Characterization and identification of *Catharanthus roseus* epigenetic-related genes that in response to peanut witches'-broom phytoplasma-mediated infection

Hsin-I Tseng^{1,4,5}, Chan-Pin Lin^{1,4,5}, Shih-Shun Lin^{1,2,3,6}

- ¹Institute of Biotechnology, College of Bioresources and Agriculture, National Taiwan University, 81, Chang-Xing St., Taipei, Taiwan 106
- ²Genome and Systems Biology Degree Program, National Taiwan University, 1, Sec. 4, Roosevelt Rd., Taipei, Taiwan 106
- ³Agricultural Biotechnology Research Center, Academia Sinica, 128, Sec. 2, Academia Rd., Taipei, Taiwan 115
- ⁴Department of Plant Pathology and Microbiology, College of Bioresources and Agriculture, National Taiwan University, 1 Sec. 4. Roosevelt Rd., Taipei, Taiwan 106
- ⁵These authors contributed equally to this work

⁶Corresponding author; E-mail: linss01@ntu.edu.tw; Tel; +886 2 33666023; Fax: +886 2 33666001

ABSTRACT

Tseng, H.I., Lin, C.P., and Lin S.S. 2014. Characterization and identification of *Catharanthus roseus* epigenetic-related genes that in response to peanut witches'-broom phytoplasma-mediated infection. Plant Pathol. Bull. 23:67-77.

The high-throughput transcriptome profile of *Catharanthus roseus* in the ContigViews database provides useful information on the interaction between phytoplasma and plant. In this study, several post-transcriptional gene silencing (PTGS)- and RNA-directed DNA methylation (RdDM)-related genes were identified in *C. roseus*, which might be involved in peanut witches'-broom (PnWB) phytoplasmamediated leafy flower formation. The evolutionary relationships and expression profiles of these genes in healthy flowers (HF) and Stage 4 (S4) PnWB-infected leafy flowers were also characterized. Notably, *ARGONAUTE 5* (*CrAGO5*) and *RNA-DEPENDENT RNA POLYMERASE 1* (*CrRDR1*) in *C. roseus* were significantly up-regulated in S4 leafy flowers, whereas other PTGS- and RdDM-related genes maintained normal expression levels. These results suggested that CrAGO5 and CrRDR1 might involve in PnWB-mediated leafy flower formation and plant defense.

Keywords: Next-generation sequencing, *Catharanthus roseus*, peanut witches'-broom phytoplasma, post-transcriptional gene silencing, RNA-directed DNA methylation

Peanut witches'-broom (PnWB) phytoplasma mediates virescence and phyllody symptoms (herein referred to as leafy flower) in *Catharanthus roseus* (1). Our previous study demonstrated that the gene expression tendency of *C. roseus* was to convert the reproductive stage to the vegetative stage during PnWB infection (2). In addition, epigenetic regulations, which are controlled by post-transcriptional gene silencing (PTGS) (3, 4) and RNA-directed DNA methylation (RdDM) (5, 6), might also be involved in leafy flower formation.

In the PTGS pathway, 21- or 22-nt microRNAs (miRNAs) were generated by the DICER-LIKE 1 (DCL1) and HYPONASTIC LEAVES 1 (HYL1) complex in the nucleus. Subsequently, the double-stranded miRNA was transferred by HASTY (HST) from the nucleus to the cytoplasm and loaded into ARGONAUTE 1 (AGO1) for mRNA recognition and cleavage (4). DCL2 and DCL4 are responsible for short-interfering RNA (siRNA) biogenesis in the cytoplasm, which has been demonstrated for plant defense mechanisms (7). DCL3 produces 24-nt siRNAs, which play an important role in mediating RdDM (8).

Total RNA was extracted from 10 healthy flowers (HF) and Stage 4 (S4) PnWB-infected leafy flowers using TRIzol reagent (Invitrogen) according to the manufacturer's protocol. The deep sequencing of small RNAs of HF and S4 PnWB-infected leafy flowers was analyzed using SOLiD 4 system (Life Technologies) by the next-generation sequencing (NGS) core facility of Techcomm at National Taiwan University (9). The small RNA profiles showed that 24-nt small RNAs were significantly decreased (approximately 62% reduction) in the S4 sample, whereas the 21- and 22-nt small RNA levels were similar in the HF and S4 samples, suggesting that PnWB specifically interferes with 24-nt small RNA biogenesis (Fig. 1A).

The whole transcriptome database of *C. roseus* were published and built on ContigViews database (www.contigviews.bioagri.ntu.edu.tw) (2). We further identified epigenetic-related genes of *C. roseus*, such as DCLs and AGOs, from the ContigViews database. Due to the large protein size of DCLs (approximately 2,000 amino acids), each DCL ortholog of C. roseus was divided into 2 contig sequences. DDS2790 belongs to the 5' end of CrDCL1, and DDS11937 belongs to the 3' end of CrDCL1. DDS44869 belongs to the 5' end of CrDCL2, whereas DDS13971 belongs to the 3' end of CrDCL2. DDS11443 belongs to the 5' end of CrDCL3, whereas DDS16373 belongs to the 3' end of CrDCL3. Finally, DDS112374 belongs to the 5' end of CrDCL4, whereas DDS3193 belongs to the 3' end of CrDCL4. To compare the amino acid sequences of the CrDCLs with Arabidopsis orthologs, 2 contig sequences were linked based on their alignment with Arabidopsis ortholog sequences. The missing amino acid residues in the junction area are denoted with an "X" in Figure 2A. CrDCL2 and CrDCL4 are represented as partial sequences because the C-terminus of CrDCL2 and a few amino acids of the N-terminus of CrDCL4 were missing (Fig. 2A). An alignment comparison of the DCLs of C. roseus and Arabidopsis showed that DEXDc, HELICc, PAZ and 2 RIBOc domains were highly conserved (Fig. 2A). In addition, CrDCL1 and AtDCL1 showed 78.5% amino acid sequence similarity between C. roseus and Arabidopsis, whereas the other DCLs showed 49.7 to 60.8% similarity (Fig. 1B). Moreover, the phylogenetic tree showed that 4 CrDCLs were classified into 4 groups with AtDCLs based on their sequence homologies (Fig. 1C).

Ten AGOs with specific biological functions exist in Arabidopsis. AtAGO1 (AT1G48410) is the most important protein and responds to most miRNA and siRNA-mediated mRNA cleavage or translational inhibition (10). In the *C. roseus* transcriptome database, only *CrAGO1* (DDS12000), *CrAGO4* (DDS13330), *CrAGO5* (DDS58064), *CrAGO7* (DDS11464, and DDS47116), *CrAGO10.1* (DDS42136) and *CrAGO10.2* (DDS42372) were identified, indicating that these genes were expressed in flower tissues of HF and S4. Note that DDS11464 belongs to the 5' end of *CrAGO7* (Fig. 3A). In addition, AOG7 is represented as a partial sequence because its N-terminus was missing (Fig. 3A). Two CrAGO10 orthologs exist in *C. roseus* and have 86.5 to 74.9% similarity with AtAGO10 (AT5G43810) (Fig. 1B). In addition, CrAGO1 has 74.5% similarity with AtAGO1, whereas other CrAGOs show approximately 60% similarity with Arabidopsis AGOs (Fig. 1B). Like DCLs, CrAGOs show high conservation in their PAZ, and PIWI domains (Fig. 3A). The phylogenetic tree shows that AGO1 and AGO10 can be classified into the same group (Fig. 1D). Indeed, AGO10 competes with AGO1 for miR165/166 binding in shoot apical meristem development, suggesting that functional redundancy exists between AGO1 and AGO10 (11). In addition, AGO4 and AGO6 have beenclassified into the same group, and both have been shown to be functionally redundant in RdDM (8, 12).

RNA-DEPENDENT RNA POLYMERASE (RDR) plays a role in synthesis of double-stranded RNA for DCL2, DCL3 and DCL4 cleavage to produce 22-, 24- and 21-nt siRNAs, respectively (5-7). In the *C. roseus* database, *CrRDR1* (DDS16060), *CrRDR2* (DDS47054) and *CrRDR6* (DDS12885) were identified as *AtRDR* orthologs of Arabidopsis and showed 53.3 to 77.4% similarity with Arabidopsis genes (Figs. 1B and 4A). The phylogenetic tree of RDRs includes 3 groups, as shown in Figure 1E. Other PTGS-related genes, such as *CrHYL1* (DDS45028) (Fig. 2B), *CrHEN1* (DDS46990) (Fig. 3B), *CrWEX* (DDS13870) (Fig. 3C), *CrHST* (DDS41673) (Fig. 4B), *CrSDE3* (DDS51099) (Fig. 4C), *CrSGS3.1* (DDS20269) and *CrSGS3.2* (DDS42452) (Fig. 4D), were identified in this study.

RNA polymerase IV and V are important components in plant 24-nt siRNA biogenesis and the RdDM mechanism plants (15, 16). In Arabidopsis, At*RDR1* can be induced by salicylic acid (SA), and *NaRDR1* also can be induced by jasmonic acid in *Nicotiana attenuata* (16). Our previous study also indicated that *CrPR1* and other SA response genes were also induced in response to S4 leafy flower (2); therefore, we presume that increasing *CrRDR1* expression triggers the defense response.

We conclude that this high-throughput genomics approach provides an advantage for studying the molecular biology of non-model organisms. In this study, the epigenetic-related genes of *C. roseus* were identified and characterized. We found that PTGS and RdDM maintained (8). CrNRPD1 (DDS47753) and CrNRPE1 (DDS16069) are the orthologs of Arabidopsis RNA polymerase IV and V large subunits, respectively. They showed 57.8% and 31.1% similarity to AtNRPD1 and AtNRPE1, respectively (Figs. 1B and 5A). Moreover, CrNRPD2 (DDS13297) is an ortholog of the second large subunit of RNA polymerase IV (AtNRPD2; AT3G23780) in Arabidopsis, with 75.9% similarity (Figs. 1B and 5B). The other RdDM-related genes, such as *CrDRD1* (DDS58), *CrDMS3* (DDS15300), *CrSPT5L* (DDS41978), *CrIDN2* (DDS19408) and *CrSUVH9* (DDS45142), were identified in the *C. roseus* database (Fig. 6) and showed 44.6 to 65.8% similarity with their Arabidopsis orthologs (Fig. 1B).

We compared the epigenetic-related gene expression levels in the HF and S4 samples based on the fragments per kilobase of transcript per million mapped reads (FPKM). The data indicated that most PTGS- and RdDM-related genes did not show different expression levels between HF and S4 (Fig. 1F), suggesting the epigenetic regulation have to keep the stabilization. However, CrAGO5 and CrRDR1 showed significantly increased expression levels during PnWB infection (Fig. 1F). The AGO phylogenetic tree indicated that AGO5 is closely related to AGO1 and AGO10 (Fig. 1D). Moreover, recent studies have indicated that AGO5 might have the same function as AGO1 to play a role in megagametogenesis during reproductive development (13, 14).

In addition, recent studies have indicated that RDR1 functions in anti-viral silencing in Arabidopsis and other normal gene expression during PnWB infection. However, CrAGO5 and CrRDR1 might involve in PnWB-mediated leafy flower formation and plant defense.

ACKNOWLEDGMENT

This study was supported by a grant from the National Science Council (NSC 102-2313-B-002-068-MY3) and National Taiwan University (103R7602B2).

70 植物病理學會刊 第23卷 第1期2014



Figure 1. Small RNA and epigenetic-related gene expression levels in healthy flowers (HF) and Stage 4 (S4) peanut witches'broom phytoplasma-infected leafy flowers of *Catharanthus roseus*. (A) The small RNA length distribution profiles of HF and S4. Total reads represents the amount of the corresponding small RNA. (B) The amino acid sequence similarity of epigeneticrelated genes in Arabidopsis and *C. roseus*. (C) The phylogenetic tree of DICER-like (DCL) between Arabidopsis and *C. roseus*. (D) The phylogenetic tree of ARGONAUTE (AGO) between Arabidopsis and *C. roseus*. (E) The phylogenetic tree of RNA-DEPENDENT RNA POLYMERASE (RDR) between Arabidopsis and *C. roseus*. (F) Comparison of epigenetic-related gene expression levels in *C. roseus*, as determined using next-generation sequencing (NGS). The fragments per kilobase of transcript per million mapped reads (FPKM) value was used to indicate the gene expression levels that were determined using the NGS of HF and S4 samples.



Figure 2. The amino acid sequence alignment of DICER-LIKE (DCL) and HYPONASTIC LEAVES 1 (HYL1). (A) The DCL alignments of Arabidopsis and *Catharanthus roseus*. The undetected amino acids are marked with an "X". DEAD-like helicases domain (DEXDc). Helicase superfamily c-terminal domain (HELICc). Piwi Argonaut and Zwille domain (PAZ).

helicases domain (DEXDc). Helicase superfamily c-terminal domain (HELICc). Piwi Argonaut and Zwille domain (PAZ). Prokaryotic ribonuclease III (RIBOc). (B) The HYL1 alignment of Arabidopsis and *C. roseus*. The specific protein domains are labeled on the sequence alignment. Double-stranded RNA binding motif (DSRM).

Characterization and identification of Catharanthus roseus epigenetic-related genes 71

72 植物病理學會刊第23卷第1期2014

Α		
AtAGO1 CrAGO1 AtAGO4 CrAGO4 AtAGO5 CrAGO5	20 40 60 80 100 120 140 160 180 EVENTSTARSEGSSBERAGYSSGG GSSBRGSTQGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	150 182 51 - 119 103
AtAGO6 CrAGO6 AtAGO7 CrAGO7 AtAGO10 CrAGO10.1 CrAGO10.2	BERTININGHSTKALP SERENTPLLIKTYNNINGGTNEP TLLIPSSIGULUUVPERERNGAPORERMASPSGENKOSSOFLEPPI-VTATSSLAPPIERERNGAPORERMASPSGENKOSSOFLEPPI-VTATSSLAPPIERERNGAPORERMASPSGENKOSSOFLAPPIERERNGAPOR	22 62 121 - 115 109 74
AtAGO1 : CrAGO1 : CrAGO4 : CrAGO5 : AtAGO5 : AtAGO5 : AtAGO5 : CrAGO5 : AtAGO7 : CrAGO7 : CrAGO7 : CrAGO10 : CrAGO10 .1 : CrAGO10.2 :	200 200 300 300 300 300 300 300 300 300	311 342 39 278 259 180 229 1161 262 215
AtAGO1 CrAGG1 AtAGO4 AtAGO5 CrAGO5 AtAGO5 CrAGO5 AtAGO6 CrAGG7 CrAGG7 CrAGG7 CrAGG7 CrAGG10.1 CrAGG10.2	0 400 420 440 460 460 500 520 540 560 - DAFGIN FOLT LPTR-VIDER MEDICAL MEDICAL FOLDANDEL STUDIE MARCEL MARCEL <td>174 505 380 141 345 386 458 428 425 378</td>	174 505 380 141 345 386 458 428 425 378
AtAGO1 CrAGO1 AtAGO4 CrAGO4 AtAGO5 CrAGO5 CrAGO5 AtAGO7 CrAGO7 CrAGO7 CrAGO7 CrAGO10 CrAGO10.1 CrAGO10.2	50 50 700 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 722 70 70 70 722 70 70 70 722 70 70 70 70 70 70 70 70 70 70 70 70 70	57 588 558 526 503 522 564 630 474 608 561
AtAGO1 CrAGO1 AtAGO4 AtAGO5 CrAGO4 AtAGO5 CrAGO6 CrAGO6 CrAGO7 CrAGO7 CrAGO7 CrAGO10 CrAGO10.1 CrAGO10.2	HO 20 POINT DEPENDENT OF DEPEND	139 17357 1557 1881 1741 1804 1892 7892 742
AtAGO1 CrAGO1 AtAGO4 AtAGO5 CrAGO5 AtAGO5 CrAGO6 CrAGO7 AtAGO10 CrAGO10.1 CrAGO10.2	940 940 940 1040 <t< td=""><td>)18)58)98)969)720)969)965)965)965)964 914</td></t<>)18)58)98)969)720)969)965)965)965)964 914
AtAGO1 CrAGO1 AtAGO4 CrAGO4 AtAGO5 CrAGO5	1140 1160 BSTRACTWINAR ME ALVER MUCUUME 1008 	

в

С

Figure 3. The amino acid sequence alignment of ARGONAUTE (AGO), HUA ENHANCER 1(HEN1) and WERNER SYNDROME-LIKE EXONUCLEASE (WEX). (A) The AGO alignment between Arabidopsis and *Catharanthus roseus*. The undetected amino acids are marked with an "X". Piwi Argonaut and Zwille domain (PAZ). P-element induced wimpy testis (PIWI). (B) The HEN1 alignment between Arabidopsis and *C. roseus*. (C) The WEX alignment between Arabidopsis and *C. roseus*. The specific protein domains are labeled on the sequence alignment.

Characterization and identification of Catharanthus roseus epigenetic-related genes 73



Figure 4. The amino acid sequence alignment of RNA-DEPENDENT RNA POLYMERASE (RDR), HASTY (HST),

THE REPORTED AN A REPORT OF AN AN AND A REPORT OF A THE REPORT OF A REPORT DD- : 625

Figure 4. The amino acid sequence alignment of RNA-DEPENDENT RNA POLYMERASE (RDR), HASTY (HST), SILENCING DEFECTIVE 3 (SDE3) and SUPPRESSOR OF GENE SILENCING 3 (SGS3). (A) The RDR alignment between Arabidopsis and *Catharanthus roseus*. (B) The HST alignment between Arabidopsis and *C. roseus*. (C) The SDE3 alignment between Arabidopsis and *C. roseus*. (D) The SGS3 alignment between Arabidopsis and *C. roseus*. The specific protein domains are labeled on the sequence alignment. Coiled coil domain (Coiled coil).

74 植物病理學會刊第23卷第1期2014

Α	2					C 0				100		100				100		***		
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	MENEGHTEQOEVAL	20 VPVGTLTSIG IPEARVRSIS DGEIVGI FA	FSISNNNDRDK FSILTEAEAGK LASHHEICIQS	40 SVLETEAPN AAKVIGVVS SESAINHPS	OVTO <mark>SR</mark> LGLP EVTO <mark>PALGIP</mark> QLTN <mark>AP</mark> LGLP	60 NPDSVCR NLSNQ L LEFGK ES	CSKDRKVCSC GKCLKT ATEPDK	80 HECVINFAY: HECINFPY HECYIQLPVI	S IN YF KOV TIN YF PRV PYH AHWNOI	AALUNKIOPG AKUUNRIOPA KOUUSLIOLK	KYIRKKOFOI KSFAADKKVP LKIKKAKGTS	IZU ITEDQPER KSAEYAIFRG SGGLADRLLG	KRNKCKYCDQ VCCEEASQIS	THU CTLNTGYPLM GKFRNSYPIN IKDRASDGAS	KFRUTTKEVF RFKUSPROMF YLEUKLPSRS	160 RRSGIVVEVNE GKTAIVAEVEL RLQPGCWNFLE	ESLMKLKKRG RGRSER RYGYRYGSDY	VLTU PPDYWS ASDYWD TRPULAREVK	FUPODS : FUPPDP : EULRRI :	177 184 187
	200		220		240		260		200		200	KFC	220		240		360		200	
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	NIDESCLKPT RI QIEKNNOSST RL PEESRKKLTA GH	THAQVYALLI SHAQVYDIV PQEGYILEYI	GOORLHKKDI ENDESVIQSFL PIPPNCISVPE	PMFNSLGMTS RRKNDIFUNC ASDGFSTVSV	FPVTPNGYR LLVTPNCHR DPSRIELKD	TEIVHQFNG TEFGHHIT- LKKVIAIK-	ARLIFDER FDASTK	IYKKLVGPEC LYKKLADFKC SGETNFESHP	280 INTLELSS ITANELSI KAEASEME	ECMQYSELFS DRIKVSEVRS DTYLQVEGTA	SUU ETVSSSKDSAN DKSSDVDSASV KAARNIDMRYC	NPYQKKSDTP /AS SVSKISDSSS	SZU KLCGLRFIED GLKYVE SKAWTEK	VL GARSDH IL AARNDH IF RAGSGFS	SHU ERTVVVGDPS ERTVVVGDPN SRSVT <mark>T</mark> GD <mark>A</mark> Y	IKLNEIGIPES IKLDEIGVPCO RHVNEVGIPIE	JAKRIQVSOH LABKIRVAOH LAQRITFEIR	NQC KERUV AKW LDE E SVH RGY Q	TSFVPT : RYLELM : KLVDDK :	368 360 370
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	LDNKEMHVRRGD IGGADRRVRRGD CLSYTQGSTTYSI	400 LVAIQVN-D LVRLSLEDK RDGSKGHTE	* QTEDKIFSSM QKEDIYSPS KPCQVIHER M	420 Deptydynre Dedtiddinre Dedtydfinre	PSTHOHSLIA PSTHOHSLIA PTTHKHSL <mark>O</mark> A	440 TRLPTT SKLPIG RYNHEDN	* SWSINPICO SVIAINPINC -TMKINPINC	460 L FRG FDGI Silrg FDGI Silsal FDGI FDGI	* DCLH <mark>EYV</mark> PQSI DCLHEYTPQSU DCVHIFYPQSL DCIHIFYPQSL	480 QERVELDELV DSRVSLIELV SARASVIELF SARASVIELF	* ALDKQL NRQI ALDKQLFNGQ SVSKQL SSH SVSKQL SSH	500 GRH ISI GQ GRH ISI SH GQL II OGGS GNFNI GLAT	* DSLTAAYLWN DSLTAAHLIL DSLLSLRWL DSLLSLKLNF	520 VEKNOTONEA GDG-VI EV ERV DOA KRY FGGE	* OMOLONYCE OMOLONYCE TACOLANYGS TACOLANYAE	540 FOL PEATINA ROLOL ATVO LSU PEATROS ALL RSAVINS	* PSS EPONT N-LAKARFNT SKSGPANT SNSGCLAT	560 GMOLFGULFP GKOLFS FLP VFOILO AFP VIOLLOTALP	* GEYT : GEYL : ERLSCK : SEDCS :	558 549 556 107
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	580 YPLNNGVSNGED FRSNGVIRKGED G-DRFDJCGSDD G-DRYDTYKSEIW	* SFSEGSAWLF TSSNGSSWLF FDFGVDAMGS VDFNREMAQS	600 DGEGNFIER DNDEN FHH SIINEI TSIF SIFIDI TSIF	* 6 KHDKGKVUDI KYCRETIDI EKGPKETIGE SKGPTIVERE	20 IY AVENISO LFAA EVICE FDSLEPHIME FDSLEPHIME	NJLARGISV NJSRGISV SIFADGESL NJYOĐGESV	640 SLAD VLSSD SLSD VLASD SLSD SVSRA CLEDFVLPKG	* LQSRKNLTE TCLRNNLID DMDVIHNLII ALNDIQ-ASI	660 SISYGLREAEQ SVSCGLQEAE IREISPNVSR IQSVSPLLYY	* VCNKQOLMVE QSYISTIMIN R RS	680 SWRDFLAVNGE CNQQYLSAN-S	* SDKEEDSVSD SDRNERMNFA	700 LARFCYEROK IEHMSIQQOA	* SAT S LAVS SAP S AFVS 	720 AFK AYRDVO AFKOVFLDHO QLE SIHKAK ELE YLRGKK	* 7 ALEYREGDOSN HUAVOAAN DN SWANEMINSY LPVSNEVLISS	40 SELIVSKAGS AFLTV RAGS SIRNU DIKS AIGHL DSKS	* 7 KGNIGKLVCH KGNULKLVCH NSÄLTKLVCQ DSÄLSKVVCQ	60 SMCIGI : SMCIGI : GFIGI : IGFIGN :	749 739 704 256
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	ONSAVSISEGFPRE ONSVVPISERIPHO DISOKKEPTETIA DISOKGEPTETIA	80 LTOPA NDPN LSOPA NDO EDMEI CKR KDVDELFLK	* 80 SPLRGAKGKDS ISDCKAQDIP- YGRISSSG YPSSGMYPSE-	0 TTTESYVPIC EHSKSYIPLA Dae Qae	* 82 VIENS LTGL VV SS LTGL IV COFFICL LV SCLFYGL	O NPLESFUHS NPLECFUHS DPYEEMAHS DPYEELUHS	* 8 V SRDSSFSG L NRDGSFSG TAAREVIVRS ISSREVIVRS	40 NADL GTI HADVSGTI SRGLADIGTI TRGLTE GTI	* 8 SRRIMFEMRD INRKIMFEMRD INRKIMATIRD	60 IYAA DGTVR LYTA DGTVR IV TNDGTVR VV CIDGTVR	* 88 NSFGNCLVOFT NAYGNCLVOFS NTCSNSVIQE NVCSNSVIQES	80 METEG SMNTQGGLN GGV SE SGGING	* 9 SCNDNKVDPH	00 PVEDIT ESSHNSDIQG -RGHQGLFEA -MSDDNEFGA	* 9 GEALGSISAC GEVGSUAC GEVGVUAT GEVGVUAT	20 Alseraysald Aiseraysald Amsnerykavl Amsnerykavl	* 94 QPISLLETSP LPCSALETSP DSSPNSNSS- DSSPSSNSS-) ILNLKNVLEC WELMKRVLLC WELMKRVLLC	* SSSK : SVSK : KNNFQN : SFIN :	918 925 870 424
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	960 : GOREQT SLYLSEY : SSG TASLELSKI : TNORR I DYDNEC : DISORRI LYDNDC	* 96 ISKOKHGPBY IGRWPHGPBY HCG RFCCPY ACG NYCRD)	10 * IGSLEIKINDEK IG LEVIC LER IA CTVRIKUNK IA YVVKILIGK	1000 ISFSEIVSTS ILFSDIISRI VSLKDTAVEF ISLKEAAMEL	* HISSESSE INEYRKOPTI LIENKPTHSV	102 SNTKVPLSP GRRSHMSP SEIFGIDSC IESSDLNIG	0 W CHEHISEK W CHEHICKE IHGHIHINKT I GHIHIDEM	* 104 VERSONSA VRSPES LODWN SESTANSE	40 SWSSLNEQY SHIDALHMRW DHQKCEDVI SEWSKCEDKM	* 106 KSRNE- YSVRE- NSLGOKKY K NFRKALK	0 LKLDIVD LKITLPO ATDDFKRTS DGYLFKKII	108 QNTNHCSS IISKGCSLA SKECSFRD ASSCCSFQQ	0 ddQamkDDNV akgnemDtel pcgskgsDmp ssask-sfDc	* 110 OTV VVEAS OTAASVSESS OTFSYNAAD LQFIWQDTSN	0 KHSVLEL AI RN-RVKV II PDLERTLIVI YKLEEISH I	1120 RIVETEFIC RIVETEFIC CONVENIEI AUTOCVILG	* PVKG QG KK IKG PEFEK IKG SR CS IKG PR SS	1140 NILWT RPKJ DILWKDAKS ANIIWNSSDM DIIWI PDT	APKRN- : S : TTWIRN : PTWIRN :	1099 1101 1061 612
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	* 1160 -GNHLA CAPYLEN KGTPNG PYLEN RHASER CAVUDY PCKSRKCODALON	VYGURGER VSEUCDAT VEKSAVEQSC VEKUAVEQPC	1180 COTALLETCL KENSILVISCL TAGRIVIISCL	* PHUDVIDWG QIWAVIDWEC SWIHLIDTK PVIDLIDTK	1200 SHEDNICCC SREDDINDIT SIEYSUUUU SIECALUUU	* SMY <mark>SI</mark> DAGR RSYSVDA V BILGISCIF BILGISCIF	1220 SIF ANGESA GLFRTSINSA EQA QRUSAS EQA QRUSAS	VSDTGROID ICDTGRIVI WRMVSRGVI VTMVTRGVI	* 1240 HLLL ADSI HVII TADCI HVII TADCI HLVL ANSM	* SVTGEFVALN SATGEFVALN TOSGTMLGEN TYSGNLIGEN	1260 AKEWSKOROVE PKELALOR OF SGEYKALT SI MGEIKALS SI	* ESTPAPTOA EGTVSPLNQA IN KAPTOA IN OVPTEA	1280 CFSSISOULL CFSSISDNLV TLIASKOE TLFARKOE	KAAKEGVRO RAAKTGAVD KAAEKCHTDS RAAEKCHVD	1300 LOGSIDALAW LOGTVDALAW LSTVVGSCSA LSSVVASCAW	* GRVPGFGTCD GRRPP GTCA GRRVD GTGS GRQVA GSCSA	1320 EETHISPAVH DIMYSWADN EELHNQKET DILLDTRKV	* GETT VD YD EPTK TD YS GLDDKEETDV ELNO DG DV	1 LISSTK : ILDNHI : YSFLQ: : YDFLHI :	1287 1288 1252 803
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	340 * TMRRATSAPKSDKA QEAPKVKLPKKOPC VISTTADAFVSSE VSSSSLGEELDTAC	1360 TVQPFGLLHS ISAKSLAQHI GFDVIEEEMP IGAEIDNIDI	* AFLKDIKVLDG FKFEDLTAQPC EWAESPERDSA DDESMGIDVSP	1380 KGIPMSLL KSLTKPKPSL LGEPK-FEDS LRDSE-KPTF	* RTIFTWKNIE FGRLSVRDIL ADFON HDEG EDGMEIDDN-	1400 LLSOSLKR NLSVTLRN KPSGANWEK	* LHSYBITELL LHKYDIDSLL SSSADNGCSG ESGNFKISSR	1420 NERDEGLVKI TEEEKTYVLJ GSEWGVSKST TAKSNVGWD(* IN QLHENS IAI WFHEOK IGGEANEESNN DAI DKAENGGN	1440 VE IGP RE IGT EKTTNVEKED DQAID AHS	* VKGIRIARSKH IVNIK GSHAH AWSSWNT KDJ GWDQA D VPC	1460 HGDS KYBES AQESSKSDSG G	CCRE VRIIG RCSVIERKOG GANGIKTKOA GGIDQVDIAQ	1480 TFEDESYHKC TVEDESYHKC DADTTENWET NAGIGEKGES	* VLGATKI AP VYNALKG NE SPAPKDS VE GVWDQAIDKA	1500 KKMNFYKSKNL DRAKLYESKNL ENNEPTSDV/G PNG	* KNGTLESGGF KSDRQMASKL HKSVSDKSWD GWDQAAQNAS	1520 SENP ASPKSF KKNWGTESAP DGPKASDSAW	AAWGST AGWGKA	1453 1459 1442 968
	* 1540		1560		1580		1600	•	1620	*	1640		1660		1680	*	1700		172	
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	DAAVWGSSDKKNSE EV	TESDAAAWGS	RDKNNSDVGSG	AGVLGPWNKK	SSETESNGAT	WGSSDKTKS	GAAAWNSWDK	KNIETDSEP	AAWGSQGKKNS	ETESGPAAWG	AWDKKKSETEI	PGPAGWGMGD	KKNSETELGP	AAMGNWDKKK	SDTKSGPAAW	GSTDAAAWGSS	DKNNSETESD	AAAWGSRNKK	TSEIES :	1633 970
	0 ·	1740		1760		1780		1800		1820		1840	*	1860		1880		1900	*	
AtNRPD1 : CrNRPD1 : AtNRPE1 : CrNRPE1 :	GAGAWGSWGQPSP1	AEDKDTNEDD	RNPWVSLKETK	SREKDDKERS	QWGNPAKKFP	SSGGWSNGG	GADWKGNRNH	TPRPPRSEDI	NLAPMFTATRO	RLDSFTSEEQ	ELLSDVEPVMF	RTLRKIMHPS	AYPDGDPISD	DDKTFVLEKI	LNFHPQKETK	LGSGVDFITVD	KHTIFSDSRC	FFVVSTDGAK	QDFSYR :	1824
	1920	*	1940		1960	*	1980		2000		2020	*	2040		2060					
AtNRPD1 CrNRPD1 AtNRPE1 CrNRPE1	KSLNNYLMKKYPDF	AEEFIDKYFT	KPRPSGNRDRN	NQDATPPGEE	QSQPPNQSIG	NGGDDFQTQ	TQSQSPSQTR	AQSPSQAQA(QSPSQTQSQSQ	SQSQSQSQSQ	ଽୢଌଽୢ୰ଽୢ୰ଽୢ୰ଽ	SQSQSQSQSP	SQTQTQSPSQ	TQAQAQSPSS	QSPSQTQT-	1976				





Figure 5. The amino acid sequence alignment of NUCLEAR RNA POLYMERASE D1 and E1 (NRPD1/E1), and NRPD2. (A) The NRPD1/E1 alignment between Arabidopsis and *Catharanthus roseus*. RNA polymerase I subunit A N-terminus (RPOLA_N). (B) The NRPD2 alignment between Arabidopsis and *C. roseus*. The specific protein domains are labeled on the sequence alignment.



Figure 6. The amino acid sequence alignment of DEFECTIVE IN MERISTEM SILENCING 3 (DMS3), DEFECTIVE IN RNA-DIRECTED DNA METHYLATION 1 (DRD1), INVOLVED IN DE NOVO 2 (IDN2), SU(VAR)3-9 HOMOLOG 9 (SUVH9) and SUPPRESSOR OF TY INSERTION 5-LIKE (SPT5L). (A) The DMS3 alignment between Arabidopsis and *Catharanthus roseus*. Coiled coil domain (Coiled coil). (B) The DRD1 alignment between Arabidopsis and *C. roseus*. DEADlike helicases domain (DEXDc). Helicase superfamily c-terminal domain (HELICc). (C) The IDN2 alignment between Arabidopsis and *C. roseus*. (D) The SUVH9 alignment between Arabidopsis and *C. roseus*. Sar1p-like members of the Rasfamily of small GTPases (SAR). N-terminal to some SET domains (PreSET). Su(var)3-9, Enhancer-of-zeste, Trithorax domain (SET). (E) The SPT5L alignment between Arabidopsis and *C. roseus*. The specific protein domains are labeled on the sequence alignment. Kyprides, Ouzounis, Woese motif (KOW).

LITERATURE CITED

- Chen, W.Y., and Lin, C.P. 2011. Characterization of Catharanthus roseus genes regulated differentially by peanut witches' broom phytoplasma infection. J. Phytopath. 159:505-510.
- Liu, L.Y.D., Tseng, H.I., Lin, C.P., Lin, Y.Y., Huang, Y.H., Huang, C.K., Chang, T.H., and Lin, S.S. 2014. High-throughput transcriptome analysis of the leafy flower transition of Catharanthus roseus induced by peanut witches'-broom phytoplasma infection. Plant Cell Physiol. doi:10.1093/pcp/pcu029.
- Vaucheret, H. 2006. Post-transcriptional small RNA pathways in plants: mechanisms and regulations. Genes Dev. 20:750-771.
- Vaucheret, H., Vazquez, F., Crété, P., and Bartel, D.P. 2004. The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. Genes Dev. 18:1187-1197.
- Law, J.A., and Jacobsen, S.E. 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. 11:204-220.
- Zhang, H., and Zhu, J.K. 2011. RNA-directed DNA methylation. Curr. Opin. Plant Biol. 14:142-147.
- Bouché, N., Lauressergues, D., Gasciolli, V., and Vaucheret, H. 2006. An antagonistic function for Arabidopsis DCL2 in development and a new function for DCL4 in generating viral siRNAs. EMBO J. 25:3347-3356.
- Daxinger, L., Kanno, T., Bucher, E., van der Winden, J., Naumann, U., Matzke, A.J., and Matzke, M. 2009. A stepwise pathway for biogenesis of 24-nt secondary siRNAs and spreading of DNA methylation. EMBO J. 28:48-57.
- 9. Tesng, H.I. 2011. High-throughput transcriptome and small RNA analysis for studying phytoplasma infection on Cantharanthus roseus using Next

Generation Sequencing. Master thesis. National Taiwan University.

- Baumberger, N., and Baulcombe, D.C. 2005. Arabidopsis ARGONAUTE1 is an RNA Slicer that selectively recruits microRNAs and short interfering RNAs. Proc. Natl. Acad. Sci. USA 102:11928-11933.
- Zhu, H., Hu, F., Wang, R., Zhou, X., Sze, S.H., Liou, L.W., Barefoot, A., Dickman, M., and Zhang, X. 2011. Arabidopsis Argonaute10 specifically sequesters miR166/165 to regulate shoot apical meristem development. Cell 145:242-256.
- Zheng, X., Zhu, J., Kapoor, A., and Zhu, J.K. 2007. Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. EMBO J. 26:1691-1671.
- Tucker, M.R., Okada, T., Hu, Y., Scholefield, A., Taylor, J.M., and Koltunow, A.M. 2012. Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in Arabidopsis. Development 139:1399-1404.
- Borges, F., Pereira, P.A., Slotkin, R.K., Martienssen, R.A., and Becker, J.D. 2011. MicroRNA activity in the Arabidopsis male germline. J. Exp. Bot. 62:1611-1620.
- Liao, Y.W., Sun, Z.H., Zhou, Y.H., Shi, K., Li, X., Zhang, G.Q., Xia, X.J., Chen, Z.X., and Yu, J.Q.
 2013. The role of hydrogen peroxide and nitric oxide in the induction of plant-encoded RNAdependent RNA polymerase 1 in the basal defense against Tobacco mosaic virus. PLoS One 8:e76090.
- Hunter, L.J., Westwood, J.H., Heath, G., Macaulay, K., Smith, A.G., Macfarlane, S.A., Palukaitis, P., and Carr, J.P. 2013. Regulation of RNA-dependent RNA polymerase 1 and isochorismate synthase gene expression in Arabidopsis. PLoS One 8:e66530.

摘要

曾欣怡^{1,4,5} 林長平^{1,4,5} 林詩舜^{1,2,3,6}.2014. 日日春基因靜默機制相關基因參與花生簇葉病菌質體 誘發花器葉化機制之探討. 植病會刊 23:67-77. (¹臺灣大學生物科技學研究所;²臺灣大學基因體 暨系統生物學學程;³中央研究院農業生物技術研究中心;⁴臺灣大學植物病理暨微生物學系;⁵ 共同第一作者;⁶聯絡作者,電子信箱: linss01@ntu.edu.tw)

ContigViews 資料庫平台提供龐大的日日春植物 (*Catharanthus roseus*) 轉錄體基因資訊。為了 研究表觀遺傳對於花生簇葉病植物菌質體 (peanut witches'-broom phytoplasma; PnWB) 誘發日日春 花器葉化的影響。其參與在後轉錄時期基因靜默 (post-transcriptional gene silencing; PTGS) 及核醣 核酸引發之去氧核醣核酸甲基化 (RNA-directed DNA methylation; RdDM) 相關基因已由此資料庫 中選殖出來,並與阿拉伯芥同源基因相互比較。這些相關基因的演化關係及在健康花 (HF) 與 PnWB 誘發第四型 (S4) 葉狀花的表現量都被詳細的研究。研究結果顯示,日日春 *ARGONAUTE 5* (*CrAGO5*) 及 *RNA-DEPENDENT RNA POLYMERASE 1* (*CrRDR1*) 基因在 S4 葉狀花中表現量有明 顯的增加。然而,大部份的 PTGS 及 RdDM 相關基因在 HF 及 S4 樣本中的表現量都維持恆定, 顯示表觀遺傳的基因表現必須維持恆定。這些證據推測 CrAGO5 與 CrRDR1 可能參與在 PnWB

關鍵字:次世代基因定序、日日春、花生簇葉病植物菌質體、後轉錄時期基因靜默、核醣核酸引 發之去氧核醣核酸甲基化