIKQLK ERVAASEGGG K--------- ---------- ---------- ---------- 182
AtTCP18  GLQDMLGFDK ASKTVEWLLT QAKPEIIKIA TTLSHHGCFS S-------G D--------- ---------- ---------- 220
AtTCP1   DLQDMLGFDK ASKTLDWLLK KSRKAIKEVV QAKNLNNDDE DF------G NI-------- ---------- ---------- 161
IaTCP1   DLQDMLGFDK ASKTLDWLLK KSRKAIKELV HEKKLTNDGD ---------- ---------- ---------- ---------- 152
AmCYC    DLQEMLGFDK PSKTLDWLLT KSKTAIKELV QSKSTKSNSS ---------- ---------- ---------- ---------- 156
AmDICH   DLQEMLGFDK PSKTLDWLLT KSKEAIKELV QSKSSKSNIS ---------- ---------- ---------- ---------- 156
LvCYC    DLQEMLGFDK PSKTLDWLLT KSKTAIKELV QSKSTKSNSS N--------- ---------- ---------- ---------- 165
SiCYC1A  DLQEMLGFDK PSKTLEWLLT KSKVAIKDLV LAKKSSSSRS ---------- ---------- ---------- ---------- 156
SiCYC1B  DLQEMLGFDK PSKTLDWLLT KSKVAIKDLV LTKKSSSSRS ---------- ---------- ---------- ---------- 157
GhCYC2   CLQDLLGFDK ASKTLDWLFT KSKPAIDELL EGTKNSSSST V--------- ---------- ---------- ---------- 163
ObCYC1   DLQEMLGFDK PSKTLEWLLT KSKVAIKDLI HTKKSSSARS ---------- ---------- ---------- ---------- 158
LjCYC2   DLQDMLGFDK ARNTLEWLFN KSKRAIKDFA RSKNSNSGGG ---------G D--------- ---------- ---------- 177
LjCYC1   DLQDMLGFDK ASNTLEWLFS KSNKAIEELF RSKHSDNNNS TGACAADGDG D--------- ---------- ---------- 193
LjCYC3   DLQDMLEFDK PSNTLEWLFT KSESAIQELA RSKNCSCSSN ---------- ---------- ---------- ---------- 183
HaCYC2C  VLQDLLGFDK ASKTLDWLFT KSKKAIKELI EETKHSSSST V--------- ---------- ---------- ---------- 208
PhCYC1C  DLQEMLGFDK PSKTLEWLLT KSKVAIKDLA HTKKSSSARS ---------- ---------- ---------- ---------- 158
PhCYC1D  DLQEMLGFDK PSKTLEWLLT KSKVAIKDLV HTKKSSSARS ---------- ---------- ---------- ---------- 158
SlCYC1   DLQDMLGYDK PSKTLDWLFT KSKLAIEDLI NDVSKKSTPL S--------- ---------- ---------- ---------- 135
SlCYC2   DLQELLGFDK PSKTIDWLFT HSELALEELT NCQLETKMFF ---------- -----VWSVQ LKLF------ ---------- 154
NbCYC1A  DLQDMLRFDK PSKTLDWLFT KSKLAIEELT TGQMKTASLN E--------- ---------- ---------- ---------- 163
NbCYC1B  DLQDMLGFDK PSKTLDWLFT NSKLAIEELT S-QMNTTSPN E--------- ---------- ---------- ---------- 161
NbCYC2A  DLQEMLGFDK PSKTLDWLFT NSKLAIEELT NWSTHQDHDS NIA------G SIPILTIRMS LSFFSREGSP NLRQSP--SC 185
NbCYC2B  DLQEMLGFDK PSKTLDWLFT HSKLAIEELT NWSTHQDHGS KIA------G SIPVLTNRMC LIRFFLGKAP PTSASPFLSC 187
SsCYC    DLQEMLGFDK PSKTLDWLLT KSKAAIKELV QAKKSGSGSA K--------- ---------- ---------- ---------- 158
```

Supplementary figure 2 continued

```
                                          340                                  360                                  380                                  400
AtTCP12  ----------  ----------  ----------  ----------  ----------  ----------  -----DEHLQ  VDEKEKDETL  197
AtTCP18  ----------  ----------  ----------  ----------  ----------  ----------  --ESHIRSMD  TSSDLC-ELA  237
AtTCP1   ----------  ----------  ----------  ----------  ----------  ----------  GGDVEQEEEK  EEDD------  175
IaTCP1   ----------  ----------  ----------  ----------  ----------  ----------  -GDGEEEEED  ADEDED----  167
AmCYC    ----------  ----------  ----------  ----------  ----------  ----------  ----------  SPCDDC-EEV  165
AmDICH   ----------  ----------  ----------  ----------  ----------  ----------  ----------  NSPSECDQEV  166
LvCYC    ----------  ----------  ----------  ----------  ----------  ----------  ---------S  HSDDDCDDEV  176
SiCYC1A  ----------  ----------  ----------  ----------  ----------  ----------  ---------T  SSPSEC-EVV  166
SiCYC1B  ----------  ----------  ----------  ----------  ----------  ----------  ---------T  SSPSEC-EVV  167
GhCYC2   ----------  ----------  ----------  ----------  ----------  ----------  ----------  --TDQS-EVG  170
ObCYC1   ----------  ----------  ----------  ----------  ----------  ----------  ---------T  SSPSEC-EVV  168
LjCYC2   ----------  ----------  ----------  ----------  ----------  ----------  -HDNNKSFSS  EEDEDC-EVV  195
LjCYC1   ----------  ----------  ----------  ----------  ----------  ----------  -GDD-DSFSS  SSSED--EVV  209
LjCYC3   ----------  ----------  ----------  ----------  ----------  ----------  --------SN  NNPSNN-NKS  194
HaCYC2C  ----------  ----------  ----------  ----------  ----------  ----------  ----------  -STNQC-EMA  216
PhCYC1C  ----------  ----------  ----------  ----------  ----------  ----------  ---------T  PLPSEC-EVV  168
PhCYC1D  ----------  ----------  ----------  ----------  ----------  ----------  ---------T  SSPSEC-EVV  168
SlCYC1   ----------  ----------  ----------  ----------  ----------  ----------  ----IHNNNN  NNNSECDEDM  151
SlCYC2   ----------  ----------  ----------  ----------  ----------  -LFYYIYVGA  ----------  ----------  163
NbCYC1A  ----------  ----------  ----------  ----------  ----------  ----------  ----PKKIPS  CINSEC-EDV  178
NbCYC1B  ----------  ----------  ----------  ----------  ----------  ----------  -SSAAKKSTS  AINSEC-EDV  179
NbCYC2A  PTL-------  ----------  ----------  ----------  -------VSF  QIFYFIYVGA  TKSSCDEIAK  DYASEC-EDP  220
NbCYC2B  LELENQNIGK  GKVCVYITFS  RSHLWDFTGF  AVVVIVVYPK  ESLENLVTTF  QIFYFIYVGA  TKSNCDEVAK  DCASEW-EDL  266
SsCYC    ----------  ----------  ----------  ----------  ----------  ----------  --------SI  SSPSEC-EVV  169

                                          420                                  440                                  460                                  480
AtTCP12  ----------  ----------  ----------  --KLRVS-KR  RTKTME----  -----SSFKT  KESRERARKR  ARERTMAKMK  235
AtTCP18  SMWTVDDRGS  N--TNTTE-T  RGNK-VDGR-  --SMRGK-RK  RPEPRT----  ---PILKKLS  KEERAKARER  AKGRTMEKMM  302
AtTCP1   ----------  ---------N  GDKSFVYGLS  PGYGEEE-VV  CEATKAGIRK  KKSELRNISS  KGLGAKARGK  AKERTKEMMA  235
IaTCP1   -----DDDGG  ---------D  GNKSFVFGLS  PEYCEEE-VV  SDVNKAA---  EKSELSNNSS  K------RSK  AKAEKAKEMI  223
AmCYC    VSV--DS-EN  -----VTDHS  KGKS---LKA  NNKCKE---A  MDSHQA----  --------AA  KESRAKARAR  ARERTKEKMC  219
AmDICH   LSA--DL-PY  -----IGS-S  KGKAAVGLNS  N-KCKG---G  RD--------  -----AVDLA  KESRAKARAR  ARERTKEKMC  220
LvCYC    VSAEGDS-EN  -----AAD-S  KGKS-VLIKG  NYKCKEAASA  MDSQQA----  -----ALNLV  KESRAKARAR  ARERTKEKMC  239
SiCYC1A  LNG--ETFED  GSCLLPPD-S  KRKS--ASTA  ----------  RDPAQS----  -----ASTLA  KESRAKARAR  ARERTKEKLC  222
SiCYC1B  LNG--EAFQD  GNCLRHPD-S  RKSS--SKNA  N-KC------  KDPAQS----  -----ASNLA  KESRAKARAR  ARERTKEKMC  226
GhCYC2   VLEIINGGSN  -----EED-K  AKKK-KTTPN  CVDEKGK-KT  TRKHKS----  -----GSFPV  NQSRAEARAR  ARERTKEKWH  233
ObCYC1   LNG--EAFEY  GNCLLPAD-S  KRKS-VLMNA  N-KCKG---A  KDPTQS----  ----TASTLA  KESRAKARAR  ARERTKEKMC  232
LjCYC2   SSNLQQDQEQ  R--LDSNH-N  KGGG-AEKEL  LNKLKR---A  QKEPSC----  -----ARAKI  KESREKARAR  ARERTSNKIC  259
LjCYC1   ----------  ----------  ----------  --SVKRA-QK  EKEPSG----  -----VQAKM  KESREKARAR  ARERTSNKMC  247
LjCYC3   LFGGGDNH--  ---------N  GTNS-IRGRG  NSKLKWT-TQ  RDQDVC----  -----VLQNK  KESRERARAR  ARERTCYKML  252
HaCYC2C  FLETIKGGSG  -----SDE-D  KGQK-MSALR  FLDGEKK-KM  TQKCKSGFH-  -----ASLAR  GQSRAEARAR  ARERTKQKLR  282
PhCYC1C  LNG--EAFEH  GSCLLPAD-S  KRKS-VLMNA  N-QCKG---A  KDPTQS----  -----ASTLA  KESRAKARAR  ARERTKEKMC  231
PhCYC1D  LNG--EAFEN  GNCLLGED-S  KRKW-VSINA  N-KCKG---A  KDPTQS----  -----ASTLA  KESRAKARAR  ARERTKEKMC  231
SlCYC1   IVP-------  ----------  ----------  --------LA  KKAKQE----  ----------  RDSRAKARAR  ARERTIKKIW  182
SlCYC2   --INNKGLE-  ---------R  NPKR------  ----------  ----------  ----AKEVTL  KELRKKARAR  ARERTIKKMW  201
NbCYC1A  VIA-------  ----------  ----------  --------RA  KQEEEA----  ----IVHL-V  RESRAKARAR  ARERTIRKVW  214
NbCYC1B  FLA-------  ----------  ----------  --------RA  KQEEEA----  ----ILNLVV  RESRAKARAR  ARERTIKKVW  216
NbCYC2A  VITTNKGLE-  ---------R  NPKR------  ---AKEKKEV  KDEATD----  -----LALVA  RESRAKARTR  ARERTIKKML  272
NbCYC2B  TITTNEDLE-  ---------R  KPKR------  --LAKEKKEV  KDEATD----  -----HALVA  RESRAKARER  ARERTIKKMW  319
SsCYC    SAGNGETFEN  GRYL-DAE-S  KKKS-LPLNP  NYKCKE--YS  KDPQQS----  -----ALNLA  KVSRAKARAR  ARERTREKMC  235

                                          500                                  520                                  540                                  560
AtTCP12  MR-LFETSET  ISDPHQE---  -TREIKITNG  VQL-----LE  KENKEQEWS-  -NTNDVHMVE  YQMDSVSIIE  KFLGLTSDSS  303
AtTCP18  MK-MKGRSQL  VKVVEEDAHD  HGEIIKNNNR  SQ------VN  RSSFEMTH--  -CEDKIEELC  K-NDRFAVCN  EFIMNKKDHI  371
AtTCP1   YD-NPETASD  ITQSEI----  -MDPFKRSIV  F---------  NEGEDMTHLF  -YKEPIEEFD  N--------Q  ESILT--NMT  289
IaTCP1   YK-HP-----  --QSET----  -MDPPKGSIT  F--------Y  EEEENMTTSF  -YKKAIEEFD  K--------Q  EYILTKKKIN  273
AmCYC    IK-QLNEAIV  LR--------  -------NHQ  F---------  ----------  ----EVS---  --GTREAFV-  ----------  244
AmDICH   IK-QLNQERN  KSYEWNP---  -------SVL  F-------QS  KS--------  --SQQFEVSG  PSTNYEELNQ  ESIMIKRKLK  272
LvCYC    IKQQLNEARN  NNKGGDWI--  -------NNP  FNNV--IQSN  RN--------  --HQQFE---  --SREAAFV-  ----------  282
SiCYC1A  IK-KLNESRN  MN--------  -------NNL  F--------A  SNSQPVLHCP  -ITNEAT---  --ATQQDLIQ  ESSVIKRMLR  272
SiCYC1B  IK-KLNESRN  MD--------  -------NNL  F--------A  SNSRPVLQCP  -ITNEAT---  --ATHEDLIQ  ESSVIKRMLR  276
GhCYC2   IK-KLDDGSK  KVLNECRC--  ---PVSDSNL  TL--------  ----------  -QSSVWS---  ----------  ----------  265
ObCYC1   IK-TLNESRN  M---------  -------NNL  FEVC--RPSA  SNSQPILHCP  -IINEATAAT  L-AAPEDLIQ  ESNVVKRMLR  291
LjCYC2   NN-NNNTTLK  KKL-------  ---PATENSL  ILQQPKSPPR  MLHHPHPHHL  -HHHLVGGEA  P-RDDFNVIE  ESIVIKRKLK  326
LjCYC1   PE-TENPQMQ  LHQF------  -------SSE  F---------  -----QPHQE  -LHR------  --DDDFKVFE  ESIVIKRKLK  290
LjCYC3   ED-QRCPANT  TQI-------  -MHQLRSSSP  I---------  QLLQPYPHLM  DNSEAVSGGG  G-GDGFNVIE  ESIMIKRNMM  313
HaCYC2C  IK-ELDNDLK  KIPDDYPC--  --HALSPSNT  TL--------  ----------  -QSNSWGQFE  SQSDYNDILH  GSMLEQRFSV  338
PhCYC1C  IK-KLNESRN  M---------  -------NNL  FEVC--RPSA  SNSQLILHCP  -ITDEATAAT  V-TATDHIIQ  ESNVVKRMLR  290
PhCYC1D  IK-KLNESRN  MGSNLNP---  -SVPIQRNNL  FEVC--RPSA  SN----IHCP  -ITNEATTAT  VAATPEDLIQ  ESNVIKRMLR  299
SlCYC1   TQ-IAPNREA  TAS-------  ----------  ----------  ----------  ----------  ----------  ----------  194
SlCYC2   TK-IETSHKS  ----------  -------ANF  F-------GK  KDISEMEQHF  -FKNKLQ---  ---ANKEIIE  GSGVTKIKIM  249
NbCYC1A  SQ-IEAKINS  SSV-------  ---VVKEQTK  I---------  ----------  ----------  --DEIKEMIH  GSVLADSKNI  252
NbCYC1B  SQ-IEAKINS  SSV-------  ---VVKEQTK  I---------  ----------  ----------  --DEIKEMIH  GSVLVERKNI  254
NbCYC2A  SR-TETSQRS  TSQ-------  ---LIK-SSS  F-------LR  KDIKEMEEQF  -CLNKLQASQ  Q-VANKEIIQ  GFGVIKFKTK  331
NbCYC2B  SR-IETSQRS  TSQ-------  ---LFR-SSF  F-------LK  KDIREMEEQF  -CKN-IQVSQ  E-VANKRIIQ  GSGVTKIKIK  377
SsCYC    IK-KLNESRN  MD--------  -------PDL  N-------PS  NQIQPTLHCP  -LTNNVPA--  --ATTEDLIQ  ESIVIKRMLK  287

                                          580                                  600                                  620                                  640
AtTCP12  ----SS----  --SIFGDSE-  E---------  ---------C  ----YTSLSS  VRGMS-----  --TPREHNTT  SIATVD----  339
AtTCP18  ----S---NE  SYDLVNYKP-  N----SSFP-  -----VINHH  ----------  --------RS  QGAANSIEQH  QFTDLHYSFG  415
AtTCP1   L--------P  TKMGQSYNQN  N---GILML-  VDQSSSSNY-  ----NTFLPQ  NLDYSYD-QN  PFHDQT--LY  VVTDKNFPKG  349
IaTCP1   L--------P  MKIDQSYNQQ  NEYGTTFTL-  VDHGCSSNY-  ----NTILPQ  NLDYDYG-QN  PFIDQP--FS  ----------  326
AmCYC    --------H  P--VFGFHQ-  Q---------  -------NY-  ------GNAS  HENWD---QS  NLSSQSNQLC  AILN------  280
AmDICH   -------QNH  P-SMFGFQP-  E---------  ----------  --------NA  TENWDY--YS  NFTSQSNQLC  AILD------  308
LvCYC    -------HH  P--VFGFHQ-  Q---------  -------NYS  ----NASASS  HENWD-----  ----QSNQLC  AILN------  316
SiCYC1A  --------HH  Q-SFFGFHC-  A---------  ---------A  ----LPSPGV  DENWD---AG  SLTSQSN-LC  DILD------  310
SiCYC1B  --------HH  QSSFFGYHC-  ----------  ---------G  ----LPSPNV  NENWD---VG  SLTSQSN-LC  DILD------  314
GhCYC2   ----------  ----------  ----------  ----------  ----------  ----------  ----------  ----------  265
ObCYC1   --------HH  P-SFFGFHC-  ----------  ---------S  ----LPSPNV  NENWD---VS  SLTSQSN-FC  DILD------  328
LjCYC2   PWLMSSSSHH  HHHQQQQHH-  N----LVIP-  KEGSFNNNSE  Y---HSFPNS  SPNWDNNTSN  SATSRSN-FC  TIASMNLSTG  396
LjCYC1   QSLMSSSHHH  Q------NQ-  N----IGIP-  KEASFNNNNN  S-ECPSSFPM  SPNWD---AS  GATPRSN-IC  AIASVNLSTG  353
LjCYC3   ----SSNSHP  -------HE-  NH--HLAIPN  KEAGFNNNYI  DYPLLQPYST  TPNWEIA---  --TMNMNLST  CFMN------  368
HaCYC2C  ----SSS---  --TLYTYNH-  N----FVVS-  DEC-------  ----------  ----------  --------IS  QISSKGMP--  366
PhCYC1C  --------HH  P-SFFGFHC-  ----------  ---------S  ----LPSPNI  NENWD---VS  SLTSQSN-FC  DILDQ-----  328
PhCYC1D  --------HH  S-SFFGFHC-  ----------  ---------S  ----LPSPNV  NENWD---VS  SLTSQSN-FC  DILD------  336
SlCYC1   ----------  -------HY-  N---------  ----------  --------NS  TRNWNHDDVN  PTIMSSMDAS  TICCTSLPIG  229
SlCYC2   ----------  PSLILGFNP-  N----PYAP-  IESG--TYY-  ----GGSSLS  Q---------  ----------  GIVL------  281
NbCYC1A  --------KS  SSSTLGFHHL  N----LSGT-  EEARANSNY-  ----NSSSSS  TRNWDHD-VN  RTITNSSLSA  VMATISTSSQ  313
NbCYC1B  --------KS  SSSTLGFHHP  N----LSGT-  EEAAANSNY-  ----NSSSSS  TGNWDHD-IN  RTIMNSSLSA  VMASISASSQ  315
NbCYC2A  ----------  PSLILGFHP-  N----FSAP-  IEPS-TTYY-  ----ASSSFN  TENCDKGTTS  G---------  ----------  370
NbCYC2B  ----------  PSLVLDFYP-  N----FSAP-  IESS-TTYY-  ----GSSSFN  AEKLG-----  ----------  ----------  410
SsCYC    --------QY  P-SFFGFQQ-  N----LIIS-  RDL----NCN  ----LPSPNI  NDNWD---IN  SLTSQSN-LC  DILD------  334

                                          660                                  680
AtTCP12  --------EE  KS------PI  SSFSLYD---  -----YLCY*  357
AtTCP18  AK--------  ----------  -PRDLMHN--  ----YQNMY*  430
AtTCP1   ----------  ----------  ----------  -------FL*  352
IaTCP1   ----------  ----------  ----------  ---------*  327
AmCYC    ----------  ----------  -QHKFIN---  ---------*  287
AmDICH   ----------  ----------  -QHKFIN---  ---------*  315
LvCYC    ----------  ----------  -QHKFIN---  ---------*  323
SiCYC1A  ----------  ----------  -QHKFI----  ---------*  316
SiCYC1B  ----------  ----------  -QHKFI----  ---------*  320
GhCYC2   ----------  ----------  ----------  ------PMI*  269
ObCYC1   ----------  ----------  -QHKFIN---  ----RCRNI*  340
LjCYC2   LQIFGKSWEE  CT------NP  RPH-------  ---------*  414
LjCYC1   LQIFGKSWEE  CT------SP  HLH-------  ---------*  371
LjCYC3   ----------  ----------  ---Q------  --------W*  371
HaCYC2C  ----------  ----------  -KFKVIHEEQ  S---HRAMI*  382
PhCYC1C  ----------  ----------  -QHKFIN---  ----RSSNI*  340
PhCYC1D  ----------  ----------  -QHKFIN---  ----RSSNI*  348
SlCYC1   KKL-------  ----------  -KYYL-----  ---------*  237
SlCYC2   ----------  ----------  -EFKLYILNI  FFLTLIDVI*  300
NbCYC1A  GKILITKV--  ----------  -ESNLVI---  ---------*  328
NbCYC1B  GNLLQKWIPI  QSFNYFHNSP  PQFSLSIVNR  K--------*  347
NbCYC2A  ----------  ----------  ----------  ---------*  371
NbCYC2B  ----------  ----------  ----------  ---------*  411
SsCYC    ----------  ----------  -QHKFIN---  ----SSSNI*  346
```

Supplementary figure 3

Promoter regions of *SsCYC* and Δ*SsCYC* genes

```
                         -740                   -720                   -700
SsCYC_promoter   TGTCGAGTTT GGTGTTCTTA TAATAGAAGA GATGAAAATA AAGCAAACCT TCAACAAACA  60
                 ACAGCTCAAA CCACAAGAAT ATTATCTTCT CTACTTTTAT TTCGTTTGGA AGTTGTTTGT
dSsCYC_promoter  TATCGAGTTT GGTGTTCTTA TAATAGAAGA GATGAAATTA AAGCAAACCT TAAACAAACA  60
                 ATAGCTCAAA CCACAAGAAT ATTATCTTCT CTACTTTAAT TTCGTTTGGA ATTTGTTTGT

                         -680                   -660                   -640
SsCYC_promoter   GTTCAACGAG AACTCAAAAA CCCAGTCTTT AAATTGTATT TAAAAATAAT TACCCACCAA  120
                 CAAGTTGCTC TTGAGTTTTT GGGTCAGAAA TTTAACATAA ATTTTTATTA ATGGGTGGTT
dSsCYC_promoter  GTTCAACAAG AACTCAAAAA CCCAGTCTTT AAATTGTATT TAAAAATAAT TACCCACCAA  120
                 CAAGTTGTTC TTGAGTTTTT GGGTCAGAAA TTTAACATAA ATTTTTATTA ATGGGTGGTT

                         -620                   -600                   -580
SsCYC_promoter   ACAAAAAA-- -AGAAAAAAA GAAAAAGAAT GCATCTGAAA ATTACCATTC ATACATAGGG  177
                 TGTTTTTT-- -TCTTTTTTT CTTTTTCTTA CGTAGACTTT TAATGGTAAG TATGTATCCC
dSsCYC_promoter  AAAAAAAATA AAAAGAAAAA GAAAAAGAAT GCATCTGAAA ATTAGTAATC ATACATAGGG  180
                 TTTTTTTTAT TTTTCTTTTT CTTTTTCTTA CGTAGACTTT TAATCATTAG TATGTATCCC

                         -560                   -540                   -520
SsCYC_promoter   CAAGGAACCC ACTCTTCAAG AATTAGGGTT TTCAGAAGTT TTCTGTAAAT CAGGACAAGC  237
                 GTTCCTTGGG TGAGAAGTTC TTAATCCCAA AAGTCTTCAA AAGACATTTA GTCCTGTTCG
dSsCYC_promoter  CAAGAAACCC ACTCTTCAAG AATTAGGGTT TTCAGAAGTT TTCT-TAAAA CAGGACGAGC  239
                 GTTCTTTGGG TGAGAAGTTC TTAATCCCAA AAGTCTTCAA AAGA-ATTTT GTCCTGCTCG

                         -500                   -480                   -460
SsCYC_promoter   TCTAAACGTT TACAAACCCA AATTCTTCGC TCTTAATTCA CACTTCAAAA TGGTTCAAAC  297
                 AGATTTGCAA ATGTTTGGGT TTAAGAAGCG AGAATTAAGT GTGAAGTTTT ACCAAGTTTG
dSsCYC_promoter  ACTAAACGTT TACAAACCCA AATTATTCGC TCTTAATTCA CACATCAAAA TGGTTAAAAC  299
                 TGATTTGCAA ATGTTTGGGT TTAATAAGCG AGAATTAAGT GTGTAGTTTT ACCAATTTTG

                         -440                   -420                   -400
SsCYC_promoter   ACAGAAACAA AAAAA----- ----GGGTTT CAAAAAAATT ACACATTCAT TCAAACTAAA  348
                 TGTCTTTGTT TTTTT----- ----CCCAAA GTTTTTTTAA TGTGTAAGTA AGTTTGATTT
dSsCYC_promoter  CCAGAAACAA AAAAAAAAAA AATGGGGTTT CAAAAAAATT ACACAAACAT TCAGACTAAA  359
                 GGTCTTTGTT TTTTTTTTTT TTACCCCAAA GTTTTTTTAA TGTGTTTGTA AGTCTGATTT

                         -380                   -360                   -340
SsCYC_promoter   AGCCAAAAAG AAAAAGGAAA AATAAAACAA AACTGTAAAT TCTTTTAATA AATTGTTATT  408
                 TCGGTTTTTC TTTTTCCTTT TTATTTTGTT TTGACATTTA AGAAAATTAT TTAACAATAA
dSsCYC_promoter  AGCCAAAA-G AAAAAGGAAA AATAAAACAA AACTGTAAAT TCTTTTAATA AATTGTTATT  418
                 TCGGTTTT-C TTTTTCCTTT TTATTTTGTT TTGACATTTA AGAAAATTAT TTAACAATAA

                         -320                   -300                   -280
SsCYC_promoter   ATTTTAACAG CTGCAAA-GA AGGAAATTAA GAAAAAGCTG TGGACCAAAA GAGAAGAGAT  467
                 TAAAATTGTC GACGTTT-CT TCCTTTAATT CTTTTTCGAC ACCTGGTTTT CTCTTCTCTA
dSsCYC_promoter  ATTTTAACAG CTGCAAAAGA AGGAAATTAA GAAAAAGCTG TGGACCAAAA GAGAAGAGAT  478
                 TAAAATTGTC GACGTTTTCT TCCTTTAATT CTTTTTCGAC ACCTGGTTTT CTCTTCTCTA

                         -260                   -240                   -220
SsCYC_promoter   TTGTGGGTTG CTTACTGTCA TACCAGAAAT GGAGAGATGA GTGTTTATTC TGAGTGTGGG  527
                 AACACCCAAC GAATGACAGT ATGGTCTTTA CCTCTCTACT CACAAATAAG ACTCACACCC
dSsCYC_promoter  TTGTGGGTTG CTTACTGTCA TACCAGAAAT GGAGAGATGA GTGTTTATTC TGAGTGTGGG  538
                 AACACCCAAC GAATGACAGT ATGGTCTTTA CCTCTCTACT CACAAATAAG ACTCACACCC

                         -200                   -180                   -160
SsCYC_promoter   TGAAGAGGAC ACGAATATAA TGGGGCCCCA TTACAAAAGC AAAGAAAATG AAGTTAAATT  587
                 ACTTCTCCTG TGCTTATATT ACCCCGGGGT AATGTTTTCG TTTCTTTTAC TTCAATTTAA
dSsCYC_promoter  TGAAGAGGAC ACGAATATAA TGGGGCCCCA TTACAAAAGC AAAGAAAATG AAGTTAAATT  598
                 ACTTCTCCTG TGCTTATATT ACCCCGGGGT AATGTTTTCG TTTCTTTTAC TTCAATTTAA

                         -140                   -120                   -100
SsCYC_promoter   TGTCCATTGC GTAGGGTGAT AGACAAAGGT ACGTACGTTT TTAGACTCTA AAAAACTGTT  647
                 ACAGGTAACG CATCCCACTA TCTGTTTCCA TGCATGCAAA AATCTGAGAT TTTTTGACAA
dSsCYC_promoter  TGTCCATTGC GTAGGGTGAT AGACAAAGGT ACGTACGTTT TTAGACTCTA AAAAACTGTT  658
                 ACAGGTAACG CATCCCACTA TCTGTTTCCA TGCATGCAAA AATCTGAGAT TTTTTGACAA

                         -80                    -60                    -40
SsCYC_promoter   TTAATCACCA GCTTTTTTTA ATTTTTATTT TCACCATGTA TACATGCAGA GGTCACACAC  707
                 AATTAGTGGT CGAAAAAAAT TAAAAATAAA AGTGGTACAT ATGTACGTCT CCAGTGTGTG
dSsCYC_promoter  TTAATCACCA GCTTTTTTTT ATTTTTATTT TCACCATGTA TACATGCAGA GGTCACACAC  718
                 AATTAGTGGT CGAAAAAAAA TAAAAATAAA AGTGGTACAT ATGTACGTCT CCAGTGTGTG

                         -20                    1                 20
SsCYC_promoter   CCTACAAAAA CCTCACGGCT CTTCTTTGGA AAACCTAACT ACACACACTC ACACACCAAA  767
                 GGATGTTTTT GGAGTGCCGA GAAGAAACCT TTTGGATTGA TGTGTGTGAG TGTGTGGTTT
dSsCYC_promoter  CCTACAAAAA CCTCACGGCT CTTCTTTGGA AAACCTAACT ACACACACTC ACACACCAAA  778
                 GGATGTTTTT GGAGTGCCGA GAAGAAACCT TTTGGATTGA TGTGTGTGAG TGTGTGGTTT
```

Supplementary figure 3 continued

```
                                   40                    60                    80
SsCYC_promoter    CTAAGAGAAA TTCACACTTC ATATCCCTCT CTCTCTCTCT CTCT----GC ACATATAGAG  823
                  GATTCTCTTT AAGTGTGAAG TATAGGGAGA GAGAGAGAGA GAGA----CG TGTATATCTC
dSsCYC_promoter   CTATGAGAAA ATCACACTTC ATATCCCTCT CTCTCTCTCT CTCTCTCTGC ACATATTGAG  838
                  GATACTCTTT TAGTGTGAAG TATAGGGAGA GAGAGAGAGA GAGAGAGACG TGTATAACTC

                                  100                   120                   140
SsCYC_promoter    ATACCATCAA ACCCTAGCTA CCCTTCTTTT TATTAGTACC TTTTTATGCT TTCAAGATTT  883
                  TATGGTAGTT TGGGATCGAT GGGAAGAAAA ATAATCATGG AAAAATACGA AAGTTCTAAA
dSsCYC_promoter   ATACCATCAA ACCCTAGCTA CCCTTCTTTT TATTAGTACC TTTTTATGCT TTCAAGATTT  898
                  TATGGTAGTT TGGGATCGAT GGGAAGAAAA ATAATCATGG AAAAATACGA AAGTTCTAAA

                                  160                   180                   200
SsCYC_promoter    TGTTTTCTCG ATCATGGATT AATTAATGGT ACCGTTGAAC CAAAATGAAT ACAATACTAA  943
                  ACAAAAGAGC TAGTACCTAA TTAATTACCA TGGCAACTTG GTTTTACTTA TGTTATGATT
dSsCYC_promoter   TGTTTTCTCG ATCATGGATT AATTAATGGT ACTGTTGAAC CAAAATGAAT ACAATACTAA  958
                  ACAAAAGAGC TAGTACCTAA TTAATTACCA TGACAACTTG GTTTTACTTA TGTTATGATT

                                  220                   240                   260
SsCYC_promoter    GCAATTCAAT TCGAAATTGT AGTAGTAATA ATTAATGGTG GTGGTAGTAG TAATGAAGAA 1003
                  CGTTAAGTTA AGCTTTAACA TCATCATTAT TAATTACCAC CACCATCATC ATTACTTCTT
dSsCYC_promoter   GCAATACAAT ACGAAATGGT AGTAGTAATA ATTAATAGTG TTGGTAGTAG TAATGAAGAA 1018
                  CGTTATGTTA TGCTTTACCA TCATCATTAT TAATTATCAC AACCATCATC ATTACTTCTT

                                  280                   300                   320
SsCYC_promoter    TTAATCATTA TTTTGAGGGA TTGACTGATT TTATCTGATG TTGCTGAAGT GGAGGAACTG 1063
                  AATTAGTAAT AAAACTCCCT AACTGACTAA AATAGACTAC AACGACTTCA CCTCCTTGAC
dSsCYC_promoter   TTAATCATTA TTTTGAGGGA GTGGCTGATT TTATCTGATG TTGCTGAAGT GGAGGAACTG 1078
                  AATTAGTAAT AAAACTCCCT CACCGACTAA AATAGACTAC AACGACTTCA CCTCCTTGAC

                                  340                   360                   380
SsCYC_promoter    TAGCATAACT GTAGATTACA TTTTGAATTG ACAATAAATT TTTGTACAGC GCTAAAAGTG 1123
                  ATCGTATTGA CATCTAATGT AAAACTTAAC TGTTATTTAA AAACATGTCG CGATTTTCAC
dSsCYC_promoter   TAGCATAACT GTAGATTACA TTTTGAATTG ACAATAAATT TTTGTACTGC GCTAAAAGTG 1138
                  ATCGTATTGA CATCTAATGT AAAACTTAAC TGTTATTTAA AAACATGACG CGATTTTCAC

                                  400                   420                   440
SsCYC_promoter    AAGAAGATAA AGTTAAACTA GGTAGTTTTT TTTATTATTA TTATCACCAA TTTAATACCC 1183
                  TTCTTCTATT TCAATTTGAT CCATCAAAAA AAATAATAAT AATAGTGGTT AAATTATGGG
dSsCYC_promoter   AAGAAGATAA AGTTAAACTA GGTAGTTTTT TTTATTATTA TTATCACCAA TTTAATACCC 1198
                  TTCTTCTATT TCAATTTGAT CCATCAAAAA AAATAATAAT AATAGTGGTT AAATTATGGG

                                  460                   480                   500
SsCYC_promoter    TATTCAGTGC ATCTGAACAA ATTTTATTTG GAGATTAAAG AAAGGGTACA ATACTTTTAC 1243
                  ATAAGTCACG TAGACTTGTT TAAAATAAAC CTCTAATTTC TTTCCCATGT TATGAAAATG
dSsCYC_promoter   TATTCAGTGC ATCTGAACAA ATTTTATTTG GAGATTAAAG AAAGGGTACA ATACTTTTAC 1258
                  ATAAGTCACG TAGACTTGTT TAAAATAAAC CTCTAATTTC TTTCCCATGT TATGAAAATG

                                  520                   540                   560
SsCYC_promoter    TCCTGAAAAC CCAAAATTTT TCCCAATTCA TCATATCTTC GTCCTCCATT TTTCACCTAC 1303
                  AGGACTTTTG GGTTTTAAAA AGGGTTAAGT AGTATAGAAG CAGGAGGTAA AAAGTGGATG
dSsCYC_promoter   TCCTGAAAAC CCAAAATTTT TCCCAATTCA TCATATCTTC GTCCTCCATT TTTCACCTAC 1318
                  AGGACTTTTG GGTTTTAAAA AGGGTTAAGT AGTATAGAAG CAGGAGGTAA AAAGTGGATG

                                  580                   600                   620
SsCYC_promoter    ACGCTAGCCT TCCAGTCTTT CTCAGGCAAA GTCATTTTCT TTGGTGTAAT ATAAAGCAAA 1363
                  TGCGATCGGA AGGTCAGAAA GAGTCCGTTT CAGTAAAAGA AACCACATTA TATTTCGTTT
dSsCYC_promoter   ACGCTAGCCT TCCAGTCTTT CTCAGGCAAA GTCATTTTCT TTGGTGTAAT ATAAAGCAAA 1378
                  TGCGATCGGA AGGTCAGAAA GAGTCCGTTT CAGTAAAAGA AACCACATTA TATTTCGTTT

                                  640                   660                   680
SsCYC_promoter    GACAAGAAAA ATTTGCATAT AACTATATAT ATACACACAC ATTTATCATC AATAATAAAT 1423
                  CTGTTCTTTT TAAACGTATA TTGATATATA TATGTGTGTG TAAATAGTAG TTATTATTTA
dSsCYC_promoter   GACAAGAAAA ATTTGCATAT AACTATATAT ACACACACAC ATTTATCATC AATAATAAAT 1438
                  CTGTTCTTTT TAAACGTATA TTGATATATA TGTGTGTGTG TAAATAGTAG TTATTATTTA

                                  700                   720                   740
SsCYC_promoter    AAGTGATGCT AGAGTTATTG ATCTCTTGAG GAAAAAAAAA GAAAAAAAAA ACCTTAGTTC 1483
                  TTCACTACGA TCTCAATAAC TAGAGAACTC CTTTTTTTTT CTTTTTTTTT TGGAATCAAG
dSsCYC_promoter   AAGTGATGCT AGAGTTAT-G ATCTCTTGAG GACAAAAAAA --AGATAAAA ACCTTAGTTC 1495
                  TTCACTACGA TCTCAATA-C TAGAGAACTC CTGTTTTTTT --TCTATTTT TGGAATCAAG

                                  760                   780                   800
SsCYC_promoter    TCATTCTGGA GAAACCTTCA AACCAGCTCT CACAGGTTGA TTGCATAAAC AATAAATATG 1543
                  AGTAAGACCT CTTTGGAAGT TTGGTCGAGA GTGTCCAACT AACGTATTTG TTATTTATAC
dSsCYC_promoter   TCATTCTGGA GAAACCTTCA AACCAGCTCT CACAGGTTGA TTGCATAAAC AATAAATATG 1555
                  AGTAAGACCT CTTTGGAAGT TTGGTCGAGA GTGTCCAACT AACGTATTTG TTATTTATAC
```

Supplementary figure 3 continued

	820		840		860	
SsCYC_promoter	GTTAAAAAAT TCAAGAACTT	AAGGGTTTCT TTCCTTCTTT	TTTTCTTTTT	TATGTAAGAA		1603
	CAATTTTTTA AGTTCTTGAA	TTCCCAAAGA AAGGAAGAAA	AAAAGAAAAA	ATACATTCTT		
dSsCYC_promoter	GTTAAAAAAT TCAAGAACTT	GAGGGTTTCT TTCCTTCTTT	TTTTCTTTTT	CATGTAAGAA		1615
	CAATTTTTTA AGTTCTTGAA	CTCCCAAAGA AAGGAAGAAA	AAAAGAAAAA	GTACATTCTT		

	880		900		920	
SsCYC_promoter	ATTAATTAGG GTTTATTAAC	CCTTCTTCCC CTCCCCTCTC	GAAAAAAAGA	AGA		1656
	TAATTAATCC CAAATAATTG	GGAAGAAGGG GAGGGGAGAG	CTTTTTTTCT	TCT		
dSsCYC_promoter	ATTAATTAGG GTTTATTAAC	CCTTCTTCCC CTCCCCTCTC	GCAAAAAAGA	AGA		1668
	TAATTAATCC CAAATAATTG	GGAAGAAGGG GAGGGGAGAG	CGTTTTTTCT	TCT		

Supplementary figure 4

Nucleotide sequence of *SsCYC* and Δ*SsCYC* genes

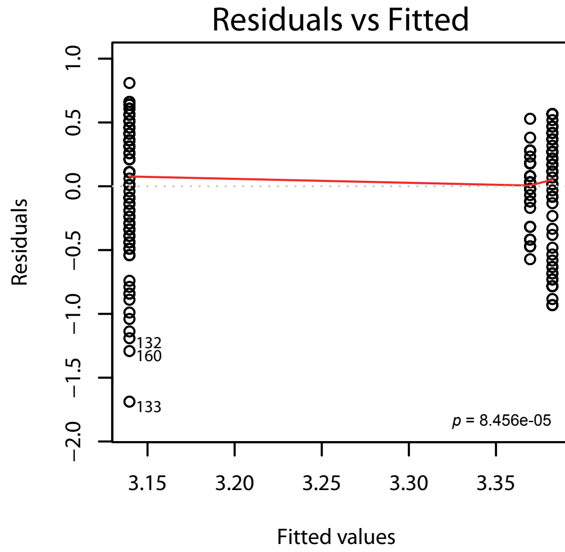
```
                                      20                   40                   60
                                       |                    |                    |
 SsCYC  ATGTTTAGCA AGAGCACATA CCTTCATGTT CCACAGGTTT CACCATCTCT TCAATCTCGT 60
dSsCYC  ATGTTTAGCA AGAGCACATA CCTTCATGTT CCACAGGTTT CACCATAT-- --------GT 50
                                      80                  100                  120
 SsCYC  GCCTCTACTT CTTTGGTTGA CCTTAATGGA GGTGAAATCT TGCTTCATCA CCACCACCAC 120
dSsCYC  GCCTCTACTT CTTTGGTTGA CCTTAATGGA GGTGAAATCT TGCTTCATCA CCACCACCAC 110
                                     140                  160                  180
 SsCYC  CATGACATGC TTTCCAGCCA TTACTTAGCC GTGAATGCCC CGTTTCTTGA GGCTTCCTCC 180
dSsCYC  CATGACATGC TTTCCAGCCA TTACTTAGCC GTGAATGCCC CGTTTCTTGA GGCTTCCTCC 170
                                     200                  220                  240
 SsCYC  TTGTATAACC AAGATGCTAT TGTTGGTCTA AATGAAGATC CTTCTGCCAT GGCCAACACG 240
dSsCYC  TTGTATAACC AAGATGCTAT TGTTGGTCTA AATGAAGATC CTTCTGCCAT GGCCAACACG 230
                                     260                  280                  300
 SsCYC  TTTCCAAGGA AGCAAACGGT GAAAAAAGAT CGGCACAGTA AAATTGTTAC AGCTCAAGGG 300
dSsCYC  TTTCCAAGGA AGCAAACGGT GAAAAAAGAT CGGCACAGTA AAATTGTTAC AGCTCAAGGG 290
                                     320                  340                  360
 SsCYC  CCGAGGGATC GGAGAGTCAG GCTTTCTATT GGCATAGCAA GAAAGTTCTT TGATCTTCAA 360
dSsCYC  CCGAGGGATC GGAGAGTCAG GCTTTCTATT GGCATAGCAA GAAAGTTCTT TGATCTTCAA 350
                                     380                  400                  420
 SsCYC  GAAATGCTAG GTTTTGACAA GCCAAGTAAA ACCCTTGACT GGTTGCTCAC TAAATCTAAA 420
dSsCYC  GAAATGCTAG GTTTTGACAA GCCAAGTAAA ACCCTTGACT GGTTGCTCAC TAAATCTAAA 410
                                     440                  460                  480
 SsCYC  GCAGCCATTA AGGAGCTGGT GCAGGCTAAG AAAAGTGGGA GTGGGAGTGC TAAGAGCATT 480
dSsCYC  GCAGCCATTA AGGAGCTGGT GCAGGCTAAG AAAAGTGGGA GTGGGAGTGC TAAGAGCATT 470
                                     500                  520                  540
 SsCYC  TCTTCCCCTT CTGAATGCGA GGTAGTGTCT GCAGGAAATG GTGAAACTTT CGAAAATGGC 540
dSsCYC  TCTTCCCCTT CTGAATGCGA GGTAGTGTCT GCAGGAAATG GTGAAACTTT CGAAAATGGC 530
                                     560                  580                  600
 SsCYC  AGCTATTTGG ATGCGGAATC AAAGAAGAAA TCACTGCCCC TGAATCCTAA TTACAAGTGT 600
dSsCYC  AGCTATTTGG ATGCGGAATC AAAGAAGAAA TCACTGCCCC TGAATCCTAA TTACAAGTGT 590
                                     620                  640                  660
 SsCYC  AAAGAATATT CAAAAGATCC ACAGCAGTCT GCATTAAATC TTGCAAAAGT ATCAAGGGCT 660
dSsCYC  AAAGAATATT CAAAAGATCC ACAGCAGTCT GCGTTAAATC TTGCAAAAGT ATCAAGGGCG 650
                                     680                  700                  720
 SsCYC  AAGGCAAGAG CAAGGGCCAG AGAAAGAACT AGAGAGAAAA TGTGCATCAA GAAGCTTAAT 720
dSsCYC  AAGGCAAGAG CAAGGGCCAG AGAAAGAACT AGAGAGAAAA TGTGCATCAA GAAGCTTAAT 710
                                     740                  760                  780
 SsCYC  GAATCAAGAA ACATGGATCC TGATTTGAAC CCTTCAAACC AAATTCAGCC TACCCTCCAC 780
dSsCYC  GAATCAAGAA ACATGGATCC TGATTTGAAC CCTTCAAACC AAATTCAGCC TACCATCCAC 770
                                     800                  820                  840
 SsCYC  TGTCCCTTAA CTAATAATGT ACCTGCTGCA ACAACTGAAG ATTTAATTCA AGAATCCATT 840
dSsCYC  TGTTCCTTAA CTAATAATGT ACCTGCTGCA ACAACTGAAG ATTTAATTCA AGAATCCATT 830
                                     860                  880                  900
 SsCYC  GTCATTAAAA GGATGTTGAA ACAGTACCCT TCATTTTTTG GATTTCAACA AAACCTTATC 900
dSsCYC  GTCATTAAAA GGATGTTGAA ACAGTACCCT TCATTTTTTG GATTTCAACA AAACCTTATC 890
                                     920                  940                  960
 SsCYC  ATTTCAAGGG ATTTGAACTG CAATCTCCCT TCTCCTAATA TCAACGATAA TTGGGATATC 960
dSsCYC  ATTTCAAGGG ATTTGAACTG CAATCTCCCT TCTCCTAATA TCAACGATAA TTGGGATATC 950
                                     980                1,000                1,020
 SsCYC  AATAGCTTAA CCTCACAATC CAACCTCTGT GACATTTTGG ATCAGCATAA GTTCATGAAT 1020
dSsCYC  AATAGCTTAA CCTCACAATC CAACCTGTGT GACATTTTGG ATCAGCACAA GTTCATGAAT 1010

 SsCYC  AGCTCTTCAA ATATATAG 1038
dSsCYC  AGCTCTTCAA ATATATAG 1028
```

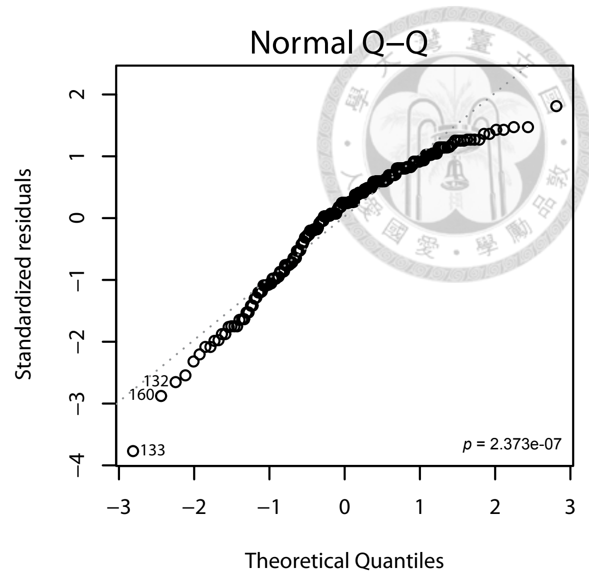
Supplementary figure 5

Amino acid sequence of *SsCYC* and Δ*SsCYC* genes

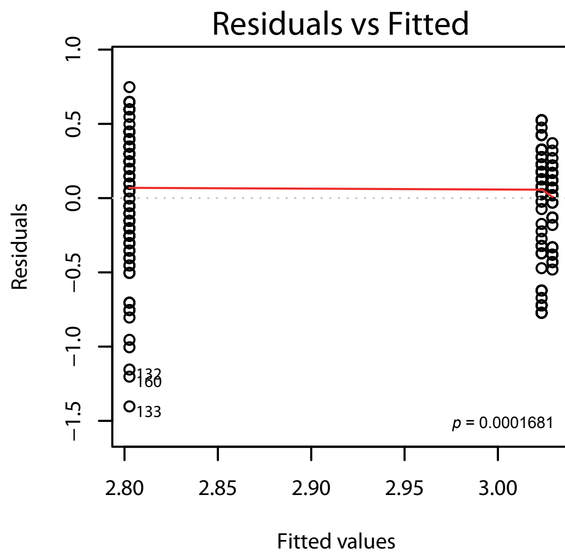
Supplementary figure 6

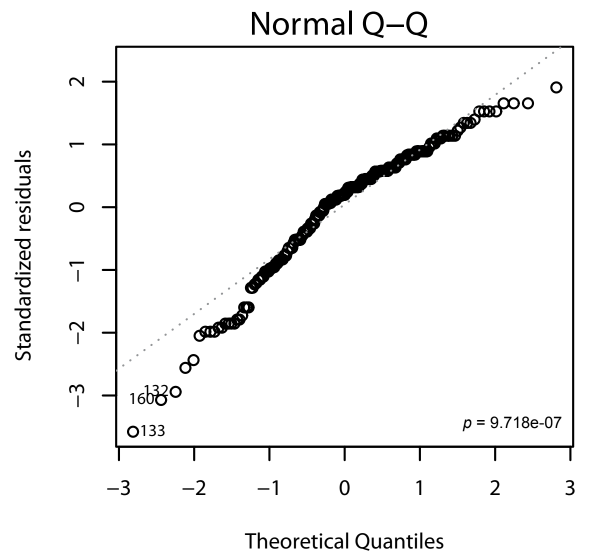
A. Total length ~ treatments



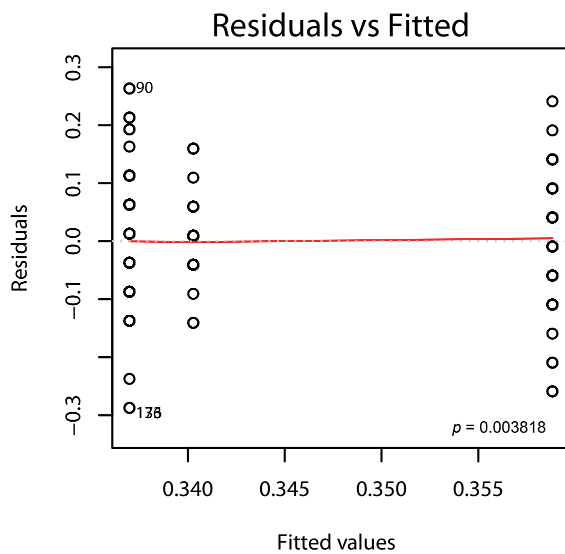


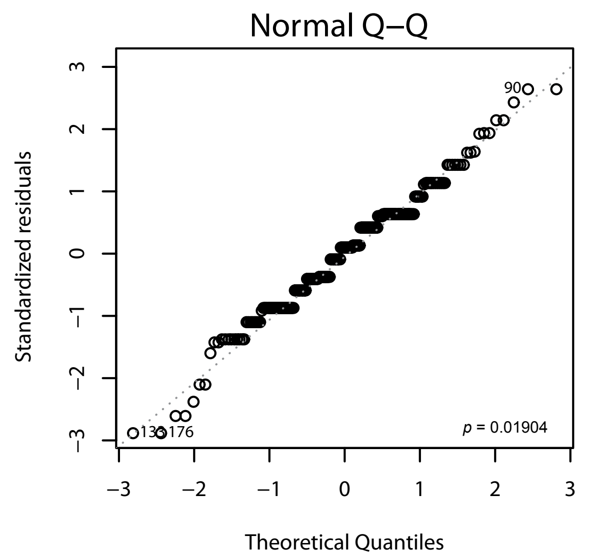
B. Tube length ~ treatments





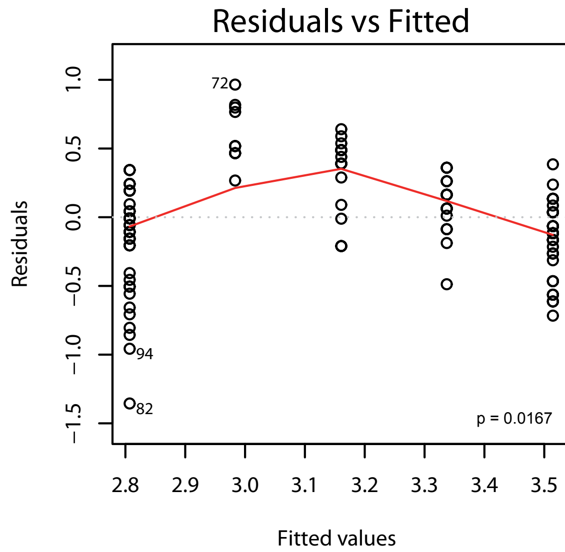
C. Lobe length ~ treatments

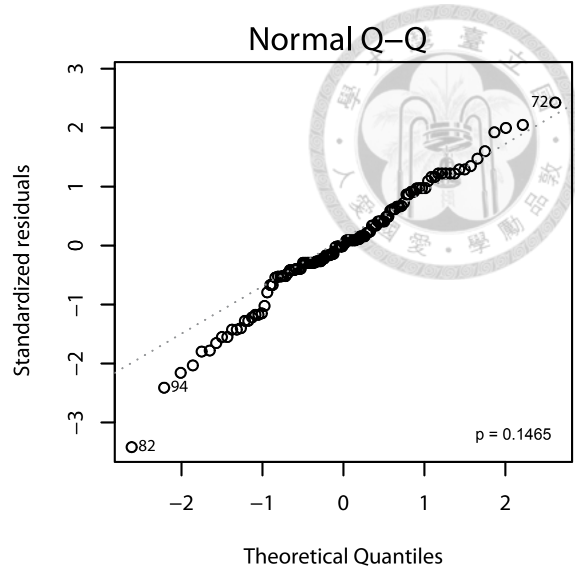




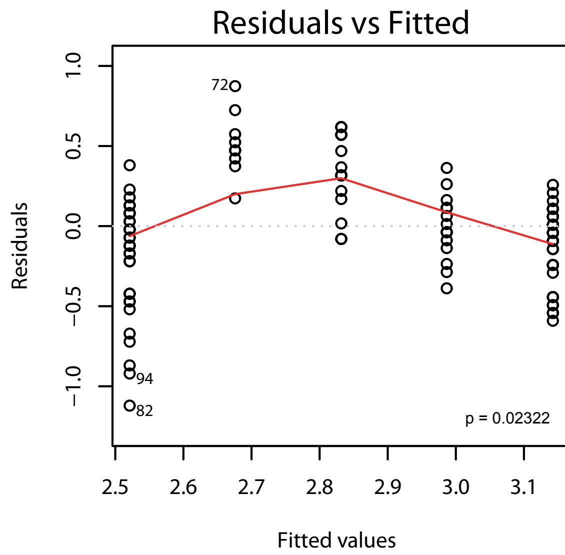
Supplementary figure 7

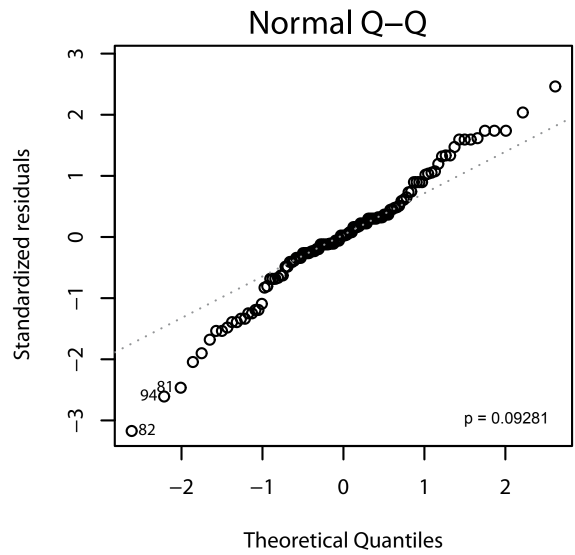
A. Total length ~ treatments



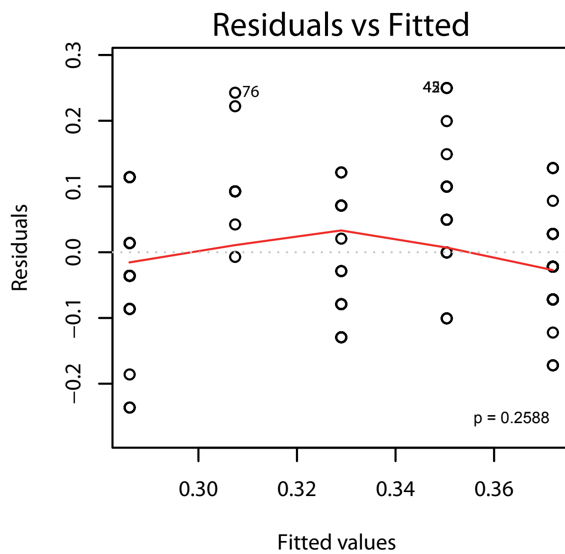


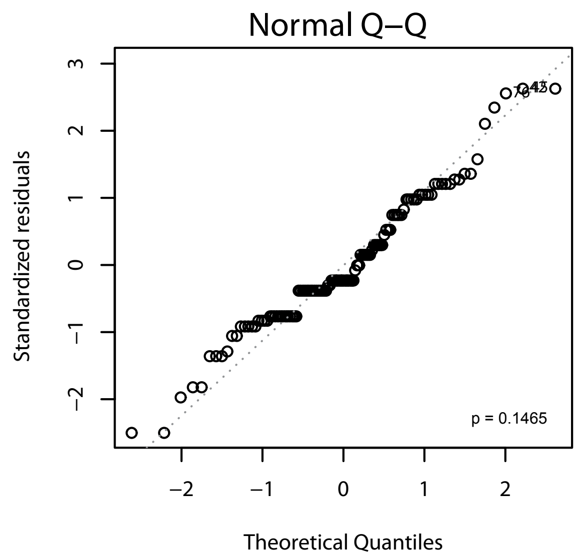
B. Tube length ~ treatments





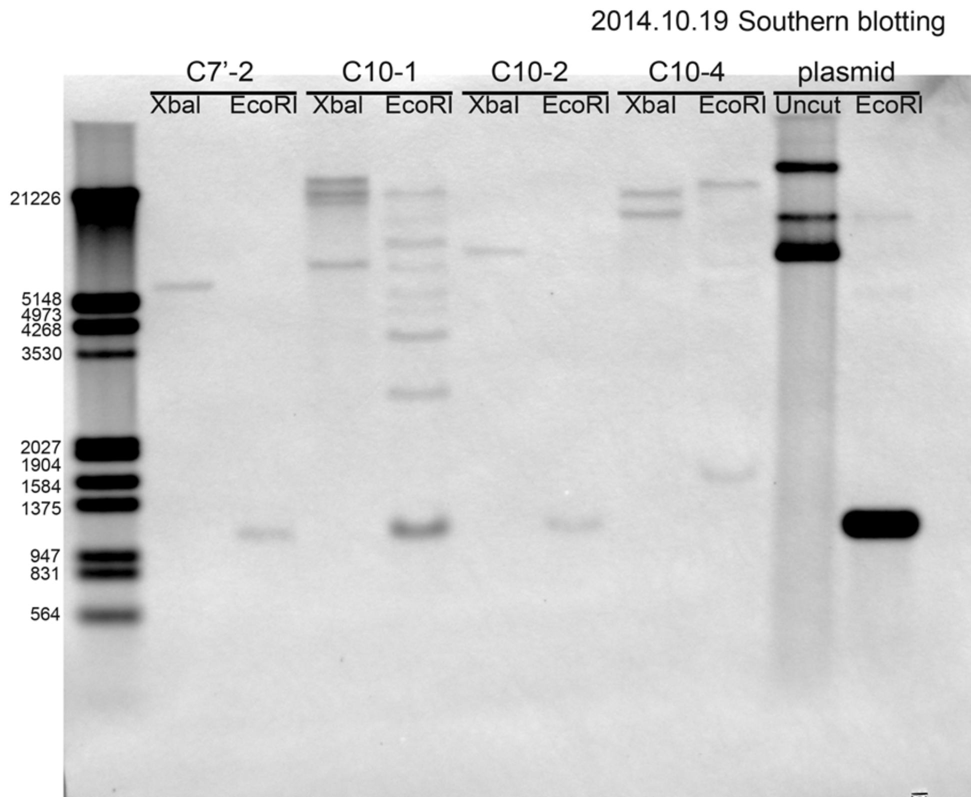
C. Lobe length ~ treatments



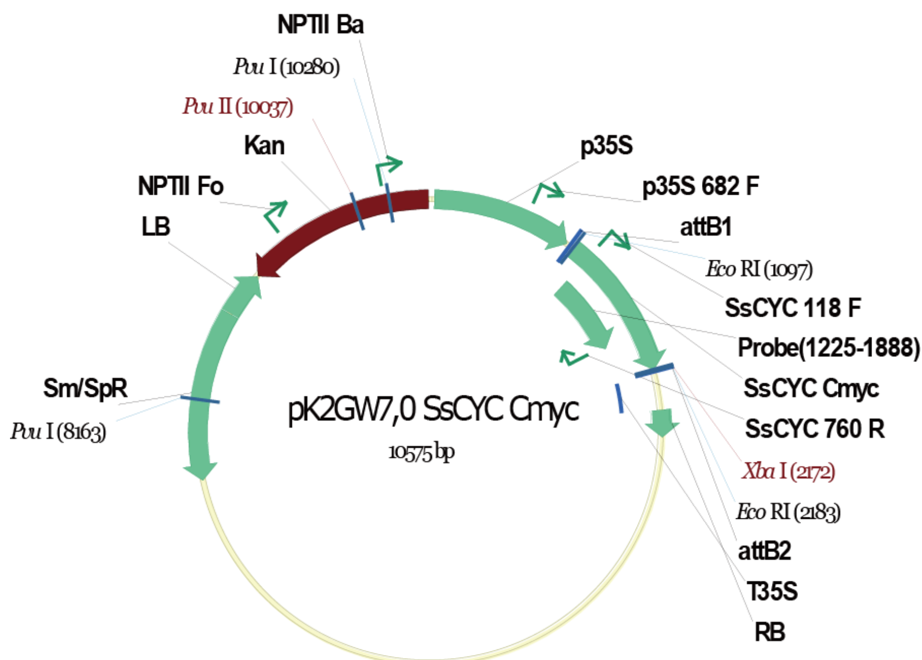


Supplementary Figure 8

A



B



Supplementary figure 8 Southern blotting of T_0 transgenic plants

A, the Southern blotting result of transgenic tobacco plant T_0. The 7.5 μg genomic DNA was digested by the labeled restriction enzymes and hybridized the *SsCYC* specific DIG-labeled probe (as shown in Fig 4B).Single copy of T DNA insertion could be seen in the C7'-2 and C10-2 lines. Double copy of T DNA insertions could be seen in the C10-4 line. Multiple bands in the C10-1 line might be resulted from incomplete digestion of genomic DNA. In plasmid control, triple bands could be seen in the uncut treatment, presenting the circular (upper), linear (middle) and supercoiled (lower) forms of secondary structure; a single band could be seen in EcoR1 digested treatment, showing the full coding sequence of *SsCYC* gene.

B, the restriction map of pK2GW7,0_SsCYC_cMyc vector. Note that there is only one XbaI cutting site inside the T DNA region and two EcoRI cutting sites beside the *SsCYC* gene.