Functional analysis of the histone modification enzymes SDG7 and SDG8 during Arabidopsis development

WANG, YICONG

Nara Institute of Science and Technology

Graduate School of Science and Technology

Laboratory of Plant Stem Cell Regulation and Floral Patterning

(Division of Biological Science)

(Professor Toshiro Ito)

Submitted on

2022.9.10

Laboratory (Supervisor)	Professor Toshiro Ito		
Name	WANG, YICONG	Date	2022.9.10
Title	Functional analysis of the histone modification enzymes SDG7 and SDG8 during Arabidopsis development		

Post-translational modifications of histones play important roles in regulating gene expression. More than one-third of all Arabidopsis genes contain histone H3 methylation at lysine 36 (H3K36me). Dimethylation and trimethylation of H3K36 are preferentially located in actively transcribed genes, but the molecular functions and recruitment mechanisms of enzymes responsible for H3K36me2 and H3K36me3 deposition remain unclear. Here, I show that two closely related Arabidopsis homologs belonging to class II of the SET DOMAIN GROUP (SDG) family, SDG7 and SDG8, are required for proper growth and development. The two genes were highly expressed throughout plant development. Consistent with this observation, the *sdg7 sdg8* double mutant displayed a dwarf phenotype with much smaller leaves and flowers, shorter primary roots and fewer lateral roots. The loss of SDG7 and SDG8 function was accompanied by a drastic decrease in H3K36me3 and H3K36me2 levels *in vivo*. Furthermore, *cis*-regulatory motif analyses using SDG-regulated genes with different H3K36me2/3 levels and yeast two-hybrid assays identified G-box motifs and basic leucine zipper (bZIP) transcription factors, respectively, as possible interaction partners of SDG7 and SDG8. Among bZIP family members, bZIP53 strongly interacted with SDG8. Taken together, these results suggest that bZIPs may recruit SDG7 and SDG8 during plant growth and development.

Table of Contents

List of Figures and Tables

Figures

Tables

I. Introduction

In eukaryotes, epigenetic regulation plays key roles in the spatiotemporal regulation of gene expression during development and in response to the environment. Epigenetic regulations such as chromatin remodeling, DNA methylation, and histone modifications affect gene expression without permanent changing the genomic sequence (Cheung et al., 2000; Gan et al., 2013; Jenuwein & Allis, 2001; Smolikova et al., 2021; Strahl & Allis, 2000). Histone modifications are covalent post-translational modifications of the N-terminal tails of histones H3 and H4. Histone modifications of different types affect transcriptional activation and inactivation, as well as chromatin structure and DNA damage/repair (Berr et al., 2011; Davie, 1996; Dinant et al., 2008). For example, trimethylation of histone H3 at lysine 4 (H3K4me3) or at lysine 36 (H3K36me3) predominantly marks the transcription start sites of genes and is often correlated with transcriptional activation, while H3K27me3 and H3K9me3 marks are associated with transcriptional repression (Barski et al., 2007; Greer & Shi, 2012; He et al., 2012). Histone methylation requires the enzymatic transfer of one, two or three methyl groups from sulfur-containing metabolites, such as S-adenosylmethionine, to lysine or arginine residues in the histone tails of core histone proteins. This enzymatic reaction is regulated by histone methyltransferases. The SET DOMAIN GROUP (SDG) family of histone methyltransferases is responsible for various types of lysine methylation in both animals and plants. In the model land plant Arabidopsis (*Arabidopsis thaliana*), SDGs are encoded by a large gene family of 32 members (Qian & Zhou, 2006; Springer et al., 2003). *Arabidopsis* SET domain-containing proteins can be divided into five classes based on sequence similarity and domain organization (Springer et al., 2003) (Table 1). Class I SET domain

proteins include Polycomb Group (PcG) proteins, such as CURLY LEAF (CLF, also named SDG1), SWINGER (SWN, also named SDG10) and MEDEA (MEA, also named SDG5). PcGs deposit H3K27me3 to silence gene expression (Hennig & Derkacheva, 2009; Makarevich et al., 2006) (Fig. 1). Class V SET domain proteins include SU(VAR)3-9 HOMOLOG 4 (SUVH4, also named SDG33), SUVH5 (also named SDG9), and SUVH6 (also named SDG23), which are redundantly required for H3K9me2 deposition (Johnson et al., 2014; Li et al., 2015). SDG4, SDG7, SDG8, SDG24, and SDG26 are class II SET domain proteins, which are characterized by an Associated With SET (AWS) domain, a SET domain, and a post-SET domain (Springer et al., 2003). A distinguishing feature of class II SET domain proteins is their diverse domain organization outside the SET domain. Hence, this class is further divided into four ortholog groups: II-1, II-2, II-3, and II-4. SDG7 (363 amino acids [aa]) and SDG24 (329 aa, although several splice variants exist) belong to group II-1 and have N-terminal and C-terminal regions of similar length. SDG4 (497 aa) and SDG26 (492 aa, with several splice variants) are members of group II-2 and II-3, respectively. Only SDG4 possesses a PHD zinc finger domain at its N terminus. Unlike other class II SET domain proteins, SDG26 has a very short N-terminal region. SDG8 (1,767 aa) belongs to group II-4 and has very long N-terminal and C-terminal regions. Class II SDG proteins have a role in the deposition of H3K4me2, H3K4me3, H3K36me1, H3K36me2, and H3K36me3 to activate gene expression for proper control of plant development and environmental responses (Berr et al., 2010; Cartagena et al., 2008; Cazzonelli et al., 2010; Dong et al., 2008; Grini et al., 2009; Kumpf et al., 2014; Lee et al., 2015; Li et al., 2015; Wang et al., 2014; Xu et al., 2008; Xu et al., 2014; Zhao et al., 2005). *sdg7* mutants flower early due to reduced expression of the floral repressor gene *FLOWERING LOCUS C* (*FLC*) (Lee et al., 2015).

Table 1. Classification of SET domain group proteins in Arabidopsis

AGI code	Gene name	Subfamily	Known function
AT2G23380	*SDG1/CLF*	Class I	Similar to the product of the Polycomb-group CLF (SDG1), MEA (SDG5) and SWN (SDG10) deposit H3K27me3 to silence gene expression
AT1G02580	*SDG5/MEA*		
AT4G02020	*SDG10/SWN*		
AT4G30860	*SDG4/ASHR3*	Class II	Encodes a member of the trxG protein family. Overexpression results in plieotrophic developmental defects.
AT2G44150	*SDG7/ASHH3*		Encodes a protein-lysine N-methyltransferase.
AT1G77300	*SDG8/ASHH2*		Inhibitor of flowering specifically involved in the autonomous promotion pathway.
AT3G59960	*SDG24/ASHH4*		histone-lysine N-methyltransferase
AT1G76710	*SDG26/ASHH1*		
AT3G62420	*SDG2/ATXR3*	Class III	ATXR3 acts redundantly with ATX4/5 to regulate histone H3-K4 methylation. Involved in bolting/flowering time together with ATX1 and ATX4.
AT3G61740	*SDG14/ATX3*		
AT4G27910	*SDG16/ATX4*		
AT5G53430	*SDG29/ATX5*		
AT5G42400	*SDG25/ATXR7*		Required for histone H3-K4 methylation and for transcriptional activation of Flowering Locus C.
AT2G31650	*SDG27/ATX1*		Putative histone methyltransferase related to the Drosophila trithorax group proteins TRX and TRR
AT5G53430	*SDG29/ATX5*		
AT1G05830	*SDG30/ATX2*		
AT5G09790	*SDG15/ATXR5*	Class IV	H3K27 monomethyltransferases required for chromatin structure and gene silencing. Play roles in cell-cycle regulation or progression.
AT5G24330	*SDG34/ATXR6*		
AT2G33290	*SDG3/SUVH2*	Class V	Encodes a SU(VAR)3-9 homolog, involved in epigenetic control of gene expression and act as histone methyltransferases. 10 SUVH genes contain additional conserved SRA domain
AT2G23740	*SDG6/SUVR5*		
AT2G35160	*SDG9/SUVH5*		
AT2G05900	*SDG11/SUVH10*		
AT1G04050	*SDG13/SUVHR1*		
AT1G17770	*SDG17/SUVH7*		
AT5G43990	*SDG18/SUVR2*		
AT1G73100	*SDG19/SUVH3*		
AT3G03750	*SDG20/SUVR3*		
AT2G24740	*SDG21/SUVH8*		
AT4G13460	*SDG22/SUVH9*		
AT2G22740	*SDG23/SUVH6*		
AT3G04380	*SDG31/SUVR4*		
AT5G04940	*SDG32/SUVH1*		
AT5G13960	*SDG33/SUVH4*		
AT1G26760	*SDG35/ATXR1*	N.A. (S-ET)[e]	Unknown function
AT3G21820	*SDG36/ASHR2*		
AT2G17900	*SDG37/ASHR4*		
AT5G06620	*SDG38/ATXR4*		
AT2G19640	*SDG39/ASHR2*		

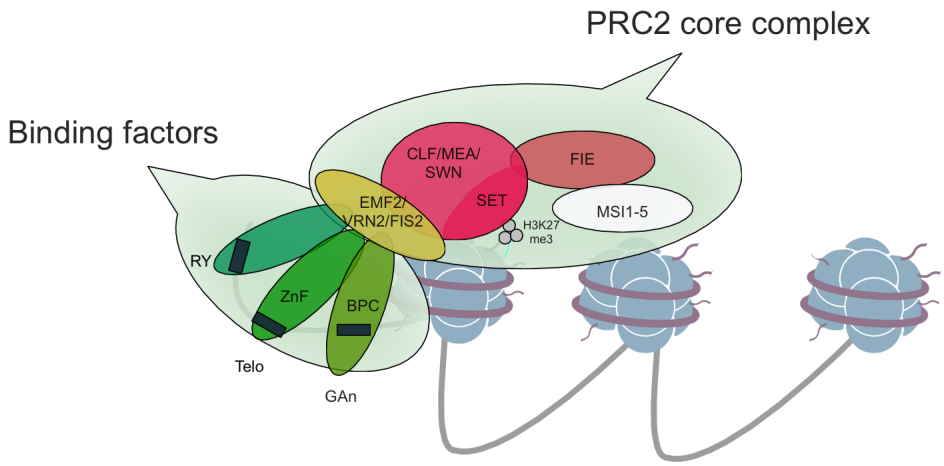

Figure. 1. Current model of PRC2-mediated H3K27me3 deposition.

PRC2-core complex, including binding factors, deposits H3K27me3 at downstream genes to repress its expression. Red circle represents H3K27 methyltransferases MEA, CLF and SWN. Brown circle represents histone H3 tail binder FIE; white circle represents nucleosome binder MSI1-5. Yellow circle represents cofactors binder EMF2/VRN2/FIS2. Gradient green circles, representative cofactors. Black boxes, cis-regulatory motifs for histone-binding nucleosome remodeling factor. Gray line, DNA. Dark blue ellipses, histones. Purple lines, histone tails.

The phenotypes of *sdg24* mutants have not been reported yet. *sdg4* mutants have shorter primary roots resulting from an abnormal organization of the root apical meristem and fewer cell divisions (Kumpf et al., 2014). *sdg4* mutants also exhibit reduced fertility (Cartagena et al., 2008). In *sdg4* flowers, H3K4me2, H3K4me3, H3K36me2, and H3K36me3 levels are globally reduced (Cartagena et al., 2008). *sdg8* mutants show an early flowering phenotype due to reduced *FLC* expression (Yang et al., 2016). *sdg8* mutants also display reduced responses to mechanical stimulation due to lower H3K4me3 levels at the *TOUCH 3* (*TCH3*) locus (Cazzonelli et al., 2014). In addition, *sdg8* mutants show pleiotropic phenotypes such as more shoot branches, less carotenoid biosynthesis, and compromised plant immunity (Cazzonelli et al., 2010; Dong et al., 2008; Lee et al., 2016). Unlike *sdg8*, *sdg26* mutants display a late flowering phenotype (Xu et al., 2008). Phenotypic analysis of *sdg8 sdg26* double mutants revealed that the *sdg8* mutation is epistatic to *sdg26* in the level of H3K36me3, H3K36me1 and distinct functional interplays of *SDG8* and *SDG26* in the regulation of plant growth and development (Liu et al., 2016). Both *sdg7* and *sdg8* single mutants show an early flowering phenotype (Kim et al., 2005; Lee et al., 2015; Soppe et al., 1999; Xu et al., 2008; Zhao et al., 2005). Furthermore, the global reduction of H3K36me3 and H3K36me2 levels seen in *sdg8* is more pronounced in the *sdg7 sdg8* double mutant (Xu et al., 2014). These observations suggested there might be partial functional redundancy between SDG7 and SDG8. However, the genetic interaction between *SDG7* and *SDG8* has yet to be tested.

Among SDG proteins from class II, the binding of SDG8 to H3K36 hypomethylated genes has been examined in a genome-wide manner (Fig. 2) (Li et al., 2015). In plants, H3K36me2 and H3K36me3 marks accumulate across the transcribed regions with a specific distribution

pattern (Xiao et al., 2016). H3K36me2 and H3K36me3 are at their highest levels in the 5’ or 3’ ends of transcribed genes, respectively. Although H3K36me2 levels have not been examined in the *sdg8* mutant, the H3K36me3 mark is reduced in the *sdg8* mutant background, suggesting that SDG8 is mainly required for maintenance of H3K36me3 along the gene body, with a bias towards the 3’ end of transcribed genes (Li et al., 2015). Genome-wide SDG8 binding is essentially observed in the genic regions. There is a significant overlap between SDG8-bound genes and H3K36 hypomethylated genes in *sdg8* mutants. In the promoter of SDG8- and H3K36me3-bound genes, the *cis*-regulatory motifs, basic leucine zipper (bZIP) binding motif G-box (CACGTG) and FORC^{A} motif (TGGGC), are significantly overrepresented (Li et al., 2015). Since SDG8 does not contain a DNA binding domain, SDG8 might interact with transcription factors such as bZIPs. However, how SDG8 or other class II SDGs recognize their targets along the genome is not understood. Further, how H3K36me2 and H3K36me3 are differentially deposited is not known, either.

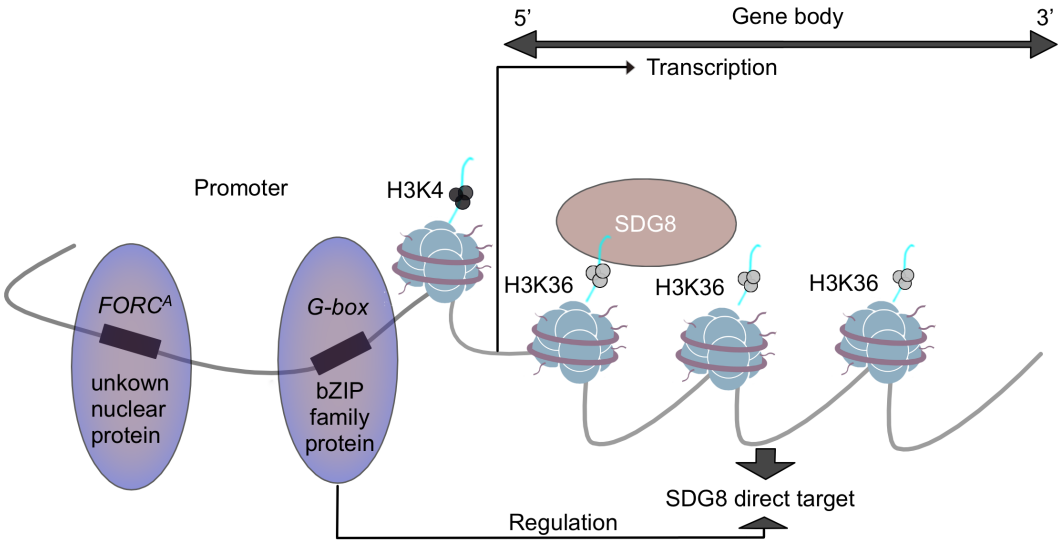

Fig. 2. Current model of SDG8-mediated H3K36me3 deposition.

SDG8 regulates downstream genes through the deposition of H3K36 methylation marks in the gene-coding region. In the promoter of SDG8-bound and H3K36me3-decorated genes, the bZIP binding motif G-box (CACGTG) and FORC^{A} motif (TGGGC) are significantly overrepresented (Li et al., 2015). Purple ellipses, transcription factor proteins. Black boxes, cis-regulatory motif. Gradient black, SDG8 protein. Gray circles, H3K36 methylation. Black circles, H3K4 methylation. Gray line, chromatin. Blue ellipses, histones. Light blue lines, histone tails.

In this work, I examined the genetic interaction between *SDG7* and *SDG8* using a *sdg7 sdg8* double mutant. The *sdg7* mutant showed synergistic interactions with *sdg8* with respect to plant height, leaf size, flowering time, flower size, primary root length, and the number of lateral roots. The genes encoding these two class II SDG proteins showed overlapping expression patterns during plant development. H3K36me3 marks toward the 3' end of transcribed genes further decreased in the *sdg7 sdg8* double mutant compared to either single mutant, and H3K36me2 marks also disappeared across the gene body in the *sdg7 sdg8* double mutant. Furthermore, *sdg8* single mutant showed an overall weak phenotype compared with *sdg7 sdg8* double mutant, while vegetative and reproductive development of *sdg7* single mutant was similar to that of wild type except for lateral root. These results indicate that SDG7 and SDG8 have partially overlapping functions via the deposition of H3K36me2 and H3K36me3 marks during plant development. I screened for putative recruiters of SDG7 and SDG8 by yeast two-hybrid assays and identified bZIP family members as potential interacting partners. These results may reveal how SDG proteins recognize their target genes for proper histone modifications.

II. Materials and Methods

Plant materials and growth conditions

All Arabidopsis (*Arabidopsis thaliana*) lines analyzed in this study are in the Columbia (Col-0) accession. *sdg8-2* (SALK_131218), *sdg7-2* (SALK_026442) (Grini et al., 2009; Xu et al., 2014). *maf3* (SALK_051357), *dcl1* (CS84714), *bzip53* (SALK_060119), *bzip11* (SALK_123068). *pLBD16:GUS* (Okushima et al., 2007), *pPLT2:PLT2-YFP* (Galinha et al., 2007), *pSHR:SHR-GFP* (Gallagher et al., 2004), *pDR5:GUS* (Ulmasov et al., 1997), *pLBD16:GUS* (Okushima et al., 2007), *pWOX5:NLS-GFP* (Bureau et al., 2010), and *pCYCB1;2:CYCB1;2- NLS-YFP* (Iwata et al., 2011) were used. Primers used for genotyping are listed in Table 1. *pSDG8:SDG8-GFP*, *pSDG7:SDG7-GUS*, and *pSDG8:GUS* lines were kindly provided by Dr. Caroline Dean, Dr. Yifeng Xu, and lab members, respectively.

Arabidopsis seedlings were grown on half-strength Murashige and Skoog medium (MS) plates. MS salts (Duchefa Biochemie), 2-(N-Morpholino) ethanesulfonic acid (Nacalai Tesque, Kyoto), and agar (Nacalai Tesque, Kyoto) were added to distilled water. The pH was adjusted from 4.3 to 5.6 with KOH. After autoclaving at 121ºC for 20 min, 50 mL medium was poured into sterile No.2 square culture dishes (Eiken chemical) on a clean bench.

Seeds were surface-sterilized with 70% (v/v) ethanol for 10 min, washed three times with sterile distilled water, and sown on half-strength MS plates. After stratification at 4ºC in the dark for up to 7 days, the plates were placed in dish drainer trays located in a growth chamber at 22ºC under constant light conditions.

Table 2. Genotyping primers.

Primer name	Sequence (5'-3')
***sdg7-2* LP**	GTTACCTCAATAGGTCGACGCTC
***sdg7-2* RP**	GACATGTCGTTGTTGGAACG
***sdg8-2* LP**	TCGTCGTCTTGTAGTGGAAGC
***sdg8-2* RP**	TGCATTTGATGTTTTTGGTTG
SALK_LBb1.3 BP	ATTTTGCCGATTTCGGAAC

Leaf phenotyping

For leaf phenotypic analyses, the 5th true leaves from plants grown on soil were used. Leaves were excised with forceps and placed into glass vials containing FAA (formalin: ice-cold acetic acid:70% ethanol at a ratio of 1:1:18) for fixation. After 30 min fixation under vacuum, the leaves were transferred to a chloral hydrate solution and incubated for at least 16 h until clear. Leaves were then placed on glass slides (Matsunami Glass Industry Co., Ltd.), and covered with coverslips (Matsunami Glass Industry Co., Ltd.). Images were taken with an AXIO Zoom V16 stereo microscope (Zeiss). Leaf area and leaf length from the leaf tip to the petiole were measured with ImageJ software (NIH). Significant differences were determined based on one-way ANOVA tests. Statistical significance was computed using post-hoc Tukey's honest significant difference (HSD) test.

To quantify palisade cell size and number, the same leaf samples were used. The cells were imaged using AxioScope A light microscope (Zeiss). The epidermal cell area was measured with ImageJ software (NIH). Cell numbers were calculated from leaf and cell areas. Significant differences were calculated based on one-way ANOVA tests. Statistical significance was computed using post-hoc Tukey's HSD test.

Petal phenotyping

For petal phenotypic analyses, plants were grown on soil. Petals from stage 15 flowers were used. The petal area was measured with ImageJ software (NIH). Significant differences were calculated based on one-way ANOVA tests. Statistical significance was computed using post-hoc Tukey's HSD test.

To quantify cell size and number, the petal epidermal cells were imaged by scanning electron microscopy (SEM). For SEM, petals were placed in FAA (45% [v/v] ethanol, 2.5% [v/v] formaldehyde, and 2.5% [v/v] acetic acid), vacuum-infiltrated until the tissues sank, and left for at least 16 h at room temperature. The fixed tissues were then passed through a gradient ethanol series (50% [ethanol:water, v/v], 60%, 70%, 80%, 90%, 95%, 100% × 2) for 20 min each, followed by a gradient acetone series (25% [acetone:ethanol, v/v], 50%, 75%, 95%, 100% × 2) for 30 min each. Then, the tissues were dried with an EM CPD300 critical point dryer with liquid CO_2 (Leica Microsystems). Samples were gold-coated with a coating time of 45 s using a gold coater Hitachi E 1010. The tissues were imaged with a S-4700 scanning electron microscope (Hitachi) with an accelerating voltage of 15 kV. At least 10 petals for each genotype were observed; representative images are shown. Cell area was measured with ImageJ software (NIH). Cell numbers were calculated from petal and cell areas. Significant differences were calculated based on one-way ANOVA tests. Statistical significance was computed using post-hoc Tukey's HSD test.

Plant height phenotyping

For plant height analyses, plants were grown on soil. The final height of the main inflorescence was measured. Images of the entire plant were taken, and the height was measured with ImageJ software (NIH).

Significant differences were calculated based on one-way ANOVA tests. Statistical significance was computed using post-hoc Tukey's HSD test.

Root phenotypic analyses

For root phenotypic analyses, seedlings were grown on vertically oriented MS plates. Root length of 7-day-old seedlings after germination (7 DAG) was measured. Images of the entire seedling were taken. The length from the base of the hypocotyl to the root tip was measured with ImageJ software (NIH).

To quantify lateral root number, 7-DAG root samples were used. The lateral root number was counted using a microscope (Olympus SZ). Significant differences were calculated based on one-way ANOVA tests. Statistical significance was computed using post-hoc Tukey's HSD test.

Plasmid construction and transformation

For yeast two-hybrid assays, the full-length coding sequences of *bZIP*s (*bZIP1*, *bZIP2*, *bZIP9*, *bZIP11*, *bZIP13*, *bZIP14*, *bZIP25*, *bZIP27*, *bZIP40*, *bZIP41*, *bZIP44*, *bZIP53*, *bZIP54*, and *bZIP55*) and the SET domain–encoding regions of *SDG7* and *SDG8* were amplified with PrimeSTAR GXL DNA Polymerase (Takara) using gene-specific primer pairs. The resulting DNA fragments were introduced into pENTR/D-TOPO vector, confirmed by sequencing, and recombined into either pDEST32 or pDEST22 by Gateway LR Clonase II Enzyme Mix. Primers used for cloning are listed in Table 2.

Table 3. Cloning primers.

Primer name	Sequence (5'-3')
***bZIP1*-CDS F**	CCACATGGCAAACGCAGAGAAG
***bZIP1*-CDS R**	TCATGTCTTAAAGGACGCC
***bZIP2*-CDS F**	CACCATGGCGTCATCTAGCAGC
***bZIP2*-CDS R**	TCAATACATATTGATATCATTAGCC
***bZIP11*-CDS F**	CACCATGGAATCGTCGTCGTCG
***bZIP11*-CDS R**	TTAATACATTAAAGCATCAGAAGA
***bZIP44*-CDS F**	CACCATGAATAATAAAACTGAAATGGGATC
***bZIP44*-CDS R**	CTAACAGTTGAAAACATCACCA
***bZIP53*-CDS F**	CACCGATCATGTTGGATCTCCTAATG
***bZIP53*-CDS R**	TCAGCAATCAAACATATCAGC
***bZIP41*-CDS F**	CACCATGGGAACGAGCGAAGACA
***bZIP41*-CDS R**	TTAATTTGTTCCTTCACCATCTT
***bZIP54*-CDS F**	CACCATGGGTAGCAACGAAGAAG
***bZIP54*-CDS R**	TCAGCTAGCCGCGACA
***bZIP55*-CDS F**	CACCATGGGAAATAGCAGCGAG
***bZIP55*-CDS R**	TCAGCCTGCAGCTACTG
***bZIP13*-CDS F**	CACCATGACGTCGTTTCAGGTGATG
***bZIP13*-CDS R**	TCACCATTCCAAAGACCGA
***bZIP40*-CDS F**	CACCATGGCGTCCTTCAAGTTGAT
***bZIP40*-CDS R**	TCACCATTCCAAGGAATGGC
***bZIP14*-CDS F**	CACCATGTTGTCATCAGCTAAGCATC
***bZIP14*-CDS R**	TCAAAATGGAGCTGTGGAAGA
***bZIP27*-CDS F**	CACCATGGAAGAAGTATGGAAAGAAATC
***bZIP27*-CDS R**	TCAAAATGGAGCTGTGGAAGAC
***SDG7*-SET-F**	AATATATACTTGACGAAGAAAG
***SDG7*-SET-R**	TTTGCTAGGTTTTACACCAAGC
***SDG8*-SET-F**	AACCAGTTCCTTCATCGCAATC
***SDG8*-SET-R**	ATTCAGAGGATCTCCCCCAATATA

Yeast two-hybrid screening

The coding sequences of the SET domain region of SDG7 or SDG8 were introduced into the pGBKT7 vector. The transcription factor library pDEST_GAD424 was used as previously reported (Mitsuda et al., 2010). First, baits were transformed into the Y2H GOLD strain, inoculated in SD medium lacking tryptophan, and cultured at 30°C for 2 days. The resulting colonies were then cultured at 30°C in liquid SD medium lacking tryptophan on a shaker at 200 rpm for 1 day. The resulting yeast strains were transferred to 50 mL YPDA medium and grown at 30°C until OD_{600} reached 0.5–0.8. These yeast pre-cultures were washed with sterile water and layered with 5 mL lithium acetate. Then 2.4 mL 50% (w/v)

polyethylene glycol (PEG), 360 μL 1 M lithium acetate, 50 μL carrier DNA (2.0 mg/mL) and the transcription factor library (1,000 ng/μL) were added, quickly mixed by vortexing and incubated at 30°C for 30 min to complete transformation. Dimethyl sulfoxide (400 μL) was then added and incubated at 42°C for 15 min before adding the mixture to 10 mL YPDA medium for growth at 30°C, with shaking at 200 rpm for 1 h. The resulting yeast cells were plated on selective SD medium lacking tryptophan, leucine, and histidine and cultured at 30°C for 2 days. The grown colonies were restreaked onto the same selective SD medium at 30°C for 2 days. The inserts of the clones harbored by these isolated colonies were sequenced; the sequence of each insert was then used as a query for BLAST to identify potential transcription factors interacting with SDG7 or SDG8.

Yeast two-hybrid assay

The pDEST22-bZIP prey and pDEST32-SDG bait constructs were co-transformed into Y2HGold Yeast Strain (Takara). After transformation, cells were plated on synthetic defined (SD) medium lacking leucine and tryptophan at 30°C for 2 days. Single colonies of double transformants were inoculated in SD medium lacking leucine and tryptophan and grown in a shaker at 30°C overnight. Equal cell density was spotted onto SD medium lacking leucine, tryptophan, and histidine and grown at 30°C for up to 4 days.

RNA analysis

Total RNA was extracted from the entire root of 5-day-old MS-grown seedlings. The plant mini kit (Qiagen) was employed to extract RNA following the manufacturer's instructions. First-strand cDNA synthesis was initiated from 500 ng total RNA per sample at 35°C for 15 min followed by 95°C for 5 min using PrimeScript RT Master Mix (Takara).

Real-time and cycling conditions were as follows: 95°C for 10 min; 40 cycles of 95°C for 20 s, 55°C for 10 s and 72°C for 30 s, with Faststart Essential DNA green master mix (Roche). The *TUBULIN BETA CHAIN 2* (*TUB2*) (*At5g62690*) gene was selected as internal control. Primer sequences are given in Table 3. Statistical analysis was performed in GraphPad Prism9.0 and Microsoft Excel. Significant differences were calculated based on one-way ANOVA tests. Statistical significance was computed using post-hoc Tukey's HSD test.

For semiquantitative RT-PCR, total RNA was extracted from 7-day-old whole seedlings. *ACTIN 2* (*ACT2*) (*At3g18780*) was used as internal control; PCR amplicons were visualized by gel electrophoresis. Primers are listed in Table 3.

Table 4. RT-PCR primers.

Primer name	Sequence (5'-3')
***MAF1* F**	GCTGGCAACAGATGATGAGA
***MAF1* R**	TAATTGAGCAGCGGGAGAGT
***MAF3* F**	AGCGAATCGAGAACAAAAGC
***MAF3* R**	CAACAGCGATGGAAGATTCA
***PSD* F**	CGTCCTGGAATTGGTCTCAT
***PSD* R**	AAGGCTCGACCTTTGGAAAT
***DCL1* F**	TTGCGAGGACATCTCTTGTG
***DCL1* R**	GCCGTTTTCATCAGTGGATT
***DCL4* F**	TTGAAAGGCTTGAGGTGCTT
***DCL4* R**	CGTCTCCGAGTCAACTCTCC
***SDG8* F**	GCTGGATGATGGAAAGGAAA
***SDG8* R**	TTGTGCCACCATAACCTTCA
***BON1* F**	AATCGACGCTCCAGTTTCAC
***BON1* R**	GTTGAAAAGGGCACTGGTGT
***BAK7* F**	CAAAAGGCTTTCGATCTTGC
***BAK7* R**	GCTGCTCCACTTCTGTTTCC
***MLO2* F**	AAGGGGAAAGTGGCTTTTGT
***MLO2* R**	ACGACTTCCACGTCCTCATC
***SDG7* F**	ACAATAAGCCGTTCCAACA
***SDG7* R**	TCTTCTTCTGCCACAATCC
***ACT2* F**	ACCTTGCTGGACGTGACCTTACTGAT
***ACT2* R**	GTTGTCTCGTGGATTCCAGCAGCTT
***TUB2* F**	GCTCATTTCCAAGATCCGCG
***TUB2* R**	GTAGTTGAGGTCACCGTAGGAGG

Chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR)

To conduct ChIP, up to 0.8 g (fresh weight) of 5-day-old seedlings grown on half-strength MS plates was harvested in vials filled with phosphate-buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 2 mM KH_2PO_4, pH 7.4). Fixation was performed with 1% (v/v) formaldehyde for 15 min. During fixation, tissues were vacuum-infiltrated three to five times until they sank. Afterwards, samples were moved to PBS with 0.125 M glycine solution for 5 min to quench the fixation. Seedlings were rinsed with ice-cold PBS twice, frozen in liquid nitrogen, and stored in a deep freezer until use.

Tissues were ground to a fine powder using mortar and pestle and dissolved in Nuclei Extraction Buffer (100 mM MOPs pH 8.0, 10 mM $MgCl_2$, 0.25 M sucrose, 0.5% [w/v] dextran T-40, 2.5% [w/v] Ficoll 400, 9.36 μL/mL β-mercaptoethanol, and 10 μL/mL protease inhibitors) to extract protein–DNA complexes. After filtration over two layers of Miracloth and centrifugation, chromatin pellets were resuspended in nuclei lysis buffer (10 mM EDTA pH 8.0, 50 mM Tris-HCl pH 8.0, and 1% [w/v] SDS). ChIP dilution buffer (16.7 mM Tris-HCl pH 8.0, 1.2 mM EDTA, 167 mM NaCl, 1.1% [v/v] Triton and 0.01% [w/v] SDS) was then added to each sample. Sonication was conducted with an ultrasonic disruptor (TOMY) to obtain 200- to 700-bp sheared DNA. ChIP dilution buffer with 22% (v/v) Triton X-100 was mixed in with the sonicated samples and centrifuged to remove debris. After preclearing by Dynabeads™ Protein A (Thermo Fisher) and collection of Input DNA, immunoprecipitation was conducted at 4°C overnight using Dynabeads and commercial antibodies (H3K36me3, ab9050; H3K36me2, ab9049 Abcam). Beads were successively washed in low salt buffer (2 mM EDTA, 20 mM Tris-HCl pH

8.1, 150 mM NaCl, 0.1% [w/v] SDS and 1% [v/v] Triton X-100), LiCl buffer (0.25 M LiCl, 1% [w/v] Nonidet P-40, 1% [w/v] deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.1), and Tris EDTA buffer (TE, 10 mM Tris-HCl, 0.1 mM EDTA pH 8.1) twice each for 10 min at 4°C. Elution of protein–DNA from beads was conducted in nuclei lysis buffer by incubation for 25 min at 65°C twice. After combining both eluates, crosslinking was reversed with NaCl overnight at 65°C. DNA was purified by a QIAquick PCR purification kit (Qiagen) according to the manufacturer's instructions with slight modifications. qPCR was performed as described above. Primer sequences are listed in Table 4. Significant differences were calculated based on one-way ANOVA tests. Statistical significance was computed using post-hoc Tukey's HSD test.

Table 5. ChIP-qPCR primers.

Primer name	Sequence (5'-3')
***MAF3*-H F**	TTTCCTCTTCTGGGTCTCACTC
***MAF3*-H R**	TTGTGGTAACGCATAGGCTCTA
***MAF3*-M F**	CCAACCGATATATACTGTATCC
***MAF3*-M R**	TCTTTGACATGCTGGAGATA
***MAF3*-T F**	CTGATAGAAGAGAACCAGATTC
***MAF3*-T R**	TTGACGCAAGTGTAATGATG
***MLO2*-H F**	GTAGTGTTGATAGTGATAGGAG
***MLO2*-H R**	AATTGGAACCGTAAGTCTCT
***MLO2*-M F**	CAGGATCTGCATAACAAGAC
***MLO2*-M R**	TCAGTATTCCAACGGTATGT
***MLO2*-T F**	GGCATCCATGTAACACTAATG
***MLO2*-T R**	CTATGTGACTCTTCCACTCTAT
***DCL1*-H F**	TGCTTGCGAGGACATCT
***DCL1*-H R**	CATCAGTGGATTCATTGACAG
***DCL1*-M F**	CACAACAGGTGGACATAATAG
***DCL1*-M R**	CACGCAATACTTCAAGACAT
***DCL1*-T F**	TTCCTCCTGGTCGGTTAA
***DCL1*-T R**	CGTGACGAAGGTACAAGT
***TA3* F**	CTGCGTGGAAGTCTGTCAAA
***TA3* R**	CTATGCCACAGGGCAGTTTT

β-glucuronidase (GUS) staining

For GUS staining, 5-day-old seedlings grown on MS plates were fixed in ice-cold 90% (v/v) acetone for 10–15 min. The samples were washed with sterile water, followed by GUS buffer twice (50 mM phosphoric acid buffer, 0.5 mM $K_3Fe[CN]_6$ [Nacalai Tesque, Kyoto], 0.5 mM $K_4Fe[CN]_6$ [Nacalai Tesque, Kyoto]). Subsequently, tissues were incubated at 37°C for 30 min to 1 h in GUS buffer with 1 mM X-Gluc (5-bromo-4-chloro-3-indolyl β-D-glucuronide) (Nacalai Tesque, Kyoto). To stop the staining reaction, samples were transferred into 70% (v/v) ethanol.

To observe GUS-stained tissues, samples were washed in sterile water three times, incubated in 4% (v/v) HCl with 20% (v/v) methanol for 15 min at 55°C. Then, the solution was changed to 70% (v/v) ethanol with 7% (w/v) NaOH and incubated for another 15 min at room temperature. Subsequently, tissues were successively incubated in 40% (v/v), 20%, 10% ethanol for 10 min. Finally, samples were incubated in 10% (v/v) ethanol with 50% (v/v) glycerol for 30 min and observed with an upright microscope (Zeiss Scope.A1) by differential interference contrast microscopy. Stages of lateral root development were defined based on a previous publication (Malamy & Benfey, 1997).

Confocal microscopy

To observe fluorescence signals, 5-DAG seedlings were stained with 10 μg/mL propidium iodide (PI) for 2 min to stain the cell wall. Roots were observed by confocal laser scanning microscopy (using either LSM710, Zeiss, or TCS SP8 SMD, Leica). Fluorescence signals were collected at 488 nm, 561 nm, and 527 nm (Ar laser) to observe green fluorescent protein (GFP), PI, and cyan fluorescent protein (CFP) signals, respectively.

RNA sequencing (RNA-seq)

To conduct RNA-seq, 5-day-old seedlings were used. About 2 cm of root tips were cut, frozen in liquid nitrogen, and stored in the deep freezer until use. The RNeasy Plant Mini Kit (Qiagen) and the RNase-Free DNase Set (Qiagen) were used to extract RNA and remove genomic DNA contamination, respectively. Libraries for RNA-seq were generated following the Breath Adapter Directional sequencing (BrAD-seq) method (Townsley et al., 2015). mRNA polyA-adapter priming was performed by PCR with 5X Thermo Scientific RT buffer and 3' priming adapter. Subsequently, cDNA was synthesized by RevertAid reverse transcriptase (Thermo Fisher Scientific) and purified by AMpure XP beads. 5' sequencing adapters were added by the breath capture method to generate strand-specific libraries using DNA Pol I. Adapters were extended by using oligonucleotides containing adaptors with various indices and Phusion High-Fidelity DNA Polymerase (New England Biolabs). After confirmation of library size distribution and titer by agarose gel electrophoresis and a microplate photometer, the resulting libraries were sequenced using NextSeq 500 (Illumina).

Gene Ontology (GO) term enrichment analysis

To perform GO enrichment analysis, 785 differentially expressed genes between wild type (WT) and the *sdg7 sdg8* double mutant were examined with the online agriGO database (http://bioinfo.cau.edu.cn/agriGO/). Network and tree maps were generated after reduction of duplicated GO terms by REVIGO (http://revigo.irb.hr/).

ChIP-seq

For ChIP-seq, about 2 cm of root tips from 5-day-old WT and *sdg7 sdg8* double mutant seedlings were used. Frozen tissues were ground

to a fine powder using mortar and pestle and transferred to nuclei isolation buffer (10 mM HEPES, 1 M sucrose, 5 mM KCl, 5 mM $MgCl_2$, 5 mM EDTA, 16% [v/v] formaldehyde, 20% [v/v] Triton X-100, β-mercaptoethanol, 100 mM Pefabloc) for 10 min to crosslink proteins and DNA. The reaction was quenched by adding 0.125 M glycine. After removal of debris by filtration through two layers of Miracloth, the resulting solution was centrifuged at 3,000g for 10 min. Pellets were resuspended in nuclei isolation buffer. Chromatin was separated using nuclei separation buffer (1 M HEPES, 3 M KCl, 1 M $MgCl_2$, 0.5 M EDTA) by centrifugation. The pellets were dissolved in SDS lysis buffer (1 M Tris-HCl, 10% [w/v] SDS, 0.5 M EDTA) and ChIP dilution buffer (50 mM Tris-HCl, 0.167 M NaCl, 1.1% [v/v] Triton-X, 0.11% [v/v] sodium deoxycholate).

Then, sonication was conducted with a Covaris M2 (M&S instruments) to obtain approximately 300-bp sheared DNA fragments. Immunoprecipitation was conducted using anti-H3K36me3, anti-H3K36me2, and anti-H3 antibodies (H3K36me3, ab9050; H3K36me2, ab9049; H3, abcam) and Dynabead Protein A (Thermo Fisher) for at least 16 h at 4℃. Beads were washed with low salt RIPA buffer (50 mM Tris-HCl, 0.15 M NaCl, 1 mM EDTA, 0.1% [w/v] SDS, 1% [v/v] Triton-X, 0.1% [v/v] sodium deoxycholate), high salt RIPA buffer (50 mM Tris-HCl, 0.3 M NaCl, 1 mM EDTA, 0.1% [w/v] SDS, 1% [v/v] Triton-X, 0.1% [v/v] sodium deoxycholate), LNDET buffer (0.25 M LiCl, 1% [v/v] IGEPAL, 1% [v/v] sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl) and 1×TE. Then, ChIP dilution buffer was added to beads and incubated at 65℃ overnight to reverse the crosslinking. RNase and Proteinase K were then added to digest residual RNA and proteins. ChIP DNA was purified by Monarch PCR & DNA Cleanup Kit (Monarc). To generate libraries for sequencing, a ThruPLEX DNA-seq Kit (RUBICON GENOMICS) was

used following the manufacturer's instructions. Libraries were purified by AMPure beads (Beckman Coulter) and sequenced on a NextSeq 500 instrument (Illumina).

Mapping of ChIP-seq reads was performed on a supercomputer as reported previously (Inagaki et al., 2017). Using Bowtie (version 1.2.2), bam files were created by selecting the same sequences matching the *Arabidopsis thaliana* TAIR10 reference genome. Based on the bam files, bed files were created by bedtools (version 2.27.0). The coverage was determined in R software (Version 3.5.2). Histone binding sites were visualized in the Integrative Genomics Viewer (IGV) browser (version 2.4.14).

III. Results

Comparison of *SDG7* and *SDG8* expression and their protein structure

SDG8 plays important roles during Arabidopsis growth and development. However, *sdg8* single mutants only have subtle phenotypic defects, regardless of importance of H3K36me3 during plant development. Since histone modification enzymes belonging to the same subfamily often function redundantly (Wang et al., 2006), I hypothesized such factors must exist. Thus, I first generated a phylogenetic tree from an alignment of Arabidopsis protein sequences for class II SDG family members using Clustal W (Fig. 3A). Although SDG8 and SDG26 share the highest degree of similarity among class II family members, *sdg8* mutation is epistatic to the *sdg26* mutation during development, indicating that SDG8 and SDG26 do not function redundantly (Liu et al., 2016; Xu et al., 2008) (Fig. 3A).

I next assumed that SDG8 and redundant factors might have overlapped expression. Thus, I generated heatmap using expression values data from the *Arabidopsis* eFP website to compare expression pattern. Based on the previous expression data registered in the website, *SDG8* was expressed throughout development. Unlike *SDG8*, *SDG4* was specifically expressed in stamens only. Furthermore, *SDG24* was expressed at very low levels in almost all the issues, suggesting that neither of them is not likely to function with *SDG8*. The remaining family member among class II SDG family is *SDG7*. Like *SDG8*, *SDG7* was expressed throughout development (Fig. 3B). Hence, *SDG7* could be a potential candidate of redundant factor of *SDG8*.

To gain insight into protein function between SDG7 and SDG8, protein structure was compared. Both proteins contain well-conserved SET, post-SET, and AWS domains (Fig. 3C). In particular, 3D structure of conserved

domains were almost identical (Fig. 3D). I thus hypothesized that redundancy between SDG7 and SDG8 may exist during growth and development.

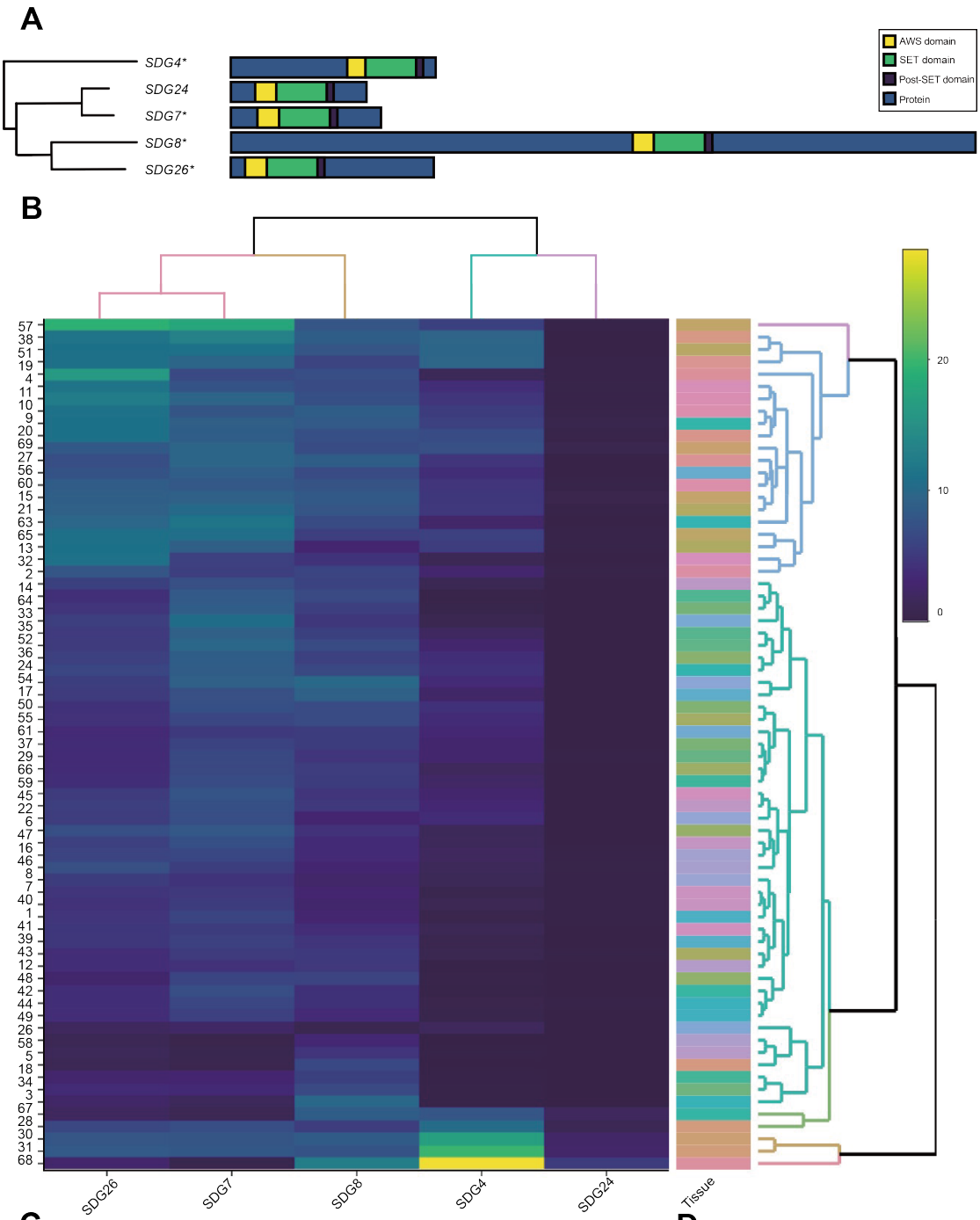
A
SDG4*
SDG24
SDG7*
SDG8*
SDG26*
AWS domain
SET domain
Post-SET domain
Protein
B
20
10
0
SDG26
SDG7
SDG8
SDG4
SDG24
Tissue
C
SDG7 ----------ESEEFELPEWLNKGKPTP----------YIFIR 48
SDG8 FADCSKSQEMSNEEINEELGIGQDEADAYDC-DAAKRGKEKEQKSKRLTGKQKACFKAIK 959
SDG7 RNIYLTKKVKRRVEDDGIFCSCSSSSPGSSST-VCGSNCHCGMLFSSCS-SSCKCGSECN 106
SDG8 TNQFLHRNRKSQTIDEIMVCHCK---PSPDGRLGCGEECLNRMLNIECLQGTCPAGDLCS 1016
SDG7 NKPFQQRHVKKMKLIQTEKCGSGIVAEEEIEAGEFIIEYVGEVIDDKTCEERLWKMKHRG 166
SDG8 NQQFQKRKYVKFERFQSGKKGYGLRLLEDVREGQFLIEYVGEVLDMQSYETRQKEYAFKG 1076
SDG7 ETNFYLCEITRDMVIDATHKGNKSRYINHSCNPNTQMQKWIIDGETRIGIFATRGIKKGE 226
SDG8 QKHFYFMTLNGNEVIDAGAKGNLGRFINHSCEPNCRTEKWMVNGEICVGIFSMQDLKKGQ 1136
SDG7 HLTYDYQFVQF--GADQDCHCGAVGCRRKLGVKP---------- 258
SDG8 ELTFDYNYVRVFGAAAKKCYCGSSHCRGYIGGDP--LNGDVIIQSDSDEEYPELVILDDD 1194
SDG7 ----------SKP----------KIASDEAFNLVAHELAQT--------- 279
SDG8 ESGEGILGATSRTFTDDADEQMPQSFEKVNGYKDLAPDNTQTQSSVSVKLPEREIPPPLL 1254
D
SDG7
SDG8

Fig. 3. Identification of SDG7 and SDG8.

(**A**) Phylogenetic tree of Arabidopsis class II SDG family proteins. Alignments of full-length sequences were conducted using Clustal W. The phylogenetic tree was generated by ETE3. (**B**) Heatmap of gene expression of Arabidopsis class II SDG family proteins in specific tissues. Heat maps display expression values change in Col-0. (**C**) Sequence alignment of SDG7 and SDG8 conserved domain. Blue, green, and red lines indicate the AWS, SET, and post-SET domains, respectively. Asterisks show conserved amino acids. (**D**) 3D structure of SDG7 and SDG8 conserved domain. Blue, green, and red lines indicate the AWS, SET, and post-SET domains, respectively.

Isolation of *sdg7* and *sdg8* mutants

To elucidate the role of SDG7 and SDG8 in Arabidopsis development, I obtained the *sdg7-2* (SALK_026442) and *sdg8-2* (SALK_131218) alleles for characterization. The T-DNA was inserted in the third intron of *SDG7* in *sdg7-2*, and in the second exon of *SDG8* in *sdg8-2* (Fig. 4A) (Grini et al., 2009; Xu et al., 2014). I conducted genotyping by PCR on individual plants. With the *SDG7* gene-specific primer set (*sdg7* LP and RP), I amplified a DNA fragment from *SDG7* using genomic DNA extracted from wild type (Col-0) as template, but not from *sdg7-2* mutant plants (Fig. 4B). PCR using the *sdg7* RP and SALK_LBb1.3 BP primer pair confirmed the presence of the T-DNA in *sdg7-2*, but yielded no amplicon in the wild type. These results indicated that the *sdg7-2* insertion line is homozygous for the insertion. PCR using *SDG8* gene-specific and T-DNA primer sets also confirmed that the *sdg8-2* insertion line is homozygous for the insertion (Fig. 4B). I further examined the expression levels of *SDG7* and *SDG8* in wild type and the single mutants by RT-PCR. Although both genes are highly expressed in the wild type, I did not detect any transcripts in the respective single mutant backgrounds (Fig. 4C), indicating that *sdg7-2* and *sdg8-2* are knock-out mutants.

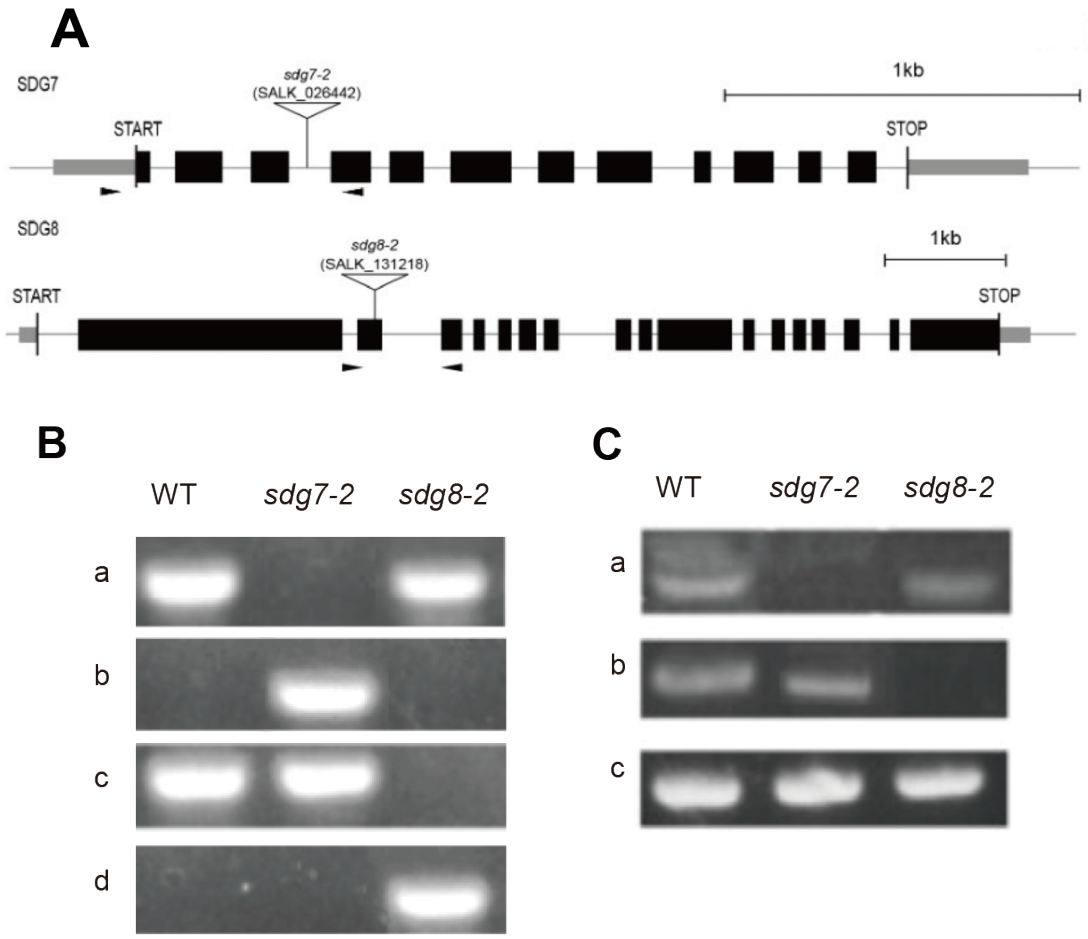

Fig. 4. Isolation of *sdg7* and *sdg8* mutants

(**A**) Schematic diagram of the *SDG7* and *SDG8* loci. The black boxes, gray boxes, and short lines inside gene bodies indicate exons, untranslated regions, and introns, qPCR amplified regions, respectively. White triangles indicate the position of T-DNA insertions; black arrowheads denote genotyping primers (*sdg7-2* LP, *sdg7-2* RP, *sdg8-2* LP, and *sdg8-2* RP). White arrowheads denote RT-PCR primer pairs. (**B**) Genotyping results for *sdg7-2* and *sdg8-2*. Primer pairs were a, *sdg7-2* LP, *sdg7-2* RP; b, *sdg7-2* RP, SALK_LBb1.3 BP; c, *sdg7-2* F, *sdg7-2* R; and d, *sdg8-2* LP, *sdg8-2* RP. (**C**) *SDG7*, *SDG8*, and *ACT2* mRNA levels based on semiquantitative RT-PCR using the primer sets; a, *SDG7* F, *SDG7* R; b *SDG8* F, *SDG8* R; and c *ACT2* F, *ACT2* R.

Expression pattern of *SDG7* and *SDG8* in aerial tissues

To uncover the spatial distribution pattern of *SDG7* and *SDG8* in aerial tissues, I used *pSDG7:SDG7-GUS* and *pSDG8:GUS* reporter lines in which the *ß-GLUCURONIDASE* (*GUS*) reporter gene was driven by the *SDG7* or *SDG8* promoter (Fig. 5). The *SDG7* construct showed a strong GUS signal in true leaves but an absence of staining in cotyledons during the vegetative phase (Fig. 5A). Based on *pSDG7:SDG7-GUS* sections, I detected intense GUS signal throughout the shoot apical meristem (SAM) and leaf primordia (Fig. 5B). During the reproductive phase, I observed GUS staining from the *pSDG7:SDG7-GUS* reporter throughout the inflorescence including the inflorescence meristem, floral meristems, young flowers, inflorescence stems, and pedicels (Fig. 5C, D). Although I detected a strong GUS signal in the *pSDG7:SDG7-GUS* reporter line in developed stamens, developed petals showed no staining (Fig. 5C, D). Like the GUS staining pattern seen for *pSDG7:SDG7-GUS*, the *SDG8* promoter drove the expression of the *GUS* reporter gene in the SAM, leaf primordia, and true leaves (Fig. 5E, F). The *pSDG8:GUS* reporter was expressed in inflorescence meristems, floral meristems, and young flowers, but signal intensity decreased at later stages of flower development (Fig. 5G, H). I also observed GUS staining driven by *pSDG8:GUS* at the base of pedicels (Fig. 5G, H). Overall, the *pSDG7:SDG7-GUS* and *pSDG8:GUS* lines exhibited an overlapping expression pattern.

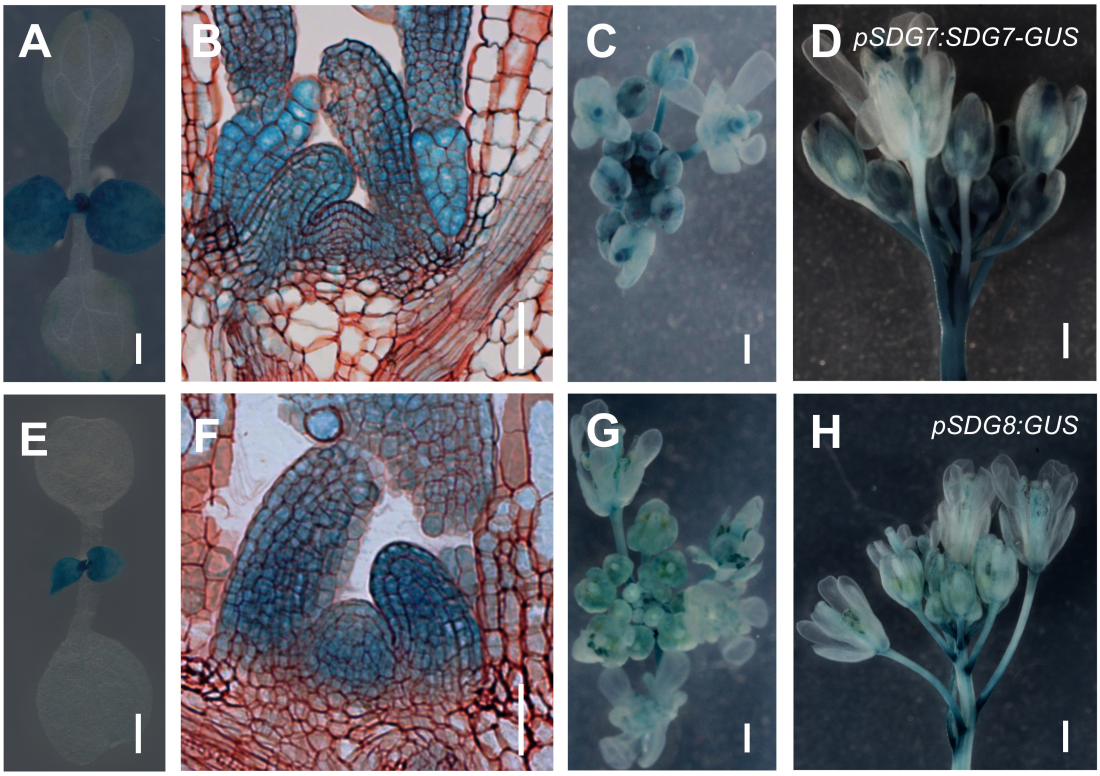

Fig. 5. *SDG7* and *SDG8* accumulation pattern in aerial parts of Arabidopsis.
(**A, B**) GUS staining pattern in *pSDG7:SDG7-GUS* transgenic lines in leaves (**A**) and the shoot apical meristem (**B**). (**C, D**) Top (**C**) and side (**D**) views of GUS staining pattern from the inflorescences of *pSDG7:SDG7-GUS* transgenic lines. (**E, F**) GUS staining pattern in *pSDG8:GUS* transgenic lines in leaves (**E**) and the shoot apical meristem (**F**). (**G, H**) Top (**G**) and side (**H**) views of GUS staining pattern from the inflorescences of *pSDG8:GUS* transgenic lines. Scale bars = 500 μm (**A**, **C**, **D**, **E, G, H**) or 50 μm (**B**, **F**).

Expression pattern of *SDG7* and *SDG8* in underground tissues

I also observed SDG7 and SDG8 accumulation patterns in underground tissues in the *pSDG7:SDG7-GUS* line and in a *pSDG8:SDG8-GFP* line (Fig. 6). In primary roots, GUS staining in the *pSDG7:SDG7-GUS* line were detected in root apical meristem basal meristem (Fig. 6A). During lateral root development, *SDG7* appeared to be expressed by stage II. The accumulation of the SDG7-GUS fusion protein continued until later stages of lateral root development (Fig. 6B, C). Furthermore, I detected GUS signal in the *pSDG7:SDG7-GUS* line in emerged lateral roots (Fig. 6D). Overall, the GFP signal from *pSDG8:SDG8-GFP* showed similar patterns as GUS staining in *pSDG7:SDG7-GUS* line that accumulated in the cortex, endodermis, pericycle, stele, and quiescent cells with a lower signal in columella cells. I observed a high GFP signal from *pSDG8:SDG8-GFP* in the meristematic root zones of primary roots (Fig. 6E). I also detected GFP signal for the *pSDG8:SDG8-GFP* line in lateral root primordia, starting in stage I until the emergence of lateral roots (Fig. 6F–H). Hence, SDG7 and SDG8 show similar accumulation patterns not only in aerial tissues, but also underground tissues.

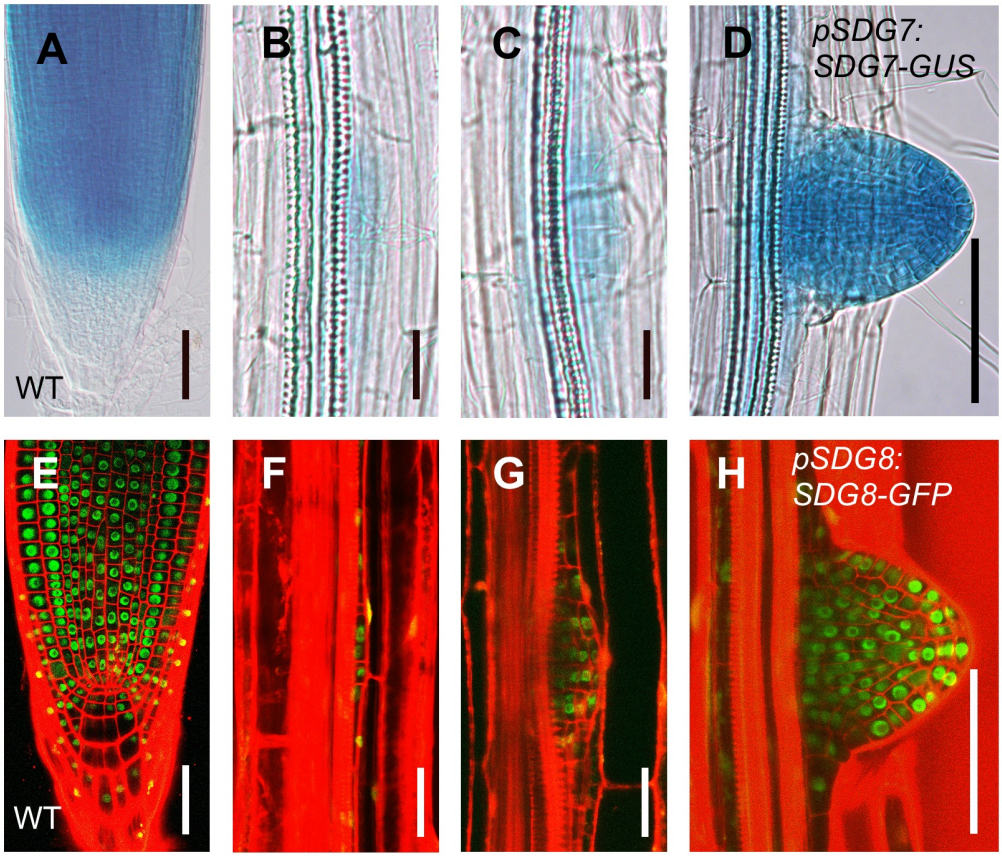

Fig. 6. *SDG7* and *SDG8* expression in Arabidopsis roots.

(**A–D**) GUS staining pattern from *pSDG7:SDG7-GUS* transgenic lines in the primary root tip (**A**), stage II lateral root (**B**), stage IV lateral root (**C**), and emerged lateral root (**D**). Scale bars = 50 μm. (**E–H**) GUS staining pattern from *pSDG8:SDG8-GFP* transgenic lines in the primary root tip (**E**), stage I lateral root (**F**), stage IV lateral root (**G**), and emerged lateral root (**H**). Scale bars = 50 μm.

Leaf phenotypes in the *sdg7 sdg8* double mutant

To explore the function of SDG7 and SDG8 during development, I generated the *sdg7 sdg8* double mutant by crossing the *sdg7* and *sdg8* single mutants. In agreement with the overlapping expression patterns of their wild type genes, the *sdg7 sdg8* double mutant exhibited a dwarf phenotype not seen in either single mutant during both vegetative and reproductive development (Figs. 7–10). I first characterized the leaf phenotypes of the wild type, the *sdg7* and *sdg8* single mutants, and the *sdg7 sdg8* double mutant (Fig. 7). I measured the final leaf blade length of the 5th true leaf in all genotypes, which had an average length in wild type of 1.5 ± 0.2 cm (Fig. 7A, C). Compared to the wild type, *sdg7* and *sdg8* had shorter leaf blades, with lengths of 1.3 ± 0.3 cm (*sdg7*, $p = 4.7 \times 10^{-3}$), and 0.9 ± 0.3 cm (*sdg8*, $p = 5.6 \times 10^{-3}$) (Fig. 7A, C). The *sdg7 sdg8* double mutant displayed significantly shorter leaf blades than either single mutant, which were only about one-third the length of wild type leaves (0.49 ± 0.17 cm; $p = 5.6 \times 10^{-20}$). I also quantified the leaf area in wild type, *sdg7*, *sdg8*, and *sdg7 sdg8* (Fig. 7A). In wild type, the average leaf size was 95.9 ± 5.2 mm^2 (Fig.7A, D). The size of the 5th true leaf in the *sdg7* single mutant was slightly smaller or similar to that of wild type at 77.6 ± 6.3 mm^2 ($p > 0.05$) (Fig. 7A, D). Leaves in the *sdg8* single mutant were less than half of the wild type (37.2 ± 4.4 mm^2; $p = 2.7\times10^{-12}$) (Fig. 7A, D). The *sdg7 sdg8* double mutant exhibited the smallest leaf size among all genotypes: the area in the double mutant was only 1.7 ± 0.2 mm^2 ($p = 2.7 \times 10^{-12}$) (Fig. 7A, D). These results suggest that SDG7 and SDG8 control leaf size.

To further understand the underlying cellular defects, I quantified palisade cell size and number in wild type, *sdg7*, *sdg8* and *sdg7 sdg8* (Fig. 7B). I detected no significant difference in cell size between the wild type (2,003 ± 93 μm^2) and the *sdg7* single mutant (1,985 ± 57 μm^2; $p > 0.05$)

(Fig. 7E). Palisade cell size in *sdg8* decreased compared to the wild type (1,552 ± 72 μm^2; $p = 2.4 \times 10^{-4}$) (Fig. 7E). Cell size in the *sdg7 sdg8* double mutant was the smallest among the four genotypes (667 ± 68 μm^2; $p = 1.1 \times 10^{-16}$) (Fig. 7E). I also inferred leaf cell number in wild type, the *sdg7* and *sdg8* single mutants and the *sdg7 sdg8* double mutant from the collected data for leaf area and cell area. There were 49,204 ± 2,769 palisade cells in wild type (Fig. 7F). The *sdg7* and *sdg8* single mutants had slightly fewer cells than the wild type (*sdg7*: 39,475 ± 3,286 cells; $p = 3.5 \times 10^{-2}$; *sdg8*: 24,077± 2,531 cells; $p = 8.7 \times 10^{-9}$) (Fig. 7F). The *sdg7 sdg8* double mutant had fewer cells than any other genotype (3,067 ± 407 cells; $p = 1.1 \times 10^{-16}$ when compared to wild type) (Fig. 7F). These results suggest that the reduced leaf size in the *sdg8* and *sdg7 sdg8* double mutant are due to defects in both cell size and cell number.

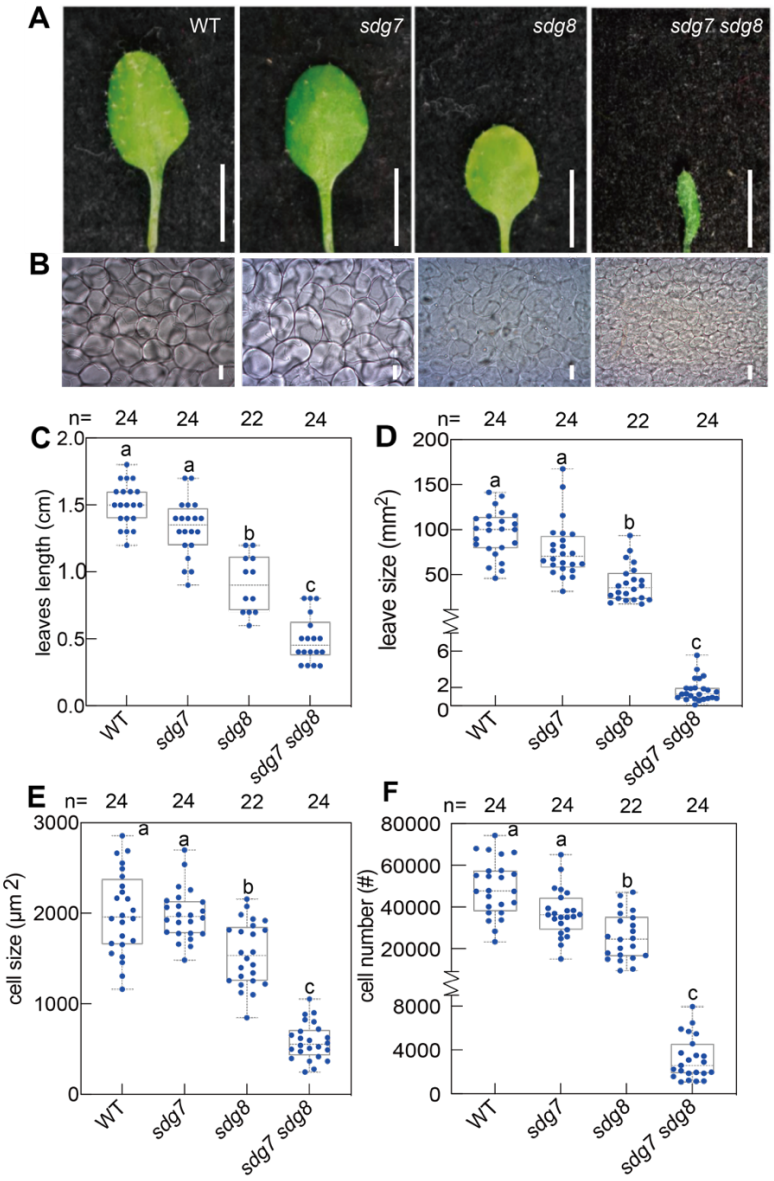

Fig. 7. Phenotypes of the *sdg7 sdg8* double mutant in true leaf.

(**A**) Top view of 5th true leaf in Col-0 (wild type, WT), *sdg7*, *sdg8,* and *sdg7 sdg8*. Scale bars = 1 cm. (**B**) Appearance of palisade cells from 5th true leaves in wild type, *sdg7*, *sdg8*, and *sdg7 sdg8*. Scale bar = 20 μm. (**C–F**) Leaf length (**C**), leaf area (**D**), cell size (**E**), and cell number (**F**) in wild type, *sdg7*, *sdg8*, and *sdg7 sdg8*. n > 10. Significant differences were calculated based on one-way ANOVA tests ($p < 0.05$). Different letters indicate significant differences based on post-hoc Tukey's HSD test.

Primary inflorescence length in the *sdg7 sdg8* double mutant

I also measured the final length of the primary inflorescence in wild type, *sdg7*, *sdg8*, and the *sdg7 sdg8* double mutant after they ceased elongating (Fig. 8). The primary inflorescences of wild type and the *sdg7* single mutant were over 30 cm in length. I observed no significant difference in the length of stem between wild type and the *sdg7* mutant ($p > 0.05$) (Fig. 8A, B). As previously reported, *sdg8* had a shorter primary inflorescence than that of wild type (Dhami & Cazzonelli, 2020). The length of the primary inflorescence in *sdg8* was 22.4 ± 8.3 cm ($p = 5.3 \times 10^{-7}$) (Fig. 8C). The *sdg7 sdg8* double mutant showed a synergistic interaction for this phenotype, with a very short primary stem (1.7 ± 0.72 cm; $p = 1.0 \times 10^{-27}$) (Fig. 8D) that terminated before reaching 2 cm in length (Fig. 8D). These results suggest that SDG7 and SDG8 are both important for proper elongation of the primary inflorescence.

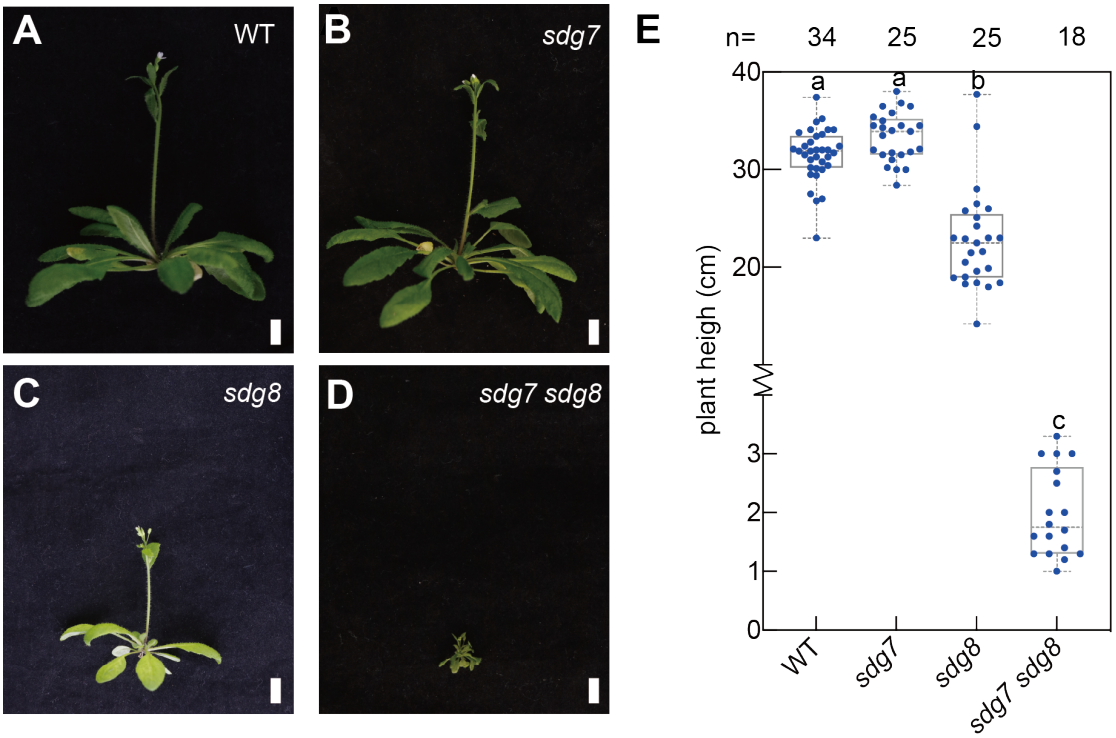

Fig. 8. Inflorescence stem phenotypes of the *sdg7 sdg8* double mutant.

(**A–D**) Side view of soil-grown wild type (**A**), *sdg7* (**B**), *sdg8* (**C**), and *sdg7 sdg8* (**D**) plants at the time of first flower opening. Scale bars = 1 cm. (**E**) Plant height in wild type, *sdg7*, *sdg8*, and *sdg7 sdg8*. n > 17. Significant differences were calculated based on one-way ANOVA tests ($p < 0.05$). Different letters indicate significant differences based on post-hoc Tukey's HSD test.

Petal phenotypes in the *sdg7 sdg8* double mutant

To understand the role of SDG7 and SDG8 during flower development, I also examined a number of flower traits in the double mutant. Flowers from the wild type and the *sdg7* single mutant contained all four properly sized floral organs at stage 15: sepals, petals, stamens and carpels (Fig. 9A). At the same stage of flower development, flowers of the *sdg8* single mutant were smaller than those of the wild type or the *sdg7* mutant (Fig. 9A). The *sdg7 sdg8* double mutant displayed even smaller flowers compared to the *sdg8* single mutant (Fig. 9A). Although sepals were visible in *sdg7 sdg8*, petals were difficult to notice when viewed from above due to their dramatic size reduction (Fig. 9A). In particular, petal size differed considerably between wild type, *sdg7*, *sdg8*, and *sdg7 sdg8* (Fig. 9B, D). Petal area in wild type and *sdg7* reached 1.03 mm^2. No significant difference was observed between wild type and *sdg7* ($p > 0.05$) (Fig. 9B, D). The petal area in *sdg8* was noticeably smaller than the wild type at 0.41 ± 0.01 mm^2 ($p = 0.1 \times 10^{-2}$). The petals from the *sdg7 sdg8* double mutant were the smallest among all four genotypes (0.03 ± 0.02 mm^2, $p = 0.1 \times 10^{-2}$).

To further understand these cellular defects, I scored epidermal cell size and cell number in wild type, *sdg7*, *sdg8* and *sdg7 sdg8* (Fig. 9C, E, F). I detected no significant difference for petal cell size between wild type, *sdg7* or *sdg8* single mutants ($p > 0.05$). Only the *sdg7 sdg8* double mutant showed a significant reduction in petal cell size (29.3 ± 9.3 μm^2; $p = 1.1 \times 10^{-16}$) (Fig. 9E). Cell numbers were decreased in *sdg8* and double mutant. In particular, wild type petals had about 14,127 ± 3,000 epidermal cells at stage 15 of flower development. The number of petal epidermal cells in *sdg7* and *sdg8* decreased significantly to 12,587 ± 2,674 (*sdg7*, $p = 1.1 \times 10^{-2}$) and 5,724 ± 1,354 (*sdg8*, $p = 1.1 \times 10^{-16}$) (Fig. 9F). The *sdg7 sdg8*

double mutant displayed a further reduction in cell number to only 1,198.7 ± 929 cells ($p = 1.1 \times 10^{-16}$) (Fig. 9F). These results indicate that the reduced size of petals in the *sdg7 sdg8* double mutant is due to defects in both cell size and cell number.

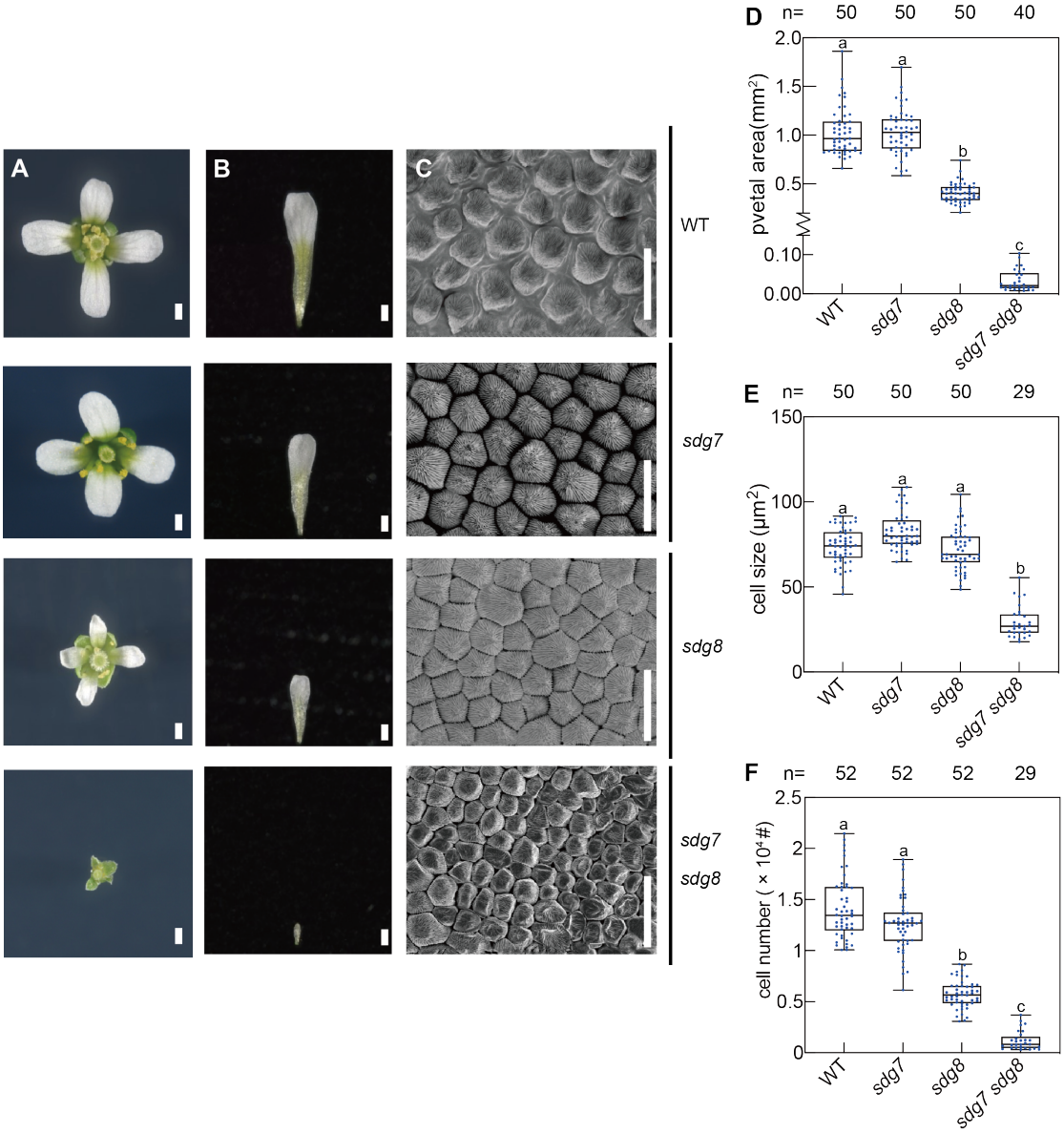

Fig. 9. Phenotypes of the *sdg7 sdg8* double mutant in flowers.

(**A**) Top view of stage 15 flowers from soil-grown wild type (WT), *sdg7*, *sdg8,* and *sdg7 sdg8*. Scale bars = 25 μm. (**B**) Shape of petals in wild type, *sdg7*, *sdg8,* and *sdg7 sdg8*. Petals were excised from flowers at stage 15. Scale bars = 25 μm. (**C**) Petal epidermal cells in wild type, *sdg7*, *sdg8*, and *sdg7 sdg8*. (**D**–**F**) Petal area (**D**), petal cell number (**E**), and petal cell size (**F**) in wild type, *sdg7*, *sdg8*, and *sdg7 sdg8*. n > 40. Significant differences were calculated based on one-way ANOVA tests ($p < 0.05$). Different letters indicate significant differences based on post-hoc Tukey's HSD test.

Root phenotypes in the *sdg7 sdg8* double mutant

Next, I examined the role of SDG7 and SDG8 in ground tissue development by measuring primary root length, lateral root number, and lateral root density in WT, *sdg7*, *sdg8*, and *sdg7 sdg8* at 7 DAG (Fig. 10). In wild type, primary root length of 7-DAG seedlings reached 6.5 cm. Primary root length in the *sdg7* single mutant was similar to that of the wild type ($p > 0.05$), while primary root length in the *sdg8* single mutant was significantly shorter ($p = 6.7\times10^{-12}$) than that of the wild type with a length of 5.3 ± 0.9 cm at 7 DAG. The *sdg7 sdg8* double mutant exhibited the shortest primary root among four genotypes at the same age, with a length of 2.9 ± 0.5 cm ($p = 2.1 \times 10^{-19}$ relative to WT) (Fig. 10B). These results suggest that SDG7 and SDG8 contribute to primary root elongation.

I also counted the number of lateral roots. Wild type at 7 DAG formed about 20 lateral roots including primordia from the primary root. Both *sdg7* and *sdg8* single mutants had fewer lateral roots compared to wild type ($p = 3.8 \times 10^{-3}$ and $p = 1.0 \times 10^{-3}$, respectively). The *sdg7 sdg8* double mutant at 7 DAG only produced three lateral roots ($p = 1.8 \times 10^{-8}$) (Fig. 10C). The reduction in lateral root number may be due to a reduction in primary root length. To test this possibility, I calculated the density of lateral roots (lateral root number divided by the root portion where they are present) for all four genotypes. The density of lateral roots in 7-DAG wild type seedlings was 3.2 lateral roots/cm. I detected a modest but significant decrease in lateral root density in the *sdg7* and *sdg8* single mutants (*sdg7*, $p = 1.0 \times 10^{-3}$; *sdg8*, $p = 1.3 \times 10^{-2}$). The *sdg7 sdg8* double mutant showed a strong reduction in lateral root density ($p = 2.7 \times10^{-10}$) (Fig. 10D), as it dropped to 1.5 ± 1.0 lateral roots/cm. Hence, the reduction in lateral root number seen in the *sdg7 sdg8* double mutant was not merely a reflection

of shorter primary roots. SDG7 and SDG8 are thus at least partially involved in lateral root formation.

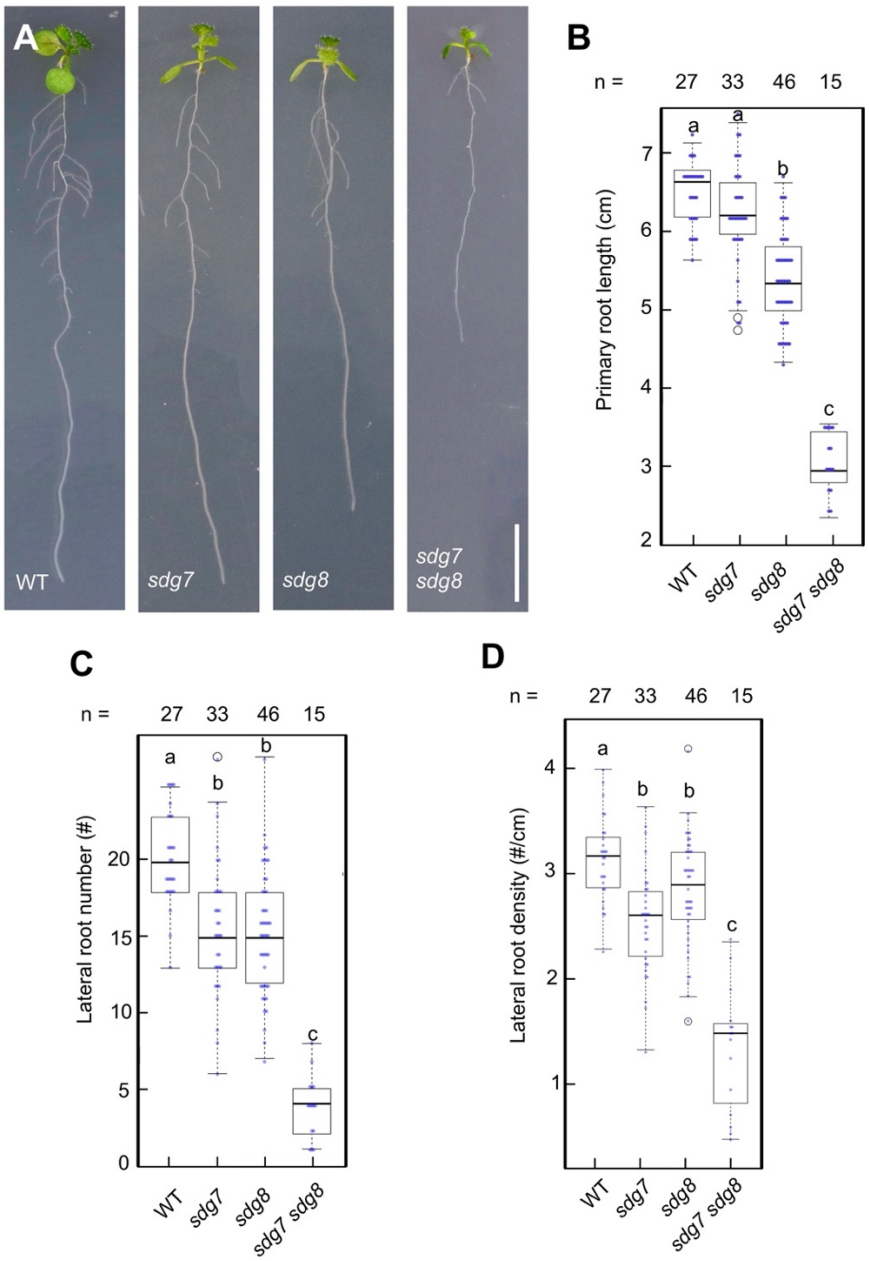

Fig. 10. Root phenotypes of the *sdg7 sdg8* double mutant.

(**A**) Root phenotypes of the wild type, *sdg7*, *sdg8*, and *sdg7 sdg8*. Scale bars = 1 cm. (**B–D**) Quantification of root phenotypes. Primary root length (**B**), lateral root number (**C**), and lateral root density (**D**) in wild type, *sdg7*, *sdg8*, and *sdg7 sdg8* 7 days after germination. n > 14. Significant differences were calculated based on one-way ANOVA tests ($p < 0.05$). Different letters indicate significant differences based on post-hoc Tukey's HSD test.

Anatomical and molecular defects in the *sdg7 sdg8* root apical meristem (RAM) and shoot apical meristem (SAM)

To examine cellular patterning in 7-DAG *sdg7 sdg8* double mutant seedlings, I characterized the structure of the RAM under the microscope (Fig. 11). In wild type, the quiescent center (QC) of the RAM controls stem cell fate in adjacent cells and organizes the cellular pattern of the stem cell niche, which consisted of four layers of columella cells located just below the QC (Fig. 11A). The *sdg7 sdg8* double mutant had a disorganized columella, and the QC was difficult to identify (Fig. 11B), as determined by staining of the root tip with PI and Lugol. The typical regular arrangement of columella cells in wild type was severely disrupted in *sdg7 sdg8* (Fig. 11C, D). The *sdg7 sdg8* double mutant presented irregular cell location, suggesting that SDG7 and SDG8 are required for organizing the typical layered structure of the RAM.

To monitor cell cycle progression and the synthetic auxin response in shoot apical meristem, I used the G2-M phase marker *pCYCB1;2:CYCB1;2-NLS-YFP* (Iwata et al., 2011) and *pDR5:GUS* (Ulmasov et al., 1997) (Fig. 12). In wild type, actively dividing cells were enriched in leaf protoderm and relative faint signal in the center of shoot apical meristem (Fig. 12A). By contrast, signal was almost dismissed in the *sdg7 sdg8* double mutant (Fig. 12B). As for auxin response, I observed strong GUS staining derived from *pDR5:GUS* activity in the apex of leaf primordia apex in wild type (Fig. 12C). Leaf apical enrichment was completely disappeared in *sdg7 sdg8* double mutant and GUS signal was overall weaker than that of WT. (Fig. 12C, D). These results suggest that loss of SDG7 and SDG8 may cause defects in cell cycle progression and weak auxin response in the SAM.

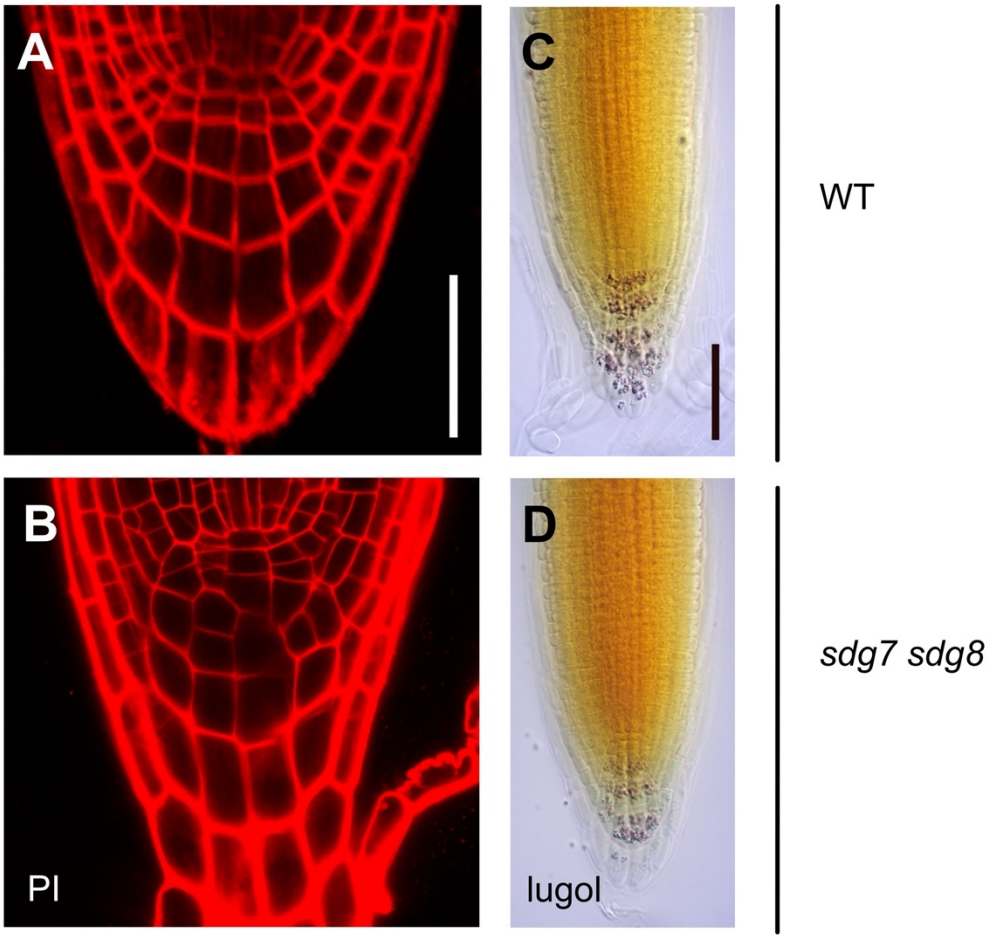

Fig. 11. Root apical meristem and columella cells in the *sdg7 sdg8* double mutant. (**A, B**) Cellular organization of the root meristem, as visualized by propidium iodide (PI) staining in wild type (**A**) and *sdg7 sdg8* (**B**). (**C, D**) Cellular organization of the columella root cap, as visualized by Lugol staining in wild type (**C**) and *sdg7 sdg8* (**D**). Scale bars = 50 μm.

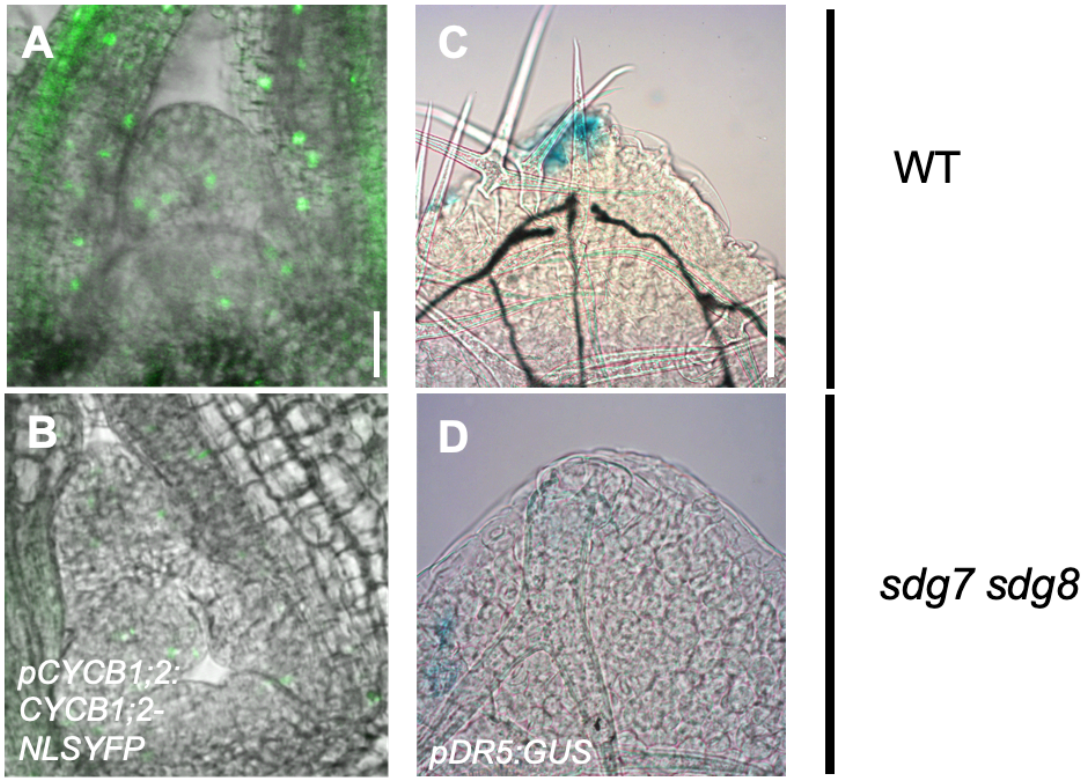

Fig. 12. Cell cycle and auxin marker gene expression in shoot apical meristem and leaf primordia apex.

(**A–F**) Expression of marker genes *pCYCB1;2:CYCB1;2-NLS-YFP* (**A, B**) and *pDR5:GUS* (**C, D**) in wild type (**A, C**) and *sdg7 sdg8* (**B, D**). Scale bars = 50 µm

Next I monitored the *pDR5:GUS* in RAM and lateral root (Fig. 13A-D). In wild type, I observed strong GUS staining derived from *pDR5:GUS* activity in the QC, as well as a lower signal intensity in the columella layer (Fig. 13A). I detected GUS staining in the columella layer of *sdg7 sdg8* double mutant similar to the wild type (Fig. 10B). In agreement with the disorganized RAM seen in the *sdg7 sdg8* double mutant, *pDR5:GUS* activity was greatly reduced in the QC of the *sdg7 sdg8* double mutant (Fig. 13B). I also observed *GUS* expression from the *pDR5:GUS* reporter construct in the lateral root primordia of the wild type. In *sdg7 sdg8,* the *pDR5:GUS* signal was weaker than in the wild type (Fig. 13C, D). The lateral root primordia marker gene *LATERAL ORGAN BOUNDARIES-DOMAIN 16* (*LBD16*) was analyzed using the *pLBD16:GUS* reporter construct (Okushima et al., 2007) (Fig 13E, F). In both wild type and the *sdg7 sdg8* double mutant, strong GUS signal was detected in the root stele and lateral root primordia, albeit at a lower intensity in the double mutant (Fig. 13E, F). These results suggest that reduced lateral root number in *sdg7 sdg8* may be due to defects in auxin response or auxin level.

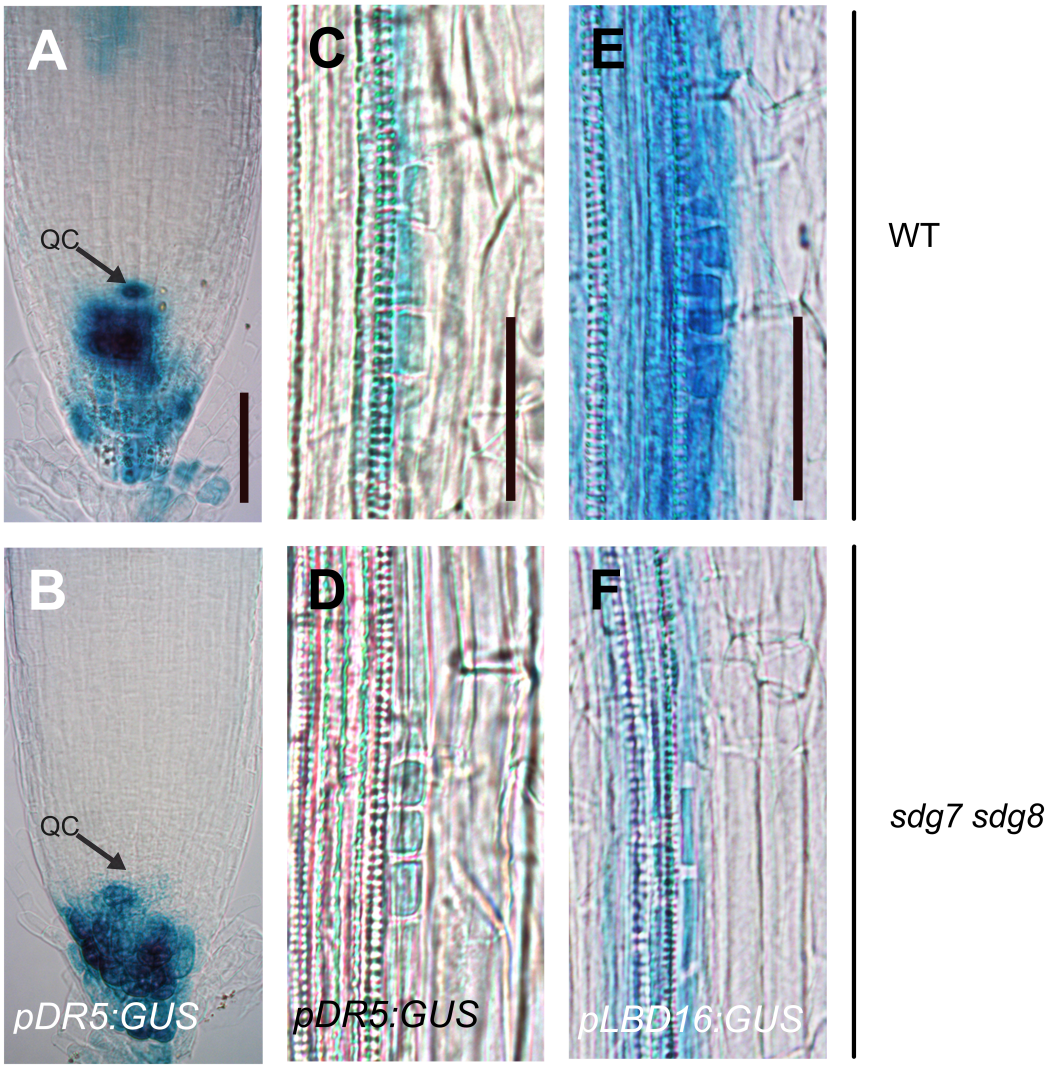

Fig. 13. Auxin and lateral root marker gene expression in primary and lateral roots.

(**A–F**) Expression of marker genes *pDR5:GUS* (**A, B**) and *pLBD16:GUS* (**C–F**) in wild type (**A, C, E**) and *sdg7 sdg8* (**B, D, F**) primary roots. Scale bars = 50 μm.

To gain more insight into the molecular basis of the observed defects in the *sdg7 sdg8* double mutant, I analyzed the expression pattern of several well-characterized reporter constructs in wild type and the *sdg7 sdg8* double mutant (Fig. 14). First, I examined the expression of *WUSCHEL RELATED HOMEOBOX 5* (*WOX5*), which is required for columella stem cell identity, using the *pWOX5:NLS-GFP* reporter. In the wild type, I detected *WOX5* promoter activity in the QC cells, as reported previously (Pi et al., 2015) (Fig. 14A). In the *sdg7 sdg8* double mutant, however, *WOX5* expression was often absent from the presumptive QC cells and occasionally observed in their neighboring cells (Fig. 14B). Another marker line made use of *PLETHORA 2* (*PLT2*), which controls cell division and differentiation around the RAM, with the *pPLT2:PLT2-YFP* reporter construct (Galinha et al., 2007) (Fig. 14C, D). In wild type, PLT2 accumulated to low levels in the stem cell area (Fig. 14C) but showed high abundance in meristematic (Fig. 14C). Although the distribution of the PLT2-YFP fusion protein was broadly the same in wild type and the *sdg7 sdg8* double mutant, fluorescence intensity decreased in the *sdg7 sdg8* double mutant (Fig. 14D). SHORT ROOT (SHR), which controls tissue patterning through asymmetric cell division, was also observed using the *pSHR:SHR-GFP* reporter (Gallagher et al., 2004) (Fig. 14E, F). In wild type, I detected strong activity in the QC, stele, and endodermis, as previously described (Fig. 14E). Correlated with the morphological changes, SHR-GFP signal was often absent in the *sdg7 sdg8* double mutant. Furthermore, the SHR-GFP signal was reduced in the stele and endodermis of the *sdg7 sdg8* double mutant compared to the wild type (Fig. 14E, F). To monitor cell cycle progression and the number of dividing cells, I used the G2-M phase marker *pCYCB1;2:CYCB1;2-NLS-YFP* (Iwata et al., 2011) (Fig. 14G, H). In wild type, G2/M-phase cells were marked by YFP fluorescence (Fig. 14G). By

contrast, I detected no signal in the *sdg7 sdg8* double mutant (Fig. 14H). Together with the short root phenotype of the double mutant, the data suggest that loss of SDG7 and SDG8 may cause defects in stem cell maintenance and cell cycle progression in the RAM.

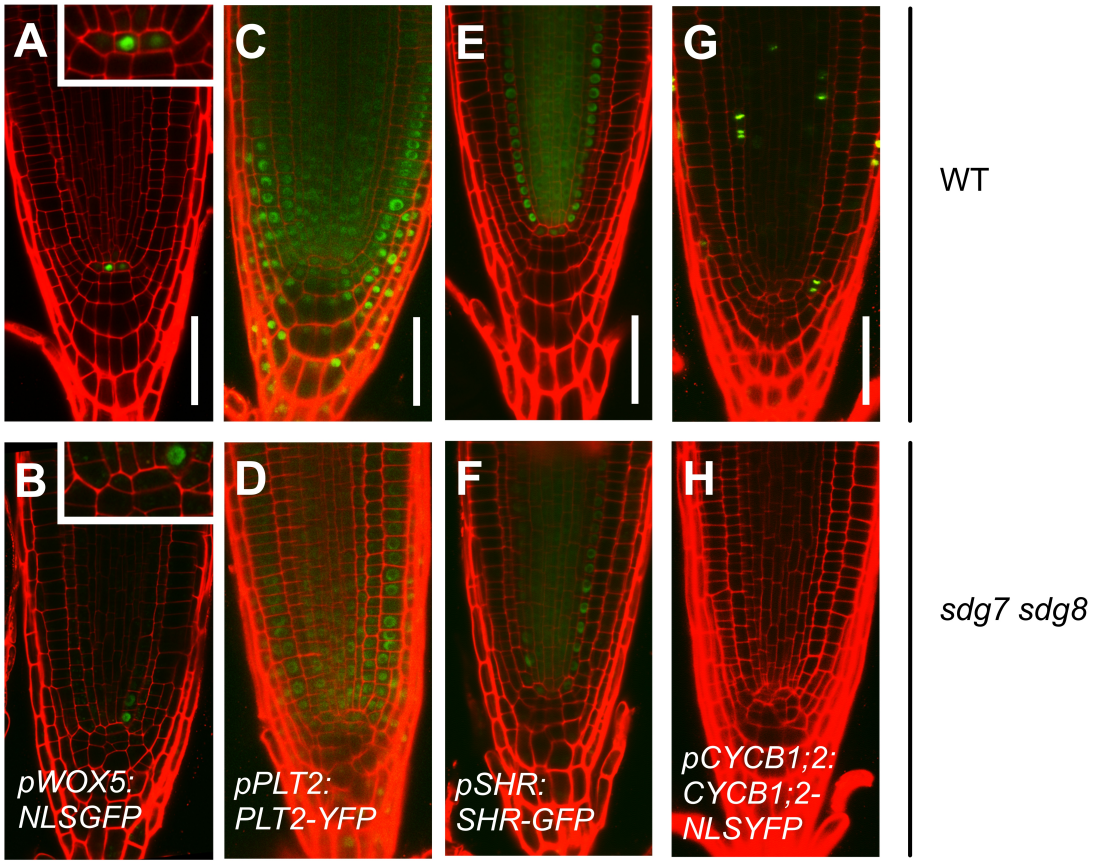

Fig. 14. Meristem and cell cycle marker gene expression in the primary root.

(**A–H**) Expression of cell markers in WT (**A**, **C**, **E**, **G**) and *sdg7 sdg8* (**B**, **D**, **F**, **H**). *pWOX5:NLS-GFP* (**A, B**), *pPLT2:PLT2-YFP* (**C, D**), *pSHR:SHR-GFP* (**E, F**), *pCYCB1;2:CYCB1;2-NLS-YFP* (**G, H**). Scale bars = 50 μm.

Positional preference and enrichment of H3K36 methylation in the *sdg7 sdg8* double mutant

To gain insight into the role of SDG7 and SDG8 on H3K36me3 and H3K36me2 deposition, 2 cm of root tips from 5-day-old WT and *sdg7 sdg8* double mutant seedlings were harvested to conduct ChIP-seq using specific antibodies. I identified 8,125 differentially H3K36me3-methylated genes and 13,005 differentially H3K36me2-methylated genes in the *sdg7 sdg8* double mutant when compared to the wild type (Fig. 15A, B). Of the 8,125 genes, 3,790 were hypomethylated, while the remaining 4,335 genes were hypermethylated for H3K36me3 in the *sdg7 sdg8* double mutant. (Fig. 15A). In the case of H3K36me2, 6,732 genes were hypomethylated and 6,273 genes were hypermethylated in the *sdg7 sdg8* double mutant relative to the wild type (Fig. 15B). Consistent with previous reports, averaged profiles in the wild type revealed that H3K36me3 and H3K36me2 reach their highest levels in the 5’ end and 3’ end of transcribed genes, respectively (Li et al., 2015; Xiao et al., 2016). The *sdg7 sdg8* double mutant showed higher levels of H3K36me3 marks at the 5’ end of genes compared to the wild type, and lower levels for marks located near the 3’ end of transcribed genes (Fig. 15C, E). By contrast, H3K36me2 marks were absent across the entire gene body of transcribed regions in the *sdg7 sdg8* double mutant (Fig. 15D, F). These results indicate that SDG7 and SDG8 are required to maintain proper distribution of H3K36me3 and H3K36me2 marks across genes.

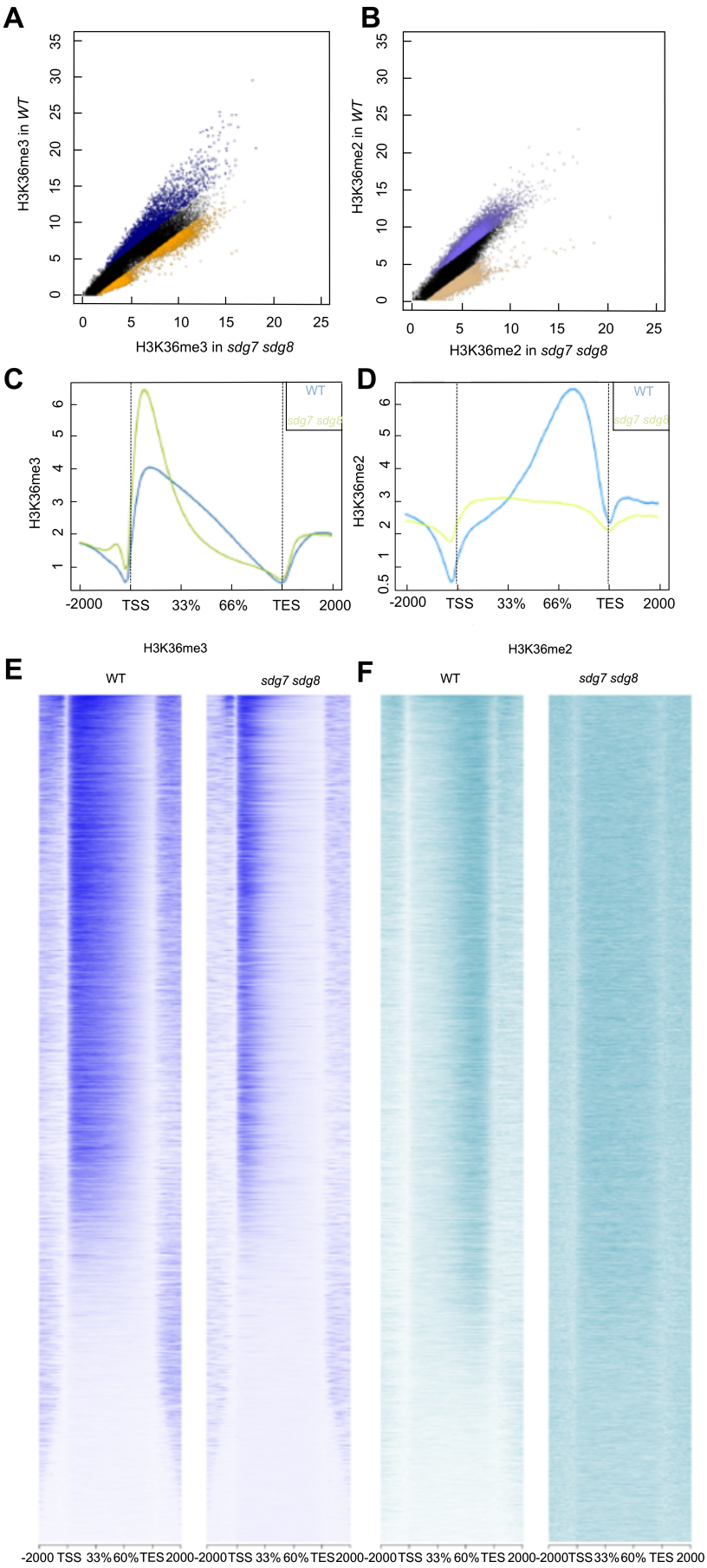
A
H3K36me3 in WT
H3K36me3 in sdg7 sdg8
B
H3K36me2 in WT
H3K36me2 in sdg7 sdg8
C
H3K36me3
WT
sdg7 sdg8
-2000 TSS 33% 66% TES 2000
D
H3K36me2
WT
sdg7 sdg8
-2000 TSS 33% 66% TES 2000
H3K36me3
H3K36me2
E
WT
sdg7 sdg8
F
WT
sdg7 sdg8
-2000 TSS 33% 60% TES 2000

Fig. 15. H3K36me3 and H3K36me2 accumulation in wild type and the *sdg7 sdg8* double mutant.

(**A, B**) Scatterplot of H3K36me3 (**A**) and H3K36me2 (**B**) levels in wild type and *sdg7 sdg8*. Hypermethylated and hypomethylated genes in wild type relative to *sdg7 sdg8* are shown in blue and yellow, respectively. (**C, D**) Averaged profiles of H3K36me3 (**C**) and H3K36me2 (**D**) signal intensity around genes in wild type and *sdg7 sdg8*. (**E, F**) Heatmaps of H3K36me3 (**E**) and H3K36me2 (**F**) accumulation across genes in wild type and *sdg7 sdg8*.

To examine the contribution of SDG-mediated H3K36 methylation on gene expression, I conducted RNA-seq analysis (Fig. 16), leading to the identification of 785 differentially expressed genes between the wild type and the *sdg7 sdg8* double mutant (false discovery rate < 0.05). Among these 785 genes, 239 genes were downregulated and 546 genes were upregulated in the *sdg7 sdg8* double mutant. I also determined that 106 of the 785 differentially expressed genes showed differential methylation in *sdg7 sdg8* for both H3K36me3 and H3K36me3 relative to the wild type (Fig. 16). Among them, 60 genes were both hypo-methylated and down regulated in *sdg7 sdg8* double mutant. These 106 genes are high-confidence genes regulated by SDG7 and SDG8 via H3K36 methylation.

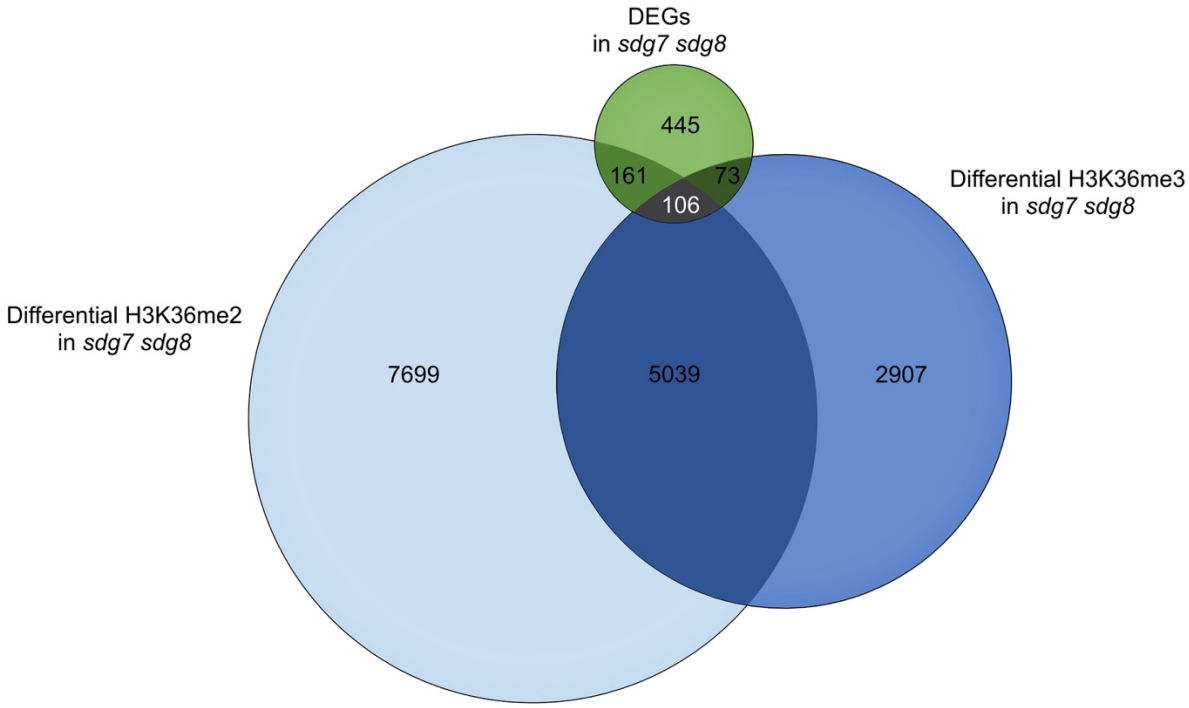

Fig. 16. H3K36me3, H3K36me2 enrichment, and differentially expressed genes in wild type and the *sdg7 sdg8* double mutant.

Venn diagram showing the overlap between differentially expressed genes (DEGS, green) and genes with differential methylation for H3K36me2 (pale blue) and H3K36me3 (blue) in *sdg7 sdg8*.

To explore the function of these 106 high-confidence genes, I performed a GO term enrichment analysis, after removing redundant GO terms with REVIGO (Fig. 17A, B). These genes were associated with plant development, such as post-embryonic development, flower development, and embryonic development, as expected based on the phenotypes of the *sdg7 sdg8* double mutant (Fig. 17A, B). For example, *MADS AFFECTING FLOWERING 1* (*MAF1*) and *MADS AFFECTING FLOWERING 3* (*MAF3*) are important regulators of flowering time (Gu et al., 2013; Ratcliffe et al., 2001). *PAUSED* (*PSD*), which is required for proper anatomical organization of the SAM and the timing of leaf formation, was among the 106 genes (Hunter et al., 2003; Telfer et al., 1997), which also showed association with various metabolic processes and response to stimulus (Fig. 17A, B). *DICER-LIKE 1* (*DCL1*) and *DCL4*, which encode RNase III ribonucleases, are involved in RNA metabolism through RNA-mediated post-transcriptional gene silencing (Gasciolli et al., 2005). *MILDEW RESISTANCE LOCUS O 2* (*MLO2*) controls responses to mold and salt stress (Acevedo-Garcia et al., 2017; Stiti et al., 2011). *BRI1-ASSOCIATED KINASE 7* (*BAK7*) and *BONZAI 1* (*BON1*) contribute to both plant development and response to stimulus (Jeong et al., 2010; Yang & Hua, 2004).

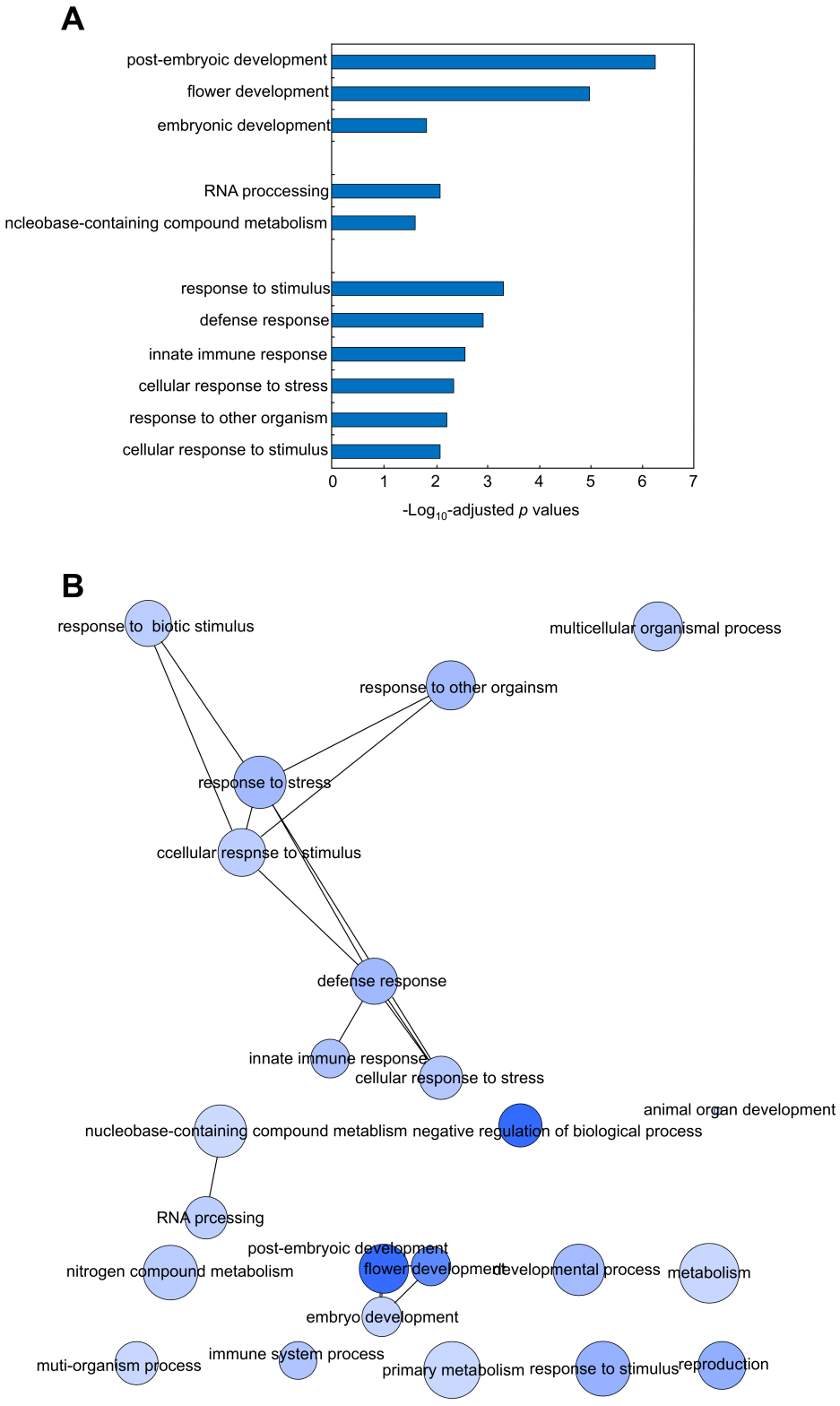

Fig. 17. Gene Ontology (GO) term enrichment analysis of 106 high-confidence genes.

(**A**) Selected GO terms determined by their –Log_{10}-adjusted *p* values are shown. (**B**) Interactive graph view of enriched GO terms, generated with REVIGO. Dark and pale colors indicate lower and higher *p* values, respectively.

Based on ChIP-seq analysis, the eight genes mentioned above (*MAF1*, *MAF3*, *PSD*, *DCL1*, *DCL4*, *MLO2*, *BAK7*, *BON1*) and *SDG8* showed different enrichment patterns for H3K36me3 and H3K36me2 in the *sdg7 sdg8* mutant background. In wild type, each gene had a higher H3K36me3 peak near the 5' end. The signal gradually decreased toward the 3' end of these genes, as often seen in H3K36me3 deposition pattern in Arabidopsis (Fig. 18A–I). H3K36me3 peaks at *MAF3*, *MAF1*, *DCL4*, *BON1*, *SDG8*, and *PSD* were significantly lower in the *sdg7 sdg8* double mutant than in the wild type (Fig. 18A–F). Although I observed comparable H3K36me3 peaks at the 5' end of *MLO2*, *BAK7* and *DCL1* in wild type and the *sdg7 sdg8* double mutant, H3K36me3 levels in the gene body greatly decreased in the double mutant (Fig. 18G–I). All nine genes showed higher H3K36me2 signals at the 3' end of genes in wild type, whereas the *sdg7 sdg8* double mutant lacked a clear signal (Fig. 18A–I).

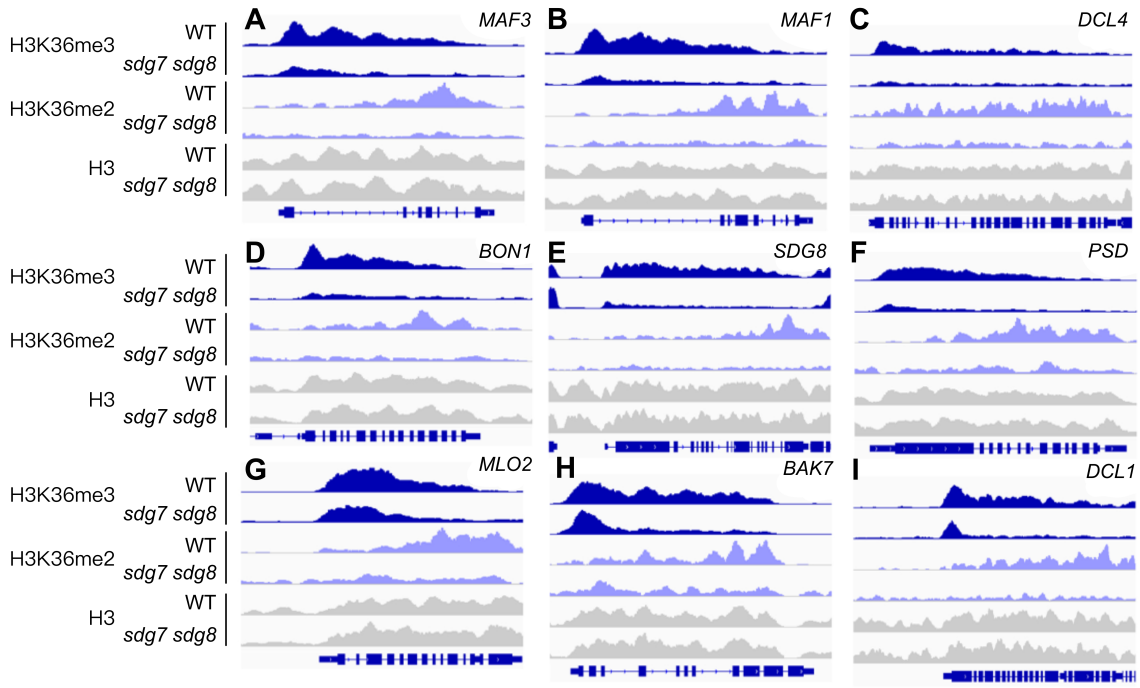

Fig. 18. H3K36me3 and H3K36me2 enrichment at nine selected loci.

(**A**–**I**) H3K36me3 and H3K36me2 enrichment over nine candidate loci: *MAF3* (**A**), *MAF1* (**B**), *DCL4* (**C**), *SDG8* (**D**), *MLO2* (**E**), *PSD* (**F**), and *BON1* (**G**) *BAK7* (**H**) and *DCL1* (**I**). For gene models: tall blue boxes, exons; blue boxes, untranslated regions; lines, introns. Dark blue, light blue and gray peaks represent the enrichment of H3K36me3, H3K36me2, and H3, respectively.

Validation of H3K36 methylation and gene expression analysis in the *sdg7 sdg8* double mutant

To validate the RNA-seq data, I conducted RT-qPCR. The nine high-confidence genes showed lower expression levels in the *sdg7 sdg8* double mutant compared to the wild type (Fig. 19A–I). I also assessed relative expression levels in the *sdg7* and *sdg8* single mutants. *MAF1*, *MAF3*, *BON1*, *MLO2*, and *DCL4* expression levels were significantly reduced in the *sdg8* single mutant (Fig. 19A–E). In fact, *MAF3*, *MAF1* and *BON1* mRNA levels were similar in the *sdg8* single mutant and the *sdg7 sdg8* double mutant (Fig 19A, B). By contrast, *MLO2* and *DCL4* expression levels were significantly lower in the *sdg7 sdg8* double mutant compared to the *sdg8* single mutant. For the remaining four genes, I detected no difference in expression between wild type and the *sdg8* single mutant (Fig. 19F–I). Furthermore, the relative expression levels for all nine genes tested were comparable between wild type and the *sdg7* single mutant (Fig. 19A–I). I concluded that *PSD*, *DCL1*, *SDG8*, and *BAK7* are redundantly regulated by SDG7 and SDG8. *MLO2* and *DCL4* may be regulated by SDG7 and SDG8 in a partially redundant manner. Moreover, *MAF1*, *MAF3* and *BON1* appear to be regulated only by SDG8.

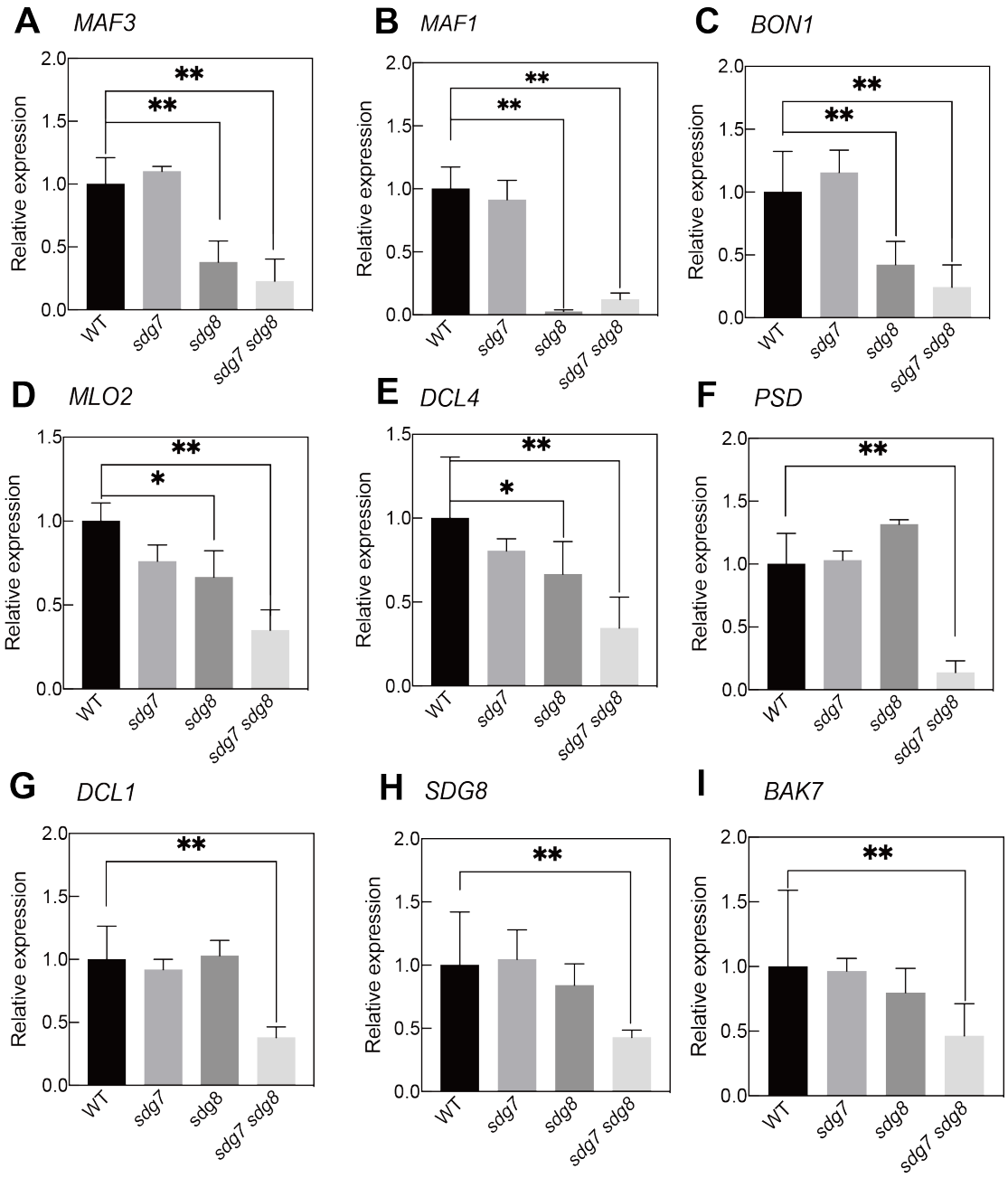

Fig. 19. Relative expression levels of selected genes in the *sdg7 sdg8* double mutant. (**A**–**I**) Relative expression levels of the candidate genes *MAF3* (**A**), *MAF1* (**B**), *BON1* (**C**), *MLO2* (**D**), *DCL4* (**E**), *PSD* (**F**), and *DCL1* (**G**) *SDG8* (**H**) and *BAK7* (**I**) in 5-DAG root tissues. n = 3. Significant differences were calculated based on one-way ANOVA tests. Asterisks indicate significant differences based on post-hoc Tukey's HSD test. (*, $p < 0.05$, **, $p < 0.001$). *TUB2* served as the reference transcript.

To further understand histone modifications through SDG7 and SDG8, I examined H3K36me3 and H3K36me2 levels by ChIP-qPCR (Figs. 20, 21). In wild type, I detected higher (for *MLO2*) and lower (for *DCL1*) H3K36me3 levels over the 5’ end, the middle region (P1 and P2) and the 3’ end (P3) (Fig. 20A–D). H3K36me3 signals at the P1 position were comparable to the wild type in the single mutants and in the double mutant (Fig. 20C, D). Consistent with the ChIP-seq results, H3K36me3 levels at the P2 position of the *MLO2* and *DCL1* loci were slightly lower in single and double mutants compared to the wild type (Fig. 20C, D). I obtained almost no signals at the P3 position of the *MLO2* and *DCL1* loci in the single and double mutants or in the wild type (Fig. 20C, D). I detected high H3K36me2 levels at the *MLO2* and *DCL1* loci over the 3’ end (P3) in the wild type (Fig. 20E, F) and lower levels in the *sdg7* and *sdg8* single mutants (Fig. 20E, F). H3K36me2 levels further decreased at the P3 position in the *sdg7 sdg8* double mutant (Fig. 20E, F). These results suggest that the loss of either SDG7 or SDG8 is sufficient to affect H3K36me2 deposition at the *MLO2* and *DCL1* loci.

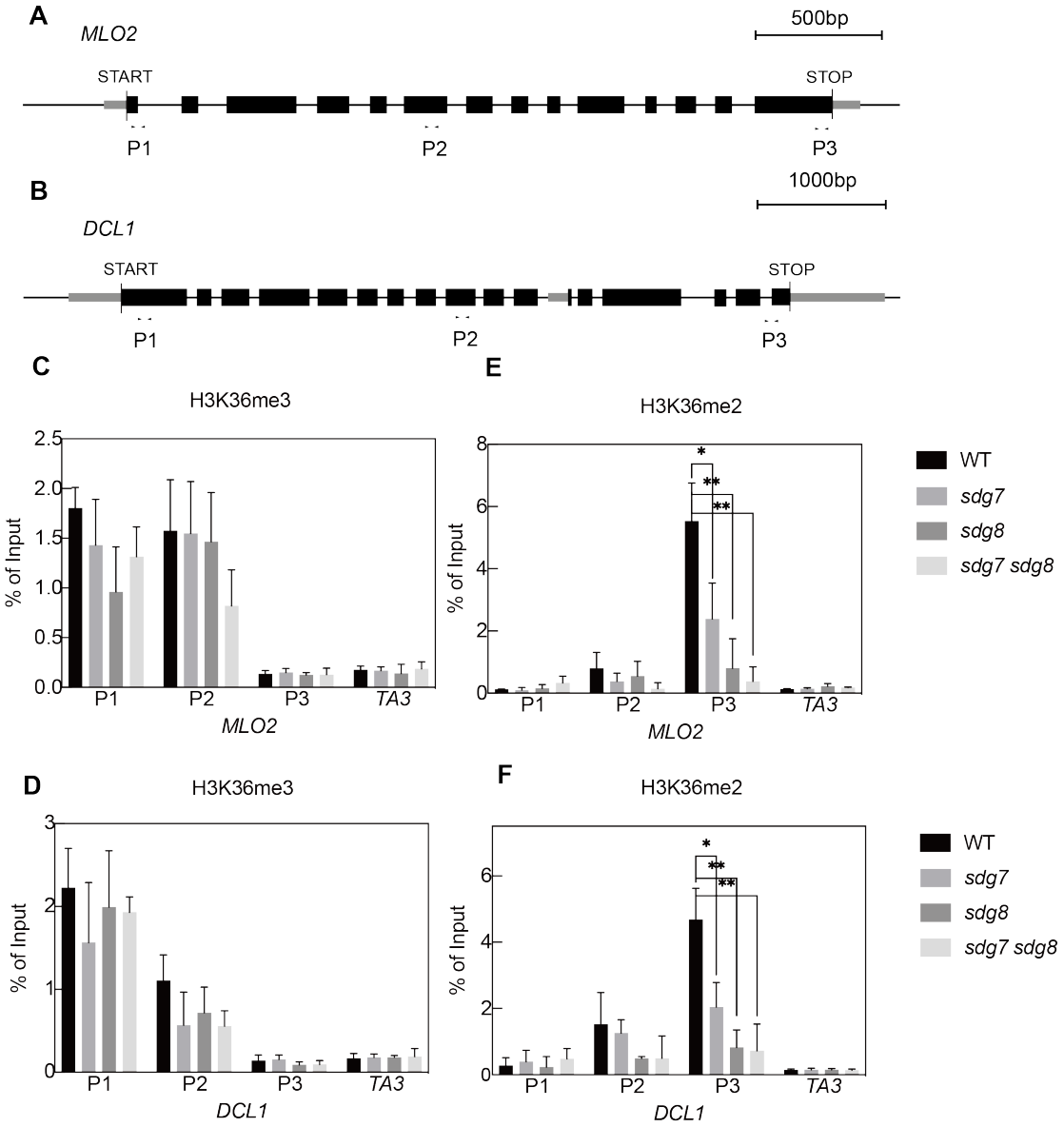

Fig. 20. H3K36me3 and H3K36me2 levels at the *MLO2* and *DCL1* genes.

(**A**, **B**) Schematic diagram of the *MLO2* (**A**) and *DCL1* (**B**) loci and PCR amplicons. The black boxes, gray boxes, short lines inside gene bodies and black arrowheads indicate introns, exons, untranslated regions, and qPCR amplified regions, respectively. H3K36me3 (**C-F**) H3K36me3 (**C, D**) and H3K36me2 (**E, F**) enrichment at the *MLO2* (**C**, **E**) or *DCL1* (**D**, **F**) locus in 5-DAG seedlings of wild type, single mutants, and double mutants. n = 3. Significant differences were calculated based on one-way ANOVA tests. Asterisks indicate significant differences based on post-hoc Tukey's HSD test. (*, $p < 0.05$, **, $p < 0.001$).

Turning to *MAF3*, I also detected higher H3K36me3 levels and lower H3K36me2 levels in wild type over the 5’ end, (P1) and 3’ end (P3) (Fig. 21A–C). Although I observed no significant change between wild type and the *sdg7* single mutant, H3K36me3 and H3K36me2 levels decreased to the same extent in the *sdg8* single mutant and the *sdg7 sdg8* double mutant (Fig. 21A–C). This result indicates that SDG8 mainly controls H3K36me3 and H3K36me2 levels at the *MAF3* locus.

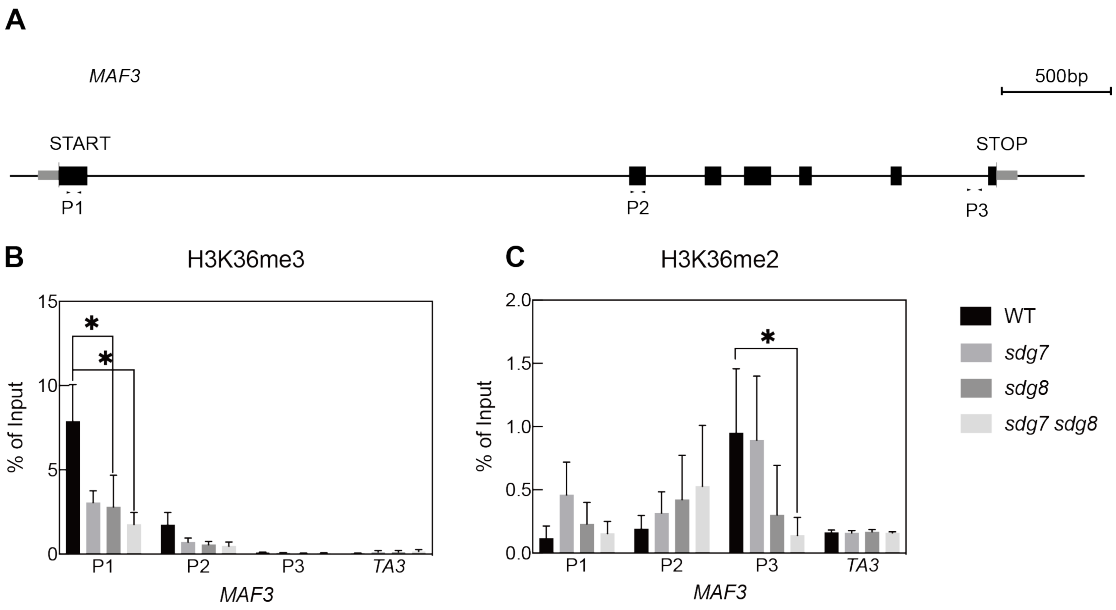

Fig. 21. H3K36me3 and H3K36me2 levels at the *MAF3* gene.

(**A**) Schematic diagram of the *MAF3* locus and PCR amplicons. The black boxes, gray boxes, short lines inside gene bodies and black arrowheads indicate introns, exons, untranslated regions, and qPCR amplified regions, respectively. (**B**, **C**) H3K36me3 (**B**) and H3K36me2 (**C**) enrichment in 5-DAG seedlings for the wild type, single mutants, and double mutants. n = 3. Significant differences were calculated based on one-way ANOVA tests. Asterisks indicate significant differences based on post-hoc Tukey's HSD test. (*, $p < 0.05$, **, $p < 0.001$).

Identification of *cis*-regulatory elements linked to SDG7 and SDG8 function

To understand how SDG7 and SDG8 control gene expression via histone modifications, I searched for conserved *cis*-regulatory elements using 2 kb of sequences upstream of the translation start site for all 106 high-confidence genes (Fig. 22). Motif analysis using DREME identified four motifs containing an ACGT core, including C-box and TGA binding motif. Based on comprehensive DNA affinity purification followed by sequencing (DAP-seq) analysis, the Arabidopsis transcription factors bZIP53 and bZIP50 recognize C-boxes (Fig. 22A, B) (O'Malley et al., 2016). Nucleotides flanking the ACGT core bound by TGACG MOTIF-BINDING PROTEIN 2 (TGA2), and TGA9/bZIP21 were also identified (Fig. 22C, D). Furthermore, I also detected an enrichment for the G-box motif recognized by bZIP69 and similar sequences with an AGCT core among the regulatory sequences of the 106 candidate genes (Fig. 22E, G). In addition, two motifs containing CAT tandem repeats were over-represented (Fig. 22F, H). As previously reported, distribution of H3K36me3 and H3K36me2 marks are highly accumulated in 5' end and 3' end of transcribed genes, respectively (Li et al., 2015). Those locations differed from these *cis*-regulatory elements in 5' UTR. These results suggest that SDG7 and SDG8 control target genes via *cis*-regulatory elements recognized by bZIP transcription factors.

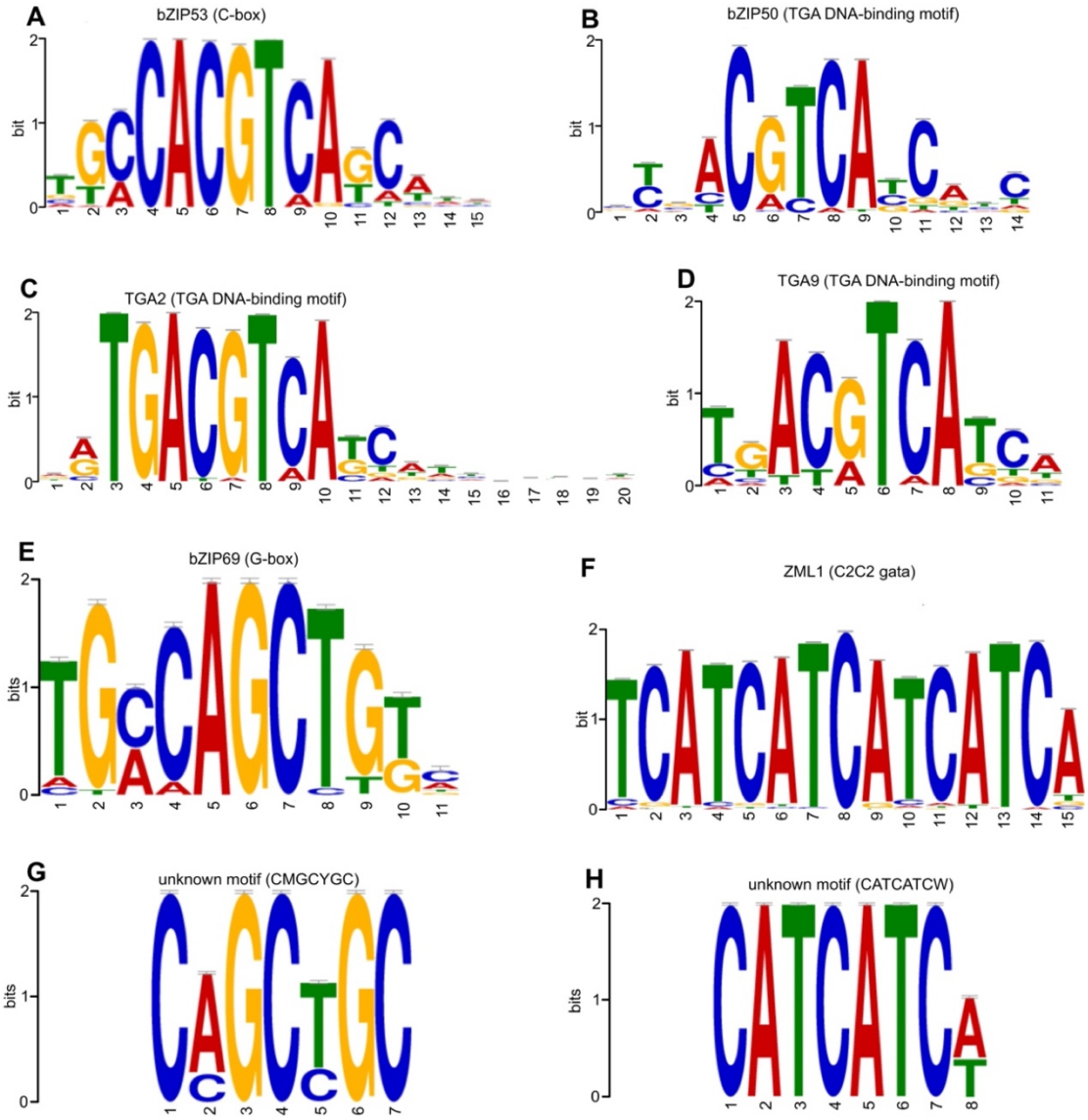

Fig. 22. Identification of *cis*-regulatory motifs from 106 genes.

(**A**–**F**) Known *cis*-regulatory motifs for bZIP53 (**A**) (bZIP S1 subfamily), bZIP50 (**B**) (bZIP S1 subfamily), TGA2 (**C**) (bZIP D subfamily), TGA9 (**D**) (bZIP D subfamily), bZIP69 (**E**) (bZIP D subfamily), and ZML1 (**F**) (GATA-type transcription factor). (**G, H**) Unknown over-represented motifs such as CATCATCW (**G**) and CMGCYG (**H**).

Identification of transcription factors interacting with SDG7 and SDG8

The above identification of *cis*-regulatory motifs suggested that SDG7 and SDG8 may interact with transcription factors to be recruited to their target loci. To identify potential interacting partners, I performed a yeast two-hybrid screen with SDG7 and SDG8 as baits in pGBKT7 against a transcription factor library pDEST_GAD424 prey library (Fig. 23, Tables 5, 6) (Mitsuda et al., 2010). Both SDG7 and SDG8 interacted with transcription factors belonging to the same family. About half of SDG7- and SDG8-interacting partners were zinc-finger transcription factors, APETALA 2 (AP2)/ETHYLENE RESPONSE FACTOR (ERF) transcription factors, or MYB domain-containing transcription factors (Fig. 23A, B). I also identified MADS-box and WRKY-type transcription factors as common interacting partners for SDG7 and SDG8 (Fig. 23A, B). Consistent with the enriched *cis*-regulatory elements identified, several members of the bZIP transcription factor family also interacted with SDG7 and SDG8 (Fig. 22A, B). I conducted a GO term enrichment analysis with those transcription factors that interact with SDG7 and SDG8, which revealed their involvement in various metabolic pathways, response to phytohormone/stress, and development (Fig. 23C, D). Although common families were identified as potential interacting partner of SDG7 and SDG8 by this screen, individual members were largely specific to each histone methyltransferase (Tables 5, 6).

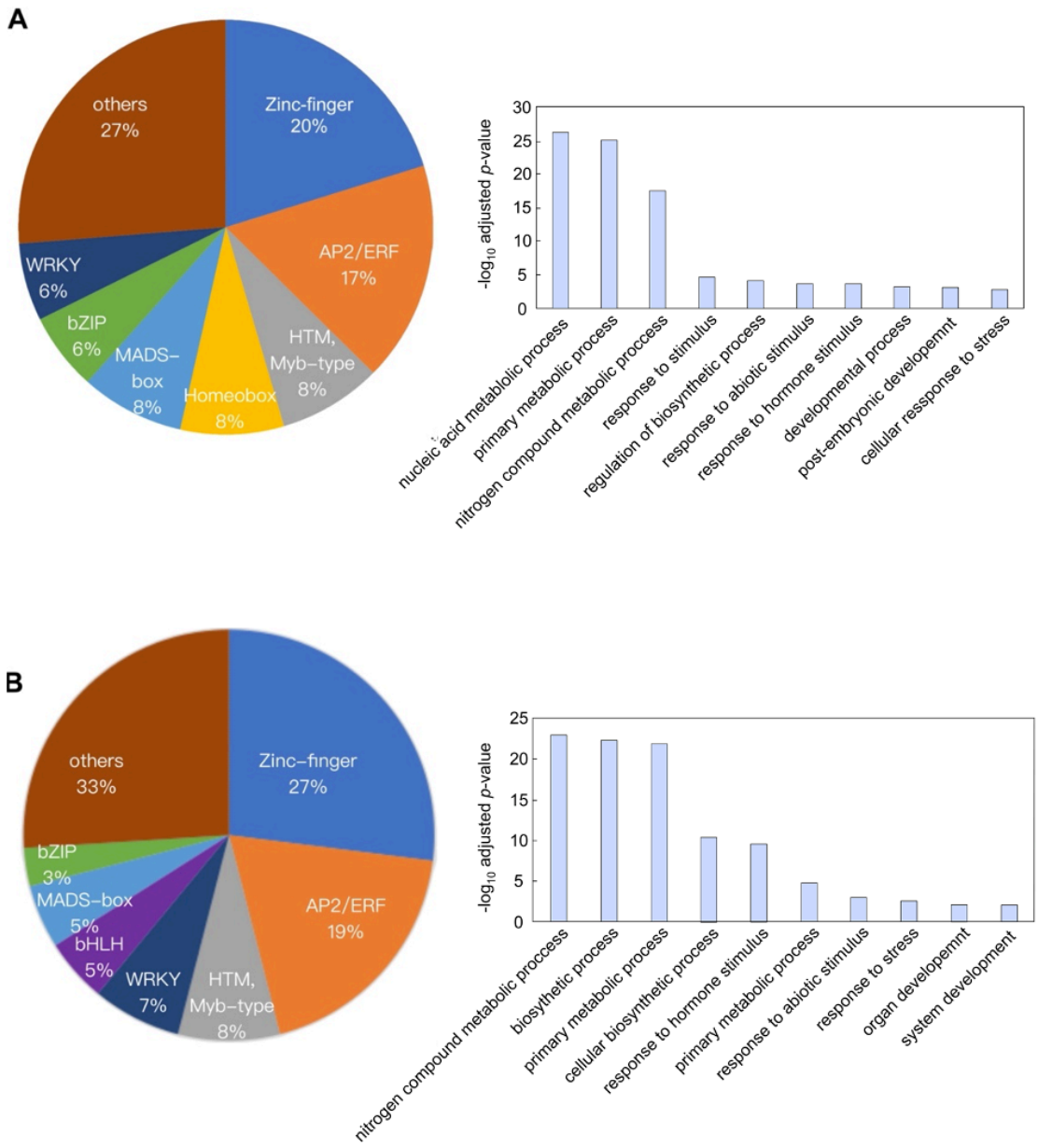

Fig. 23. Screening of potential proteins recruiting SDG7 and SDG8.

(**A, B**) Percentage of transcription factor families that interact with SDG7 (**A**) and SDG8 (**B**). (**C**, **D**) Top 10 Gene Ontology terms using 35 transcription factors that interact with SDG7 (**C**) and 40 transcription factors that interact with SDG8 (**D**).

Table 6. Identified transcription factors interacting with SDG7 by yeast two-hybrid screening.

AGI code	Gene name	DNA binding domain	Known function
At1g03970	*GBF4*	bZIP	
At3g62420	*BZIP53*		Osmotic stress response
At1g07640	*OBP2*	Zinc-finger	Wounding response
At3g02790	*MBS1*		Hyperoxide response
At3g19580	*ZF2*		Response to stress
At3g24050	*GATA1*		
At3g58040	*SINAT2*		
At5g04340	*CZF2*		Phosphate ion homeostasis
At5g49300	*GATA16*		
At1g04370	*ERF14*	AP2 /ERF	Stress response
At1g22985	*CRF7*		Cytokinin response
At1g43160	*RAP2.6*		Abscisic acid / salt / osmotic stress response
At5g11590	*TINY2*		Stress response
At5g13910	*LEP*		Gibberellin promotion of germination
At5g61590	*ERF107*		Stress response
At1g26780	*MYB117*	MYB	Meristem formation and maintenance
At1g74840			
At5g58850	*MYB119*		Gametophyte formation
At1g70510	*KNAT2*	Homeobox	
At1g70920	*HB18*		
At4g40060	*HB16*		Negative regulator of flowering
At3g61120	*AGL13*	MADS-box	Pollen and ovule of promotion and development
At3g66656	*AGL91*		
At4g11250	*AGL52*		
At2g21900	*WRKY59*	WRKY	
At4g31550	*WRKY11*		
At1g07890	*APX7*	Peroxidases heme-ligand binding site	Selenium response
At1g10200	*WLIM1*	LIM domain	Actin cytoskeleton formation
At1g14410	*WHY1*		Telomere activity suppression
At2g31380	*STH*		Increase salt tolerance
At2g37000		TCP subgroup	
At3g15270	*SPL5*	SBP-box	Conversion from vegetative to reproductive growth
At2g22670	*IAA8*		
At3g17600	*IAA31*		Regulation of, DNA-templated
At4g17600	*LIL3.1*		Synthesis of α-tocopherol and phytylated chlorophyll

Table 7. Identified transcription factors interacting with SDG8 by yeast two-hybrid screening.

AGI code	Gene name	DNA binding domain	Known function
At5g28770	*ATBZIP63*	bZIP	Control of glucose and abscisic acid pathways
At1g01250			
At1g13260	*EDF4*		Leaf aging acceleration
At1g16060	*ADAP*		Abscisic acid response
At1g21910	*DREB26*		Dehydration / high temperature / drying response
At1g22985	*CRF7*	AP2 /ERF	Cytokinin response
At1g43160	*RAP2.6*		Abscisic acid / salt / osmotic stress response
At1g50640	*ERF3*		
At2g33710			
At3g23230	*ERF98*		Salt stress response
At5g60120	*TOE2*		Photoperiodic sensing
At5g61590	*ERF107*		Stress response
At1g74840			
At2g47190	*MYB2*		Salt / dehydration response
At3g12730		MYB	
At3g49850	*TRB3*		
At3g50060	*MYB77*		Promotion of lateral root formation
At5g40360	*MYB115*		Suppression of glucosinolate synthesis
At3g61230	*PLIM2C*		Actin cytoskeleton establishment
At4g21030	*AtDOF4.2*	Zinc finger	
At4g26030			
At5g04340	*CZF2*		Phytokeratin biosynthesis
At1g69310	*WRKY57*		Suppression of leaf aging by jasmonic acid
At2g46130	*WRKY43*	WRKY	Germination / growth after germination
At4g23550	*WRKY29*		Congenital immune response
At1g03040	*PFB2*	bHLH	
At5g62610			
At1g65300	*AGL38*	MADS-box	
At1g77950	*AGL67*		
At1g66420			
At1g68800	*TCP12*	CYC/TB, R domain	Suppression of axillary bud elongation
At4g00220	*LBD30*	LOB	Formation of tubular elements
At4g16780	*HB-2*	HD-ZIP	Inhibition of cell elongation and proliferation
At4g26170			
At4g36960		RRM/RBD/RNP motifs	
At5g05500	*MOP10*		Root hair growth
At5g08190	*NF-YB12*		
At5g14000	*NAC084*		
At5g43250	*NF-YC13*		
At5g60100	*PRR3*	CCT domain	Circadian clock progress

Protein–protein interaction between SDG7/SDG8 and bZIP transcription factors

cis-regulatory motif analyses using SDG-regulated genes by H3K36me2/3 identified G-box motifs, while yeast two-hybrid screening isolated a number of bZIP transcription factors, which could bind to the motifs, as interactors for SDG7 and SDG8. To further understand the relationship between SDG7/SDG8 and bZIPs, I systematically examined the interaction potential between SDG7/SDG8 and members of the bZIP family.

The Arabidopsis bZIP family consists of 78 members that can be classified into 13 groups. Members of the same group share similar basic regions, additional conserved motifs and common features (Jakoby et al., 2002). Based on my *cis*-regulatory motif analysis, yeast two-hybrid results and previous publications, I selected the S1, C, G and A groups for the bZIP family for the following reasons. One of the S1 family members, bZIP53, was identified by both independent analyses (Fig. 22, Table 6). Heterodimerization of S1 and C groups is key for plant growth and environmental responses (Weltmeier et al., 2006). I also identified the group C member, bZIP63, as a potential interacting partner for SDG7 and SDG8 (Table 6). Yeast two-hybrid screening also identified the group A member, bZIP40 (also named G-BOX BINDING FACTOR 4 [GBF4]), as a potential interacting partner. Furthermore, the G group member bZIP41 (also named GBF1) positively regulates lateral root formation (Mallappa et al., 2006), which is impaired in the *sdg7 sdg8* double mutant. Since G group members recognize G-boxes, which were over-represented in my *cis* motif analysis, I also included this family for further characterization (Schindler et al., 1992).

Empty vectors and bait or pray vectors alone did not rescue growth on selective medium under the growth conditions used here, which constituted

an important negative control (Fig. 24A). I observed protein–protein interaction between SDG7 and several bZIP members from all four groups tested, with some preference. SDG7 interacted with the S1 group (bZIP1, bZIP11, and bZIP44), the C group (bZIP9 and bZIP25), the A group (bZIP13, bZIP14, bZIP27, and bZIP40), and the G group (bZIP41, bZIP54, and bZIP55) (Fig. 24B). Among then, S1 group protein bZIP11 has been established that it controls auxin-dependent primary root growth in maize (Weiste et al., 2014) which is similar to *sdg7 sdg8* double mutant. However, SDG7 did not interact with bZIP53 (S1), bZIP2 (S1), bZIP10 (C), or bZIP63 (C) in yeast cells (Fig. 24B). SDG8 interacted with only one S1 group member, bZIP53 (Fig. 24C) in the assay with high 3-amino-1,2,4-triazole (3-AT) which can reduce false positive signals. These results suggest that different bZIP transcription factors may regulate the recruitment of SDG7 and SDG8.

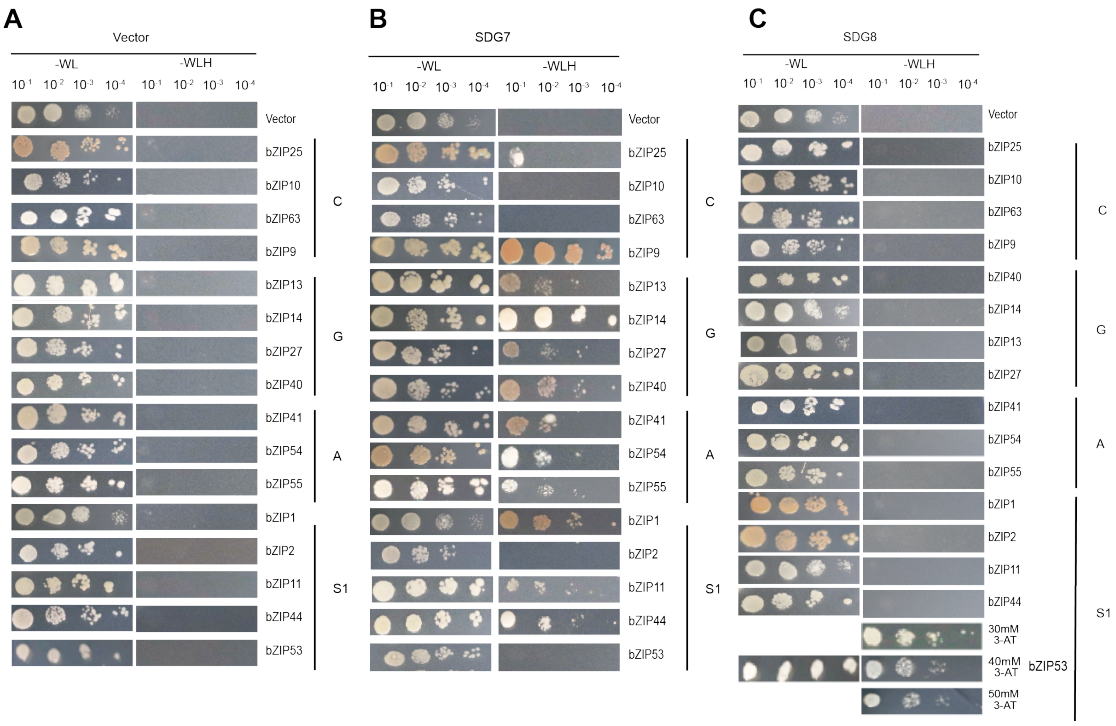

Fig. 24. Yeast two-hybrid assays between SDG7, SDG8, and bZIP members.

(**A**) bZIP members do not show autoactivation. (**B**) Interaction between SDG7 and bZIP members. (**C**) Interaction between SDG8 and bZIP members. Yeast colonies were spotted onto medium lacking Leu and Trp (–WL) or Leu, Trp, and His (–WLH).

Root phenotype in the *maf3*, *dcl1, bzip53, bzip11* single and *bzip53 bzip11* double mutants

To explore the function of MAF3, DCL1, bZIP11 and bZIP53, in root development. I measured root length, lateral root number and lateral root density using 5-DAG *maf3* (SALK_051357), *dcl1* (CS84714), *bzip53* (SALK_060119), *bzip11* (SALK_123068) single mutant, and *bzip53 bzip11* double mutant (Fig. 25). In both *bzip53*, *dcl1*, *bzip11,* and *bzip53 bzip11* mutant, primary root length of 5-DAG seedlings reached approximately 3 cm. They were similar to that of the wild type ($p > 0.05$). On the other hand, primary root length in the *maf3* single mutant was significantly shorter ($p = 1.05 \times 10^{-3}$ relative to WT) than that of the wild type with a length of 5.3 ± 0.9 cm at 5 DAG. (Fig. 25A). These results suggest that MAF3 may contribute to primary root elongation.

As to number of lateral roots, wild type at 5 DAG formed about 7 lateral roots including primordia from the primary root. Both *maf3*, *dcl1* single mutants and *bzip11 bzip53* double mutant had fewer lateral roots compared to wild type ($p = 1.01 \times 10^{-3}$). *maf3* and *dcl1* single mutants at 5 DAG only produced 5 lateral roots ($p = 1.01 \times 10^{-3}$) (Fig. 25B). Although the *bzip53* and *bzip11* single mutants produced about 6 lateral roots similar to that of wild type ($p > 0.05$), *bzip53 bzip11* double mutant had less lateral root than either single mutant. The density of lateral roots in 5-DAG wild type seedlings was 2.2 lateral roots/cm. I found significant decreases in lateral root density in the *maf3* as 1.7 lateral roots/cm, *dcl1* as 1.5 lateral roots/cm and *bzip53 bzip11* as 1.6 lateral root /cm (*maf3*, $p = 2.38 \times 10^{-2}$; *dcl1*, $p = 1.30 \times 10^{-2}$, *bzip53 bzip11*, $p = 6.17 \times 10^{-2}$, respectively) (Fig. 25C). Hence, the reduction in lateral root number seen in the *maf3, dcl1* and *bzip53 bzip11* double mutant was not merely a reflection of shorter primary roots Hence, MAF3 and DCL1 are

involved in lateral root formation. Furthermore, bZIP11 and bZIP53 may function redundantly during lateral root formation.

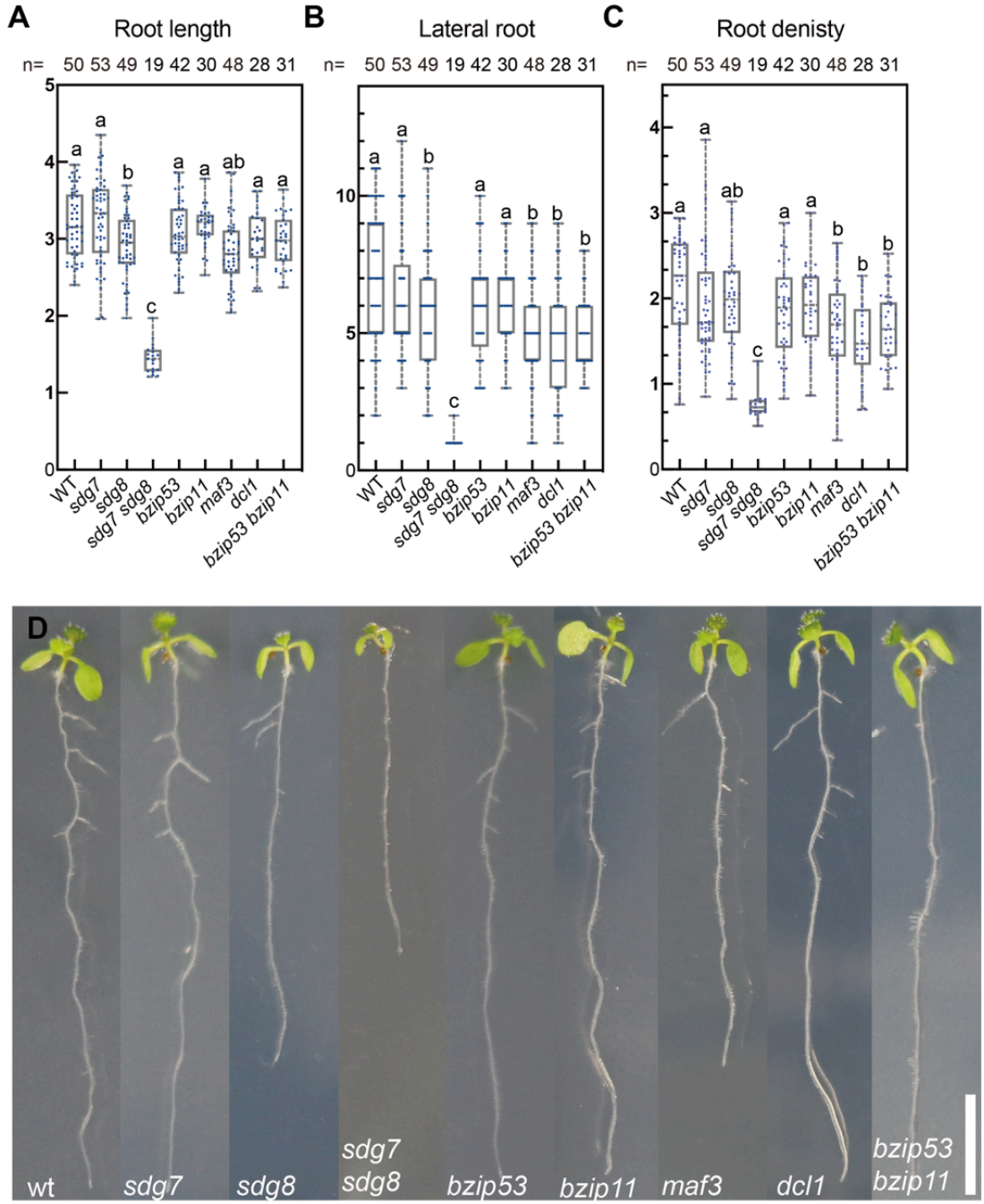

Fig. 25. Root phenotypes of the *maf3*, *dcl1*, *bzip11*, and *bzip53* single mutants, and *bzip11 bzip53* double mutants.

(**A-C**) Quantification of root phenotypes. Primary root length (**A**), lateral root number (**B**), and lateral root density (**C**) in wild type, *sdg7*, *sdg8*, *sdg7 sdg8*, *bzip53*, *bzip11*, *maf3*, *dcl1* and *bzip11 bzip53* 5 days after germination. n > 19. Significant differences were calculated based on one-way ANOVA tests ($p < 0.05$). Different letters indicate significant differences based on post-hoc Tukey's HSD test. (**D**) Root phenotypes of

the wild type, *sdg7*, *sdg8*, *sdg7 sdg8*, *bzip53*, *bzip11*, *maf3*, *dcl1* and *bzip11 bzip53* Scale bars = 0.5 cm.

IV. Discussion

In this study, I established that the Arabidopsis histone methylases SDG7 and SDG8 control growth and development through epigenetic regulation of gene expression in a partially redundant manner (Fig 26). Five pieces of evidence support this hypothesis.

1) The two proteins showed overlapping expression patterns throughout plant development (Figs. 5, 6).

2) The *sdg7 sdg8* double mutant displayed more severe developmental defects than either single mutant (Figs. 7–14).

3) H3K36me2 levels were dramatically lower in the *sdg7 sdg8* double mutant (Figs. 15–18, 20, 21).

4) H3K36me3 levels in the gene body were lower in the *sdg7 sdg8* double mutant relative to the wild type (Figs. 15–18, 20, 21).

5) Lower levels of H3K36me2 and H3K36me3 marks and repression of gene expression in the *sdg7 sdg8* double mutant were correlated (Figs. 15–21). Furthermore, motif analysis and yeast two-hybrid assays suggested that SDG7 and SDG8 recruitment to their target loci may be regulated by bZIP transcription factors (Figs. 22–24).

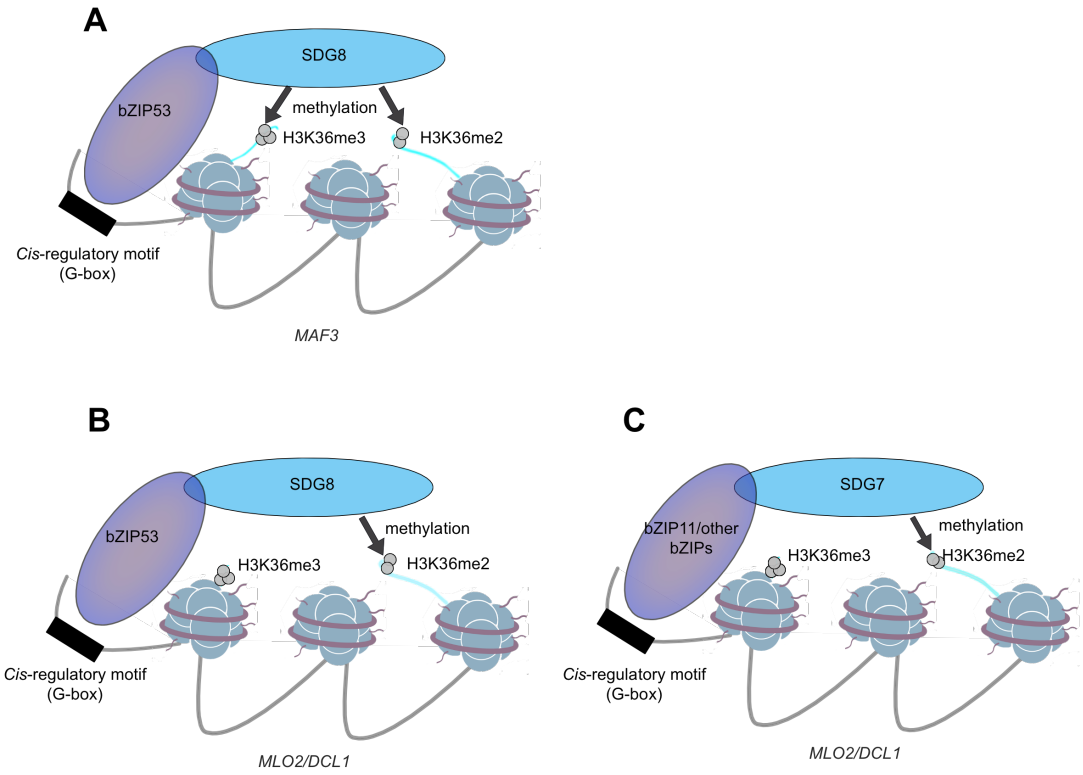

Fig 26. Potential mechanism of SDG8 and SDG7 in H3K36 methylation.

Two regulation patterns of SDG7 SDG8 in H3K36 methylation. (**A**) Recruiter bZIP53 bind with SDG8 to methylate H3K36me3 and H3K36me2 in 5' end and 3' end of downstream genes as *MAF3*, respectively. (**B, C**) Partial functional redundancy exist between SDG7 and SDG8 in methylating downstream genes as *MLO2* and *DCL1*. bZIP53 works as SDG8 recruiter and bZIP11 with other bZIPs works as SDG7 recruiters in regulating H3K36me2 in 3' end of encoding region, respectively. Black boxes, *cis*-regulatory motifs for histone-binding nucleosome remodeling factor. Gray line, DNA. Dark blue ellipses, histones. Purple lines, histone tails.

Redundant roles for SDG7 and SDG8 during growth and development

Consistent with the pleiotropic defects observed in the *sdg7 sgd8* double mutant, both *SDG7* and *SDG8* were expressed throughout plant growth and development. In agreement with the strong expression levels seen for *SDG7* and *SDG8* during early stages of plant development, phenotypic changes were also very severe, especially at early stages of plant development. Even though neither *SDG7* nor *SDG8* was expressed in sepals or petals of stage 15 flowers, the *sdg7 sdg8* double mutant exhibited clear phenotypes in sepals and petals of stage 13 flowers. I observed the same phenomena during leaf and stem development: although *SDG7* and *SDG8* expression was not detected in developed leaves or stems, I recorded clear cell expansion and elongation defects in the *sdg7 sdg8* double mutant. Although I cannot exclude the possibility of secondary phenotypic effects, this discrepancy between endogenous gene expression levels and phenotypic changes in the mutants may be explained by epigenetic regulation. Also, GUS/GFP reporter genes in double mutant could also reflect the endogenous expression patterns. One way to explain the abnormal reproductive growth is that aberrant tissues occurred by SDG7 and SDG8 in early vegetative growth was the main reason. Also, it can't be rescued even though SDG7 and SDG8 were absent during later stages of development. Differential histone modification patterns generated at early stages of plant development in the *sdg7 sdg8* mutants would be inherited after cell division, which might lead to phenotypic changes at later stages of plant growth and development.

Redundant roles for SDG7 and SDG8 in H3K36me2 and H3K36me3 deposition

Genome-wide identification of H3K36me2 and H3K36me3 marks in wild type and *sdg7 sdg8* revealed the role of SDG7 and SDG8 in their deposition. H3K36me2 was almost absent from the 3' end of 13,005 differentially H3K36me2-methylated genes in the *sdg7 sdg8* double mutant. Furthermore, H3K36me2 levels at selected genes in the *sdg7 sdg8* double mutant were lower than in either single mutant. These results suggest that SDG7 and SDG8 play a major role in the deposition of H3K36me2 in Arabidopsis. Most of the 8,125 differentially H3K36me3-methylated genes maintained their H3K36me3 peaks at 5' end in both the wild type and the *sdg7 sdg8* double mutant, suggesting that additional histone methylases contribute to H3K36me3 deposition near the transcription start site in the double mutant background. Group II-1 of histone methylases consists of five members, including SDG7, SDG8, and SDG26. Although SDG26 and SDG8 have distinct functions during development, functional interaction between SDG26, SDG7/SDG8, and the other two members SDG4 and SDG24 has not been assessed (Liu et al., 2016). SDG4 plays an important role for the deposition of both H3K36me2 and H3K36me3 (Berr et al., 2010; Cartagena et al., 2008; Cazzonelli et al., 2010; Dong et al., 2008; Grini et al., 2009; Kumpf et al., 2014; Lee et al., 2015; Li et al., 2015; Wang et al., 2014; Xu et al., 2008; Xu et al., 2014; Zhao et al., 2005). Thus, SDG4 might also contribute to the deposition in the absence of SDG7 and SDG8. Hence, functional characterization of triple mutants would be interesting. Furthermore, the precise function of SDG24 needs to be determined in the future to understand how H3K36 methylation is regulated.

Recruitment of SDG7 and SDG8 via bZIP transcription factors

My *cis* motif analysis and yeast two-hybrid screen suggested that bZIP transcription factors may recruit SDG7 and SDG8 to target promoters.

These findings are consistent with previous results that the bZIP binding motif G-box (CACGTG) is significantly over-represented in the promoter of SDG8-bound genes (Li et al., 2015). Despite the redundant roles of SDG7 and SDG8 during growth and development as well as in the regulation of target gene expression, SDG7 and SDG8 interacted with different bZIP partners in my experiments. Although *in vitro* pull-down assays and *in vivo* protein–protein interaction experiments are absolutely required for validation, my results suggest that SDG7 and SDG8 may be recruited by distinct bZIP transcription factors for the regulation of common target genes. I also identified transcription factors other than bZIPs. Thus, shared, unique, and as yet unknown recruiter(s) for SDG7 and SDG8 might also exist.

To confirm the role of bZIPs for SDG recruitment, I am planning to conduct several experiments. First, *in vivo* protein–protein interaction experiments are needed to confirm the interactions detected in yeast. I prepared constructs to conduct bimolecular fluorescence complementation assays and generated transgenic plants accumulating each protein with a tag for co-immunoprecipitation (Co-IP) *in planta*. Another key experiment will be to conduct ChIP-seq for SDG7, SDG8, and bZIPs to identify shared direct targets in a genome-wide manner, as all proteins should bind to shared target loci if they form a single complex. ChIP for SDG7 and SDG8 in a *bzip* mutant background will further test bZIP-dependent SDG recruitment. Finally, phenotypic analyses using *bzip* mutants are also needed. Currently, the role of bZIPs during plant growth and development is poorly understood, even though single mutant of *bzip53* did not observe significant root defect (Fig. 25), I observed bZIP53 accumulated in root apical meristem and elongation zone. perhaps due to functional redundancy between related family members. Even phenotype in *bzip53 bzip11* double mutant is much weaker than that of *sdg7 sdg8* double mutants (Fig. 25).

Therefore, higher-order mutants of SDG recruiters may result in severe growth and developmental defects, as those seen in *sdg7 sdg8* double mutants. For example, *bzip41* has defects in the formation of lateral roots (Mallappa et al., 2006). These analyses will illuminate the relationship between SDG7, SDG8, and bZIP transcription factors.

V. Acknowledgements

I would like to thank Dr. Toshiro Ito and Dr. Nobutoshi Yamaguchi for supervision. I also would like to thank all lab members in the Plant Stem Cell Regulation and Floral Patterning Lab for their support and encouragement throughout my doctoral studies. I would also like to thank my advisors, Dr. Masaaki Umeda and Dr. Yusuke Saijo, for their valuable feedback and comments. Finally, I want to thank my parents and family for their continuous support.

VII. References

Acevedo-Garcia, J., Gruner, K., Reinstädler, A., Kemen, A., Kemen, E., Cao, L., Takken, F. L., Reitz, M. U., Schäfer, P., & O'Connell, R. J. (2017). The powdery mildew-resistant Arabidopsis mlo2 mlo6 mlo12 triple mutant displays altered infection phenotypes with diverse types of phytopathogens. *Scientific reports*, *7*(1), 1-15.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., & Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, *129*(4), 823-837.

Berr, A., McCallum, E. J., Alioua, A., Heintz, D., Heitz, T., & Shen, W.-H. (2010). Arabidopsis histone methyltransferase SET DOMAIN GROUP8 mediates induction of the jasmonate/ethylene pathway genes in plant defense response to necrotrophic fungi. *Plant physiology*, *154*(3), 1403-1414.

Berr, A., Shafiq, S., & Shen, W.-H. (2011). Histone modifications in transcriptional activation during plant development. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1809*(10), 567-576.

Bureau, M., Rast, M. I., Illmer, J., & Simon, R. (2010). JAGGED LATERAL ORGAN (JLO) controls auxin dependent patterning

during development of the Arabidopsis embryo and root. *Plant molecular biology*, *74*(4-5), 479-491.

Cartagena, J. A., Matsunaga, S., Seki, M., Kurihara, D., Yokoyama, M., Shinozaki, K., Fujimoto, S., Azumi, Y., Uchiyama, S., & Fukui, K. (2008). The Arabidopsis SDG4 contributes to the regulation of pollen tube growth by methylation of histone H3 lysines 4 and 36 in mature pollen. *Developmental biology*, *315*(2), 355-368.

Cazzonelli, C. I., Nisar, N., Roberts, A. C., Murray, K. D., Borevitz, J. O., & Pogson, B. J. (2014). A chromatin modifying enzyme, SDG8, is involved in morphological, gene expression, and epigenetic responses to mechanical stimulation. *Frontiers in plant science*, *5*, 533.

Cazzonelli, C. I., Roberts, A. C., Carmody, M. E., & Pogson, B. J. (2010). Transcriptional control of SET DOMAIN GROUP 8 and CAROTENOID ISOMERASE during Arabidopsis development. *Molecular Plant*, *3*(1), 174-191.

Cheung, P., Allis, C. D., & Sassone-Corsi, P. (2000). Signaling to chromatin through histone modifications. *Cell*, *103*(2), 263-271.

Davie, J. R. (1996). Histone modifications, chromatin structure, and the nuclear matrix. *Journal of cellular biochemistry*, *62*(2), 149-157.

Dhami, N., & Cazzonelli, C. I. (2020). Prolonged cold exposure to Arabidopsis juvenile seedlings extends vegetative growth and

increases the number of shoot branches. *Plant Signaling & Behavior*, *15*(9), 1789320.

Dinant, C., Houtsmuller, A. B., & Vermeulen, W. (2008). Chromatin structure and DNA damage repair. *Epigenetics & chromatin*, *1*(1), 1-13.

Dong, G., Ma, D.-P., & Li, J. (2008). The histone methyltransferase SDG8 regulates shoot branching in Arabidopsis. *Biochemical and biophysical research communications*, *373*(4), 659-664.

Galinha, C., Hofhuis, H., Luijten, M., Willemsen, V., Blilou, I., Heidstra, R., & Scheres, B. (2007). PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. *Nature*, *449*(7165), 1053-1057. https://doi.org/10.1038/nature06206

Gallagher, K. L., Paquette, A. J., Nakajima, K., & Benfey, P. N. (2004). Mechanisms regulating SHORT-ROOT intercellular movement. *Current Biology*, *14*(20), 1847-1851.

Gan, E.-S., Huang, J., & Ito, T. (2013). Functional roles of histone modification, chromatin remodeling and microRNAs in Arabidopsis flower development. *International review of cell and molecular biology*, *305*, 115-161.

Gasciolli, V., Mallory, A. C., Bartel, D. P., & Vaucheret, H. (2005). Partially redundant functions of Arabidopsis DICER-like enzymes

and a role for DCL4 in producing trans-acting siRNAs. *Current Biology*, *15*(16), 1494-1500.

Greer, E. L., & Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. *Nature Reviews Genetics*, *13*(5), 343-357.

Grini, P. E., Thorstensen, T., Alm, V., Vizcay-Barrena, G., Windju, S. S., Jørstad, T. S., Wilson, Z. A., & Aalen, R. B. (2009). The ASH1 HOMOLOG 2 (ASHH2) histone H3 methyltransferase is required for ovule and anther development in Arabidopsis. *PloS one*, *4*(11), e7817.

Gu, X., Le, C., Wang, Y., Li, Z., Jiang, D., Wang, Y., & He, Y. (2013). Arabidopsis FLC clade members form flowering-repressor complexes coordinating responses to endogenous and environmental cues. *Nature Communications*, *4*(1), 1-10.

He, C., Chen, X., Huang, H., & Xu, L. (2012). Reprogramming of H3K27me3 is critical for acquisition of pluripotency from cultured Arabidopsis tissues.

Hennig, L., & Derkacheva, M. (2009). Diversity of Polycomb group complexes in plants: same rules, different players? *Trends Genet*, *25*(9), 414-423. https://doi.org/10.1016/j.tig.2009.07.002

Hunter, C. A., Aukerman, M. J., Sun, H., Fokina, M., & Poethig, R. S. (2003). PAUSED encodes the Arabidopsis exportin-t ortholog. *Plant physiology*, *132*(4), 2135-2143.

Inagaki, S., Takahashi, M., Hosaka, A., Ito, T., Toyoda, A., Fujiyama, A., Tarutani, Y., & Kakutani, T. (2017). Gene-body chromatin modification dynamics mediate epigenome differentiation in Arabidopsis. *The EMBO journal*, *36*(8), 970-980.

Iwata, E., Ikeda, S., Matsunaga, S., Kurata, M., Yoshioka, Y., Criqui, M.-C., Genschik, P., & Ito, M. (2011). GIGAS CELL1, a novel negative regulator of the anaphase-promoting complex/cyclosome, is required for proper mitotic progression and cell fate determination in Arabidopsis. *The Plant Cell*, *23*(12), 4382-4393.

Jakoby, M., Weisshaar, B., Dröge-Laser, W., Vicente-Carbajosa, J., Tiedemann, J., Kroj, T., & Parcy, F. (2002). bZIP transcription factors in Arabidopsis. *Trends in plant science*, *7*(3), 106-111.

Jenuwein, T., & Allis, C. D. (2001). Translating the histone code. *Science*, *293*(5532), 1074-1080.

Jeong, Y. J., Shang, Y., Kim, B. H., Kim, S. Y., Song, J. H., Lee, J. S., Lee, M. M., Li, J., & Nam, K. H. (2010). BAK7 displays unequal genetic redundancy with BAK1 in brassinosteroid signaling and early senescence in Arabidopsis. *Molecules and cells*, *29*(3), 259-266.

Johnson, L. M., Du, J., Hale, C. J., Bischof, S., Feng, S., Chodavarapu, R. K., Zhong, X., Marson, G., Pellegrini, M., & Segal, D. J. (2014). SRA-and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature*, *507*(7490), 124-128.

Kim, S. Y., He, Y., Jacob, Y., Noh, Y.-S., Michaels, S., & Amasino, R. (2005). Establishment of the vernalization-responsive, winter-annual habit in Arabidopsis requires a putative histone H3 methyl transferase. *The Plant Cell*, *17*(12), 3301-3310.

Kumpf, R., Thorstensen, T., Rahman, M. A., Heyman, J., Nenseth, H. Z., Lammens, T., Herrmann, U., Swarup, R., Veiseth, S. V., & Emberland, G. (2014). The ASH1-RELATED3 SET-domain protein controls cell division competence of the meristem and the quiescent center of the Arabidopsis primary root. *Plant physiology*, *166*(2), 632-643.

Lee, J., Yun, J.-Y., Zhao, W., Shen, W.-H., & Amasino, R. M. (2015). A methyltransferase required for proper timing of the vernalization response in Arabidopsis. *Proceedings of the National Academy of Sciences*, *112*(7), 2269-2274.

Lee, S., Fu, F., Xu, S., Lee, S. Y., Yun, D.-J., & Mengiste, T. (2016). Global regulation of plant immunity by histone lysine methyl transferases. *The Plant Cell*, *28*(7), 1640-1661.

Li, Y., Mukherjee, I., Thum, K. E., Tanurdzic, M., Katari, M. S., Obertello, M., Edwards, M. B., McCombie, W. R., Martienssen, R. A., & Coruzzi, G. M. (2015). The histone methyltransferase SDG8 mediates the epigenetic modification of light and carbon responsive genes in plants. *Genome biology*, *16*(1), 1-15.

Liu, B., Berr, A., Chang, C., Liu, C., Shen, W.-H., & Ruan, Y. (2016). Interplay of the histone methyltransferases SDG8 and SDG26 in the regulation of transcription and plant flowering and development. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1859*(4), 581-590.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., & Köhler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO reports*, *7*(9), 947-952.

Malamy, J. E., & Benfey, P. N. (1997). Organization and cell differentiation in lateral roots of Arabidopsis thaliana. *Development*, *124*(1), 33-44.

Mallappa, C., Yadav, V., Negi, P., & Chattopadhyay, S. (2006). A basic leucine zipper transcription factor, G-box-binding factor 1, regulates blue light-mediated photomorphogenic growth in Arabidopsis. *Journal of Biological Chemistry*, *281*(31), 22190-22199.

Mitsuda, N., Ikeda, M., Takada, S., Takiguchi, Y., Kondou, Y., Yoshizumi, T., Fujita, M., Shinozaki, K., Matsui, M., & Ohme-Takagi, M. (2010). Efficient yeast one-/two-hybrid screening using a library composed only of transcription factors in Arabidopsis thaliana. *Plant and Cell physiology*, *51*(12), 2145-2151.

O'Malley, R. C., Huang, S.-s. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., Galli, M., Gallavotti, A., & Ecker, J. R. (2016). Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell*, *165*(5), 1280-1292.

Okushima, Y., Fukaki, H., Onoda, M., Theologis, A., & Tasaka, M. (2007). ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in Arabidopsis. *The Plant Cell*, *19*(1), 118-130.

Pi, L., Aichinger, E., van der Graaff, E., Llavata-Peris, C. I., Weijers, D., Hennig, L., Groot, E., & Laux, T. (2015). Organizer-Derived WOX5 Signal Maintains Root Columella Stem Cells through Chromatin-Mediated Repression of CDF4 Expression. *Dev Cell*, *33*(5), 576-588. https://doi.org/10.1016/j.devcel.2015.04.024

Qian, C., & Zhou, M.-M. (2006). SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cellular and molecular life sciences CMLS*, *63*(23), 2755-2763.

Ratcliffe, O. J., Nadzan, G. C., Reuber, T. L., & Riechmann, J. L. (2001). Regulation of flowering in Arabidopsis by an FLC homologue. *Plant physiology*, *126*(1), 122-132.

Schindler, U., Menkens, A. E., Beckmann, H., Ecker, J. R., & Cashmore, A. R. (1992). Heterodimerization between light-regulated and ubiquitously expressed Arabidopsis GBF bZIP proteins. *The EMBO journal*, *11*(4), 1261-1273.

Smolikova, G., Strygina, K., Krylova, E., Leonova, T., Frolov, A., Khlestkina, E., & Medvedev, S. (2021). Transition from Seeds to Seedlings: Hormonal and Epigenetic Aspects. *Plants (Basel)*, *10*(9). https://doi.org/10.3390/plants10091884

Soppe, W., Bentsink, L., & Koornneef, M. (1999). The early-flowering mutant efs is involved in the autonomous promotion pathway of Arabidopsis thaliana. *Development*, *126*(21), 4763-4770.

Springer, N. M., Napoli, C. A., Selinger, D. A., Pandey, R., Cone, K. C., Chandler, V. L., Kaeppler, H. F., & Kaeppler, S. M. (2003). Comparative analysis of SET domain proteins in maize and Arabidopsis reveals multiple duplications preceding the divergence of monocots and dicots. *Plant physiology*, *132*(2), 907-925.

Stiti, N., Missihoun, T. D., Kotchoni, S., Kirch, H.-H., & Bartels, D. (2011). Aldehyde dehydrogenases in Arabidopsis thaliana:

biochemical requirements, metabolic pathways, and functional analysis. *Frontiers in plant science*, *2*, 65.

Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, *403*(6765), 41-45.

Telfer, A., Bollman, K. M., & Poethig, R. S. (1997). Phase change and the regulation of trichome distribution in Arabidopsis thaliana. *Development*, *124*(3), 645-654.

Townsley, B. T., Covington, M. F., Ichihashi, Y., Zumstein, K., & Sinha, N. R. (2015). BrAD-seq: Breath Adapter Directional sequencing: a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction [Methods]. *Frontiers in plant science*, *6*(366). https://doi.org/10.3389/fpls.2015.00366

Ulmasov, T., Murfett, J., Hagen, G., & Guilfoyle, T. J. (1997). Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *The Plant Cell*, *9*(11), 1963-1971.

Wang, D., Tyson, M. D., Jackson, S. S., & Yadegari, R. (2006). Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in <em>Arabidopsis</em&gt. *Proceedings of the National Academy of Sciences*, *103*(35), 13244. https://doi.org/10.1073/pnas.0605551103

Wang, X., Chen, J., Xie, Z., Liu, S., Nolan, T., Ye, H., Zhang, M., Guo, H., Schnable, P. S., & Li, Z. (2014). Histone lysine methyltransferase SDG8 is involved in brassinosteroid-regulated gene expression in Arabidopsis thaliana. *Molecular Plant*, *7*(8), 1303-1315.

Weiste C, Dröge-Laser W. The Arabidopsis transcription factor bZIP11 activates auxin-mediated transcription by recruiting the histone acetylation machinery[J]. Nature communications, 2014, 5(1): 1-12.

Weltmeier, F., Ehlert, A., Mayer, C. S., Dietrich, K., Wang, X., Schütze, K., Alonso, R., Harter, K., Vicente-Carbajosa, J., & Dröge-Laser, W. (2006). Combinatorial control of Arabidopsis proline dehydrogenase transcription by specific heterodimerisation of bZIP transcription factors. *The EMBO journal*, *25*(13), 3133-3143.

Xiao, J., Lee, U.-S., & Wagner, D. (2016). Tug of war: adding and removing histone lysine methylation in Arabidopsis. *Current opinion in plant biology*, *34*, 41-53.

Xu, L., Zhao, Z., Dong, A., Soubigou-Taconnat, L., Renou, J.-P., Steinmetz, A., & Shen, W.-H. (2008). Di-and tri-but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and

other processes in Arabidopsis thaliana. *Molecular and cellular biology*, *28*(4), 1348-1360.

Xu, Y., Gan, E.-S., Zhou, J., Wee, W.-Y., Zhang, X., & Ito, T. (2014). Arabidopsis MRG domain proteins bridge two histone modifications to elevate expression of flowering genes. *Nucleic acids research*, *42*(17), 10960-10974.

Yang, H., Howard, M., & Dean, C. (2016). Physical coupling of activation and derepression activities to maintain an active transcriptional state at FLC. *Proceedings of the National Academy of Sciences*, *113*(33), 9369-9374.

Yang, S., & Hua, J. (2004). A haplotype-specific Resistance gene regulated by BONZAI1 mediates temperature-dependent growth control in Arabidopsis. *The Plant Cell*, *16*(4), 1060-1071.

Zhao, Z., Yu, Y., Meyer, D., Wu, C., & Shen, W.-H. (2005). Prevention of early flowering by expression of FLOWERING LOCUS C requires methylation of histone H3 K36. *Nature cell biology*, *7*(12), 1256-1260.