

**Functional analysis of an AP2/ERF transcription factor  
STEMIN1 that induces reprogramming of gametophore  
leaf cells to chloronema apical stem cells  
in the moss *Physcomitrella patens***

Morishita, Mio

PhD Dissertation

2019

## TABLE OF CONTENTS

1. General introduction.....	4
1.1. <i>de novo</i> stem cell formation during post-embryonic development in land plants.....	4
1.2. Stem cell formation during regeneration in land plants.....	5
1.3. Changes in histone modifications during stem cell formation.....	7
2. Introduction.....	9
3. Materials and Methods.....	13
3.1. Plant material.....	13
3.2. Screening of factors to induce reprogramming of leaf cells.....	13
3.3. Construction of $\beta$ -estradiol-inducible <i>STEMIN</i> plasmids.....	14
3.4. Plasmid construction for spatial expression analysis.....	15
3.5. Plasmid construction for the deletion of <i>STEMIN</i> genes and generation of the <i>STEMIN</i> -deletion mutants.....	15
3.6. Phylogenic analysis.....	16
3.7. RNA preparation and qRT-PCR analysis.....	18
3.8. GUS staining.....	19
3.9. Microscopy analysis.....	19
3.10. Scanning electron microscopy.....	19
3.11. Measurement of frequency of side-branching in chloronemata.....	20
3.12. DNA gel-blot and immunoblot analyses.....	20
3.13. Transcriptome analysis.....	21
3.14. ChIP-sequencing (ChIP-seq) .....	21
3.15. ChIP-seq data analysis.....	23
3.16. Accession numbers.....	24
4. Results.....	27
4.1. <i>STEMIN1</i> changes leaf cells into stem cells without cutting.....	27
4.2. Function of <i>STEMIN</i> genes in stem cell formation in cut leaves.....	27
4.3. The function of <i>STEMIN1</i> and its homologs in regular development.....	30
4.4. Identification of <i>STEMIN1</i> direct target genes.....	32
4.5. <i>STEMIN1</i> regulates the histone H3K27me3 level of genes involved in stem cell formation.....	33

5. Discussion.....	80
5.1. Cell fate changes by AP2/ERF transcription factors in land plants.....	80
5.2. Functions of STEMINs in stem cell formation.....	81
5.3. Locus-specific reprogramming of H3K27me3 by STEMIN1.....	83
6. General discussion.....	86
7. Acknowledgements.....	89
8. References.....	90

# **1. General introduction**

## **1.1 *de novo* stem cell formation during post-embryonic development in land plants**

Stem cells can self-renew and give rise to other types of cells to differentiate into new tissues or organs (Potten et al., 1990; Slack, 2008; Weigel et al., 2002; Laux, 2003; Weissman, 2000). In metazoans except for some animals such as amphibians, embryonic stem cells are lost during development. On the other hand, land plants maintain meristematic stem cells committed to differentiation into specific cell types to make new organs or tissues through their life span (Birnbaum and Sánchez Alvarado, 2008). In flowering plants, multicellular stem cells are maintained at specialized multicellular zones known as meristems.

In some mammals and most of land plants, differentiated cells can revert to stem cells under appropriate conditions (Birnbaum and Sánchez Alvarado, 2008). In mammals, while stem cells do not appear to emerge after embryogenesis, induction of four transcription factors, Oct4, Sox2, Nanog, and c-Myc (Takahashi and Yamanaka, 2006; Takahashi et al., 2007), and another set of factors, Oct4, Sox2, Nanog, and Lin28 (Yu et al., 2007), reprograms chromatin states and induces conversion of differentiated cells to induced pluripotent stem (iPS) cells. However, in large part due to cell-specific epigenetic modifications, efficiency of reprogramming remains low (Mansour et al., 2012). Besides, this direct cell reprogramming appears to be a stochastic process that depends on the amount, balance, continuity, and silencing of the expression of the

factors (Gaspar-Maia et al., 2011; Orkin and Hochedlinger, 2011; Yamanaka, 2009).

On the other hand, in land plants, *de novo* stem cell formation is widely observed naturally during post-embryonic development. Under natural conditions, shoot apical meristems repeatedly produce functional units of a plant called phytomers consisting of a leaf, an axillary shoot meristem, and an internode (Yang and Jiao, 2016). In addition, *de novo* organogenesis to form shoots or roots is often observed either from detached organs or the original plant, giving rise to new whole plant body. For example, Karanchoe has reproductive leaves that form new plantlets on the leaf margins, whose detach allows asexual reproduction. In *Arabidopsis thaliana* (Arabidopsis), pro-cambium or cambium appears to function in *de novo* root formation, which is triggered by the wounding signal (Da Rocha Correa et al., 2012).

## **1.2. Stem cell formation during regeneration in land plants**

Upon injury, metazoans and land plants can undergo proliferation and differentiation to restore the missing tissues. In metazoans, pre-existing stem cells or remaining differentiated cells in the damaged tissues can respond to injury by reentering the cell cycle and differentiating into one or more cell types to restore the missing tissues (Sugimoto et al., 2011).

Land plants have an ability to regenerate lost tissues from differentiated cells, even single isolated differentiated cells, with multicellularization process, which in many cases do not include population of stem cells (Steward et al., 1958; Takebe et al.,

1971; Atta et al., 2009; Sena et al., 2009; Sugimoto et al., 2010). For instance, a root in *Zea mays* from which the root cap and quiescent center are removed is able to regenerate a complete root apical meristem in three days (Feldman, 1976). Similarly, in *Arabidopsis*, cutting off root tip that includes stem cell niche triggers embryonic-like development program in accordance with cell proliferation, in which transit-amplifying cells can reform stem cells (Sena et al., 2009; Efroni et al., 2016).

Furthermore, in flowering plants, treatment with auxin and cytokinin stimulates the cells of excised organs to proliferate, forming a callus, a mass of growing cells that has lost the differentiated cell fate, which subsequently can be fated to form shoot or root apical meristem (Skoog and Miller, 1957; Che et al., 2007; Atta et al., 2009; Sugimoto et al., 2010). Several transcription factors involved in the reprogramming via callus have been characterized in *Arabidopsis* (Banno et al., 2001; Gordon et al., 2007; Sugimoto et al., 2010; Iwase et al., 2011). WOUNDING INDUCED DEDIFFERENTIATION1 (*WIND1*), a member of AP2/ERF transcription factor, promotes shoot regeneration through direct transcriptional activation of *ENHANCER OF SHOOT REGENERATION1/DORNRÖSCHEN* (*ESR1/DRN*), which encodes another AP2/ERF transcription factor and enhances shoot regeneration through the upregulation of key regulators for shoot apical meristems, such as *WUSCHEL* (*WUS*) encoding a homeobox transcription factor, and *CUP-SHAPED COTYLEDON* (*CUC*) encoding a NAC (NAM, ATAF, and CUC) transcription factor (Iwase et al., 2011; Iwase et al., 2015; Iwase et al., 2017). *WUS* and *CUC* function in stem cell maintenance in shoot apical meristem (Laux et al., 1996) and specification of

meristem-organ boundary zones (Aida et al., 1999).

However, since stem cells form inside meristem tissue, it remains unclear whether these factors directly promote stem cell fate from differentiated cells or induce the formation of meristem tissue. Therefore, the inherent mechanism underlying reprogramming of differentiated cells to stem cells is still unknown.

### **1.3. Changes in histone modifications during stem cell formation**

Epigenetic modifications, including histone modifications and DNA methylation, function in stabilization of such cell-specific gene expression programs to maintain cellular identities (Birnbaum and Roudier, 2017; Ikeuchi et al., 2015a; Wutz, 2013). For instance, in both metazoans and land plants, trimethylation at lysine 27 of histone H3 (H3K27me3) catalyzed by polycomb repressive complex 2 (PRC2) leads to maintenance of transcriptional repressive states and functions in stabilization of developmental programs (Okano et al., 2009; Bouyer et al., 2011; Mozgova, 2017; Ikeuchi et al., 2015b). Similarly, Trithorax group (TrxG) proteins catalyze trimethylation at lysine 4 of histone H3 (H3K4me3) leading to maintenance of transcriptional active states with its histone methyltransferases activity (Springer et al. 2003; Ng et al. 2007; Ringrose and Paro, 2004).

Notwithstanding the existence of these stable cell states, in some mammals and most land plants, differentiated somatic cells can change to stem cells under the appropriate conditions (Birnbaum and Sánchez Alvarado, 2008). Many genes encoding

regulators acting during Arabidopsis regeneration, including *ESR1*, are epigenetically silenced by polycomb-mediated histone modifications (Iwase et al., 2017). In addition, the up-regulation of *WUS* gene in a callus on shoot-inducing medium requires changes in histone modifications (Li et al., 2011).

H3K27me3 demethylases act in the opposite way to the function of PRC2. In Arabidopsis, RELATIVE OF EARLY FLOWERING 6 (*REF6*, also known as Jumonji domain-containing protein 12 [*JMJ12*]) that functions as a H3K27me3 demethylase recognizes specific DNA sequences at genes enriched with the H3K27me3 and demethylates these loci to activate their gene expression in flowering (Lu et al., 2011; Gan et al., 2014; Li et al., 2016; Cui et al., 2016). However, mutation in the *REF6/JMJ12* gene does not affect regeneration process from leaves (He et al., 2012). Therefore, it remains to be elucidated what factors function in erasing the memory to permit the induction of a new cell fate and how epigenetic modifications are locally regulated to induce specific genes necessary for cellular changes without affecting other genes in a genome in land plants (Birnbaum and Roudier, 2017). A study on changes in histone modifications at specific loci during post-embryonic development and regeneration could provide new insights into molecular mechanisms of stem cell formation in land plants.

## 2. Introduction

Cells generated from stem cells acquire a new fate with their characteristic gene expression patterns. Epigenetic modifications, including histone modifications, stabilize cell-specific gene expression programs to maintain cell identities in both metazoans and land plants. Unlike metazoans, in land plants, *de novo* stem cell formation is widely observed during post-embryonic development (Malamy and Benfey, 1997; De Smet et al., 2006; McSteen and Leyser, 2005), and during regeneration process in a callus (Skoog and Miller, 1957; Sugimoto et al., 2010; Atta et al., 2009), and in a dissected root (Efroni et al., 2016; Sena et al., 2009). On the other hand, epigenetic modifications have the potential to stabilize cell identity to maintain tissue organizations in land plants (Birnbaum and Roudier, 2017; Mozgova et al., 2017; Ikeuchi et al., 2015a; Okano et al., 2009). Hence, unlike metazoans, land plants are supposed to have an inherent mechanism underlying epigenetic reprogramming of differentiated cells into stem cells during development and regeneration.

For instance, a repressive chromatin mark H3K27me<sub>3</sub>, at specific loci contributes to determining tissue-specific gene expression (Zheng and Chen, 2011; Lafos et al., 2011) and maintaining differentiated states (Mozgova et al., 2017; Ikeuchi et al., 2015a). Leaf to callus formation in *Arabidopsis* through the genetic pathway overlapped with a program of lateral root initiation, including process of stem cell formation (Sugimoto et al., 2010), requires reprogramming of H3K27me<sub>3</sub> to suppress expression of genes involved in leaf cell identity and to activate the auxin signaling

pathway (He et al., 2012). Therefore, locus-specific reprogramming of H3K27me3 should be required for the rewriting of gene regulatory networks during plant development and regeneration, which could be one of the intrinsic mechanisms underlying stem cell formation. However, since it is difficult to apply a tractable approach to assay for changes in chromatin modification during reprogramming, such as iPS cell induction in mammals (Birnbaum and Roudier, 2017), and stem cells in angiosperms are formed inside meristem tissues, it remains to be elucidated how epigenetic modifications are locally regulated to induce specific genes necessary for cellular changes without affecting other genes in a genome (Birnbaum and Roudier, 2017).

In the moss *Physcomitrella patens* (Physcomitrella), after germination from spores, hypha-like bodies, called protonemata, and subsequently shoot-like bodies called gametophores are formed. Protonemata comprise two cell types: chloronema and caulonema cells. A single stem cell is situated at the tips of chloronema and caulonema filaments, which is named chloronema and caulonema apical stem cell, respectively (Kofuji and Hasebe, 2014). Since protonemata are filamentous tissues and gametophore leaves are formed as a single cell layer, cellular changes are easily observed. In particular, when a gametophore leaf is cut and incubated on culture medium without exogenous phytohormones, leaf cells facing the cut are reprogrammed to chloronema apical stem cells (Fig. 1) (Ishikawa et al., 2011; Ishikawa and Hasebe, 2015).

In general, major limitations to understand initial molecular reprogramming

events arise from the complexity of the model plants. In particular, differentiated cells in flowering plants produce stem cells via formation of a proliferating mass of cells or a callus (Sugimoto et al., 2011; Ikeuchi et al., 2013). On the other hand, *Physcomitrella* leaf cells are composed of a single cell layer, being able to specify the reprogramming cells and to observe their cellular events from leaf cell state to chloronema apical stem cell state under a microscope. In addition, this process does not require any exogenous phytohormones.

Using this system, some factors that function in reprogramming of leaf cells have been identified, such as CYCLIN-DEPENDENT KINASE A (CDKA) for cell cycle reactivation (Ishikawa et al., 2011), WUSCHEL-RELATED HOMEODOMAIN PROTEIN 13-LIKE (WOX13L) for initiation of tip growth (Sakakibara et al., 2014), and COLD-SHOCK DOMAIN PROTEIN 1 (CSP1) for enhancement of reprogramming (Li et al., 2017). CSP1 shares highest sequence similarity and domain structure with Lin28 and appears to have a similar function with Lin28 in reprogramming (Li et al., 2017). *SQUAMOSA PROMOTER BINDING PROTEIN 4* (*SBP4*; *Pp3c3\_31330*) has been identified as a negative regulator of stem cell formation in cut leaves (Sato et al., unpublished data). In screens for factors involved in the SBP4-mediated stem cell formation, a gene encoding a member of an uncharacterized subgroup of AP2/ERF transcription factors, *STEM CELL-INDUCING FACTOR 1* (*STEMINI*; *Pp3c1\_27440*; Fig. 2A) was identified. In addition, induction of *STEMINI* in gametophore leaf cells could change leaf cells to chloronema apical stem cells even in the absence of cutting and phytohormones (Fig. 3 and Fig.4A-D).

The aim of this study is to elucidate the mechanism of cell-fate change from gametophore leaf cells to chloronema apical stem cells via the function of STEMIN1. For this purpose, STEMIN1-direct target genes were identified using a combination of RNA-sequencing (RNA-seq) and chromatin immunoprecipitation-sequencing (ChIP-seq). Furthermore, STEMIN1 induction specifically reduced the H3K27me3 levels at those loci and activated expression of these genes. These results indicate that STEMIN1 functions to remove the H3K27me3 from specific genes involved in stem cell formation to activate a genetic regulatory network underlying stem cell formation. Since STEMIN1 has homologs in other land plants, these findings could shed light on a mechanism potentially underlying the plasticity and regeneration ability of plant cells.

### **3. Materials and Methods**

#### **3.1. Plant material**

The Gransden 2004 strain of *Physcomitrella patens* (Rensing et al., 2008) (Physcomitrella) was used as the wild-type strain and cultured on BCDAT medium under continuous white light at 25°C (Nishiyama et al., 2000). Polyethylene glycol-mediated transformation and preparation of gametophores were performed as described previously (Ishikawa et al., 2011; Sato et al., 2017; Nishiyama et al., 2000).

#### **3.2. Screening of factors to induce reprogramming of leaf cells**

*SQUAMOSA PROMOTER BINDING PROTEIN 4* (*SBP4*: AJ968319) (Riese et al., 2008) has been found to negatively regulate stem cell formation in cut leaves (Sato, Higuchi, Kabeya et al., unpublished data). To identify factors that function downstream of SBP4 during reprogramming, SBP4 and SBP4 fused with the repression domain SRDX (Hiratsu et al., 2003) were overexpressed in cut leaves. Genes with decreased and increased expression levels were screened, respectively, and then 30 candidate genes were selected (Sato, Higuchi et al., unpublished data). The coding sequences of these genes were cloned into the pENTR/D-TOPO vector (Invitrogen) and were subjected to the LR reaction using the destination vector pLGZ1. The generated constructs were digested with the restriction enzyme *PmeI* for gene targeting and introduced into the GX6:NGG#63 plants (Ishikawa et al., 2011; Kubo et al., 2013).

Gametophores of the resultant transgenic plants were cultivated in 10 mM MES buffer (pH 6.0) with 1  $\mu$ M  $\beta$ -estradiol for 5 days to examine whether induction of these genes induced stem cell formation in leaves without their being cut.

### **3.3 Construction of $\beta$ -estradiol-inducible *STEMIN* plasmids**

Primers used to construct the plasmids are listed in Table 1. To construct the LGZ1-*STEMIN*1, LGZ1-*STEMIN*2, LGZ1-*STEMIN*3, and PTA2GX6-*STEMIN*1 plasmids, the *STEMIN*1, *STEMIN*2, and *STEMIN*3 coding sequences were cloned into the pENTR/D-TOPO vector (Invitrogen) to generate the plasmids pENTR:*STEMIN*1, pENTR:*STEMIN*2, and pENTR:*STEMIN*3, respectively. The three plasmids were subjected to the LR reaction using the destination vector pLGZ1. The generated constructs were digested with the restriction enzyme *Pme*I for gene targeting and introduced into the GX6:NGG#63 plant (Ishikawa et al., 2011; Kubo et al., 2013) to generate several independent GX6:NGG>*STEMIN*1, GX6:NGG>*STEMIN*2, and GX6:NGG>*STEMIN*3 plants.

For generation of the PTA2GX6-*STEMIN*1 plants, the pENTR-*STEMIN*1 plasmid was subjected to the LR reaction using the destination vector pT2GX6. The generated constructs were digested with the restriction enzyme *Pme*I for gene targeting and introduced into the RM09#35 and ProCYCD;1:NLS-sGFP-GUS#263 plants (Ishikawa et al., 2011), respectively.

To construct the GX6:*STEMIN*1-Myc plasmid, the *STEMIN*1 coding

sequence without the stop codon was cloned into pENTR/D-TOPO and subjected to the LR reaction using the destination vector pGX6M.

### **3.4. Plasmid construction for spatial expression analysis**

Primers used for plasmid construction are given in Table 1. To insert the *sGFP* gene in frame with the *STEMIN1* coding sequence, a DNA fragment of *sGFP* was amplified and inserted into pBluescript SK+. Subsequently, a genomic DNA fragment including a partial sequence of the *STEMIN1* promoter and 5' untranslated region was amplified from wild-type genomic DNA and inserted at the 5' end of the *sGFP* gene. A genomic DNA fragment from the start codon to the stop codon of *STEMIN1* was fused with the 3' end of the coding region of the *sGFP* gene.

For the promoter-reporter lines, DNA fragments including partial sequences of the *STEMIN1*, *STEMIN2*, and *STEMIN3* promoters were amplified and inserted into the *SmaI* site of pPIG1bNGGII (Ishikawa et al., 2011). The generated constructs were digested with *PmeI* for gene targeting and introduced into wild-type *Physcomitrella*.

### **3.5. Plasmid construction for the deletion of *STEMIN* genes and generation of the *STEMIN*-deletion mutants**

Primers used for plasmid construction are given in Table 1. To delete *STEMIN1*, *STEMIN2*, and *STEMIN3* in wild-type *Physcomitrella*, genomic fragments containing

the 5' and 3' flanking regions of each gene were inserted into the 5' end and 3' region of the blasticidin S deaminase (BSD) (Tamura et al., 1995) expression cassette of the plasmid p35S-loxP-BSD, of the zeo expression cassette (Sakakibara et al., 2008) of the plasmid p35S-loxP-Zeo, and of the neomycin phosphotransferase II expression cassette (nptII) (Nishiyama et al., 2000) of the plasmid pTN182, respectively. The generated constructs were digested with suitable restriction enzymes for gene targeting.

To generate  $\Delta$ stemin1 $\Delta$ stemin2 $\Delta$ stemin3 triple deletion mutants, the *STEMIN1*-deletion construct was introduced into the wild type to generate  $\Delta$ stemin1 plants. Two independent deletion mutants,  $\Delta$ stemin1#47 and  $\Delta$ stemin1#54, were used for further transformation. The *STEMIN3*-deletion construct and subsequently the *STEMIN2*-deletion construct were introduced into the  $\Delta$ stemin1 plants to generate the  $\Delta$ stemin1 $\Delta$ stemin2 $\Delta$ stemin3 triple deletion mutants.

To generate  $\Delta$ stemin1 $\Delta$ stemin2 double deletion mutants, the *STEMIN2*-deletion construct was introduced into the  $\Delta$ stemin1#47 and  $\Delta$ stemin1#54 plants, respectively.

### **3.6. Phylogenic analysis**

Amino acid sequences collected from NCBI nr database as well as cDNA sequences determined here were used to reconstruct the phylogeny using an updated procedure based on that described in Banks et al (Banks et al., 2011). A total of 1000 sequences collected with the BLASTP search against the NCBI nr database as of September 3,

2018 and the proteome dataset of *Pinus taeda* (v2.01) (Neale et al., 2014), *Azolla filiculoides* (*Azolla\_asm\_v1.1*), and *Salvinia cucullata* (*Salvinia\_asm\_v1.2*), (<https://www.nature.com/articles/s41477-018-0188-8>) were gathered. The BLASTP was performed with `word_size` set to 2 using the core STEMIN1 sequence as query. The search was separately performed for *Arabidopsis thaliana*, *Amborella tricarpha*, *Marchantia polymorpha*, *Physcomitrella patens*, and *Selaginella moellendorffii*, and green algae (Viriplantae excluding Embryophyta). About top 20 sequences (for *Selaginella* 40 as it is diploid, for green algae to 40 sequences as the category includes many species) from each run were collected. The collected sequences were merged in a single FASTA format file containing 244 sequences and aligned using MAFFT 7.407 (Kato et al., 2005) with the `einsi` setting. The alignment was converted to a nexus format file with sites that were present in at least 95% of the taxa set as included and others as excluded. The alignment was examined with MacClade version 4.08 (Maddison et al., 1989) and sequences having indels within the region were removed and 225 sequences were retained. A conserved block consisting of 56 amino acid starting (F/Y) and ending with (F/L/Y) was kept and further analysis was performed on this region. Sequences identical in this region were treated as a single OTU and thus 168 distinct OTUs remained. RAxML 8.2.12 was used to search for the ML tree. First, the best fit model was chosen with ProteinModelSelection.pl and LG was selected. The ML was searched with `-f a` option with `-#100 -m PROTGAMMALG` and `-p, -x` parameters chosen randomly, taking 24 bits from `/dev/urandom`. Bootstrap resampling was done using SEQBOOT from PHYLIP package

(<http://evolution.genetics.washington.edu/phylip.html>), and 100 bootstrap data were individually analyzed with same manner, and the consensus was counted using CONSENSE. The bootstrap values were put on the original tree.

### **3.7. RNA preparation and qRT-PCR analysis**

To analyze transcript accumulation in response to *STEMINI* induction, 3-week-old gametophores of the GX6:NGG>STEMIN1#23 lines were soaked in liquid BCDAT medium with or without 1  $\mu$ M  $\beta$ -estradiol for 24 hours. Cut leaves were prepared as described (Ishikawa et al., 2011).

Total RNA was purified from whole gametophores, cut leaves, and protonemata with the RNeasy Micro Kit (Qiagen). First-strand cDNA was synthesized using the ReverTra Ace qPCR RT Master Mix (TOYOBO). qRT-PCR was performed using an ABI PRISM 7500 (Applied Biosystems) or a QuantStudio 3 Real-Time PCR System (Applied Biosystems) with the QuantiTect SYBR Green PCR Kit (Qiagen) or THUNDERBIRD SYBR qPCR Mix (TOYOBO). The sequences of primers for qRT-PCR are listed in Table 1. Results were analyzed using the comparative critical threshold method. Each transcript level determined by qRT-PCR analysis was normalized to that of  $\alpha$ -tubulin (*TUAI*) (Ishikawa et al., 2011) during stem cell formation in cut leaves and that of *Pp3c9\_17670* encoding a thiosulfate sulfurtransferase during side blanch formation in chloronemata.

### **3.8. GUS staining**

The histochemical detection of  $\beta$ -glucuronidase (GUS) activity followed a previous report (Nishiyama et al., 2000).

### **3.9. Microscopy analysis**

Bright and fluorescent images of wild-type and transgenic plants were observed under fluorescence microscopes BX51 and SZX16 (Olympus).

For time-lapse analyses, the distal halves of the fifth to tenth youngest leaves were excised with a razor blade. To isolate single leaf cells, cut leaves were further divided into small pieces with a razor blade and all leaf cells except a single cell targeted for reprogramming were ablated by the PALM microdissection system (Zeiss) (Sato et al., 2017). Time-lapse analyses of cut leaves and singly isolated leaf cells were performed as reported previously (Sato et al., 2017). Images of leaves reprogramming were recorded at 20-min intervals with a digital camera ORCA-AG (Hamamatsu Photonics K. K.) using an inverted microscope IX81 (Olympus). The images were reconstructed to create a movie with ImageJ (<http://rsb.info.nih.gov/ij/>). Image intensities of sGFP in transgenic plants were adjusted by Photoshop CS6 (Adobe).

### **3.10. Scanning electron microscopy**

Four-week-old gametophores of the GX6:NGG>STEMIN1#23 lines were soaked in

liquid BCDAT medium with 1  $\mu\text{M}$   $\beta$ -estradiol (WAKO) for 3 days. The gametophores were frozen in liquid nitrogen and immediately observed with a scanning electron microscope XL30 (FEI, Hillsboro), with an accelerating voltage of 10 kV.

### **3.11. Measurement of frequency of side-branching in chloronemata**

Under unilateral red light, chloronema apical stem cells produce almost no side branch initial cells, whereas subsequent incubation under white light induces the formation of secondary chloronema apical stem cells (Aoyama et al., 2012; Uenaka et al., 2005). Chloronemata of wild-type, *STEMIN*-deletion, and inducible *STEMIN* lines grown on BCDAT agar medium under unilateral red light (10 to 15  $\mu\text{mol}/\text{m}^2/\text{s}$ ) for 10 days were transferred to white light (approximately 50  $\mu\text{mol}/\text{m}^2/\text{s}$ ). Concurrently, the inducible *STEMIN* lines were treated with or without 1  $\mu\text{M}$   $\beta$ -estradiol. These lines were further cultivated under continuous white light. The number of side-branch initial cells was counted 12 hours (wild type and deletion mutants) or 24 hours (inducible *STEMIN* mutants) after the transfer. Chloronemata were stained with 10  $\mu\text{M}$  of the lipophilic dye *N*-(3-triethylammoniumpropyl)-4-(6-(4-(diethylamino)phenyl)hexatrienyl)-pyridinium dibromide (FM4-64; ThermoFisher Scientific).

### **3.12. DNA gel-blot and immunoblot analyses**

DNA gel-blot and immunoblot analyses followed previous reports (Ishikawa et al.,

2011; Okano et al., 2009).

### **3.13. Transcriptome analysis**

Four-week gametophores of the GX6:STEMIN1-Myc#11 and GX6:NGG#63 plants were soaked in BCDAT medium with 1  $\mu$ M  $\beta$ -estradiol or DMSO for 24 hours. Only gametophores were collected with forceps. RNA was extracted and purified from the gametophores using a RNeasy Plant Mini Kit (Qiagen) and DNase I (Qiagen). Libraries for transcriptome analysis were prepared from the total RNA of the gametophores using a TruSeq Stranded mRNA LT Sample Prep Kit Set A (Illumina). Sequencing was conducted using a HiSeq 2500 instrument (Illumina) and 50-bp single-end reads were obtained. Adapter and low-quality sequences were removed from the reads using cutadapt version 1.9.1 (DOI; 10.14806/ej.17.1.200). The reads were analyzed using transcriptome sequence of *Physcomitrella patens* version 3.3 (Ppatens\_251\_v3.3.transcript.fa) obtained from Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>) and calculated using RSEM (version 1.3.0) with Bowtie2. The transcriptome analysis was performed in biological triplicate. Differentially expressed genes (DEGs) with  $q < 0.01$  were extracted with a R package, TCC (Sun et al., 2013), and then DEGs with a fold change  $>1.5$  were selected.

### **3.14. ChIP-sequencing (ChIP-seq)**

Four-week gametophores with rhizoids and protonemata of the GX6:STEMIN1-Myc#11 plant were soaked into BCDAT liquid medium with 1  $\mu$ M  $\beta$ -estradiol or DMSO for 24 hours and were treated with 1% formaldehyde for 20 min at room temperature. Cross-linking reaction was stopped with 0.1 M glycine. Subsequently, the tissues were washed with Milli-Q-purified water and treated with 2% Driselase for 20 min to remove rhizoids and protonemata.

ChIP was performed as reported previously (Gendrel et al., 2005) with several modifications. In short, genomic DNA were extracted and sheared to an average size of approximately 100–200 bp with 20 U/ $\mu$ l micrococcal nuclease (No.2910A; Takara) for 15 min at 37°C. The reaction was stopped with 20 mM EDTA and 1% SDS, and the lysates were incubated for 1 hour on ice. After the input samples were collected, the lysates were incubated for 2 hours at 4 °C with 10  $\mu$ g of anti-Myc (4A-6; Millipore), anti-H3 (ab1791, Abcam), anti-H3K27me3 (07-449, Millipore) or anti-H3K4me3 (07-473, Millipore) antibodies. Immunocomplexes were recovered with protein G agarose (Roche), and then washed and eluted with elution buffer (1% SDS, 0.84% NaHCO<sub>3</sub>). After cross-links were reversed, the samples were treated with RNase A and proteinase K, and the DNA was recovered after phenol-chloroform extraction. The recovered DNA was dissolved in 5 mM Tris-HCl (pH 8.5). ChIP libraries were then prepared using a DNA SMART ChIP-seq Kit (Clontech). The libraries were then subjected to an Illumina Hiseq 1500 sequencer and approximately 126-bp reads were obtained. Biological duplicate ChIP-seq experiments for STEMIN1-Myc, H3, H3K27me3, and H3K4me3 were performed.

### 3.15. ChIP-seq data analysis

Quality control of the ChIP-sequencing output was undertaken using FASTQC version 0.11.3 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). To map the sequence reads on the *Physcomitrella* genome, adapter sequences of ChIP DNA sequences in the FASTQ format were removed using cutadapt version 1.9.1 (DOI: 10.14806/ej.17.1.200). The resultant sequences were mapped to the *Physcomitrella patens* V3.0 genome (Ppatens\_251\_v3.fa) of Phytozome (DOE-JGI; <https://phytozome.jgi.doe.gov/>) using Bowtie2 version 2.2.6 (Langmead and Salzberg, 2012).

For STEMIN1-Myc binding sites, DMSO-treated gametophores were used as a control to identify the STEMIN1-Myc binding regions. Peaks appearing significantly higher in ChIP-DNA from  $\beta$ -estradiol-treated gametophores compared to DMSO-treated gametophores were detected as STEMIN1-Myc binding sites using MACS (MACS) version 1.4.2 (Zhang et al., 2008) with the default parameters. The peaks were annotated for *Physcomitrella* genetic information using the R packages, ChIPpeakAnno version 3.0.0 (DOI: 10.18129/B9.bioc.ChIPpeakAnno) and GenomicRanges version 1.22.4 (DOI: 10.18129/B9.bioc.GenomicRanges).

For distribution of histone H3, H3K27me3, and H3K4me3, the reads from immunocomplexes with anti-H3, anti-H3K27me3, and anti-H3K4me3 antibodies were aligned to a *Physcomitrella* genome using Subread version 1.5.2 (Liao et al., 2013). The

enriched levels of H3, H3K27me3, and H3K4me3 were calculated as reads per kilobase per million mapped sequence reads (RPKM) based on the formula: (number of tags mapped to each region  $\times$  1,000,000,000) / (length of each region  $\times$  number of total mapped tags).

The normalized average distribution of STEMIN1-Myc binding sites, histone H3, H3K27me3, and H3K4me3 was analyzed using ngs.plot (Shen et al., 2014). The distributions of histone H3 and H3K27me3 on the *CYCD;1* and *BETA-EXPANSIN (Pp3c17\_12980)* genes were visualized using Integrated Genomics View (IGV) (Robinson et al., 2011; Thorvaldsdottir et al., 2013).

### **3.16. Accession numbers**

Sequence data for genes and plasmids discussed in this article can be found in DDBJ/GenBank/EMBL data libraries under the following accession numbers: *STEMIN1* (LC042086), *STEMIN2* (LC042087), *STEMIN3* (LC042088), pLGZ1 (AB602442), pGX6M (LC388570), pT2GX6 (LC388571), pPIG1bNGGII (AB537478), p35S-loxP-BSD (AB537973), p35S-loxP-Zeo (AB540628), pTN182 (AB267706).

The RNA-seq and ChIP-seq data are deposited in the DDBJ Sequence Read Archive (DRA) with accession numbers DRA007364 and DRA007365, respectively.

**Table 1. Primer sequences used for qRT-PCR, RT-PCR, and plasmid construction**

Primers used for qRT-PCR

Gene name		Sequence
<i>STEMIN1</i>	(F)	5'-TCCGCCACCCTCAATTGA-3'
	(R)	5'-GCATCTTCAGCAGTCTCGAATG-3'
<i>STEMIN2</i>	(F)	5'-AAGTGCAATTCCGCGACAAT-3'
	(R)	5'-TTTCCGGTTGCGGTGAA-3'
<i>STEMIN3</i>	(F)	5'-CCTGGACGCGGAGCAAT-3'
	(R)	5'-TTGCCTCCCCGTCTTCTTC-3'
<i>TUAI</i>	(F)	5'-CGTAGGAGGGACCAGTTTGG-3'
	(R)	5'-TGCATTCATCCCCGAGTCA-3'

Primers used for RT-PCR

Gene name		Sequence
<i>STEMIN1</i>	(F)	5'- ATGGGCAGGCCGCGAGAAATACCGAG-3'
	(R)	5'- GTATCCCGAAGCTGTAAGGACAGA-3'
<i>TUAI</i>	(F)	5'- GCCGTGTTGTTGCTGATCATGCAC-3'
	(R)	5'- GAGCCATACAACAGCGTGCTGTC-3'

Primers used for plasmid construction

Construct		Sequence
LGZ1-STEMIN1	(F)	5'-CACCATGGGCAGGCCGCGAGAAATACCGAG-3'
	(R)	5'-GTATCCCGAAGCTGTAAGGACAGAG-3'
LGZ1-STEMIN2	(F)	5'-CACCATGGGCAGGCCACAGAGATA-3'
	(R)	5'-TCAAGACGAAATGCAATCGAAATGCCGA-3'
LGZ1-STEMIN3	(F)	5'-CACCATGGGCAGGCCGCGAGAGATA-3'
	(R)	5'-CTACTCGAAAGCTGTGGTGACAGC-3'
sGFP-STEMIN1	(F)	5'-CCCTCGAGCCCTCATATGTTATCCGA-3'
	(R)	5'-GGTCTGTGGATGCTATTCTCCTC-3'
	(F)	5'-ATGGGCAGGCCGCGAGAAATACCGAG-3'
	(R)	5'-TCGCGGCCGCGTATCCCGAAGCTGTAAG-

		GAC-3'
sGFP	(F)	5'-GATATCGGACTGACCACCCGCTGCTCC-3'
sGFP	(R)	5'-CCCTCGAGGTCGACTCTAGAGGATCC-3'
STEMIN1pro	(F)	5'-CTGTCGACTTGACGTACATCATAGAACA-
:NLS-sGFP-GUS		AAATACCT-3'
	(R)	5'-CTGCGGCCTGCCCATGGTCTGTGGAT-3'
STEMIN2Pro	(F)	5'-GGGGATCCCAAGTCTAGTTACTGGGAAA
: NLS-sGFP-GUS		TGTGG-3'
	(R)	5'-GGGATATCCATGATCTCCACCAACTCAA-
		ACAC-3'
STEMIN3Pro	(F)	5'-GGAGATCTCCTTTAGAGTGTTTCTCGGC-
: NLS-sGFP-GUS		ATCGA-3'
	(R)	5'-GGGATATCCATGATTTCTCTGCAAACCTC-
		AGCGACAC-3'
Δstem1	5'(F)	5'-ATTTGCGCCGAGGATGGGTAGAG-3'
	5'(R)	5'-CCAAGCTCCCAGAGTTCAGAGAGTA-3'
	3'(F)	5'-TCCCACGGTCGTGGTTTTTAC-3'
	3'(R)	5'-CAGCTCTGAAACACCCCAATTAATATC-3'
Δstem2	5'(F)	5'-GGGCCCCGTATAATGAGGGGCTTTGCC-3'
	5'(R)	5'-AAGCGGCCGCGATTTCTCTGCAAACCTC-
		AGCG-3'
	3'(F)	5'-CGATCGATAGGAAAGAGCGTTTATTG-
		ATCTG-3'
	3'(R)	5'-CGCTCGAGTGTTGCAAACCCTTTCTGG-3'
Δstem3	5'(F)	5'-CAGGTACCATTTTCCTAGGGTCCTTTGACT-3'
	5'(R)	5'-TCCTCGAGCTGGCACCAATGTCCATA-3'
	3'(F)	5'-CAGCACTATCTCAACCCCCCCCCCTACA-3'
	3'(R)	5'-TCGCGGCCGCGTGCTGATCCATCCCATGT-3'
GX6:STEMIN1	(F)	5'-GGAGAGGACACGCTGAAGCTAG-3'
-Myc	(R)	5'-AAGTCGACGTATCCCGAAGCTGTAAGGACA-
		GAGCAG-3'

---

(F), forward primer; (R), reverse primer.

## 4. Results

### 4.1. *STEMIN1* changes leaf cells into stem cells without cutting

*STEMIN1* was fused with a  $\beta$ -estradiol-inducible promoter to generate plants carrying an inducible *STEMIN1* gene (GX6:NGG>*STEMIN1*; Fig. 3A,B; Kubo et al., 2013). When gametophores of the transgenic plants were cultivated with 1  $\mu$ M  $\beta$ -estradiol for five days, a subset of the leaf cells changed to chloronema apical stem cells without being cut (Fig. 4A-D). The efficiency of protrusion was positively correlated with the level of *STEMIN1* transcript (Fig. 3C,D).

Another  $\beta$ -estradiol-inducible *STEMIN1* expression construct was introduced into the *PTA2* neutral site (Kubo et al., 2013) in the protonema-specific reporter RM09#35 line (Ishikawa et al., 2011) and the cell-cycle reactivation reporter line CYCD;1pro:NLS-sGFP-GUS#263 (CYCD;1pro:NGG#263) (Ishikawa et al., 2011), respectively (Fig. 4E,F). After induction of *STEMIN1* in gametophores of these lines with 1  $\mu$ M  $\beta$ -estradiol, the GFP signal was detected at leaf cells 24 hours after induction, prior to the onset of tip growth (Fig. 4E,F and Fig. 5). These results indicate that *STEMIN1* induces reprogramming of leaf cells to chloronema apical stem cells.

### 4.2. Function of *STEMIN* genes in stem cell formation in cut leaves

To investigate *STEMIN1* function during stem cell formation in cut leaves, a quantitative reverse transcription PCR (qRT-PCR) analysis was performed using cut

leaves of gametophores. The result indicates that *STEMIN1* mRNA kept increasing up to 30 hours after the leaves were cut and started to decrease at 36 hours (Fig. 6A). To investigate spatiotemporal expression patterns of *STEMIN1* protein, GUS-*STEMIN1* and sGFP-*STEMIN1* transgenic plants, in which *uidA* (a  $\beta$ -glucuronidase gene: GUS; Jefferson et al., 1987) and *sGFP* genes were inserted just before the *STEMIN1* initiation codon to produce the GUS-*STEMIN1* and sGFP-*STEMIN1* fusion proteins, respectively (Fig. 7 and Fig. 8) were observed. GUS-signal was detected in leaf cells facing the cut but not detected in protruded cells at 30 hours (Fig. 6B and Fig. 7), suggesting that the signal disappeared when reprogrammed cells are protruded. On the other hand, the sGFP-signal was not detected at any time point (Fig. 8C), probably because of the lower sensitivity of sGFP-*STEMIN1* detection than GUS-*STEMIN1*.

For this reason, Pro*STEMIN1*-NGG transgenic plants, which carry a 3-kb *STEMIN1* promoter sequence attached to the fusion gene *NGG*—consisting of the SV40 nuclear localization signal (*NLS*; Kalderon et al., 1984), *sGFP*, and *GUS*—inserted into the *PIG1* neutral site (Fig. 9A,B) were generated, and were investigated the promoter activity of *STEMIN1*. While *STEMIN1* promoter activity was not detected in intact leaves (Fig. 9C,D), sGFP signal was detected 24 hours after cutting in leaf cells facing the cut (Fig. 6C, and Fig. 9E). Some of the sGFP-positive cells protruded and reentered the cell cycle to become chloronema apical stem cells (Fig. 6C, arrowheads at 36 h). After approximately two rounds of cell division, the sGFP signal in the chloronema apical stem cells diminished (Fig. 6C, arrowheads at 48 hours). Other sGFP-positive cells did not protrude, and in these cells sGFP signals diminished by 36 hours (Fig.

6C,D and Fig. 9E). Cells without sGFP signals did not become stem cells (Fig. 6C,D and Fig. 9E). When a leaf cell is singly isolated from a leaf as previously described (Sato et al., 2017), the *STEMIN1* promoter was also activated before tip growth (Fig. 6E), indicative of cell-autonomous *STEMIN1* induction by wounding. Down-regulation of the *STEMIN1* promoter activity might take place dependently on positional control mechanisms from surrounding cells after *STEMIN1* expression, possibly under the control of an unknown mechanism by which stem cell inhibits reprogramming of neighboring somatic cells (Sato et al., 2017).

To examine roles of *STEMIN1* in stem cell formation in cut leaves, *STEMIN1* deletion mutant plants ( $\Delta$ stemin1; Fig. 10) were constructed. However, there were no difference in the timing of initiation of tip growth between wild-type and  $\Delta$ stemin1 plants after cutting (Fig. 6F). Phylogenetic analysis of *STEMIN*-related protein had identified two other genes in *Physcomitrella* closely related to *STEMIN1*, designated *STEMIN2* and *STEMIN3* (Fig. 2). Therefore, I made Pro*STEMIN2*-NGG and Pro*STEMIN3*-NGG transgenic plants, which carry a 3-kb *STEMIN2* and *STEMIN3* promoter sequences attached to the fusion gene *NGG*, respectively (Fig. 11). I detected the promoter activities of *STEMIN2* and *STEMIN3* at leaf cells facing the cut, as well as those of Pro*STEMIN1*-NGG (Fig. 12), although the transcripts of these genes were less abundant than those of *STEMIN1* during stem cell formation (Fig. 12E). These suggested that they function redundantly in stem cell formation in cut leaves.

To address it, triple deletion mutant plants ( $\Delta$ stemin1 $\Delta$ stemin2 $\Delta$ stemin3; Fig. 13) were established. In these mutant plants, the number of cells exhibiting tip growth

in cut leaves at 36 and 48 hours after cutting was significantly lower than in wild-type plants (Fig. 6G and Fig. 13E). In addition, the levels of *RM09* and *CYCD;1* transcripts in cut leaves 24 hours after cutting were lower in the triple mutant plants (Fig. 6H).

On the other hand, unlike *STEMIN1*, when I induced *STEMIN2* and *STEMIN3* in gametophores, respectively, I did not find any morphological changes of the gametophores (Fig. 14 and Fig. 15). These results indicate that *STEMIN2* and *STEMIN3* have functions overlapping those of *STEMIN1*, but they lack the ability of *STEMIN1* to change intact leaf cells into stem cells.

### **4.3. The function of *STEMIN1* and its homologs in protonema development**

The functions of *STEMIN1* and its homologs were further investigated in regular protonema development. Chloronema apical stem cells at the tip of chloronema filamentous tissue repeatedly divide to produce chloronema cells (Cove and Knight, 1993). Chloronema cells two cells away from the apical stem cells are usually reprogrammed and protrude, with tip growth, to form side-branch initial cells that subsequently become secondary chloronema apical stem cells (Aoyama et al., 2012). Promoter activities of *STEMIN1* and *STEMIN2* were analyzed using *STEMIN1*pro:NGG#7 and *STEMIN2*pro:NGG#238 line, respectively, and the sGFP signal was detected in the chloronema cells before protrusion, and the signal diminished before the first cell division (Fig. 16A-D and Fig. 17A). These activities were not

detected in chloronema apical stem cells.

When chloronemata are cultivated under unidirectional red light conditions, they change to become pale green protonemata and do not form side branches (Fig. 1). When those protonemata are moved to white light conditions, side-branch initial cells are simultaneously induced (Fig. 1) (Aoyama et al., 2012; Uenaka et al., 2005). In this system, the *STEMIN1* and *STEMIN2* promoters were activated and *STEMIN1* and *STEMIN2* transcript abundances increased in response to white light before side-branch initiation (Fig. 16E and Fig. 17B). *STEMIN3* transcripts were less abundant than *STEMIN1* and *STEMIN2* transcripts (Fig. 17C), and *STEMIN3* promoter activity was not detected (Fig. 16C and Fig. 17B). When red-light-grown protonemata of the triple deletion mutant plants were moved to white light conditions, the frequency of side-branch formation at the 5th and 6th cells was significantly lower than that in the wild type (Fig. 16F), while the frequency in the  $\Delta$ stemin1 and  $\Delta$ stemin1 $\Delta$ stemin2 double deletion mutants was not significantly different from that in wild type (Fig. 18 and Fig. 19). When each gene was induced in red-light-grown protonemata with  $\beta$ -estradiol just after moving the plants to white light conditions, the number of side branches increased significantly in the *STEMIN1*-overexpressing plants (Fig. 16G) but not in the *STEMIN2*- or *STEMIN3*-overexpressing plants (Fig. 20). Thus, the *STEMIN1*, *STEMIN2*, and *STEMIN3* gene products function redundantly in the initiation of side-branch initial cells from protonema cells, but only *STEMIN1* has the ability to induce side-branch initial cells that subsequently change into secondary chloronema apical stem cell formation. Since AP2/ERF gene family contains 171 genes in

Physcomitrella genome (Rensing et al., 2008; Li et al., 2017), other AP2/ERF transcription factors may have redundant functions to STEMINs.

#### **4.4 Identification of STEMIN1 direct target genes**

To elucidate mechanisms how STEMIN1 induces stem cell formation, I next focused on identification of the STEMIN1 direct target genes using a combination of RNA-sequencing (RNA-seq) and Chromatin immunoprecipitation-sequencing (ChIP-seq) analyses. First, I generated a transgenic line in which *STEMIN1* fused with a triple Myc-tag sequence is inducible by  $\beta$ -estradiol (GX6:STEMIN1-Myc#11, Fig. 21). Gametophores of the GX6:STEMIN1-Myc#11 plant were cultivated with or without  $\beta$ -estradiol for 24 hours and collected, respectively. The RNA-seq analysis demonstrated that transcript levels of 2,871 and 3,890 genes increased and decreased, respectively, at 24 hours after induction of STEMIN1-Myc in gametophores. In addition, the ChIP-seq analysis using anti-Myc antibody identified 6,733 genes that ChIP DNA from the GX6:STEMIN1-Myc#11 plant treated with  $\beta$ -estradiol were significantly enriched in their promoter regions compared with those without  $\beta$ -estradiol. Of the up- and down-regulated genes, 1,416 (pvalue=1.57821e-290) and 716 (p value=0.999687), respectively, were found to be directly targeted by STEMIN1-Myc in the ChIP-seq analysis (Fig. 22A and Fig. 24, See also Materials and Methods). In addition, 357 genes of the upregulated genes displayed at least 50% reduction of the H3K27me3 level and 242 of these genes (67.8 %) were

directly targeted by STEMIN1-Myc ( $p = 2.94388e-246$  by the hypergeometric test).

These included *CYCD;1* (Ishikawa et al., 2011) and *EXPANSIN* genes (Sakakibara et al., 2014), which function in stem cell formation and are upregulated in cut leaves.

#### **4.5 STEMIN1 regulates the histone H3K27me3 level of genes involved in stem cell formation**

H3K27me3 catalyzed by PRC2 facilitates maintenance of transcriptional repressive states and function in stabilization of developmental programs in both animals and land plants (Okano et al., 2009; Bouyer et al., 2011; Mozgova, 2017; Ikeuchi et al., 2015b). Therefore, establishment and removal of H3K27me3 at specific genes could be crucial for development and regeneration.

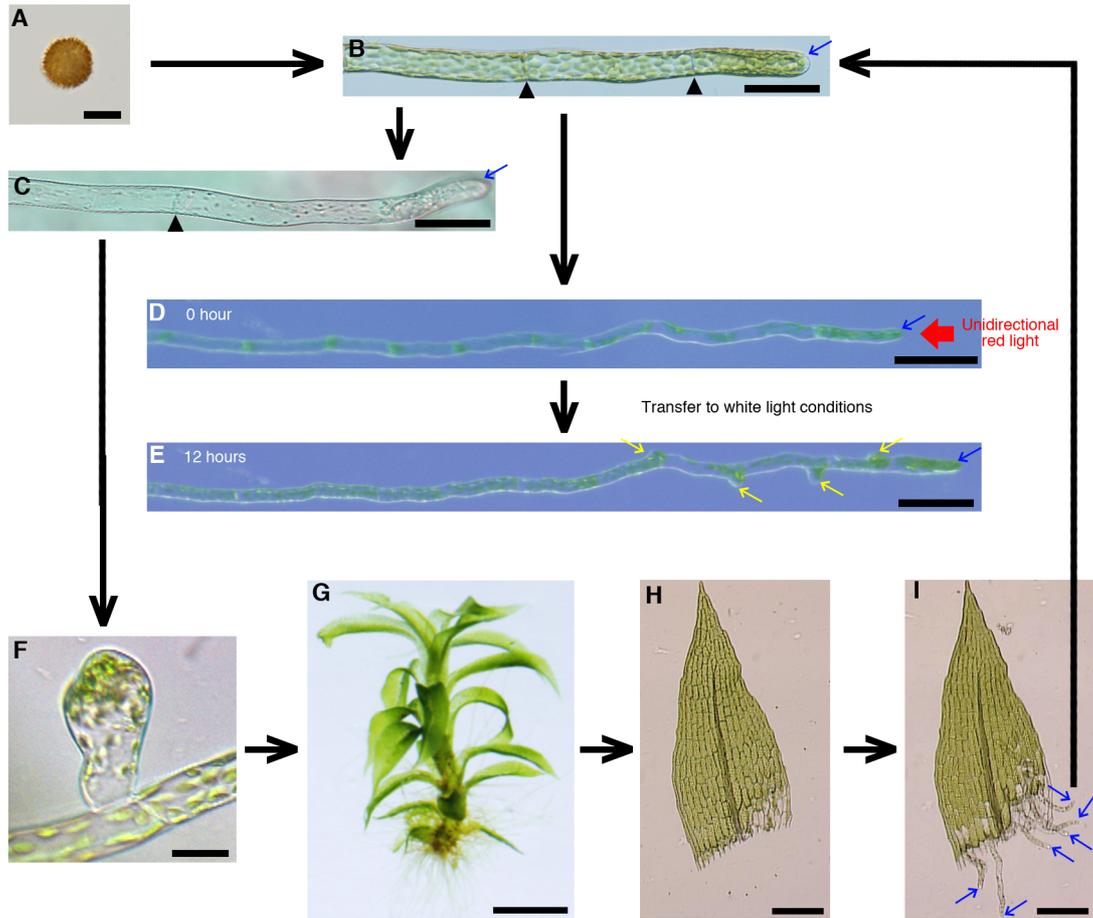
To examine changes in the H3K27me3 levels after STEMIN1 induction, I performed a ChIP-seq analysis of the GX6:STEMIN1-Myc#11 plant with or without  $\beta$ -estradiol treatment using anti-histone H3 and H3K27me3 antibodies. Among the upregulated genes, H3K27me3 levels in all 1,416 targeted genes were higher than those in the 1,455 non-targeted genes (Fig. 22C and Fig. 23A). Decreases in H3K27me3 levels after *STEMINI* induction were more conspicuous in the targeted than in the non-targeted genes (Fig. 22C and Fig. 23A). Correlations between changes in H3K27me3 and transcript levels of the directly targeted genes were stronger than for the non-targeted genes (Fig. 25). On the other hand, H3K4me3 levels, which are associated with active transcription in land plants and metazoans (Zhang et al., 2009; Barski.,

2007), increased in both targeted and non-targeted genes (Fig. 22D and Fig. 23B), suggesting that H3K4me3 levels are not specifically regulated by STEMIN1. Upregulated genes with decreased H3K27me3 levels included *CYCD;1* (Fig. 22E) and *BETA-EXPANSIN* (Fig. 26), which encode proteins with functions in stem cell formation and are upregulated in cut leaves (Ishikawa et al., 2011; Sakakibara et al., 2014). The decrease in H3K27me3 was specific to the *CYCD;1* and *BETA-EXPANSIN* genes, but the increase in H3K4me3 was not. Thus, *STEMIN1* induction functions to decrease H3K27me3 levels at certain target genes and thereby increase the abundance of their transcripts, at least some of which are involved in stem cell formation.

Since the H3K27me3 levels of 5,317 STEMIN1-targeted but non-upregulated genes did not conspicuously change (Fig. 27), STEMIN1 appears to cooperate with other factors to decrease H3K27me3 levels of 1,416 upregulated targeted genes. In the 3,890 down-regulated genes, H3K27me3 and H3K4me3 levels somewhat increased and decreased, respectively, and to similar degrees in the 716 targeted and the 3,174 non-targeted genes (Fig. 28).

Moreover, I found that STEMIN1-Myc preferentially interacted with annotated transcriptional start sites (Fig. 22B), and genes targeted by STEMIN1-Myc were more enriched among the upregulated genes (49%: Fig. 22A) than among the down-regulated genes (18%: Fig. 24). In general, AP2/ERF transcription factors have capacity to bind a wide range of cis-regulatory elements in promoters of target genes (Sasaki et al., 2007), such as GCC-box (GCCGCC element) and DRE/CRT (dehydrationresponsive element; TACCGACAT/C-repeat; TGGCCGAC) (De Boer et

al., 2011; Fujimoto et al., 2000; Ohme-Takagi and Shinshi, 1995; Hao et al., 1998, 2002; Oñate-Sánchez et al., 2006; Wang et al., 2012). Using MEME version 4.12.0 (Bailey et al., 1994), I found that STEMIN1-Myc bound to a site that resembles the GCC-box (GCCGCC element) at transcriptional starting sites (Fig. 29), corresponding to the position of the nucleosome-depleted zone (Fig. 22B,C). These results suggest that STEMIN1 could readily bind to the specific target genes with the silent chromatin status. At 24 hours after STEMIN1 induction in gametophores, cell division did not occur yet. Thus, STEMIN1-binding at the specific genes with H3K27me3 could activate an intrinsic mechanism underlying loss of the H3K27me3 on the genes in a cell division-independent manner, thereby erasing the epigenetic memory of leaf cells and rewriting a gene regulatory network for *de novo* formation of stem cells.



**Fig. 1. Formation of chloronema apical stem cells in protonemata and an excised leaf.**

(A) A spore.

(B) A chloronema filament with a chloronema apical stem cell (a blue arrow) and two produced chloronema cells. Cell septa are indicated by black arrowheads.

(C) A caulonema filament with a caulonema apical stem cell (a blue arrow) and a produced caulonema cell. A cell septum is indicated by a black arrowhead.

(D) A protonema filament cultivated under unidirectional red-light conditions for 10 days. Side branches are not formed. A protonema apical stem cell is indicated by a blue arrow.

(E) Formation of side-branch initial cells in a protonema filament at 12 hours after transfer to white light conditions. Blue and yellow arrows indicate an original chloronema apical stem and side-branch initial cells, respectively.

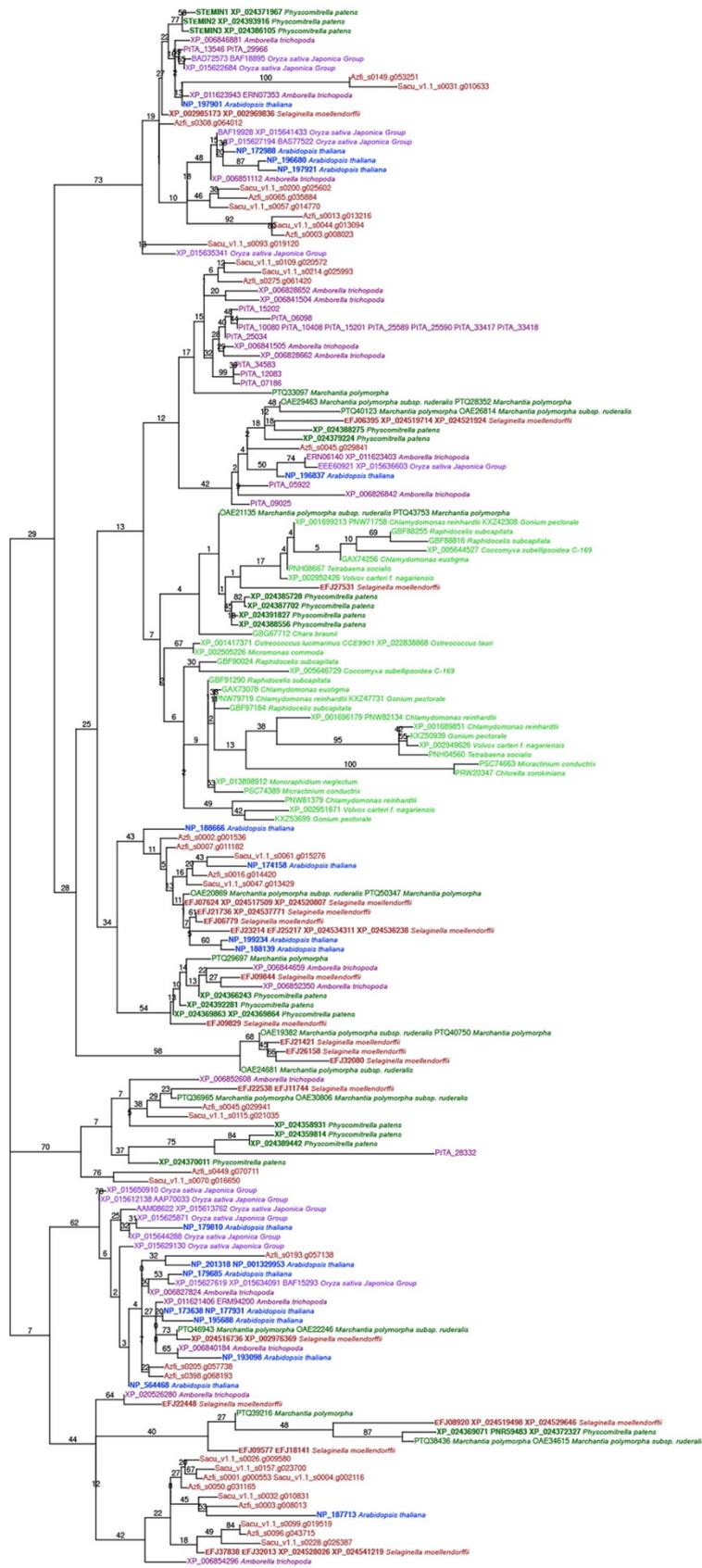
(F) A young gametophore.

(G) A gametophore.

(H,I) A excised leaf from a gametophore at 0 (H) and 48 (I) hours. Chloronema apical stem cells are indicated by blue arrows.

Scale bars, 10  $\mu\text{m}$  (A), 50  $\mu\text{m}$  (B,C), 100  $\mu\text{m}$  (D,E), 20  $\mu\text{m}$  (F), 500  $\mu\text{m}$  (G) and 200  $\mu\text{m}$  (H,I).

A



**B**

```

STEMIN1 MGRPQYRQVQRHGWGSHVSEIRHPQLKTRVWLGTFTAEDAHAHAYDEAA 50
STEMIN2 MGRPQYRQVQRHGWGSHVSEIRHPQLKTRVWLGTFTAEDAHAHAYDEAA 50
STEMIN3 MGRPQYRQVQRHGWGSHVSEIRHPQLKTRVWLGTFTAEDAHAHAYDEAA 50
AP2/ERF DNA binding domain
STEMIN1 RLMCGAKARTNFPYDPHASKGPKSOLLSATLHAKLQRYWLSLQDPDAQQG 100
STEMIN2 RLMCGARARTNFPYDPNAPKGLNSQLSATLSAKLQRYWVQSQRGGAAQG 100
STEMIN3 RLMCGVRARTNFPYDPNASKRPNSQLSATLSAKLHRWYLSQQRDGGQEG 100

STEMIN1 KKSSSRMSQSSLPCLCOLKHAEQSNLGSQCNQISSSQPLESNLVRIEHPD 150
STEMIN2 KITKDARMTQS-LTCLG-L-DAEQSNLGIWQKK--TGKQAEANWVRKVQFR 145
STEMIN3 KSKDARMTQS-LTCLG-L-DAEQSNLGIWQKK--TGRQAEANWVRKVQFG 145

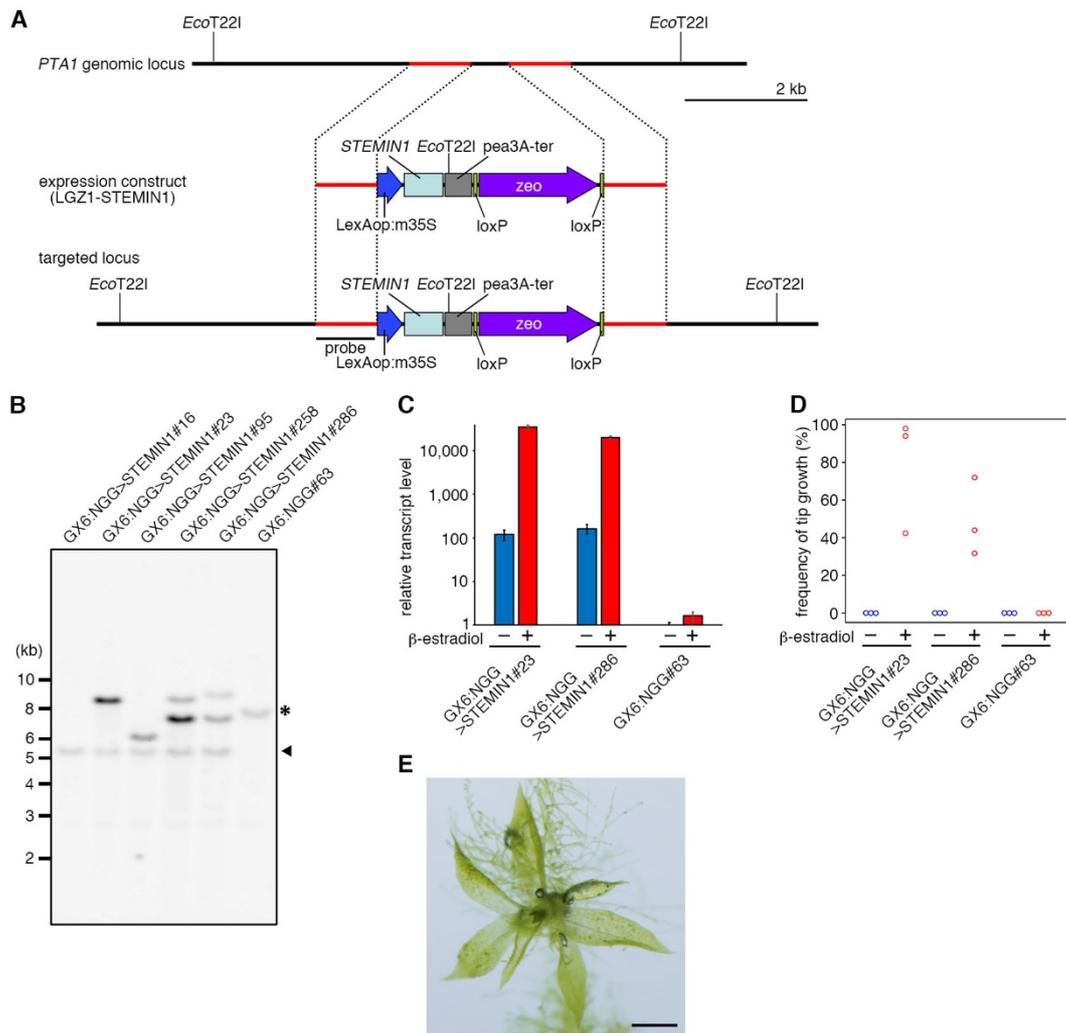
STEMIN1 DNADSFQAQPOPASSSDSCISEEDKFAAEMIEELLQSWQTSSQFYGSFNSP 200
STEMIN2 DNASLHADSPQPENSSGSSMSEEDKFAAEMIEELLG-HSPQGF-SDFGSP 193
STEMIN3 DNNSPQTDSPQPENSSGSCVSEEDKFAAEMIEELLG-YSPQGF-SNFGSP 193
cm domain
STEMIN1 TMSDSSGSSSP-CSVLTASGY----- 219
STEMIN2 ALSDSSCSSSP-CSVTAASKWDVLSLHFDICSS 226
STEMIN3 AMSDSSCSSS-CSVITAFE----- 212

```

**Fig. 2. Phylogeny of STEMIN and related proteins.**

(A) A phylogenetic tree of STEMIN and STEMIN-related proteins. The tree was constructed with the Maximum likelihood method using the LG model with gamma correction implemented in RAxML. Bootstrap values of >70% are shown on the branches. Horizontal branch length is proportional to the estimated evolutionary distance. Blue, eudicots; violet, monocots; dark magenta, basal angiosperms and gymnosperms; vermillion, ferns and lycopods; green, bryophytes. See supplement for the whole tree including outgroups.

(B) Amino acid sequence alignment of the STEMIN1, STEMIN2, and STEMIN3 proteins. All three proteins contain a single AP2/ERF DNA-binding domain (Riechmann and Meyerowitz, 1998) at the N-terminus and a C-terminal domain (cm) found in SHINE proteins (Aharoni et al., 2004).



**Fig. 3. Construction of GX6:NGG>STEMIN1 plants.**

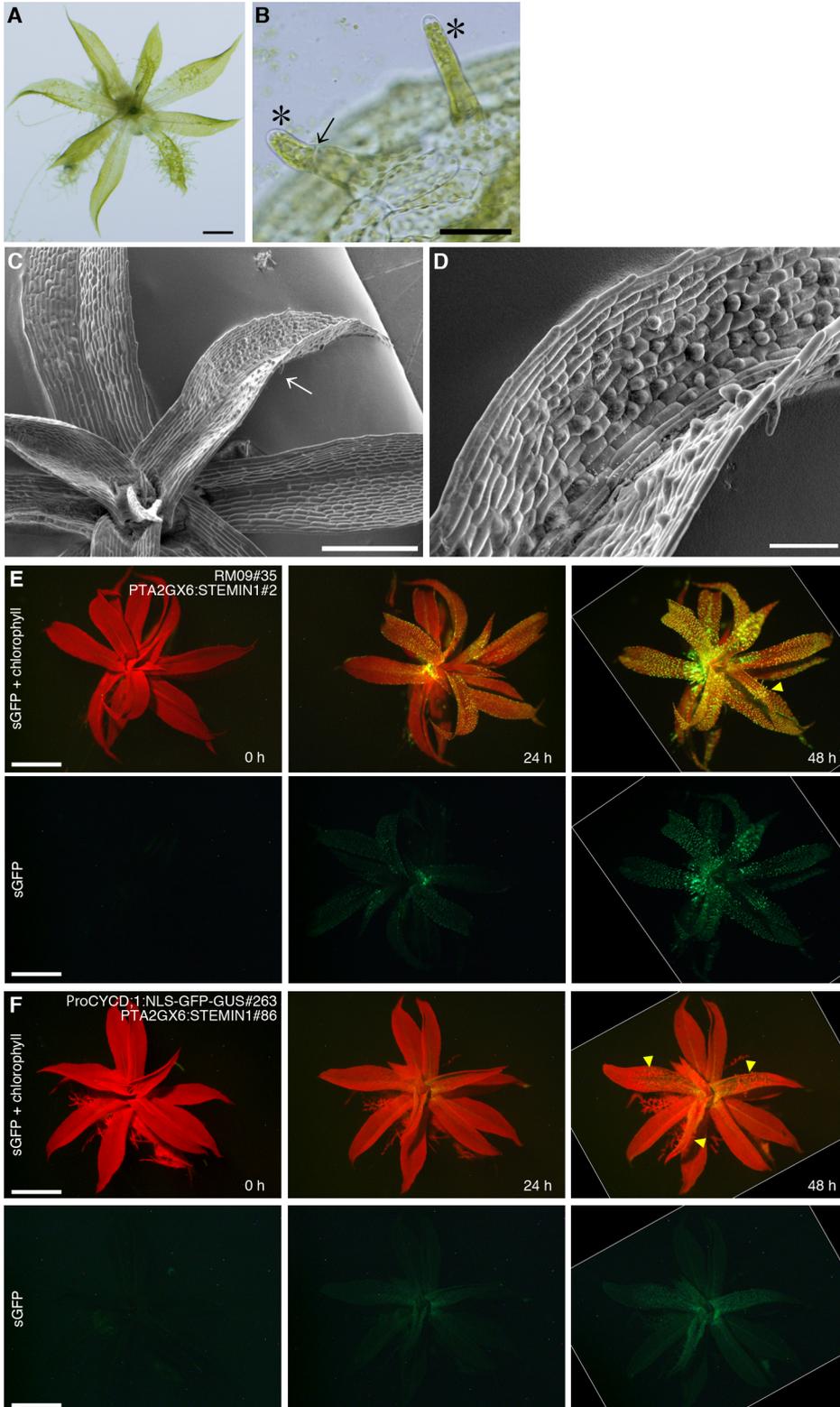
(A) A schematic showing the insertion of the  $\beta$ -estradiol-inducible *STEMIN1* expression construct (LGZ1-*STEMIN1*) into the putative neutral site PTA1. Blue and purple arrows denote a connected DNA fragment of the LexA operator and minimal 35S promoter (LexAop:m35S) (Zuo et al., 2000) and the zeocin resistance cassette (*zeo*) (Sakakibara et al., 2008), respectively. Light blue, gray, and green boxes designate DNA fragments encoding *STEMIN1*, the pea *rbcS3A* terminator (*pea3A-ter*) (Fluhr et al., 1986), and *loxP* (Odell et al., 1990), respectively. The probe used in (B) is indicated. GX6:NGG#63 (Kubo et al., 2013) was used as a host strain, in which a fusion protein of the SV40 nuclear localization signal (NLS) (Kalderon et al., 1984), sGFP (Chiu et al., 1996), and  $\beta$ -glucuronidase (*GUS*) (Jefferson et al., 1987) (NLS-sGFP-*GUS* [NGG]) is induced in response to  $\beta$ -estradiol (13). The LGZ1-*STEMIN1* construct was introduced into GX6:NGG#63 (GX6:NGG>*STEMIN1*) and the resulting plants express both NLS-sGFP-*GUS* (NGG) as a marker and *STEMIN1* proteins in the same cells in response to  $\beta$ -estradiol.

(B) DNA gel-blot analysis of targeted plants. Genomic DNA of the GX6:NGG>*STEMIN1* plants (#16, #23, #95, #258, and #286) and GX6:NGG#63 plants was digested with *EcoT22I*. An asterisk and an arrowhead indicate DNA fragments specific to the host strain GX6:NGG#63 and GX6:NGG>*STEMIN1* plants, respectively.

(C) Accumulation of *STEMIN1* transcripts in gametophores with the XVE-inducible system. Three-week-old gametophores of the GX6:NGG#63 and GX6:NGG>*STEMIN1* (#23 and #286) plants were incubated without or with 1  $\mu$ M  $\beta$ -estradiol for 24 hours. Total RNA was purified from the gametophores and subjected to qRT-PCR analysis. Transcript levels were normalized with respect to the *TUAI* transcript (Kubo et al., 2013) and the value for that transcript at 0 hour (i.e., without  $\beta$ -estradiol) in the GX6:NGG#63 line was taken as 1.0. Error bars indicate SE of the mean ( $n = 3$ ).

(D) Percentage of gametophores with at least one cell acquiring tip growth in the GX6:NGG#63 and GX6:NGG>*STEMIN1* (#23 and #286) plants 5 days after induction without or with 1  $\mu$ M  $\beta$ -estradiol, as an indicator of reprogramming ( $n > 50$ ). Three independent experiments were performed.

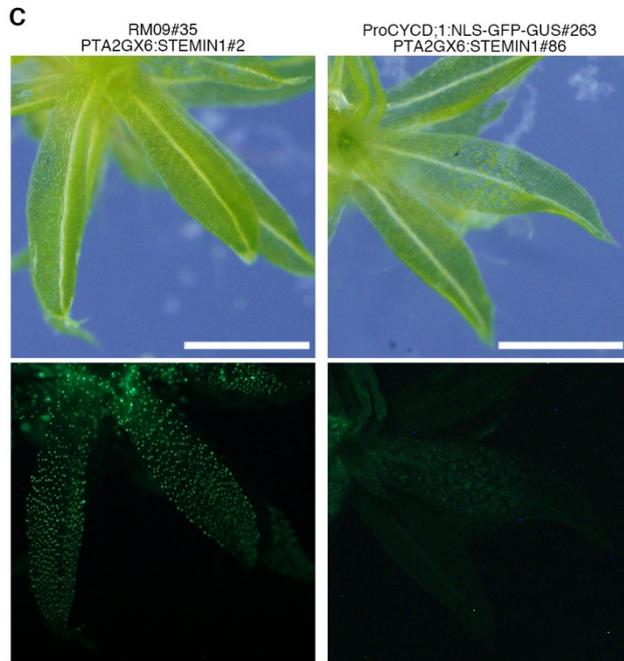
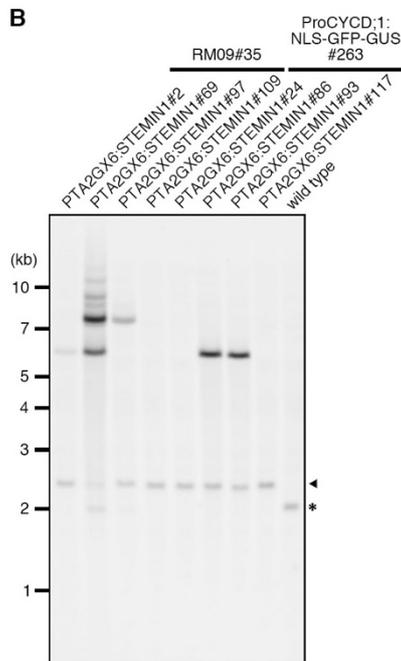
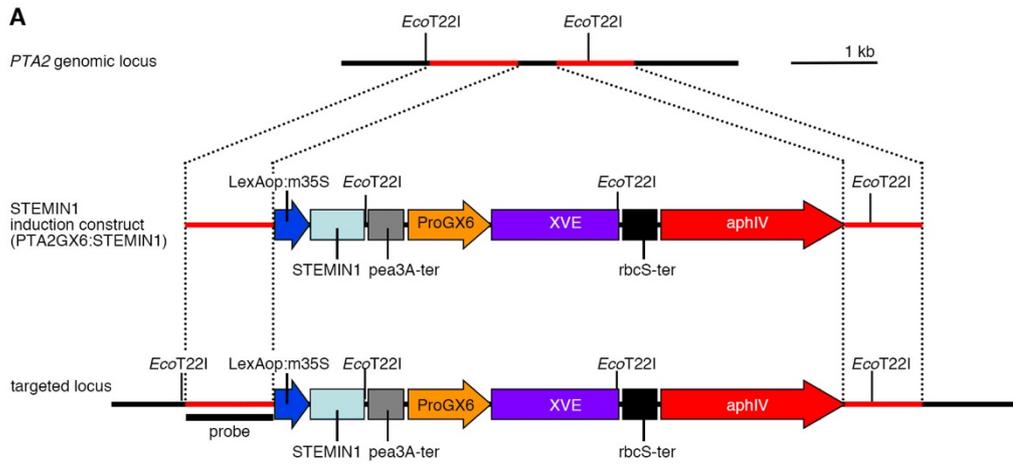
(E) Induction of *STEMIN1* in gametophores of GX6:NGG>*STEMIN1*#286 for 5 days. Scale bar, 0.5 mm.



**Fig. 4. Stem cell induction by STEMIN1 expression.**

(A to D) Induction of *STEMIN1* in gametophores of GX6:NGG>STEMIN1#23 for 5 (A and B) and 3 days (C and D). In (B), asterisks and an arrow represent chloronema apical stem cells and a newly formed cell septum, respectively. (C and D) Scanning electron micrographs of the STEMIN1-expressing gametophore. A magnification of the leaf indicated by an arrow in (C) is shown in (D).

(E and F) Induction of *STMEINI* in RM09#35 (E) and ProCYCD;1:NLS-sGFP-GUS#263 plants (F). Autofluorescence of chlorophyll (red) and sGFP fluorescence (green) images was recorded at 0, 24, and 48 hours after the STEMIN1 induction. Yellow arrowheads indicate leaves with protruded cells. Scale bars: 500  $\mu\text{m}$  in (A), and (C); 100  $\mu\text{m}$  in (D); 50  $\mu\text{m}$  in (B); 1 mm in (E) and (F).

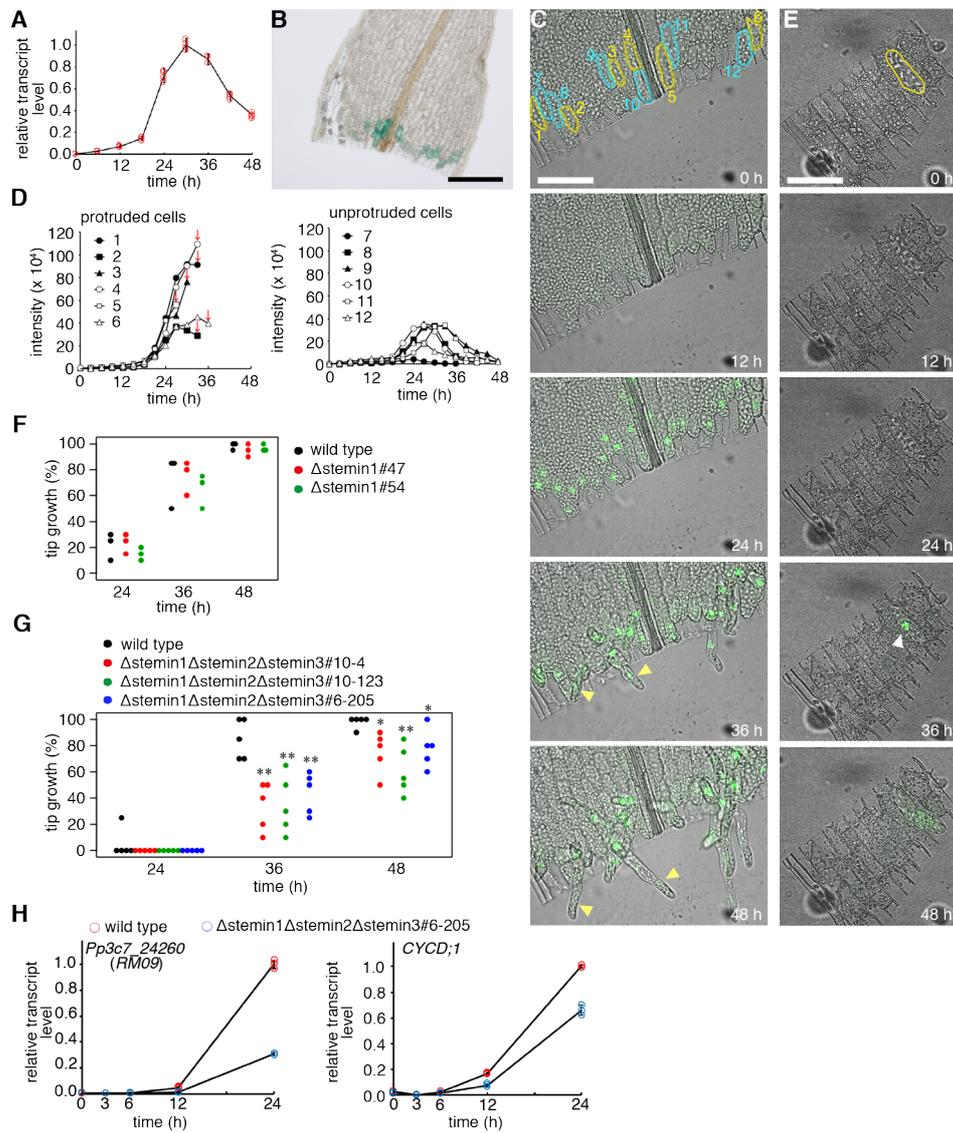


**Fig. 5. Construction of RM09 PTA2GX6:STEMIN1 and ProCYCD;1:NLS-sGFP-GUS PTA2GX6:STEMIN1 plants.**

(A) A schematic showing the insertion of PTA2GX6-STEMIN1 construct into the putative neutral site PTA2. Blue, orange, and red arrows denote a connected DNA fragment of the LexA operator and minimal 35S promoter (LexAop:m35S), the putative promoter region of the *Pp3c8\_3020* gene (ProGX6) (Kubo et al., 2013), and the aminoglycoside phosphotransferase IV expression cassette (*aphIV*) (Rensing et al., 2008), respectively. Light blue, gray, and black boxes designate a DNA fragment encoding STEMIN1, pea *rbcS3A* terminator (*pea3A-ter*) (Fluhr et al., 1986), and *rbcS* terminator (*rbcS-ter*) (Fluhr et al., 1986), respectively. The probe used in (B) is indicated. PTA2GX6-STEMIN1 construct was introduced into the host strain RM09#35 (Ishikawa et al., 2011) and ProCYCD;1:NLS-sGFP-GUS#263 (Ishikawa et al., 2011), respectively.

(B) DNA gel-blot analysis of targeted plants. Genomic DNA of the RM09#35 PTA2GX6:STEMIN1 (#2, #69, #95, #97, and #109) and ProCYCD;1:NLS-sGFP-GUS#263 PTA2GX6:STEMIN1 (#24, #86, #95, #93, and #117) plants was digested with *EcoT22I*. An asterisk and an arrowhead indicate DNA fragments specific to wild-type and transgenic plants, respectively.

(C) Magnified images of gametophores after induction of *STEMINI* in RM09#35 and CYCD;1pro:NGG#263 lines. Brightfield and sGFP fluorescence (green) images were recorded at 2 days after *STEMINI* induction. Scale bars, 1 mm.



**Fig. 6. STEMIN has a positive function in stem cell formation in cut leaves.**

(A) Accumulation of STEMIN1 transcripts in gametophore leaves after cutting. Values of transcripts at 30 hours in cut leaves was taken as 1.0. Error bars indicate SE of the mean (n = 4). Individual data are shown in circles.

(B) An excised leaf of the GUS-STEMIN1#53 plant at 30 hours with GUS-signals. Scale bar, 200  $\mu$ m.

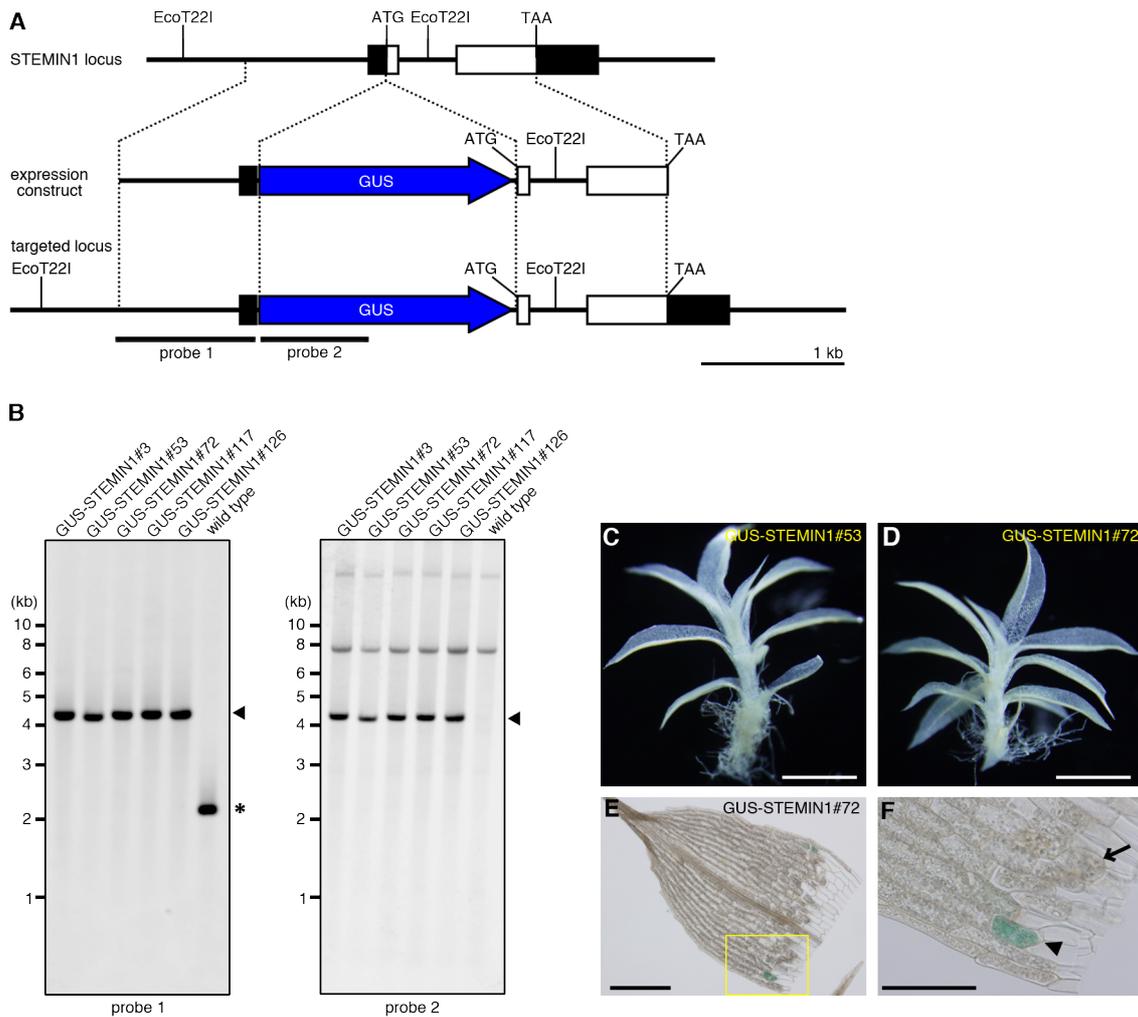
(C) STEMIN1 promoter activity in an excised leaf of STEMIN1pro:NGG#7 plant. Yellow arrowheads indicate protruded chloronema apical stem cells.

(D) Intensities of sGFP-signals in each cell facing the cut (1 to 12 in the top panel of [C]) were measured with ImageJ 4.0. Data for protruded and unprotruded leaf shells are shown on the left and right graphs, respectively. Red arrows indicate time points at which cells underwent cytokinesis.

(E) STEMIN1 promoter activity in an individual isolated leaf cell (outlined in yellow). An arrowhead indicates sGFP-signal before tip growth. Scale bars, 100  $\mu$ m (b,d).

(F,G) Percentages of excised leaves with at least one cell initiating tip growth in wild-type,  $\Delta$ stemin1, and  $\Delta$ stemin1 $\Delta$ stemin2 $\Delta$ stemin3 plants (n = 20). \*p < 0.05, \*\*p < 0.01 by a two-tailed Welch's t-test between wild type and mutant.

(H) Relative transcript levels of Pp3c7\_24260 (RM09) and CYCD;1 in excised leaves of wild-type and  $\Delta$ stemin1 $\Delta$ stemin2 $\Delta$ stemin3 plants after leaf cutting. All points have error bars indicating SE of the mean (n = 3) in (a,h) but some small error bars are hidden by the data points. Individual data are shown in circles.



**Fig. 7. Construction of GUS-STEMIN1 plants.**

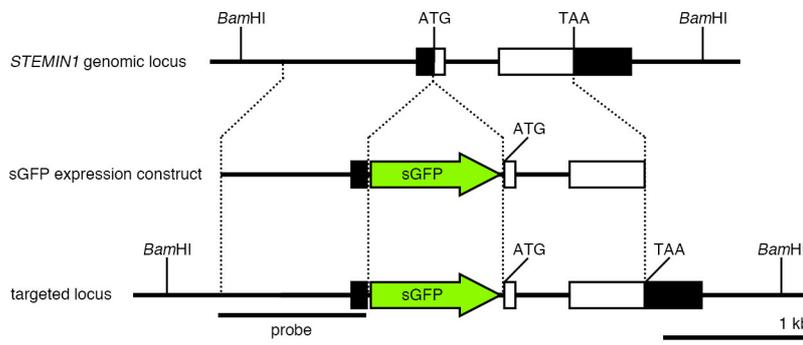
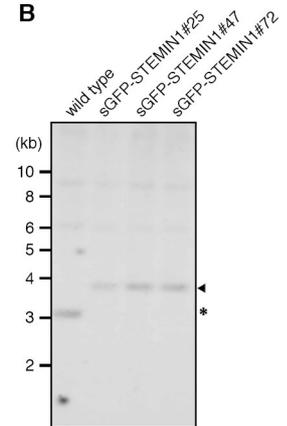
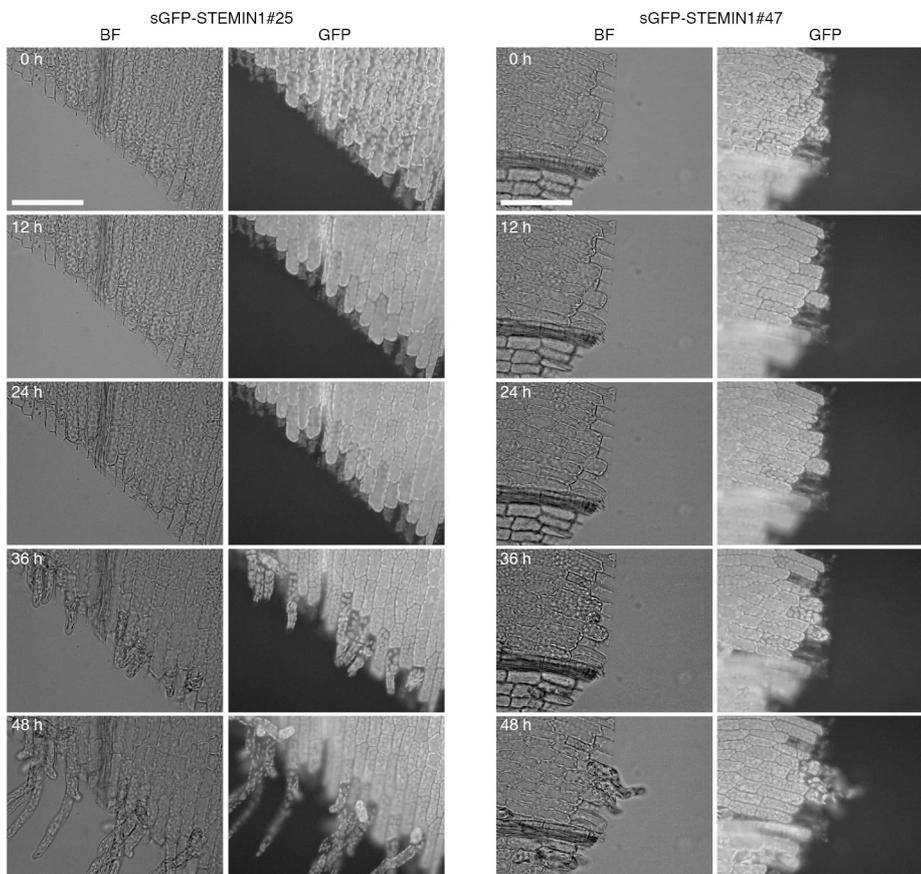
(A) Schematic showing the insertion of a GUS expression cassette into the *STEMIN1* locus. White boxes represent the *STEMIN1* coding regions. Closed boxes indicate the 5'- and 3'-untranslated regions. Blue arrow denotes the GUS expression cassette<sup>44</sup>. The probes used in B are indicated.

(B) DNA gel-blot analysis of targeted plants. Genomic DNA of wild type and GUS-STEMIN1 (#3, #53, #72, #117, and #126) was digested with *EcoT221*.

(C,D) GUS activity in a 4-week gametophores of GUS-STEMIN1#53 and #72 plants.

(E,F) Excised leaves of the #72 plant at 30 hours with GUS signals. F, The cut edge indicated by a yellow box in (E) is magnified. An arrow and an arrowhead indicate a GUS-positive cell and a chloronema apical stem cell, respectively.

Scale bar, 500  $\mu\text{m}$  (C,D), 200  $\mu\text{m}$  (E) and 100  $\mu\text{m}$  (F).

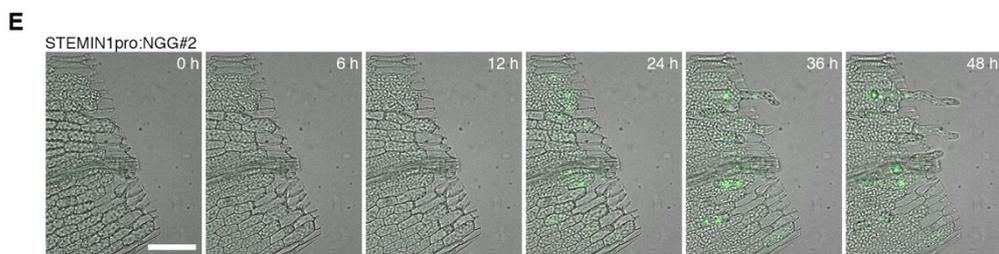
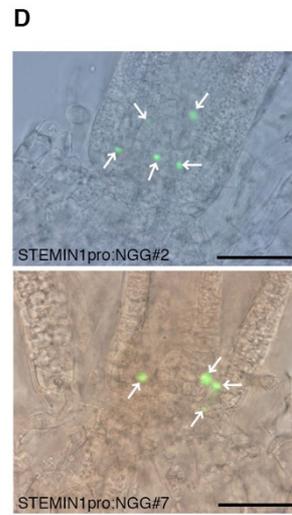
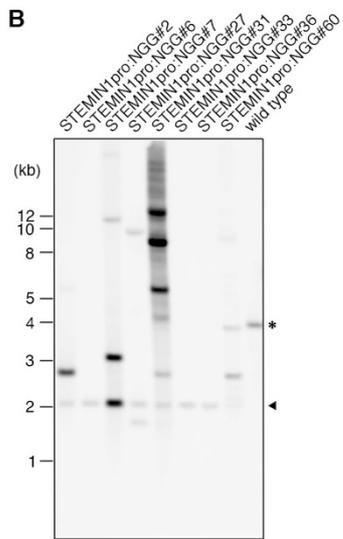
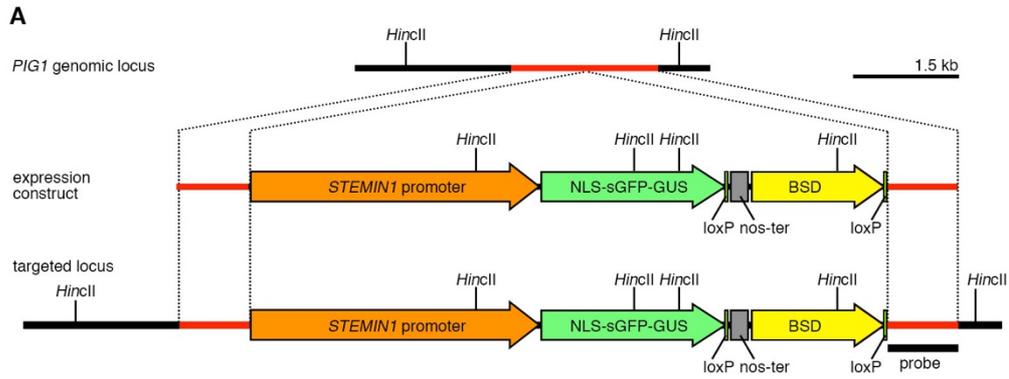
**A****B****C**

**Fig. 8. Construction of sGFP-STEMIN1 plants.**

(A) A schematic showing the insertion of a sGFP expression construct into the *STEMIN1* locus. White boxes represent the *STEMIN1* coding regions. Closed boxes indicate the 5' and 3' untranslated regions. Light green arrows denote *sGFP* (Chiu et al., 1996). The probe used in (B) is indicated.

(B) DNA gel-blot analysis of targeted plants. Genomic DNA of wild-type and sGFP-STEMIN1 (#25, #47, and #72) plants was digested with *Bam*HI. An asterisk and an arrowhead indicate DNA fragments specific to wild-type and transgenic plants, respectively.

(C) Brightfield and sGFP-filter images of an excised leaf of sGFP-STEMIN1#25 and sGFP-STEMIN1#47 at 0, 12, 24, 36, and 48 hours after cutting. sGFP signal was not detected even under exposure conditions in which chlorophyll autofluorescence was recorded in the band pass sGFP-filtered images. Scale bars, 100  $\mu$ m.



**Fig. 9. Construction of the STEMIN1pro:NLS-sGFP-GUS (STEMIN1pro:NGG) plants.**

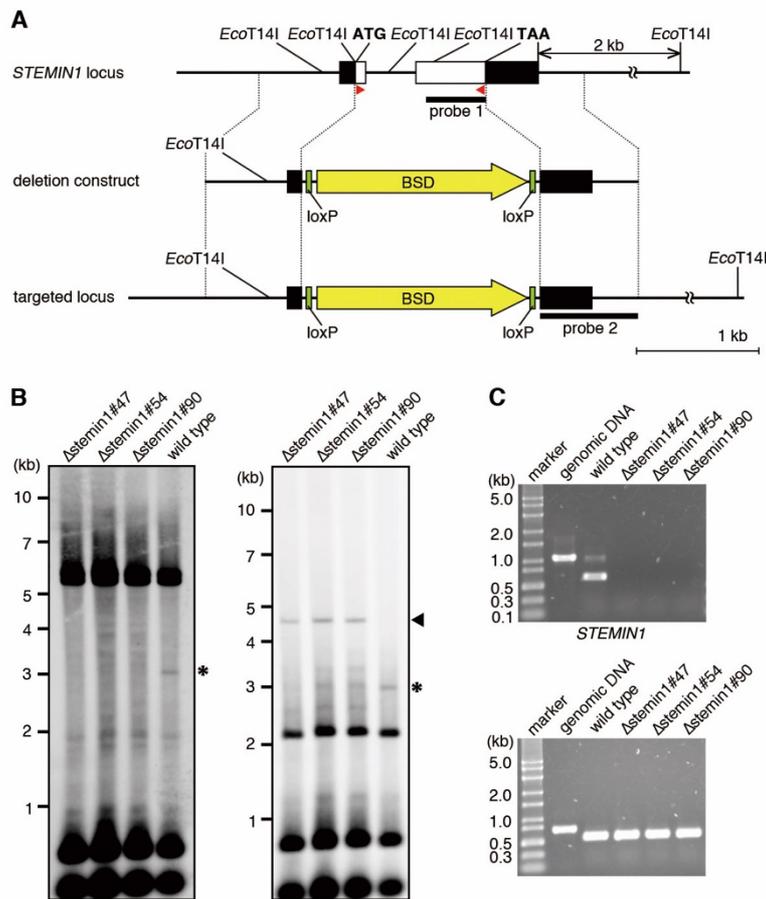
(A) A schematic showing the insertion of the STEMIN1pro:NLS-sGFP-GUS construct into the putative neutral site PIG1. Arrows denote the 3-kb *STEMIN1* promoter sequence (orange), a connected DNA fragment of a synthetic nucleotide sequence encoding the SV40 nuclear localization signal (Kalderon et al., 1984), *sGFP* gene (Chiu et al., 1996), and *uidA* gene (GUS) (Jefferson et al., 1987) (green), and the BSD expression cassette (Tamura et al., 1995) (yellow). Green and gray boxes denote loxP (Odell et al., 1990) and the terminator of the nopaline synthase gene (nos-ter) (Nishiyama et al., 2000), respectively. The probe used in (B) is indicated.

(B) DNA gel-blot analysis of targeted plants. Genomic DNA of wild-type and STEMIN1pro:NGG (#2, #6, #7, #27, #31, #33, #36, and #60) plants was digested with *HincII*. An asterisk and an arrowhead indicate DNA fragments specific to wild-type and transgenic plants, respectively.

(C) GUS activity in a 4-week gametophore of the STEMIN1pro:NGG (#2 and #7) plants. Arrows indicate GUS signals in shoot apices. A shoot apex of STEMIN1pro:NGG#2 is shown at higher magnification in the inset. Scale bars, 500  $\mu\text{m}$ ; 50  $\mu\text{m}$  in the inset.

(D) Brightfield and sGFP fluorescence images in shoot apices of STEMIN1pro:NGG#2 and STEMIN1pro:NGG #7 plants. Arrows indicate sGFP signals in axillary hairs. Scale bars, 50  $\mu\text{m}$ .

(E) *STEMIN1* promoter activity in an excised leaf of a STEMIN1pro:NGG#2 plant. sGFP signals of STEMIN1pro:NGG#2, as well as those of a STEMIN1pro:NGG#7 plant (Fig. 6C), were observed at 24 hours. Scale bar, 100  $\mu\text{m}$ .

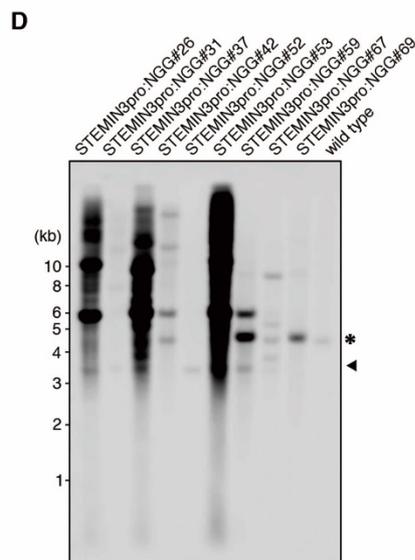
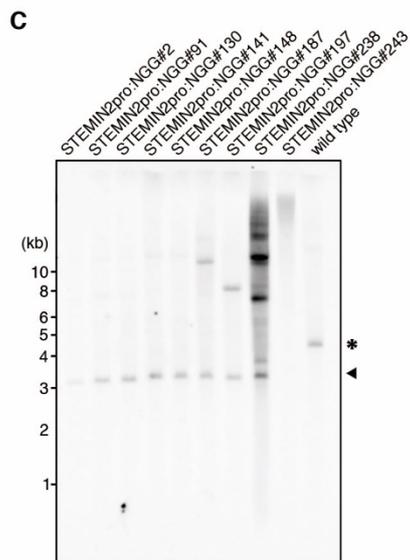
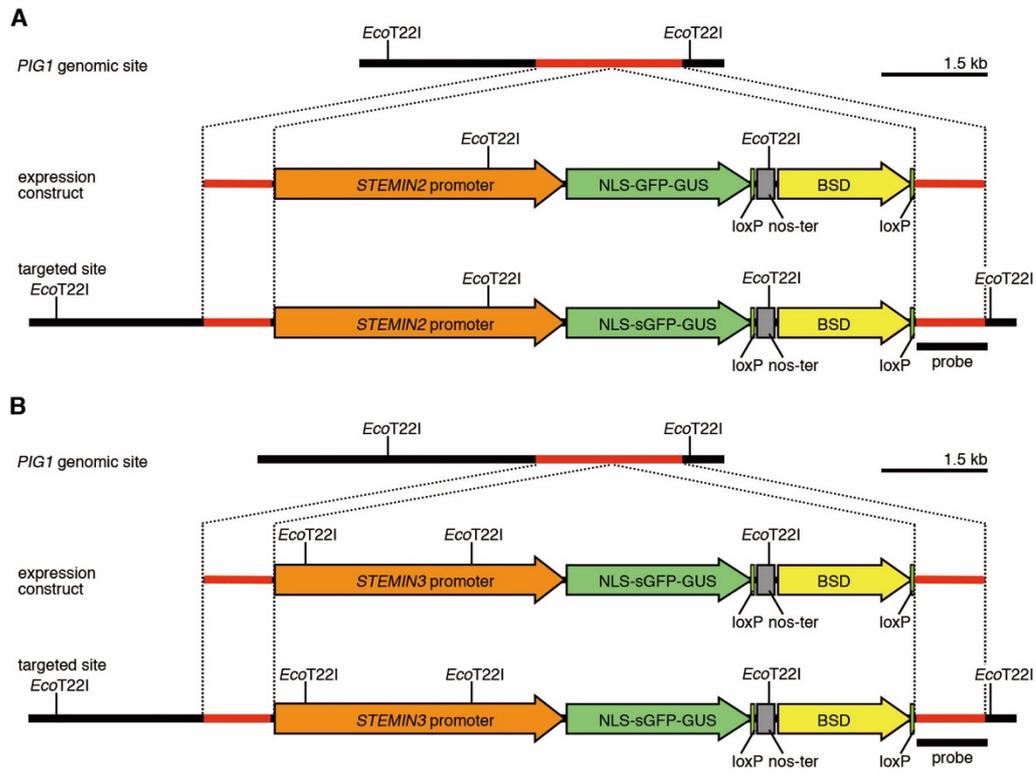


**Fig. 10. Construction of  $\Delta$ stemin1 deletion plants.**

(A) A schematic of construct targeting *STEMIN1* locus. White boxes represent the *STEMIN1* coding region. Black boxes indicate the 5' and 3' untranslated regions. Yellow arrow denotes the BSD expression cassette (Tamura et al., 1995). Green boxes denote loxP (Odell et al., 1990). Probes in (B) are indicated. Primers in (C) are indicated by red arrowheads.

(B) DNA gel-blot analyses of targeted plants with probes indicated in (A). Genomic DNA of wild-type and  $\Delta$ stemin1 (#47, #54, and #90) plants was digested with *EcoT14I*. Asterisks and arrowheads indicate DNA fragments specific to wild-type and  $\Delta$ stemin1 plants, respectively.

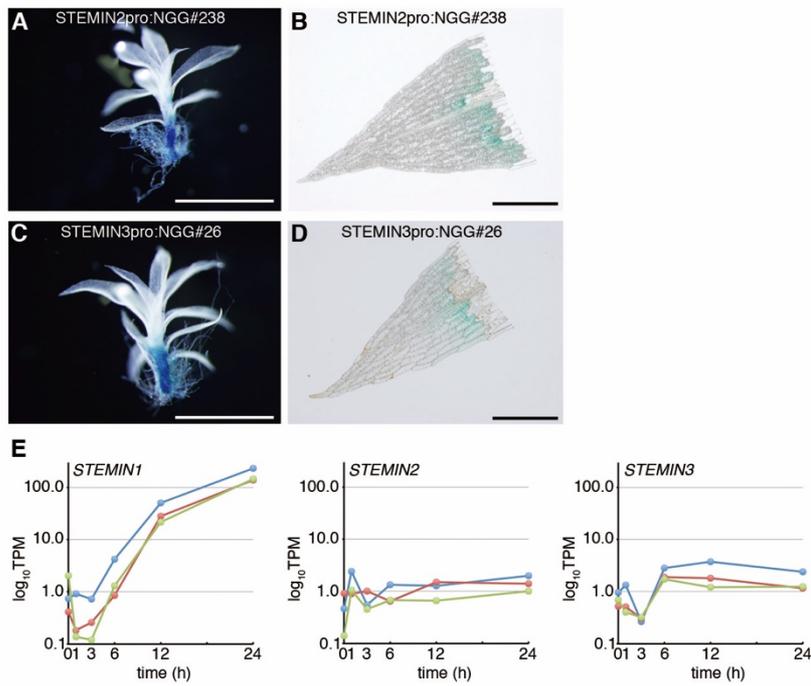
(C) RT-PCR analysis of  $\Delta$ stemin1 deletion plants. Total RNA from protonemata of wild-type and  $\Delta$ stemin1 (#47, #54, and #90) plants was purified and subjected to RT-PCR analysis with primers indicated in (A). Genomic DNA of the wild type was used as a control. Detection of the  $\alpha$ -tubulin transcript (*TUAI*) served as an internal control.



**Fig. 11. Construction of the STEMIN2pro:NLS-sGFP-GUS (STEMIN2pro:NGG) and STEMIN3pro:NGG plants.**

(A and B) Schematics for insertion of STEMIN2pro:NLS-sGFP-GUS (A) or STEMIN3pro:NLS-sGFP-GUS (B) construct into the putative neutral site PIG1. Arrows denote *STEMIN2* and *STEMIN3* promoters (orange), a connected DNA fragment of a synthetic nucleotide sequence encoding the SV40 nuclear localization signal (Riese et al., 2008), *sGFP* gene (Chiu et al., 1996), and *uidA* gene (GUS) (Hiratsu et al., 2003) (a green arrow), and the blasticidin S deaminase (BSD) expression cassette (Tamura et al., 1995) (a yellow arrow). Green and gray boxes denote loxP (Odell et al., 1990) and the terminator of nopaline synthase gene (*nos-ter*) (2), respectively. The probe used in (C) and (D) is indicated.

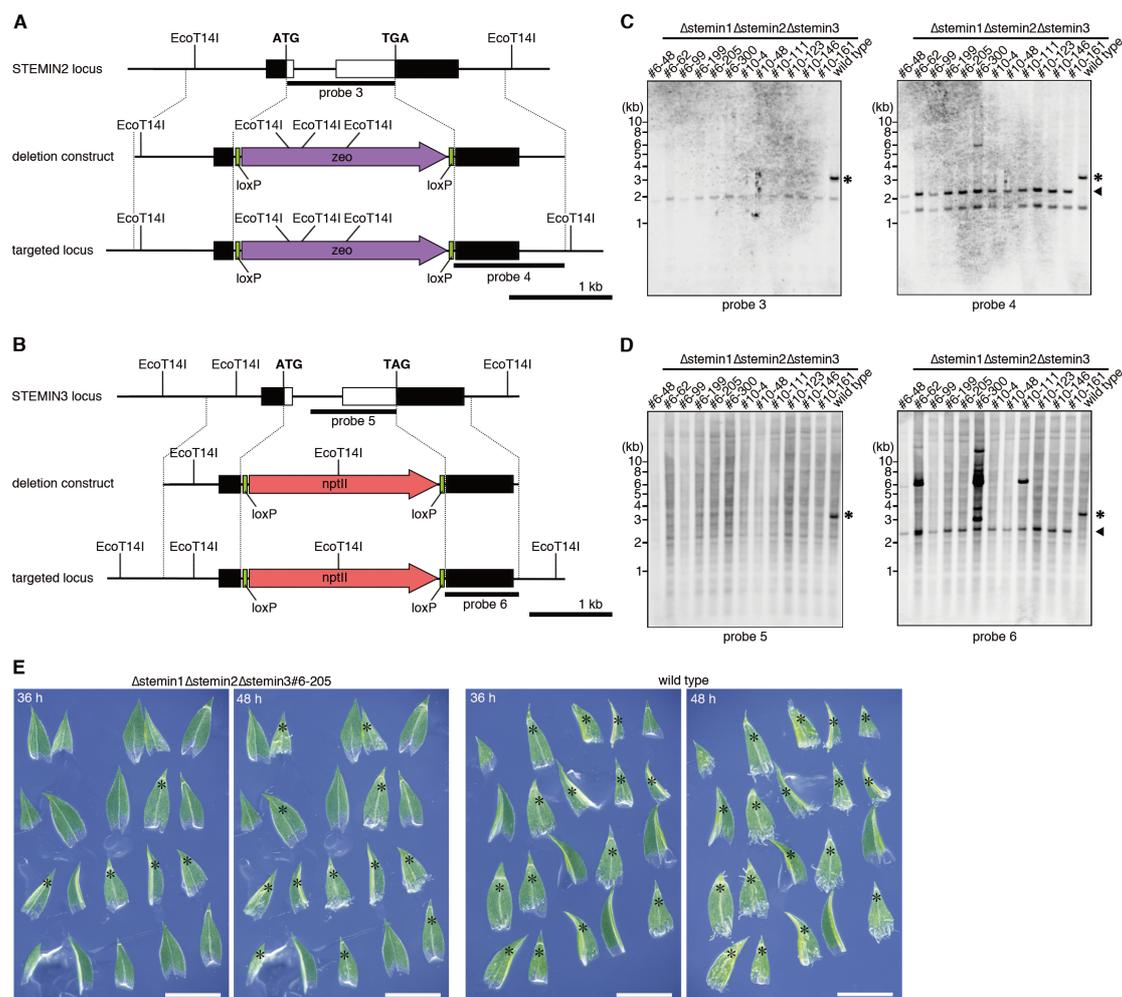
(C and D) DNA gel-blot analysis of targeted plants. Genomic DNA of wild-type, STEMIN2pro:NGG (#2, #91, #130, #141, #148, #187, #197, #238, and #243; [C]), and STEMIN3pro:NGG (#26, #31, #37, #42, #52, #53, #59 #67, and #69; [D]) plants was digested with *EcoT22I*. Asterisks and arrowheads indicate DNA fragments specific to wild-type and transgenic plants, respectively.



**Fig. 12. Promoter activities and transcript amounts of *STEMIN2* and *STEMIN3* in gametophores and excised leaves.**

(A to D) GUS activity in 3-week-old gametophores (A, C) and excised leaves of STEMIN2pro:NGG#238 and STEMIN3pro:NGG#26 plants. Leaves were fixed at 24 hours after cutting (B, D). Scale bars, 1 mm (A, C) and 200  $\mu$ m (B, D).

(E) Relative *STEMIN1*, *STEMIN2*, and *STEMIN3* transcript levels (tags per million, TPM) in cut leaves. The levels were determined by 5' digital gene expression analysis (Nishiyama et al., 2012). Results of three independent experiments are shown.

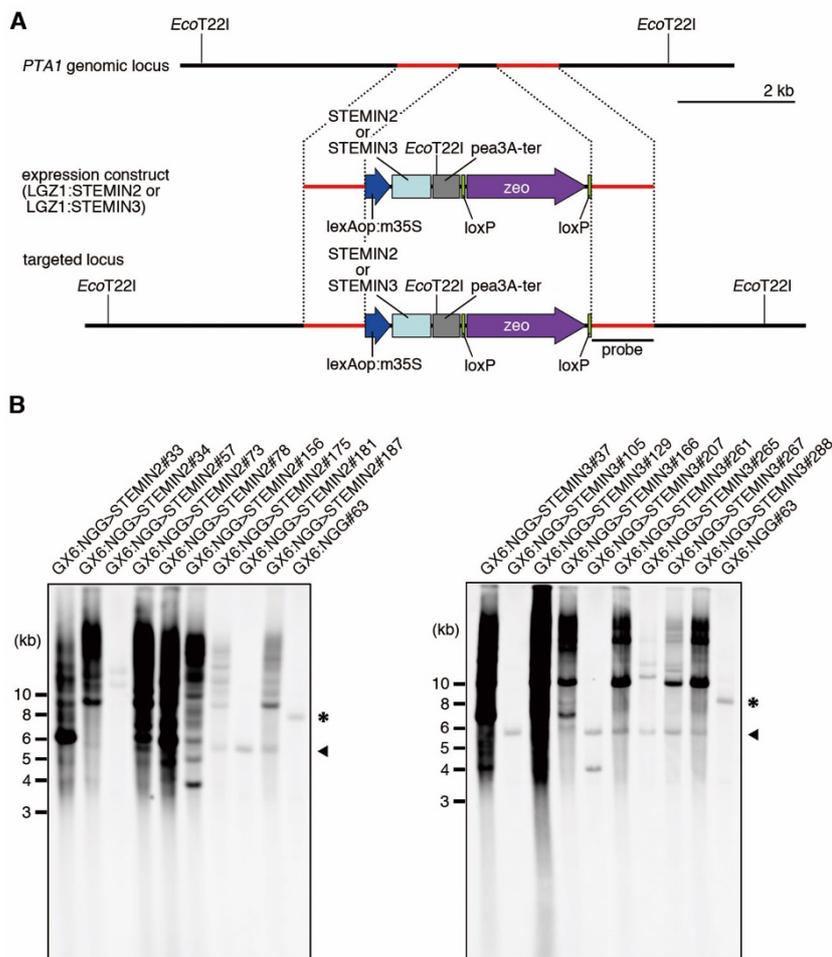


**Fig. 13. Construction of the  $\Delta$ stemin1 $\Delta$ stemin2 $\Delta$ stemin3 triple deletion plants.**

(A and B) Schematics of constructs targeting the *STEMIN2* (A) and *STEMIN3* (B) loci. White boxes represent the *STEMIN2* and *STEMIN3* coding regions. Black boxes indicate the 5' and 3' untranslated regions. Purple and red arrows denote the zeocin resistance cassette (*zeo*) (Sakakibara et al., 2008) and the neomycin phosphotransferase II expression cassette (*nptII*) (Nishiyama et al., 2000), respectively. Green boxes denote loxP (Odell et al., 1990). Probes used in (C) and (D) are indicated. Procedures to make triple deletion mutants are described in Materials and Methods.

(C and D) DNA gel-blot analyses of targeted plants. Genomic DNAs of wild-type and  $\Delta$ stemin1 $\Delta$ stemin2 $\Delta$ stemin3 (#6-48, #6-62, #6-99, #6-199, #6-205, #6-300, #10-4, #10-48, #10-111, #10-123, #10-146, and #10-161) plants were digested with *EcoT14I*. Asterisks and arrowheads indicate DNA fragments specific to wild-type and  $\Delta$ stemin1 plants, respectively.

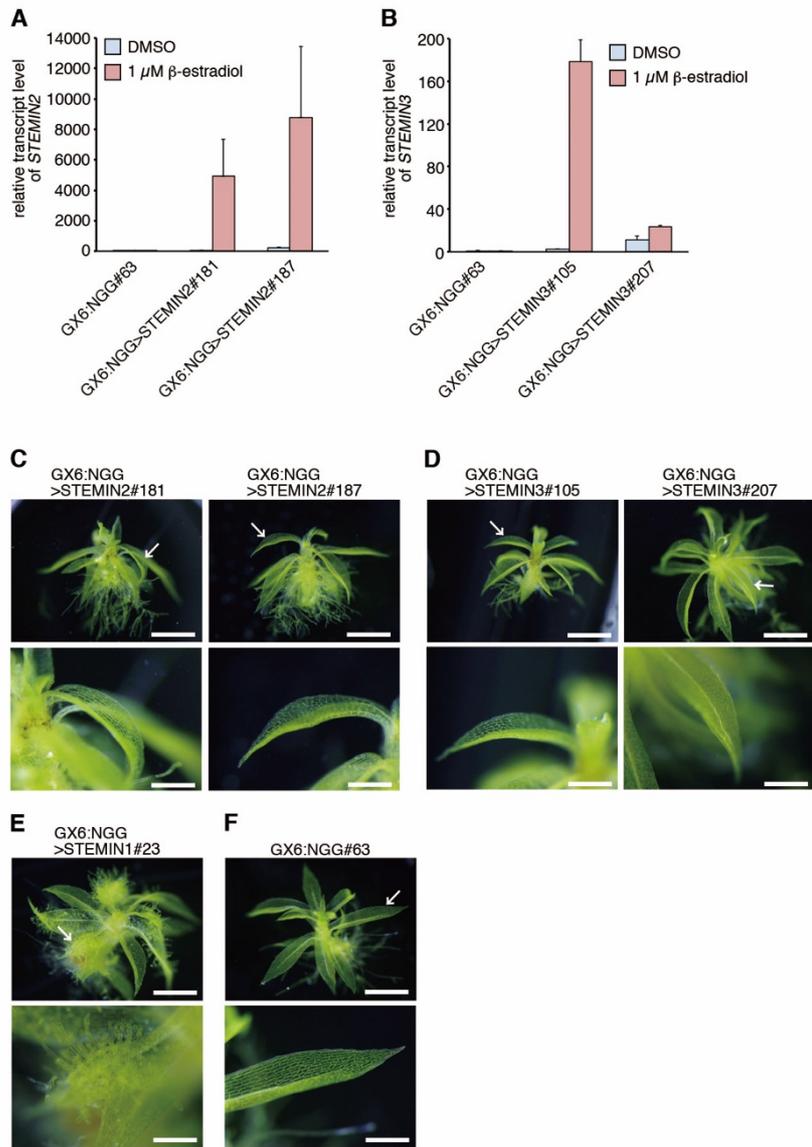
(E) Excised leaves of  $\Delta$ stemin1 $\Delta$ stemin2 $\Delta$ stemin3 #6-205 and wild-type plants at 36 and 48 hours after excision. Asterisks indicate leaves with protruded cells. Scale bars, 1 mm.



**Fig. 14. Construction of GX6:NGG>STEMIN2 and GX6:NGG>STEMIN3 plants.**

(A) A schematic showing insertion of an LGZ1-STEMIN2 or LGZ1-STEMIN3 construct into the putative neutral site PTA1. A blue arrow denotes a connected DNA fragment of the LexA operator and minimal 35S promoter (*lexAop:m35S*) (Zuo et al., 2000). A purple arrow denotes the zeocin resistance cassette (*zeo*) (Sakakibara et al., 2008). Light blue, gray, and green boxes designate the *STEMIN2* or *STEMIN3* gene, *pea rbcS3A* terminator (*pea3A-ter*) (Fluhr et al., 1986), and *loxP* (Odell et al., 1990), respectively. The probe used in (B) is indicated. The LGZ1-STEMIN2 or LGZ1-STEMIN3 construct was introduced into the GX6:NGG#63 line (Kubo et al., 2013).

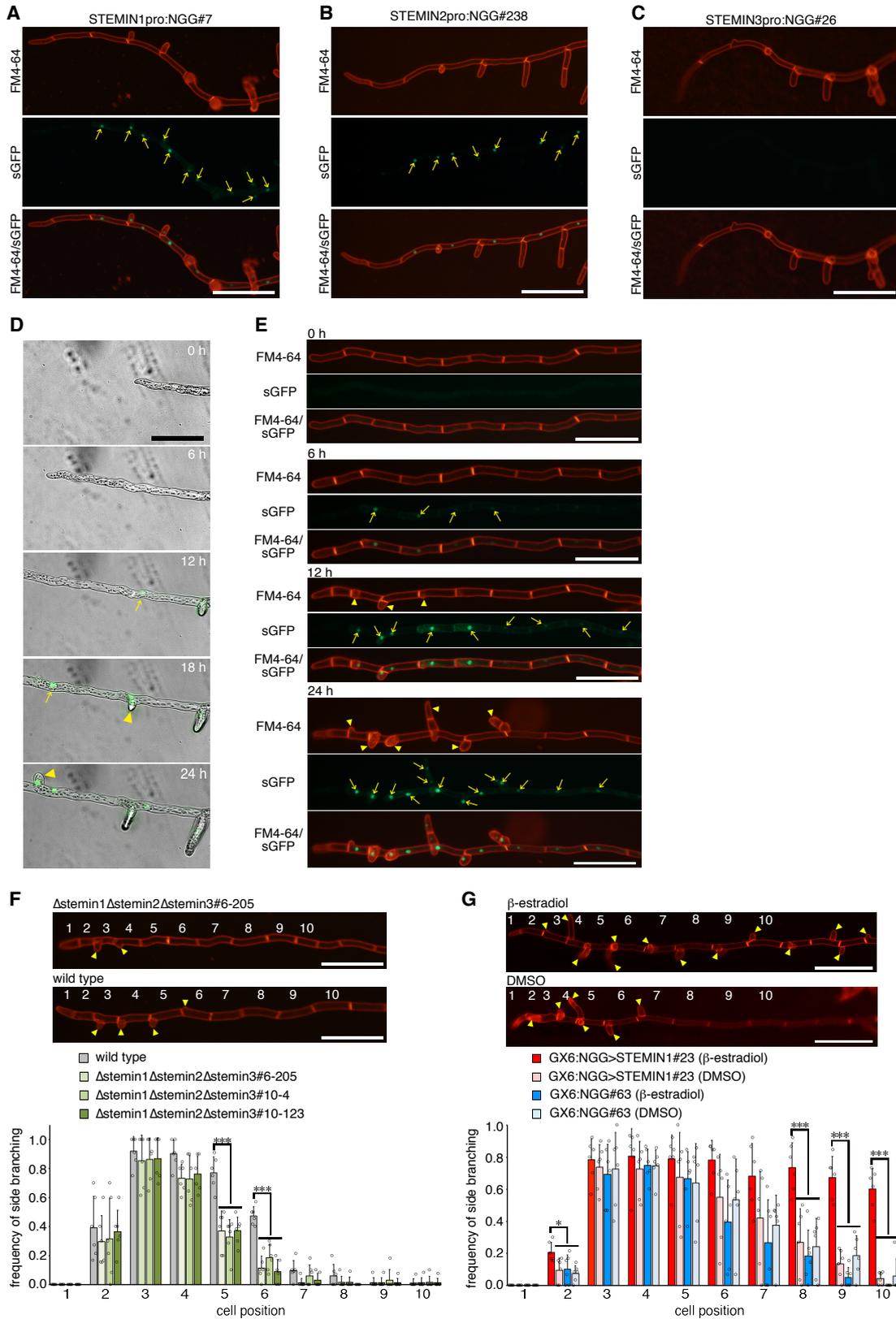
(B) DNA gel-blot analysis of targeted plants. Genomic DNA of the GX6:NGG>STEMIN2 (#33, #34, #57, #73, #78, #156, #175, #181, and #187), GX6:NGG>STEMIN3 (#37, #105, #129, #166, #207, #261, #265, #267, and #288), and GX6:NGG#63 plants was digested with *EcoT22I*. Asterisks and arrowheads indicate DNA fragments specific to wild-type and transgenic plants, respectively.



**Fig. 15. Induction of *STEMIN2* or *STEMIN3* in intact gametophores.**

(A and B) Accumulation of *STEMIN2* (A) or *STEMIN3* (B) transcripts in gametophores harboring the XVE-inducible system. Four-week-old gametophores of the GX6:NGG#63, GX6:NGG>*STEMIN2* (#181 and #187), and GX6:NGG>*STEMIN3* (#105 and #207) plants were incubated with 1  $\mu$ M  $\beta$ -estradiol for 24 hours. Total RNA was purified from the gametophores and subjected to qRT-PCR analysis. Transcript levels were normalized with *TUA1* transcript and the value of transcripts at 0 hours (without  $\beta$ -estradiol) in the GX6:NGG#63 line was taken as 1.0. Error bars indicate SE of the mean ( $n = 3$ ).

(C to F) Three-week-old gametophores of the GX6:NGG>*STEMIN2* (#181 and #187; [C]) and GX6:NGG>*STEMIN3* (#105 and #207; [D]), GX6:NGG>*STEMIN1*#23 (E), and GX6:NGG#63 (F) plants were incubated with 1  $\mu$ M  $\beta$ -estradiol for 8 days. Leaves indicated by arrows in the upper panels are magnified in the lower panels. Scale bars, 1 mm (upper panels) and 200  $\mu$ m (lower panels).



**Fig. 16. STEMIN1 induces reprogramming of protonema cells.**

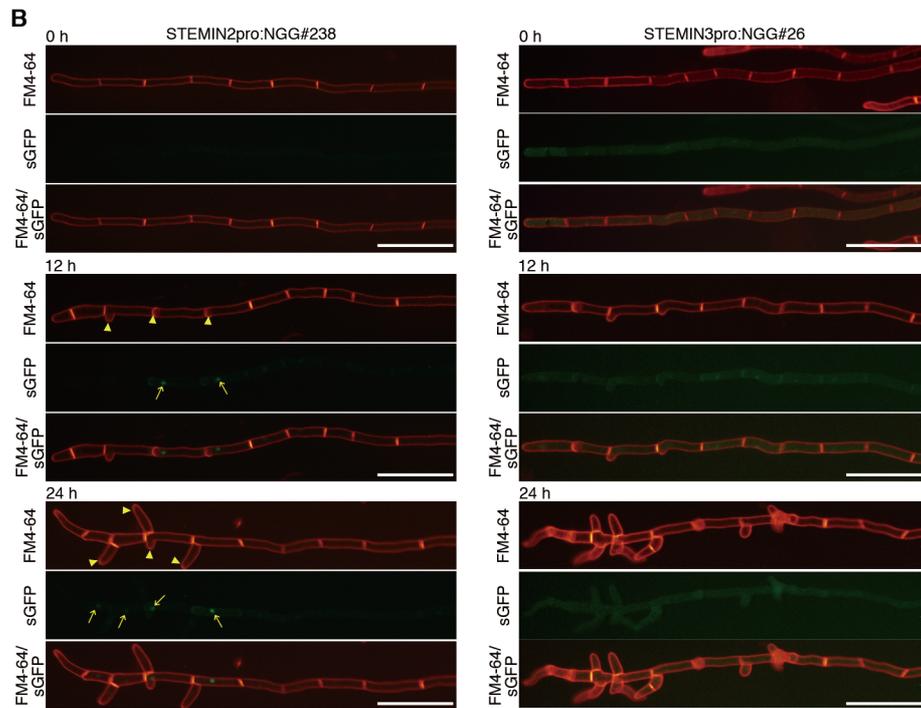
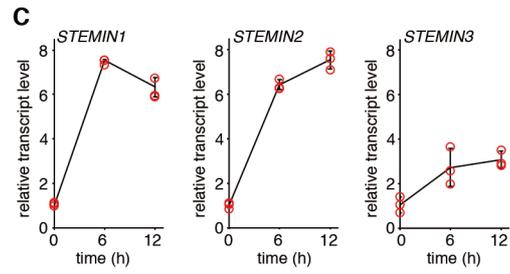
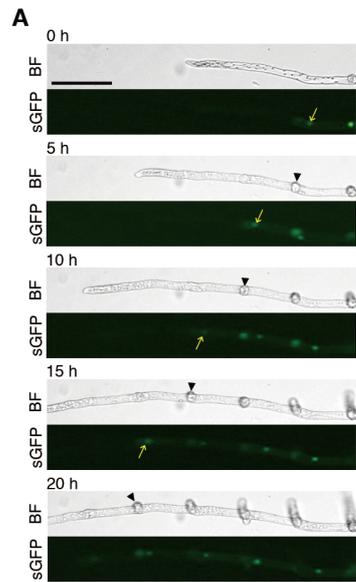
(A to C) sGFP signals in chloronemata of STEMIN1pro:NGG#7 (A), STEMIN2pro:NGG#238 (B), and STEMIN3pro:NGG#26 (C) plants grown under white light conditions.

(D) sGFP signals in growing chloronemata of STEMIN1pro:NGG#7 under white light conditions. Arrows and arrowheads indicate representative sGFP signals and side-branch initial cells, respectively.

(E) sGFP signals in protonemata of STEMIN1pro:NGG#7 after transfer from unilateral red light to white light conditions. Arrows indicate sGFP-positive cells. Time after irradiation with white light is indicated.

(F) Frequencies of side branches at each cell position of protonemata of wild-type and  $\Delta\text{stemin1}\Delta\text{stemin2}\Delta\text{stemin3}$  plants ( $n = 10-15$  protonemata) 12 hours after they were moved from unilateral red light to white light conditions. Representative chloronemata of wild-type and triple deletion mutant plants are shown. Arrowheads indicate side branches.

(G) Frequency of side branches at each cell position of GX6:NGG>STEMIN1#23 and GX6:NGG#63 lines with or without  $\beta$ -estradiol induction ( $n = 10-15$  protonemata). Representative chloronemata of the transgenic plants 24 hours after they were moved to white light conditions and supplied with either DMSO or  $\beta$ -estradiol are shown. Arrowheads indicate side branches.  $*p < 0.05$ ,  $**p < 0.001$  by two-tailed Dunnett's test for multiple comparisons. All scale bars, 100  $\mu\text{m}$ .

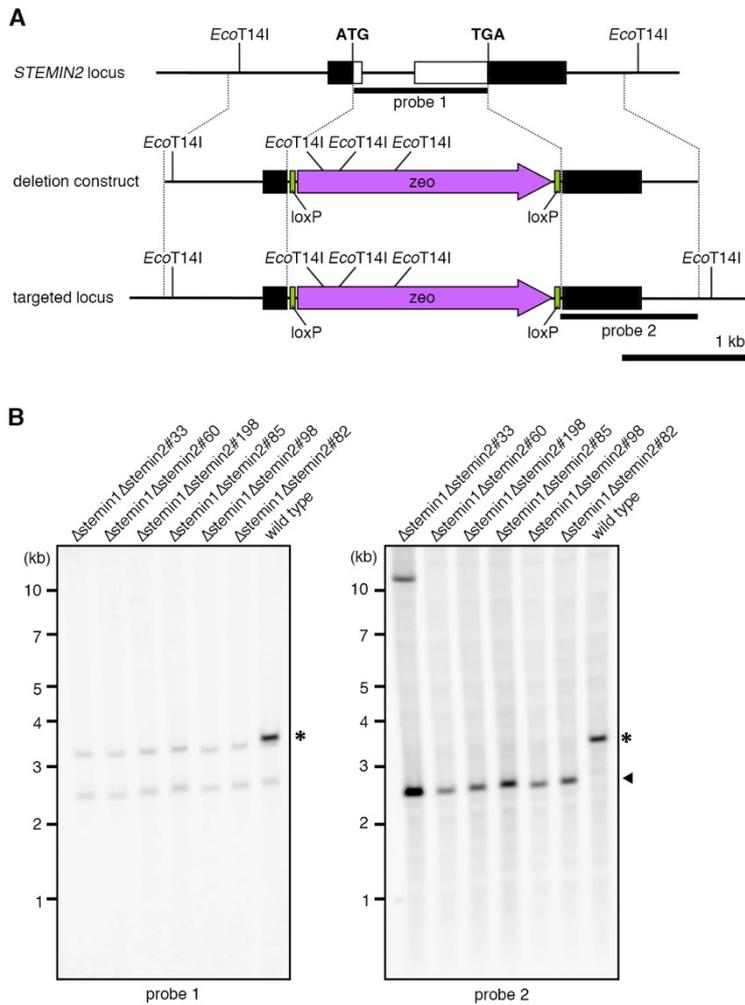


**Fig. 17. Promoter activities and transcript amounts of *STEMIN2* and *STEMIN3* in protonemata.**

(A) Activation of *STEMIN2* promoter in *STEMIN2*pro:NGG#238 line under white light conditions. Arrows and arrowheads indicate representative sGFP-signals and secondary chloronema apical stem cells, respectively.

(B) sGFP signals in chloronemata of *STEMIN2*pro:NGG#238 (left panels) and *STEMIN3*pro:NGG#26 (right panels) after transfer from unilateral red light to white light conditions. Yellow arrows indicate sGFP-positive cells. Time after irradiation with white light is indicated. Scale bars, 100  $\mu$ m.

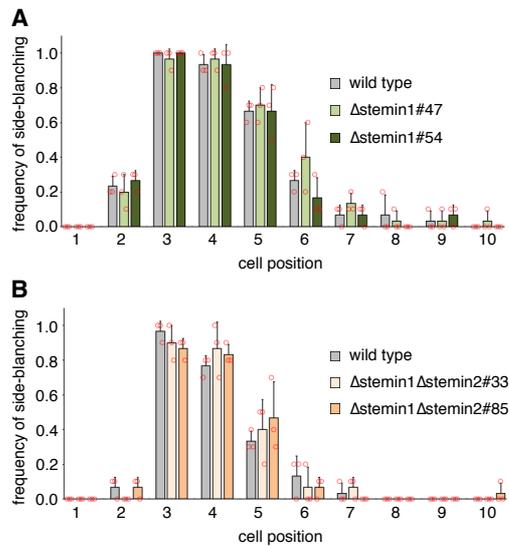
(C) Accumulation of transcripts of three *STEMIN* genes during side branch formation in chloronemata. Red-light-grown chloronemata of wild type were moved to white light conditions and collected at the indicated time points (0, 6, and 12 h). Error bars indicate SE of the mean (n = 3). Individual data are shown in circles.



**Fig. 18. Construction of  $\Delta$ stemin1 $\Delta$ stemin2 double deletion mutant plants.**

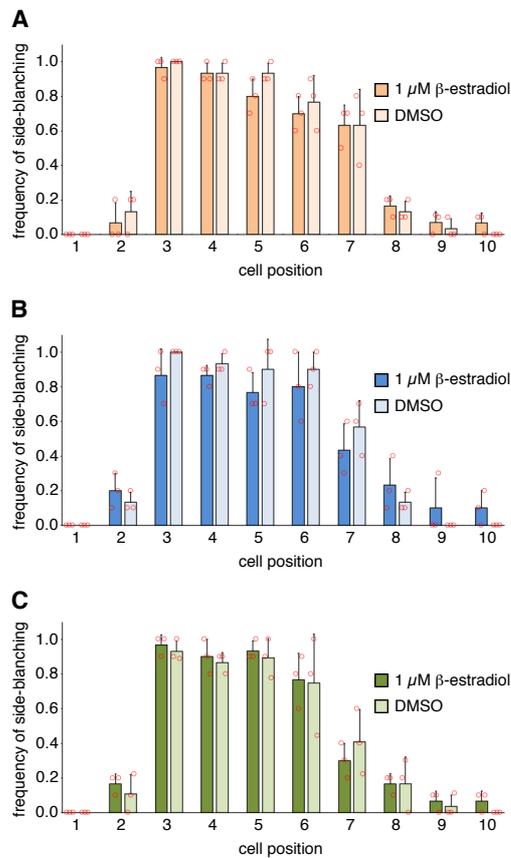
(A) A schematic of the construct targeting the *STEMIN2* locus. White boxes represent a *STEMIN2* coding region. Black boxes indicate 5' and 3' untranslated regions. Purple arrows denote zeocin resistant cassette (zeo) (Sakakibara et al., 2008). Green boxes denote loxP (Odell et al., 1990). Probes used in (B) are indicated.

(B) DNA gel-blot analyses of targeted plants. Genomic DNA of wild-type and  $\Delta$ stemin1 $\Delta$ stemin2 (#33, #60, #198, #85, #98, and #82) plants was digested with *EcoT14I*. Asterisks and arrowheads indicate DNA fragments specific to wild-type and transgenic plants, respectively.



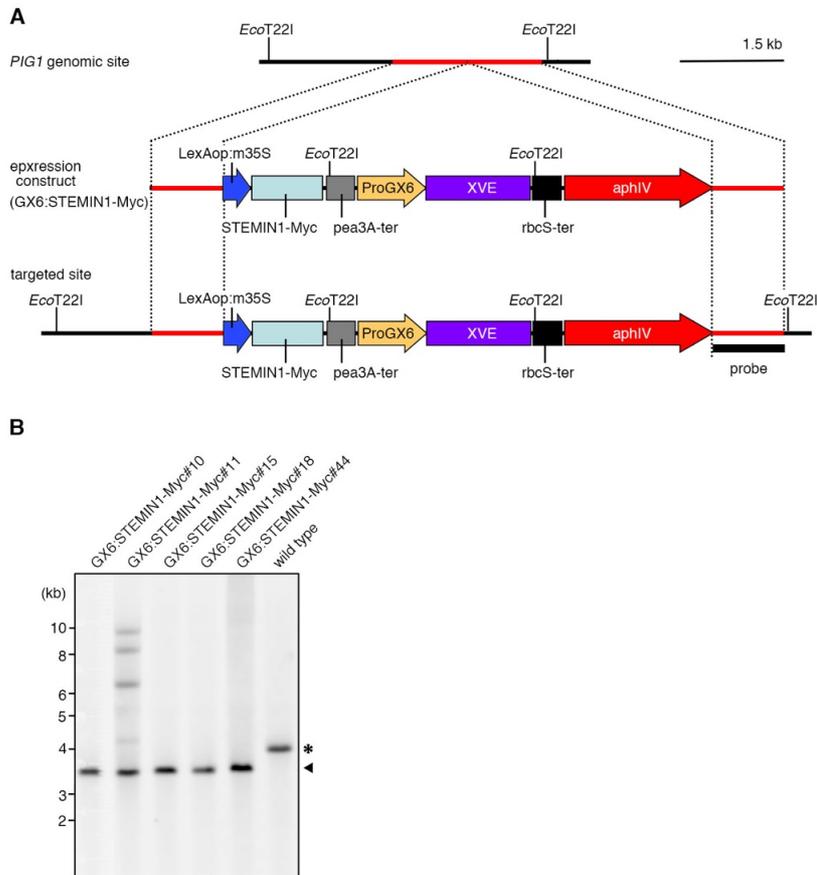
**Fig. 19. Side branching of the  $\Delta$ stemin1 $\Delta$ stemin2 double deletion mutant plants.**

(A and B) Frequency of side branching at each cell position of  $\Delta$ stemin1 single deletion plants (A) and  $\Delta$ stemin1 $\Delta$ stemin2 double deletion plants (B) ( $n = 10$  protonemata) 12 hours after they were moved from unilateral red light to white light conditions. Individual data of the frequencies are shown in circles. Center values and error bars indicate the mean of biological triplicates and SD, respectively. A two-tailed Dunnett test for multiple comparisons indicates no significant differences between the wild-type and the two deletion mutant plants at each cell position.  $p > 0.05$  (A, B).



**Fig. 20. Induction of *STEMIN2* and *STEMIN3* in red-light-grown protonemata.**

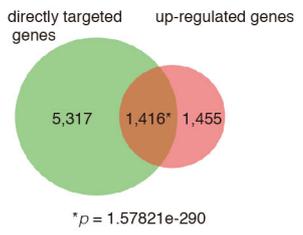
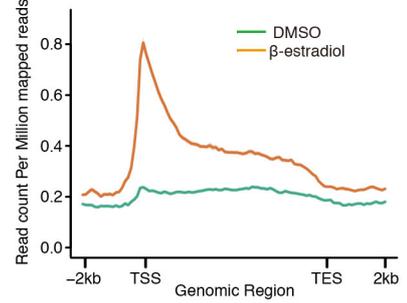
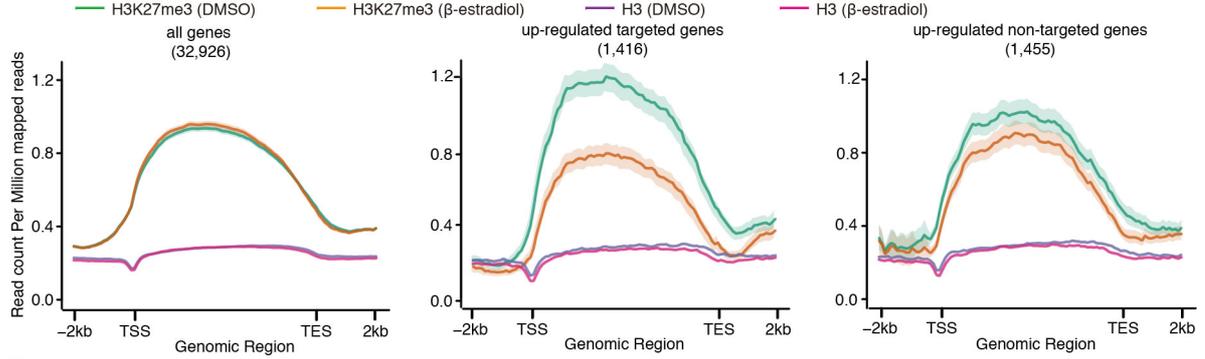
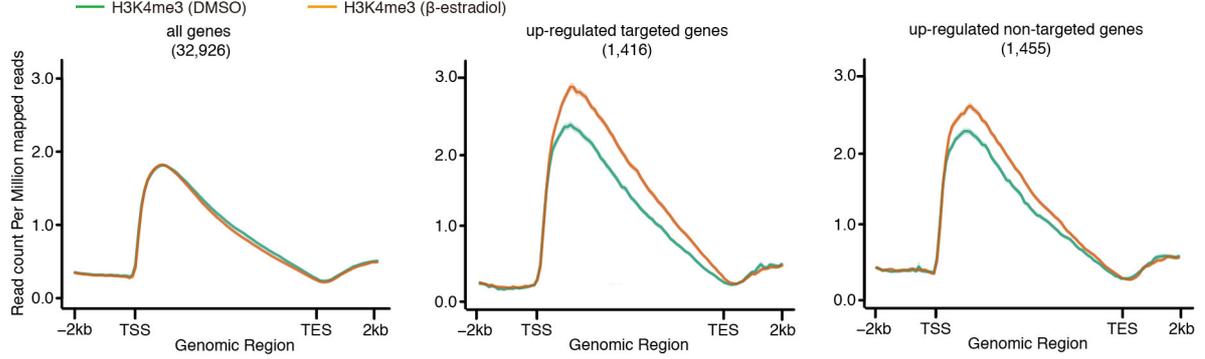
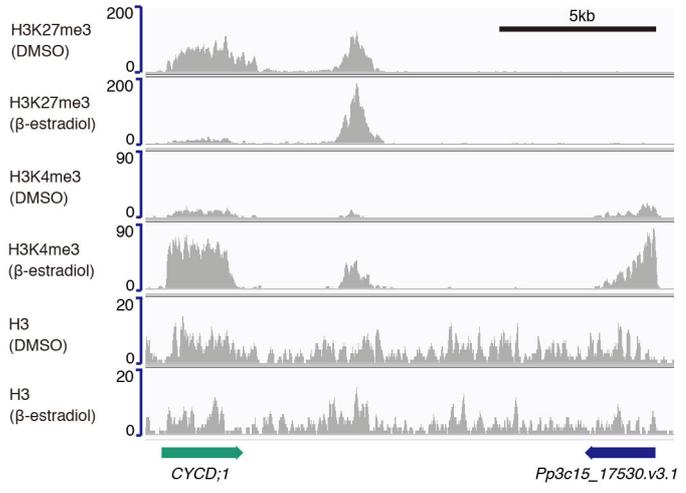
(A to C) Frequencies of side branching at each cell position of GX6:NGG>STEMIN2#81 (A), GX6:NGG>STEMIN3#105 (B), and GX6:NGG#63 (C) chloronemata with or without 1  $\mu\text{M}$   $\beta$ -estradiol ( $n = 10$  protonemata) 24 hours after they were moved from unilateral red light to white light conditions. Individual data of the frequencies are shown in circles. Center values and error bars indicate the mean of biological triplicates and SD, respectively. *F*-test and a two-tailed student's *t*-test indicate no significant differences between the frequencies of side branching with and without  $\beta$ -estradiol at each cell position.  $p > 0.05$  (A-C).



**Fig. 21. Construction of GX6:STEMIN1-Myc plants.**

(A) A schematic showing insertion of GX6:STEMIN1-Myc construct into the putative neutral site *PIG1*. Blue, orange, and red arrows denote a connected DNA fragment of the LexA operator and the minimal 35S promoter (LexAop:m35S) (Zuo et al., 2000), the *KIN1D1a* promoter (ProGX6) (Kubo et al., 2013), and an aminoglycoside phosphotransferase IV cassette (aphIV) (Hiwatashi et al., 2008), respectively. Light blue, gray, and black boxes designate the DNA fragment encoding STEMIN1 fused with a triple Myc-tag sequence at the 3' end, *pea3A-ter* (Fluhr et al., 1986), and *rbcS* terminator (*rbcS-ter*) (Fluhr et al., 1986), respectively. The probe used in (B) is indicated.

(B) DNA gel-blot analysis of targeted plants. Genomic DNA of the GX6:STEMIN1-Myc (#10, #11, #15, #18, and #44) and wild-type plants was digested with *EcoT221*. Asterisks and arrowheads indicate DNA fragments specific to wild-type and transgenic plants, respectively.

**A****B****C****D****E**

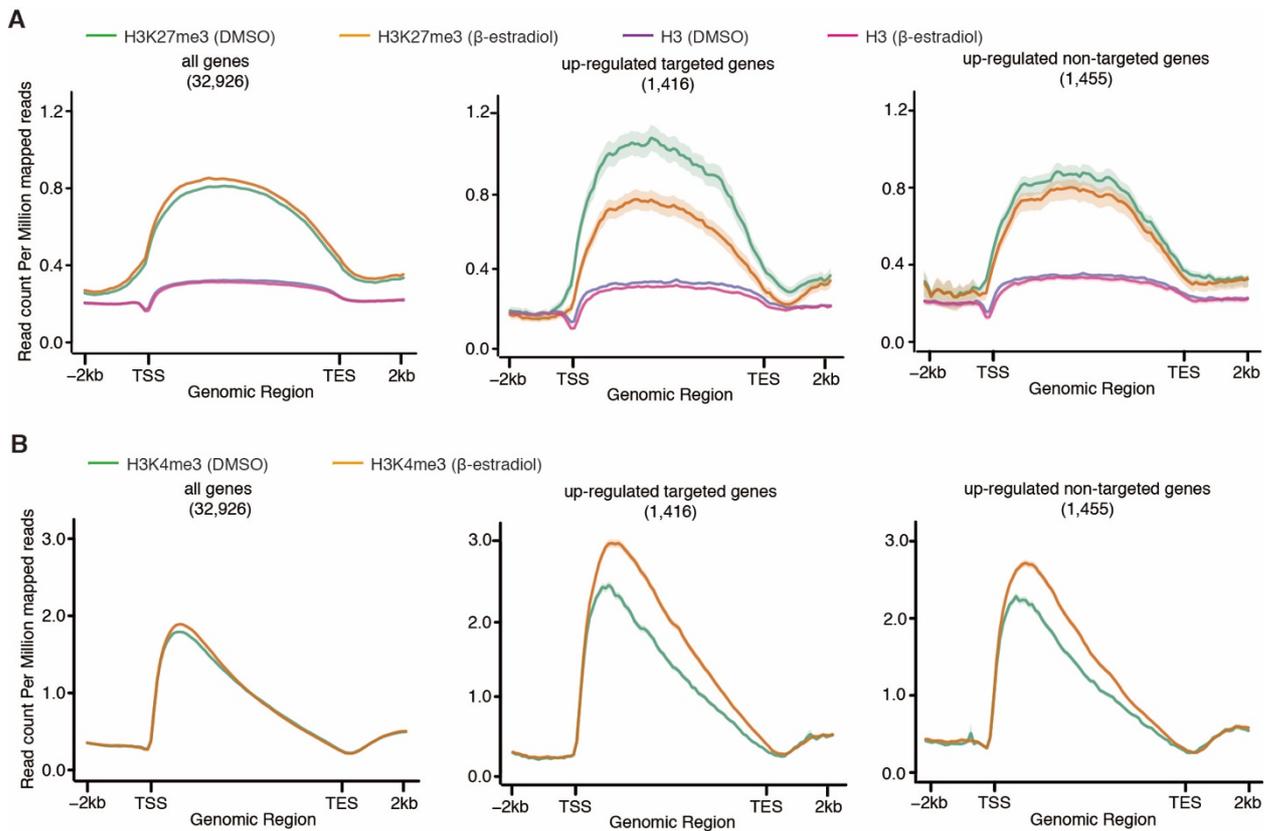
**Fig. 22. Changes in histone modifications on STEMIN1 target genes.**

(A) Venn diagram showing overlaps between STEMIN1-upregulated genes and STEMIN1-targeted genes identified by RNA-seq and ChIP-seq, respectively. *p* values are based on the hypergeometric test.

(B) Normalized average distribution of STEMIN1-binding sites in the 1416 upregulated targeted genes with or without induction of *STEMIN1* with  $\beta$ -estradiol. TSS, transcriptional start site; TES, transcriptional end site.

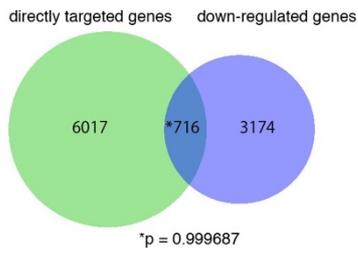
(C and D) Normalized average distribution of H3 (C), H3K27me3 (C), and H3K4me3 (D) with or without STEMIN1-Myc induction in all genes, 1416 upregulated targeted genes, and 1455 upregulated nontargeted genes. Semitransparent shades show the standard error of the mean.

(E) Integrated Genomics View visualization of genomic sequence including the *CYCD;1* gene, showing the distribution of H3K27me3, H3K4me3, and H3 with ( $\beta$ -estradiol) or without (DMSO) STEMIN1 induction. The *y* axis represents mapped read counts.



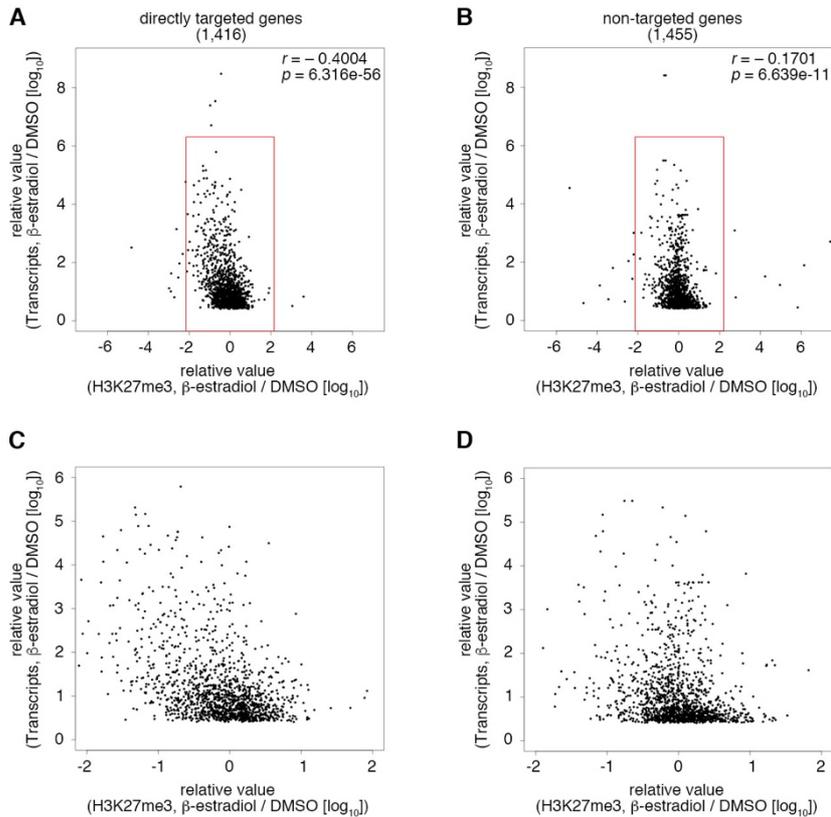
**Fig. 23. Changes in histone modifications of the upregulated genes after the induction of STEMIN1-Myc in gametophores in a separate experiment from the experiment in Fig. 22C,D.**

(A and B) Normalized average distribution of H3 (A), H3K27me3 (A), and H3K4me3 (B) in all genes, the upregulated targeted genes, and the upregulated nontargeted genes. The changes in the levels of H3K27me3 on the target genes and the nontarget genes were 54% and 20% decreases, respectively, while those of H3K4me3 were 26% and 20% increases. TSS, transcriptional start site; TES, transcriptional end site. Semitransparent shading shows the standard error of mean.



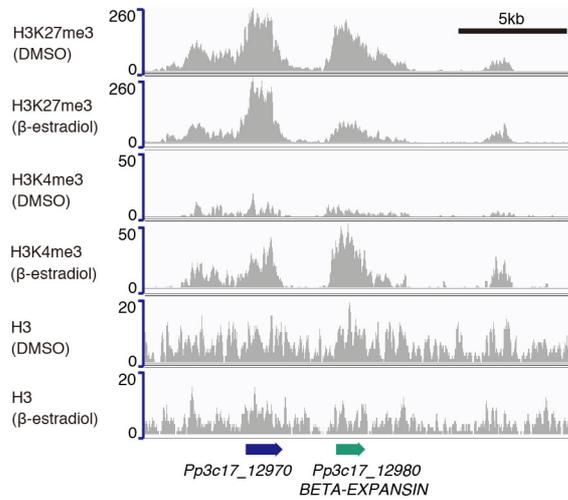
**Fig. 24. Venn diagram showing overlaps between STEMIN1-downregulated genes and directly targeted genes identified by RNA-seq and CHIP-seq analyses, respectively.**

The numbers of co-occurrences between the directly targeted genes (6,733) and downregulated genes (3,890) among all genes (32,926) were smaller than the probabilities assuming independent variables. *P* value is based on the hypergeometric test.



**Fig. 25. Relationship between changes in H3K27me3 levels and changes in transcript levels in STEMIN1-upregulated genes.**

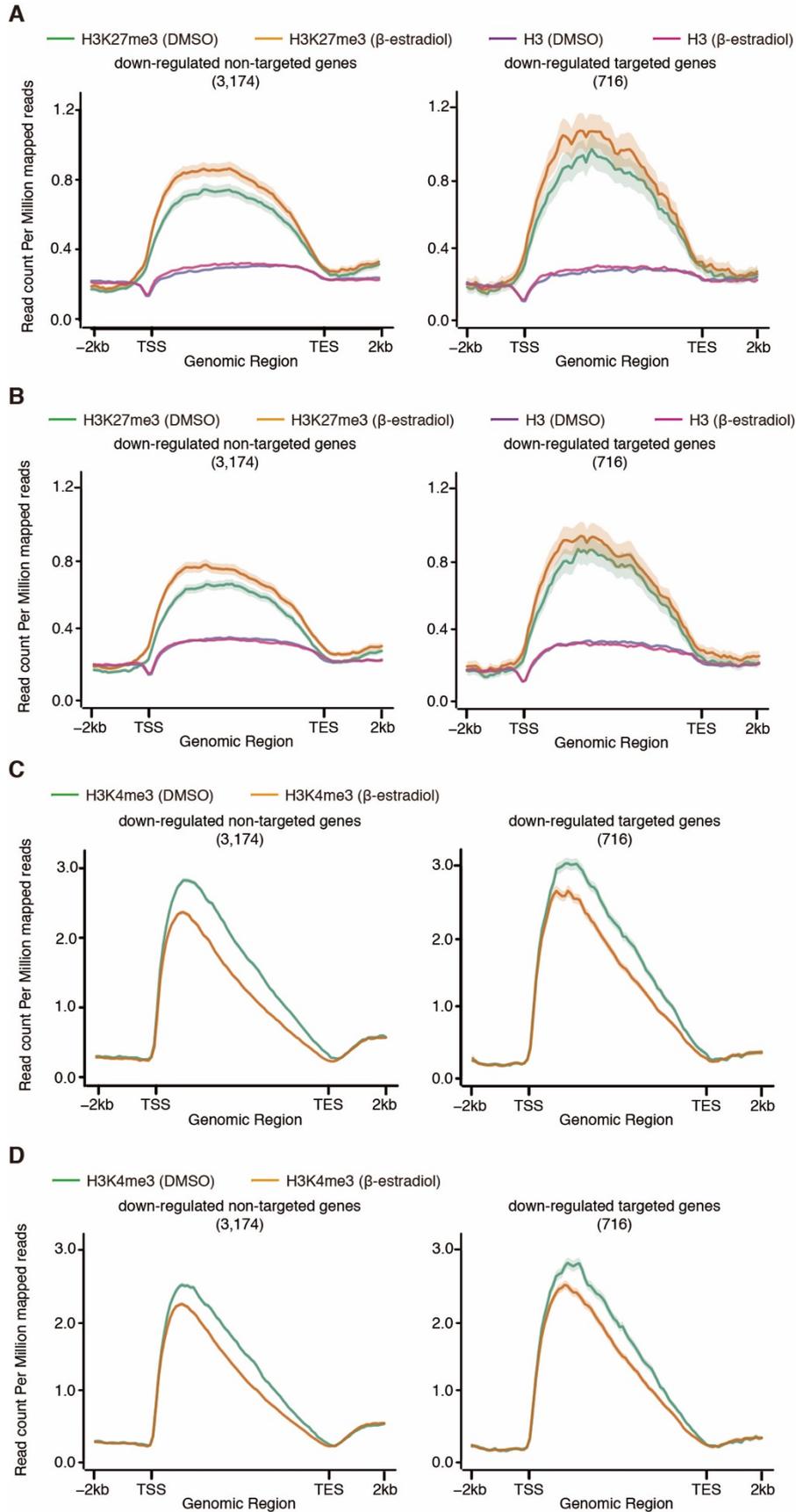
Scatter plots of changes in the H3K27me3 level of each gene in  $\beta$ -estradiol-treated gametophores compared to DMSO-treated gametophores (x axis) and fold changes in its transcript level (y axis) for the 1,416 upregulated targeted genes (**A**) and 1,455 upregulated nontargeted genes (**B**). Each dot represents one gene model from  $-2$  kb upstream of the transcriptional start site to  $+2$  kb downstream of the transcriptional end site. Pearson's correlation coefficients ( $r$ ) and  $P$  values ( $P$ ) were calculated using the R package. The red-boxed regions in (**A**) and (**B**) are magnified in (**C**) and (**D**), respectively.



**Fig. 26. Decrease in H3K27me3 levels and increase in H3K4me3 levels in *BETA-EXPANSIN* gene by induction of STEMIN1-Myc.**

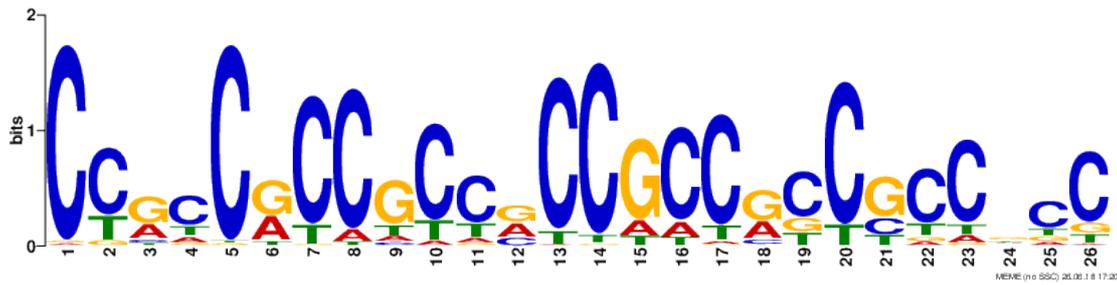
Integrated Genomics View of genomic sequence including *BETA-EXPANSIN* gene (*Pp3c17\_12980*) that shows distributions of H3K27me3, H3K4me3, and H3 with ( $\beta$ -estradiol) or without (DMSO) STEMIN1 induction. The y axis represents mapped read counts.





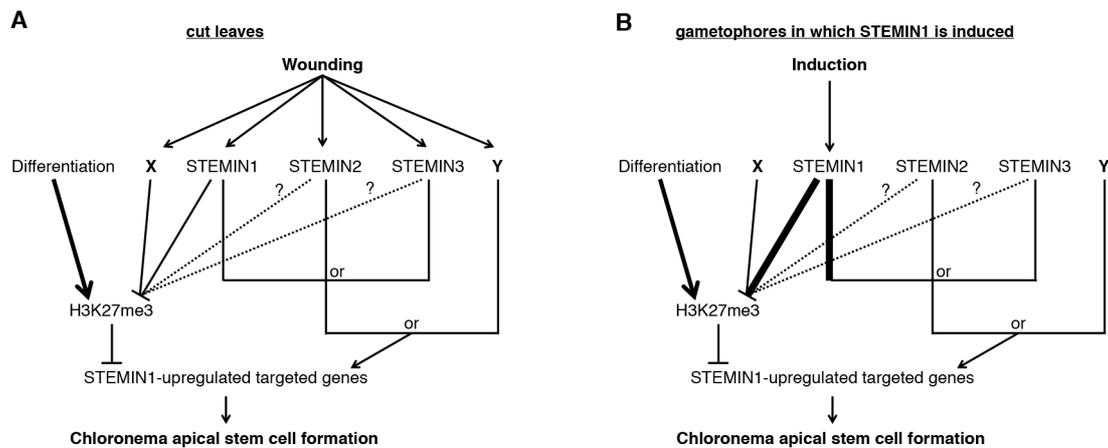
**Fig. 28. Changes in histone modifications of STEMIN1 down-regulated genes after induction of STEMIN1-Myc in gametophores.**

(A to D) Normalized average distribution of H3 (A and B), H3K27me3 (A and B), and H3K4me3 (C and D) in the downregulated nontargeted genes and downregulated targeted genes in the first (A and C) and second (B and D) ChIP-seq experiments. TSS, transcriptional start site; TES, transcriptional end site. Semitransparent shading shows the standard error of the mean.



**Figure 29. Genomic DNA motif targeted by STEMIN1-Myc**

Consensus sequence created from the 100 genes out of the 1,416 upregulated targeted genes. Enriched sequences in the DNA fraction immunoprecipitated from gametophores with STEMIN1-Myc induction were searched using the MEME suite (Bailey et al., 2009). GCC box, which has been described in promoters bounded by AP2/ERF transcription factors (Ohme-Takagi and Shinshi, 1995), is present in the sense sequences.



**Fig. 30. Hypothetical models of STEMIN functions in stem cell formation**

(A) A hypothetical model in cut leaves. It was postulated that STEMIN1-upregulated targeted genes are repressed by H3K27me3 modifications under the control of a putative differentiation signal in matured leaves. Wounding by leaf cutting induces expression of three STEMINs. STEMIN1 diminishes H3K27me3 levels of STEMIN1-upregulated targeted genes. It is a future study whether STEMIN2 and STEMIN3 are involved in diminishing H3K27me3. Since chloronema apical stem cells are formed without STEMIN1, STEMIN2, and STEMIN3, another factor X to diminish H3K27me3 is postulated. Three STEMIN proteins redundantly induce transcription of STEMIN1-upregulated targeted genes. It is a future study whether three STEMIN proteins regulate the same targets. Since reprogramming is induced in triple deletion mutants, it is postulated another factor Y to induce expression of STEMIN1-upregulated targeted genes.

(B) A hypothetical model in gametophores in which STEMIN1 is induced. STEMIN1 induction diminishes H3K27me3 levels. Since induction of each STEMIN2 and STEMIN3 does not result in stem cell formation, it is postulated that ability to diminish H3K27me3 and to activate transcription of targeted genes by STEMIN1 is stronger than those by other STEMINs. Regulations of H3K27me3 levels and transcription levels by STEMIN2 and STEMIN3 induction are future studies.

## 5. Discussion

### 5.1. Cell fate changes by AP2/ERF transcription factors in land plants

This study demonstrated that *Physcomitrella* AP2/ERF transcription factor *STEMIN1* and its homologs function in initiation of reprogramming of differentiated leaf cells and chloronema cells into chloronema apical stem cells, rather than in maintenance of chloronema apical stem cells. Wounding activated the expression of *STEMIN1*, *STEMIN2* and *STEMIN3* genes at leaf cells facing the cut, leading to reprogramming of the leaf cells into chloronema apical stem cells. Moreover, *STEMIN1* positively regulated the formation of secondary chloronema apical stem cells from chloronema cells during development, together with *STEMIN2* and *STEMIN3*. These results indicate that the gene regulatory network under the control of *STEMIN1* is an inherent mechanism to initiate stem cell formation in differentiated cells during regeneration and regular development.

During *Arabidopsis* organ regeneration by wounding, an AP2/ERF transcription factor *WIND1* is upregulated, thereby leading to expression of another AP2/ERF transcription factor *ESR/DRN* to initiate shoot regeneration. Subsequently, *CUC1* and *CUC2* transcription factors are upregulated by *ESR1* and function redundantly in shoot apical meristem formation (Hibara et al., 2003; Iwase et al., 2011a; Iwase et al., 2017; Souer et al., 1996; Aida, 1997). *Arabidopsis* homologous gene of *STEMIN1* is upregulated during shoot formation in a callus, but not during callus formation (Che et al., 2007), although it remains to be elucidated whether the *STEMIN1*

homolog is involved in the WIND1-ESR1/DRN pathway.

Physcomitrella homologous genes of *ESR1/DRN* and *CUC2*, *PpESR1/DRN* (*Pp3c3\_19630*) and *PpCUC2* (*Pp3c12\_18020*), were STEMIN1-direct target genes and upregulated by the STEMIN1 induction. *PpWIND1* is not activated by the STEMIN1 induction, while it is transiently upregulated before *STEMIN1* expression during reprogramming in cut leaves (Nishiyama et al., 2012). Thus, several components involved in formation of Arabidopsis shoot apical meristem are likely conserved in Physcomitrella stem cell formation, suggesting a common gene regulatory network underlying cellular reprogramming. Furthermore, many regulators acting during shoot regeneration in Arabidopsis, including *ESR1*, are epigenetically silenced by the polycomb-mediated histone modification, but they are rapidly induced after wounding (Ikeuchi et al., 2015b; Iwase et al., 2017). Hence, the STEMIN1-mediated transcriptional regulatory network may be a part of common mechanisms underlying the control of cell fate changes in land plants. The distinct evolution of such common mechanisms in each lineage might contribute to a distinct mode of reprogramming for postembryogenesis and regeneration. Therefore, our understanding of STEMIN1-mediated reprogramming could illuminate common mechanisms underlying the plasticity inherent in plant cells.

## **5.2. Functions of STEMINs in stem cell formation**

In this study, I found that *STEMIN2* and *STEMIN3* might function redundantly with

*STEMIN1* in stem cell formation in cut leaves (Fig. 12), but unlike *STEMIN1*, induction of *STEMIN2* or *STEMIN3* could not induce stem cell formation in intact gametophores (Fig. 15). These results indicate that *STEMIN2* and *STEMIN3* have functions overlapping those of *STEMIN1*, but they lack the ability of *STEMIN1* to change intact leaf cells into stem cells. Moreover, I found that *STEMIN1* induces stem cell formation in intact gametophores through reduction of H3K27me3 on the *STEMIN1*-target genes (Fig. 22C and Fig 23A).

Based on these results, I propose a hypothetical models about the *STEMIN* functions for stem cell formation (Fig. 30). In gametophores, *STEMIN1*-upregulated targeted genes are repressed by H3K27me3 modifications under the control of a putative differentiation signal. In cut leaves, wounding by leaf cutting induces expression of *STEMIN1*, *STEMIN2* and *STEMIN3* at leaf cells facing the cut. *STEMIN1* diminishes H3K27me3 levels of *STEMIN1*-upregulated targeted genes, although it is still unknown whether *STEMIN2* and *STEMIN3* are involved in diminishing H3K27me3 (Fig. 30A). Since chloronema apical stem cells are formed in  $\Delta\text{stemin1}\Delta\text{stemin2}\Delta\text{stemin3}$  triple deletion mutant (Fig. 13), an unidentified factor X might diminish H3K27me3. Three *STEMIN* proteins redundantly induce transcription of *STEMIN1*-upregulated targeted genes in cut leaves, suggesting another factor Y might induce expression of *STEMIN1*-upregulated targeted genes redundantly with *STEMINs* (Fig 30A)

In gametophores in which *STEMIN1* is induced, *STEMIN1* diminishes H3K27me3 levels (Fig. 22). Since induction of each *STEMIN2* and *STEMIN3* does not

result in stem cell formation (Fig. 15), abilities of STEMIN1, which diminishes H3K27me3 and activates transcription of targeted genes, are stronger than those of other STEMINs. Future studies will reveal regulations of H3K27me3 levels and transcription levels by STEMIN2 and STEMIN3 induction.

### **5.3. Locus-specific reprogramming of H3K27me3 by STEMIN1**

New cell fate establishment requires chromatin modifications to reset the previous cellular memory (Costa and Shaw, 2007). Presence of a repressive mark H3K27me3 at a specific locus contributes to the determinants of tissue-specific expression and the repression of developmental programs. Therefore, locus-specific reprogramming of the H3K27me3 should be crucial for rewriting somatic gene regulatory networks during plant development and regeneration, which could be one of mechanisms underlying formation of stem cells.

In metazoans, H3K27me3 is removed by JMJD3 (lysine-specific demethylase 6B: KDM6B) and ubiquitously transcribed tetra-tricopeptide repeat X (UTX: KDM6A), two Jumonji domain-containing histone demethylases, which are essential for regulation of homeotic gene expression, in inflammatory response and germ cell development (Agger et al., 2007; Santa et al., 2007; Lan et al., 2007; Lee et al., 2007). However, it remains unknown how the activity of these proteins is regulated during development.

In Arabidopsis, REF6/JMJ12 functions as a H3K27me3 demethylase (Lu et

al., 2011; Cui et al., 2016; Li et al., 2016), and that has two close homologs: EARLY FLOWERING 6 (ELF6) and JMJ13, which are potential H3K27me3 demethylases and act redundantly with REF6 (Lu et al., 2011). These demethylases recognize specific DNA sequences at genes enriched with the H3K27me3 and demethylate these loci to activate their gene expression in flowering (Lu et al., 2011; Cui et al., 2016; Li et al., 2016). However, it remains obscure whether these factors execute a program to induce stem cell fate through local epigenetic reprogramming.

When a differentiated cell undergoes reprogramming, in many cases the cell requires the cell cycle reentry and progression (Jopling et al., 2011). For example, efficient reprogramming of mouse postnatal neurons into iPS cells is achieved after the cell cycle reentry by suppression of the p53 tumor suppressor gene, together with expression of the four reprogramming factors (Kim et al., 2011). Likewise, in *Arabidopsis*, shoot formation in a root explant and root regeneration in a dissect root require initiation of the cell cycle progression (Che et al., 2007; Sena et al., 2009). In *Arabidopsis* floral development, the floral homeotic protein AGAMOUS (AG) in the shoot apical meristem induces expression of *KNUCKLES* (*KNU*) encoding a zinc finger protein, which in turn repress expression of *WUS* (Sun et al., 2009; Sun et al., 2014). The induction of *KNU* by AG requires dilution of H3K27me3 levels on *KNU* gene by repeated cell divisions to activate its genes expression (Sun et al., 2014). Thus, the cell cycle machinery could provide cells with a window of opportunity to change their epigenetic states dependent on the cell cycle progression.

On the other hand, in *Physcomitrella*, ectopic induction of STEMIN1 in leaf

cells decreased a repressive chromatin mark, H3K27me3, on its direct target genes before cell division, resulting in the changes of leaf cells to chloronema apical stem cells. This indicates that reprogramming by the STEMIN1 induction does not require cell cycle reactivation and progression, further suggesting that STEMIN1 activates an intrinsic mechanism underlying local H3K27me3 reprogramming and transcriptional regulation of its target genes resulting in cell fate changes and cell cycle reactivation. Since STEMIN1 protein does not contain a domain that is capable of functioning as histone demethylases, STEMIN1 appears to cooperate with other factors to decrease H3K27me3 levels of their target genes. Further studies should be directed to the identification of such factors to understand local reprogramming of H3K27me3 mediated by STEMIN1.

## 6. General discussion

During development in both metazoans and land plants, cell fate of each somatic cell is precisely specified with a tissue-specific epigenetic modification and a consequent gene expression pattern (Kaufmann et al., 2010; Chen and Dent, 2014). Therefore, conversion from a different state to a stem cell state requires a mechanism for changing the epigenetic marks to switch from an expression pattern typical of a somatic cell to a new one (Feng et al., 2010).

While plasticity of differentiated cells is currently observed in both metazoans and land plants, the inefficiency of conversion of differentiated cells into stem cells has presented considerable barriers to define the rules of reprogramming. Depletion of the nucleosome remodeling and histone deacetylation (NuRD) complex member Mbd3, that can mediate gene repression through histone deacetylation and chromatin remodeling activities, allows for nearly 100% reprogramming efficiency (Lai and Wade, 2011; Rais et al., 2013). This indicates that the Mbd3/NuRD repressor complex is the predominant molecular block preventing deterministic induction of pluripotency. Likewise, the H3K27 demethylase, UTX, functions as a critical regulator acting at molecular switches during reprogramming to ground state pluripotency to safeguard an efficient, timely and authentically demethylates H3K27me3 (Mansour et al., 2012). On the other hand, global loss of H3K27me3 causes a severe decline in the efficiency of iPS formation. Thus, locus-specific epigenetic modifications appear to create a barrier to the cellular changes and changes in the epigenetic modification on the specific genes are required

for rewriting cell-specific gene expression patterns to initiate a new developmental program.

Similarly, in *Arabidopsis* callus formation, genome-wide H3K27me3 reprogramming suppresses expression of genes involved in leaf cell identity and activates the auxin-signaling pathway, through the PRC2-mediated H3K27 methylation and the H3K27 demethylation pathways, respectively (He et al., 2012). On the other hand, mutation in genes of PRC2 components lacks the potential to form callus from leaf tissue (He et al., 2012). Therefore, it has been expected that unidentified factors or mechanisms can function in reprogramming through modulation of the H3K27me3 levels in specific genes in land plants (Birnbaum and Roudier, 2017).

This study suggests that *STEMIN1* functions in activation of an intrinsic mechanism underlying local H3K27me3 reprogramming and transcriptional regulation of its target genes for stem cell formation. In addition, *STEMIN1* bound around at transcriptional starting sites, likely corresponding to the position of the nucleosome-depleted zone, suggesting no limitation of access of *STEMIN1* to the specific target genes with the repressed chromatin status. Therefore, the *STEMIN1*-activating mechanism could readily overcome an epigenetic barrier to stabilize cell identity, which may reflect high cellular plasticity in the moss. Since other land plants have orthologs of the *STEMIN* genes, our understanding of *STEMIN1*-mediated reprogramming will shed light on a mechanism potentially underlying the plasticity and regeneration ability of plant cells.

## **Future prospects**

How differentiated cells acquire stem cell fates during plant development and regeneration is a fundamental question in biology. Molecular genetic studies using angiosperms over the last few decades have identified several factors involved in stem cell formation and revealed gene regulatory networks governing cellular reprogramming (Sugimoto et al., 2010; Ikeuchi et al., 2013; Efroni et al., 2016; Ikeuchi et al., 2016; Iwase et al., 2017; Ikeuchi et al., 2018). Furthermore, the moss reprogramming system allows *in vivo* visualization of cellular activities. Future studies will be also directed towards the identification of key regulators participating in removal of H3K27me3 together with STEMIN1 in the moss stem cell formation. Identification of such factors could be expected to uncover the general and specific molecular mechanisms on spatiotemporal specification of the cells undergoing proliferation and coordination of the acquisition of new cellular state. Such studies on molecular mechanisms of stem cell formation in the moss will highlight conservation and divergence of reprogramming into stem cells in land plants and could provide insight in evolution of stem cell regulation.

## **7. Acknowledgements**

I am deeply indebted to Professor Mitsuyasu Hasebe for his generous supports, advice, and helpful discussion. I am grateful to Dr. Masaki Ishikawa for helpful discussion, guidance, and experiment support for my project. I am also grateful to Dr. Yosuke Tamada for helpful discussion.

I would like to thank my collaborators, Dr. Shuji Shigenobu for RNA-seq and ChIP-seq analysis, Dr. Tomoaki Nishiyama for phylogenetic analysis, ChIP-seq analysis and helpful discussion, and Dr. Yukiko Kabeya for producing transgenic lines and technical supports, and E. Aoki, M. Goto, T. Nishi, and S. Ooi for technical assistance.

I also would like to thank all members of the Division of Evolutionary Biology for discussion and supports, the Model Plant Research Facility, and Data Integration and Analysis Facility in National Institute for Basic Biology.

Finally, I thank my family and friends for their continuous support.

## 8. References

- Agger, K. *et al.* UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**, 731–735 (2007).
- Aharoni, A. *et al.* The SHINE clade of AP2 domain transcription factors activates wax biosynthesis, alters cuticle properties, and confers drought tolerance when overexpressed in Arabidopsis. *Plant Cell* **16**, 2463–2480 (2004).
- Aida, M. Genes Involved in Organ Separation in Arabidopsis: An Analysis of the cup-shaped cotyledon Mutant. *Plant Cell Online* **9**, 841–857 (1997).
- Aida, M. *et al.* Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis : interaction among the *CUP-SHAPED COTYLEDON* and *SHOOT MERISTEMLESS* genes. *Development* **126**, 1563–1570 (1999).
- Aoyama, T. *et al.* AP2-type transcription factors determine stem cell identity in the moss *Physcomitrella patens*. *Development* **139**, 3120–3129 (2012).
- Atta, R. *et al.* Pluripotency of Arabidopsis xylem pericycle underlies shoot regeneration from root and hypocotyl explants grown in vitro. *Plant J.* **57**, 626–644 (2009).
- Bailey, T. L. & Elkan, C. Fitting a Mixture Model by Expectation Maximization to Discover Motifs in Bipolymers. *Proc. Second Int. Conf. Intell. Syst. Mol. Biol.* 28–36 (1994).
- Bailey, T. L. *et al.* MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Res.* **37**, 202–208 (2009).
- Banks, J. A. *et al.* The Selaginella genome identifies genetic changes associated with the evolution of vascular plants. *Science.* **332**, 960–963 (2011).
- Banno, H. *et al.* Overexpression of Arabidopsis ESR1 induces initiation of shoot regeneration. *Plant Cell* **13**, 2609–2618 (2001).
- Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
- Birnbaum, K. D. & Roudier, F. Epigenetic memory and cell fate reprogramming in plants. *Regen.* **4**, 15–20 (2017).
- Birnbaum, K. D. & Sánchez Alvarado, A. Slicing across kingdoms: regeneration in plants and animals. *Cell* **132**, 697–710 (2008).
- Bouyer, D. *et al.* Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. *PLoS Genet.* **7**, e1002014 (2011).

- Che, P. *et al.* Developmental steps in acquiring competence for shoot development in Arabidopsis tissue culture. *Planta* **226**, 1183–1194 (2007).
- Chen, T. & Dent, S. Y. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat. Rev. Genet.* **15**, 93–106 (2014).
- Chiu, W. -I. *et al.* Engineered GFP as a vital reporter in plants. *Curr. Biol.* **6**, 325–330 (1996).
- Costa, S. & Shaw, P. ‘Open minded’ cells: how cells can change fate. *Trends Cell Biol.* **17**, 101–106 (2007).
- Cove, D. J. & Knight, C. D. The Moss *Physcomitrella patens*, a Model System with Potential for the Study of Plant Reproduction. *Plant Cell* **5**, 1483–1488 (1993).
- Cui, X. *et al.* REF6 recognizes a specific DNA sequence to demethylate H3K27me3 and regulate organ boundary formation in Arabidopsis. *Nat Genet* **48**, 694–699 (2016).
- da Rocha Correa, L. *et al.* Distinct modes of adventitious rooting in Arabidopsis thaliana. *Plant Biol.* **14**, 100–109 (2012).
- De Boer, K. *et al.* APETALA2/ETHYLENE RESPONSE FACTOR and basic helix-loop-helix tobacco transcription factors cooperatively mediate jasmonate-elicited nicotine biosynthesis. *Plant J.* **66**, 1053–1065 (2011).
- De Smet, I. *et al.* Lateral root initiation or the birth of a new meristem. *Plant Mol. Biol.* **60**, 871–887 (2006).
- Efroni, I. *et al.* Root Regeneration Triggers an Embryo-like Sequence Guided by Hormonal Interactions. *Cell* **165**, 1721–1733 (2016).
- Feldman, L. J. The de novo origin of the quiescent center regenerating root apices of *Zea mays*. *Planta* **128**, 207–212 (1976).
- Feng, S. *et al.* Epigenetic Reprogramming in Plant and Animal Development. *Science* **330**, 622–627 (2010).
- Fluhr, R. *et al.* Expression dynamics of the pea *rbcS* multigene family and organ distribution of the transcripts. *Embo J.* **5**, 2063–2071 (1986).
- Fujimoto, S. *et al.* Arabidopsis Ethylene-Responsive Element Binding Factors Act as Transcriptional Activators or Repressors of GCC Box-Mediated Gene Expression. *Plant Cell* **12**, 393 (2000).
- Gan, E. S. *et al.* Jumonji demethylases moderate precocious flowering at elevated temperature via regulation of FLC in Arabidopsis. *Nat Commun* **5**, 5098 (2014).

- Gaspar-Maia, A. *et al.* Open chromatin in pluripotency and reprogramming. *Nat. Rev. Mol. Cell Biol.* **12**, 36–47 (2011).
- Gendrel, A. *et al.* Profiling histone modification patterns in plants using genomic tiling microarrays. *Nat Methods* **2**, 213–218 (2005).
- Gordon, S. P. *et al.* Pattern formation during de novo assembly of the Arabidopsis shoot meristem. *Development* **134**, 3539–3548 (2007).
- Hao, D. *et al.* Unique mode of GCC box recognition by the DNA-binding domain of ethylene-responsive element-binding factor (ERF domain) in plant. *J. Biol. Chem.* **273**, 26857–26861 (1998).
- He, C. *et al.* Reprogramming of H3K27me3 is critical for acquisition of pluripotency from cultured Arabidopsis tissues. *PLoS Genet.* **8**, e1002911 (2012).
- Hibara, K. *et al.* *CUC1* gene activates the expression of SAM-related genes to induce adventitious shoot formation. *Plant J.* **36**, 687–696 (2003).
- Hiratsu, K. *et al.* Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in Arabidopsis. *Plant J.* **34**, 733–739 (2003).
- Hiwatashi, Y. *et al.* Kinesins are indispensable for interdigitation of phragmoplast microtubules in the moss *Physcomitrella patens*. *Plant Cell* **20**, 3094–3106 (2008).
- Ikeuchi, M. *et al.* Control of plant cell differentiation by histone modification and DNA methylation. *Curr. Opin. Plant Biol.* **28**, 60–67 (2015a).
- Ikeuchi, M. *et al.* PRC2 represses dedifferentiation of mature somatic cells in Arabidopsis. *Nat. Plants* **1**, 1–7 (2015b).
- Ikeuchi, M. *et al.* Plant regeneration: cellular origins and molecular mechanisms. *Development* **143**, 1442–1451 (2016).
- Ikeuchi, M. *et al.* A Gene Regulatory Network for Cellular Reprogramming in Plant Regeneration. *Plant Cell Physiol.* **59**, 765–777 (2018).
- Ikeuchi, M. *et al.* Plant callus: mechanisms of induction and repression. *Plant Cell* **25**, 3159–3173 (2013).
- Ishikawa, M. & Hasebe, M. Cell cycle reentry from the late S phase: implications from stem cell formation in the moss *Physcomitrella patens*. *J. Plant Res.* **128**, 399–405 (2015).
- Ishikawa, M. *et al.* *Physcomitrella* cyclin-dependent kinase A links cell cycle

- reactivation to other cellular changes during reprogramming of leaf cells. *Plant Cell* **23**, 2924–2938 (2011).
- Iwase, A. *et al.* WIND1 Promotes Shoot Regeneration through Transcriptional Activation of ENHANCER OF SHOOT REGENERATION1 in Arabidopsis. *Plant Cell* **29**, 54–69 (2017).
- Iwase, A. *et al.* WIND1-based acquisition of regeneration competency in Arabidopsis and rapeseed. *J. Plant Res.* **128**, 389–397 (2015).
- Iwase, A. *et al.* The AP2/ERF transcription factor WIND1 controls cell dedifferentiation in Arabidopsis. *Curr. Biol.* **21**, 508–514 (2011a).
- Iwase, A. *et al.* WIND1: A key molecular switch for plant cell dedifferentiation. *Plant Signal. Behav.* **6**, 1943–1945 (2011b).
- Jefferson, R. A. *et al.* GUS fusions: b-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *Embo J.* **6**, 3901–3907 (1987).
- Jopling, C. Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. *Nat. Rev. Mol. Cell Biol.* (2011).
- Kalderon, D. *et al.* Sequence requirements for nuclear location of simian virus 40 large-T antigen. *Nature* **311**, 33–38 (1984).
- Katoh, K. *et al.* Improvement in the accuracy of multiple sequence alignment program MAFFT. *Genome Inf.* **16**, 22–33 (2005).
- Kaufmann, K. *et al.* Regulation of transcription in plants: mechanisms controlling developmental switches. *Nat. Rev. Genet.* **11**, 830–842 (2010).
- Kim, J. *et al.* Reprogramming of postnatal neurons into induced pluripotent stem cells by defined factors. *Stem Cells* **29**, 992–1000 (2011).
- Kofuji, R. & Hasebe, M. Eight types of stem cells in the life cycle of the moss *Physcomitrella patens*. *Curr. Opin. Plant Biol.* **17C**, 13–21 (2014).
- Kubo, M. *et al.* System for stable beta-estradiol-inducible gene expression in the moss *Physcomitrella patens*. *PLoS One* **8**, e77356 (2013).
- Lafos, M. *et al.* Dynamic regulation of H3K27 trimethylation during Arabidopsis differentiation. *PLoS Genet.* **7**, e1002040 (2011).
- Lai, A. Y. & Wade, P. A. Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nat. Rev. Cancer* **11**, 588–596 (2011).
- Lan, F. *et al.* ARTICLES A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **449**, 689–695 (2007).

- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
- Laux, T. *et al.* The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development* **122**, 87–96 (1996).
- Lee, J. & Shilatifard, A. A site to remember: H3K36 methylation a mark for histone deacetylation. *Mutat. Res.* (2007).
- Li, C. *et al.* Concerted genomic targeting of H3K27 demethylase REF6 and chromatin-remodeling ATPase BRM in Arabidopsis. *Nat Genet* **48**, 687–693 (2016).
- Li, C. *et al.* A Lin28 homologue reprograms differentiated cells to stem cells in the moss *Physcomitrella patens*. *Nat Commun* **8**, 14242 (2017).
- Li, W. *et al.* DNA methylation and histone modifications regulate de novo shoot regeneration in Arabidopsis by modulating WUSCHEL expression and auxin signaling. *PLoS Genet.* **7**, e1002243 (2011).
- Liao, Y. *et al.* The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108 (2013).
- Lu, F. *et al.* Arabidopsis REF6 is a histone H3 lysine 27 demethylase. *Nat Genet* **43**, 715–719 (2011).
- Maddison, W. P. & Maddison, D. R. Interactive analysis of phylogeny and character evolution using the computer program MacClade. *Folia Primatol (Basel)* **53**, 190–202 (1989).
- Malamy, J. E. & Benfey, P. N. Organization and cell differentiation in lateral roots of Arabidopsis thaliana. *Development* **124**, 33–44 (1997).
- Mansour, A. A. *et al.* The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409–413 (2012).
- McSteen, P. & Leyser, O. Shoot branching. *Annu. Rev. Plant Biol.* **56**, 353–374 (2005).
- Mozgova, I. *et al.* PRC2 Represses Hormone-Induced Somatic Embryogenesis in Vegetative Tissue of Arabidopsis thaliana. *PLoS Genet.* **13**, e1006562 (2017).
- Neale, D. B. *et al.* Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biol.* **15**, R59 (2014).
- Ng, D. W. *et al.* Plant SET domain-containing proteins : Structure , function and regulation. *Biochim. Biophys. Acta* **1769**, 316–329 (2007).
- Nishiyama, T. *et al.* Tagged mutagenesis and gene-trap in the moss, *Physcomitrella*

- patens by shuttle mutagenesis. *DNA Res.* **7**, 9–17 (2000).
- Nishiyama, T. *et al.* Digital gene expression profiling by 5'-end sequencing of cDNAs during reprogramming in the moss *Physcomitrella patens*. *PLoS One* **7**, e36471 (2012).
- Odell, J. *et al.* Site-directed recombination in the genome of transgenic tobacco. *Mol Gen Genet* **223**, 369–378 (1990).
- Ohme-Takagi, M. & Shinshi, H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell* **7**, 173–182 (1995).
- Okano, Y. *et al.* A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. *Proc Natl Acad Sci U S A* **106**, 16321–16326 (2009).
- Onate-Sanchez, L. *et al.* AtERF14, a Member of the ERF Family of Transcription Factors, Plays a Nonredundant Role in Plant Defense. *Plant Physiol.* **143**, 400–409 (2006).
- Orkin, S. H. & Hochedlinger, K. Chromatin connections to pluripotency and cellular reprogramming. *Cell* **145**, 835–850 (2011).
- Potten, C. S. & Loeffler, M. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development* **110**, 1001–1020 (1990).
- Rais, Y. *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**, 65–70 (2013).
- Rensing, S. A. *et al.* The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science.* **319**, 64–69 (2008).
- Riechmann, J. L. & Meyerowitz, E. M. The AP2/EREBP family of plant transcription factors. *Biol Chem* **379**, 633–646 (1998).
- Riese, M. *et al.* SBP-domain transcription factors as possible effectors of cryptochrome-mediated blue light signalling in the moss *Physcomitrella patens*. *Planta* **227**, 505–515 (2008).
- Ringrose, L. & Paro, R. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38**, 413–443 (2004).
- Robinson, J. T. *et al.* Integrative genomics viewer. *Nat Biotechnol* **29**, 24–26 (2011).
- Sakakibara, K. *et al.* Class 1 KNOX genes are not involved in shoot development in the moss *Physcomitrella patens* but do function in sporophyte development. *Evol.*

- Dev.* **10**, 555–566 (2008).
- Sakakibara, K. *et al.* WOX13-like genes are required for reprogramming of leaf and protoplast cells into stem cells in the moss *Physcomitrella patens*. *Development* **141**, 1660–1670 (2014).
- Santa, F. *et al.* The Histone H3 Lysine-27 Demethylase Jmjd3 Links Inflammation to Inhibition of Polycomb-Mediated Gene Silencing. *Cell* **130**, 1083–1094 (2007).
- Sasaki, K. *et al.* Two novel AP2/ERF domain proteins interact with cis-element VWRE for wound-induced expression of the Tobacco *tpoxN1* gene. *Plant J.* **50**, 1079–1092 (2007).
- Sato, Y. *et al.* Cells reprogramming to stem cells inhibit the reprogramming of adjacent cells in the moss *Physcomitrella patens*. *Sci Rep* **7**, 1909 (2017).
- Sena, G. *et al.* Organ regeneration does not require a functional stem cell niche in plants. *Nature* **457**, 1150–1153 (2009).
- Shen, L. *et al.* ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**, 284 (2014).
- Skoog, F. & Miller, C. O. Chemical regulation of growth and organ formation in plant tissues cultured in vitro. *Symp. Soc. Exp. Biol.* **54**, 118–130 (1957).
- Slack, J. M. Origin of stem cells in organogenesis. *Science*. **322**, 1498–1501 (2008).
- Souer, E. *et al.* The No Apical Meristem Gene of *Petunia* Is Required for Pattern Formation in Embryos and Flowers and Is Expressed at Meristem and Primordia Boundaries. *Cell* **85**, 159–170 (1996).
- Springer, N. M. *et al.* Comparative Analysis of SET Domain Proteins in Maize and Arabidopsis Reveals Multiple Duplications Preceding the Divergence of Monocots and Dicots 1 [ w ]. *Plant Physiol.* **132**, 907–925 (2003).
- Steward, F. *et al.* Growth and organized development of cultured cell. II. Organization in cultures grown from freely suspended cells. *Am. J. Bot.* 705–708 (1958).
- Sugimoto, K. *et al.* Regeneration in plants and animals: dedifferentiation, transdifferentiation, or just differentiation? *Trends Cell Biol.* **21**, 212–218 (2011).
- Sugimoto, K. *et al.* Arabidopsis regeneration from multiple tissues occurs via a root development pathway. *Dev. Cell* **18**, 463–471 (2010).
- Sun, B. *et al.* Timing mechanism dependent on cell division is invoked by Polycomb eviction in plant stem cells. *Science*. **343**, 1248559 (2014).
- Sun, B. *et al.* A timing mechanism for stem cell maintenance and differentiation in the

- Arabidopsis floral meristem. *Genes Dev.* **23**, 1791–1804 (2009).
- Sun, J. *et al.* TCC: an R package for comparing tag count data with robust normalization strategies. *BMC Bioinformatics* **14**, 219 (2013).
- Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
- Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
- Takebe, I. *et al.* Regeneration of whole plants from isolated mesophyll protoplasts of tobacco. *Naturwissenschaften* **58**, 318–320 (1971).
- Tamura, K. *et al.* Blasticidin S deaminase gene (BSD): a new selection marker gene for transformation of *Arabidopsis thaliana* and *Nicotiana tabacum*. *Biosci. Biotechnol. Biochem.* **59**, 2336–2338 (1995).
- Thorvaldsdottir, H. *et al.* Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Br. Bioinform* **14**, 178–192 (2013).
- Uenaka, H. *et al.* Four distinct photoreceptors contribute to light-induced side branch formation in the moss *Physcomitrella patens*. *Planta* **222**, 623–631 (2005).
- Wang, Y. *et al.* An ethylene response factor OsWR1 responsive to drought stress transcriptionally activates wax synthesis related genes and increases wax production in rice. *Plant Mol. Biol.* **78**, 275–288 (2012).
- Weigel, D. & Jurgens, G. Stem cells that make stems. *Nature* **415**, 751–754 (2002).
- Weissman, I. L. Stem cells: units of development, units of regeneration, and units in evolution. *Cell* **100**, 157–168 (2000).
- Wutz, A. Epigenetic regulation of stem cells : the role of chromatin in cell differentiation. *Adv. Exp. Med. Biol.* **786**, 307–328 (2013).
- Yamanaka, S. Elite and stochastic models for induced pluripotent stem cell generation. *Nature* **460**, 49–52 (2009).
- Yang, M. & Jiao, Y. Regulation of Axillary Meristem Initiation by Transcription Factors and Plant Hormones. *Front. Plant Sci.* **7**, 1–6 (2016).
- Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science.* **318**, 1917–1920 (2007).
- Zhang, X. *et al.* Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. *Genome Biol.* **10**, R62 (2009).

- Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
- Zheng, B. & Chen, X. Dynamics of histone H3 lysine 27 trimethylation in plant development. *Curr. Opin. Plant Biol.* **14**, 123–129 (2011).
- Zuo, J. *et al.* Technical advance: An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. *Plant J.* **24**, 265–273 (2000).