The Polycomb group protein CLF emerges as a specific trimethylase of H3K27 regulating gene expression and development in *Physcomitrella patens*

Idan Pereman^{1*}, Assaf Mosquna^{1,9*}, Aviva Katz^{1*}, Gertrud Wiedemann², Daniel Lang², Eva L. Decker², Yosuke Tamada^{3,4}, Takaaki Ishikawa³, Tomoaki Nishiyama^{5,6}, Mitsuyasu Hasebe^{3,4,6}, Ralf Reski^{2,7,8} and Nir Ohad^{1,7,10}.

¹ Department of Molecular Biology and Ecology of Plants, Tel-Aviv University, Tel-Aviv 69978, Israel.

² Plant Biotechnology, Faculty of Biology, University of Freiburg, 79104 Freiburg, Germany.

³ National Institute for Basic Biology, Okazaki 444-8585, Japan

⁴ Department of Basic Biology, The Graduate University for Advanced Studies, Okazaki 444-8585, Japan

⁵ Advanced Science Research Center, Kanazawa University, Kanazawa 920-0934, Japan

⁶ ERATO, Okazaki 444-8585, Japan

⁷ FRIAS – Freiburg Institute for Advanced Studies, 79104 Freiburg, Germany.

⁸ BIOSS – Centre for Biological Signaling Studies, 79104 Freiburg, Germany.

⁹ Robert H. Smith Institute of Plant Sciences and Genetics in Agriculture, Faculty of Agriculture, Hebrew University of Jerusalem, Rehovot 7610001, Israel.

¹⁰ The Manna Center Program for Food Safety & Security, Tel Aviv University, 69978 Israel.

* These authors have contributed equally to the work.

Corresponding Authors Nir Ohad (N.O.) niro@tauex.tau.ac.il.
and Ralf Reski (R.R.) ralf.reski@biologie.uni-freiburg.de

Key words:

Curly leaf, histone code, *Physcomitrella patens*, SET domain, H3K27 methylation, epigenetic regulation.

Abstract

Packaging of eukaryotic DNA largely depends on histone modifications that affect the accessibility of DNA to transcriptional regulators, thus controlling gene expression. The Polycomb group (PcG) chromatin remodeling complex deposits a methyl group on lysine 27 of histone 3 leading to repressed gene expression. Plants encode homologs of the Enhancer of zeste (E(z)), a component of the PcG complex from *Drosophila*, one of which is a SET domain protein designated CURLY LEAF (CLF). Although this SET domain protein exhibits a strong correlation with the presence of the H3K27me3 mark in plants, the methyl-transferase activity and specificity of its SET domain has not been directly tested *in-vivo*.

Using the evolutionary early-diverged land plant model species *Physcomitrella patens* we show that abolishment of a single copy gene *PpCLF*, as well as an additional member of the PcG complex, *FERTILIZATION-INDEPENDENT ENDOSPERM* (*PpFIE*), results in a specific loss of tri-methylation of H3K27. Using site-directed mutagenesis of key residues, we revealed that H3K27 tri-methylation is mediated by the SET domain of the CLF protein. Moreover, the abolishment of H3K27me3 led to enhanced expression of transcription factor genes. This in turn led to the development of fertilization-independent sporophyte-like structures, as observed in *PpCLF* and *PpFIE* null mutants. Overall, our results demonstrate the role of *PpCLF* as a SET protein in tri-methylation of H3K27 *in-vivo* and the importance of this modification in regulating the expression of transcription factor genes involved in developmental programs of *P. patens*.

Introduction

Post-translational modifications of histones, the proteinous components of eukaryotic nucleosomes, affect the accessibility to DNA by transcriptional regulators, thus contributing to the control of gene expression. To date, combinations of multiple types of post-translational modifications of residues in all of the five histone classes are postulated to mediate an epigenetic control of gene regulation by implementing the so called “histone code”¹. Modifications comprise methylation, acetylation, phosphorylation, ubiquitination and ribosylation of specific amino acid residues². Although the complexity of this code is substantially multiplied by coaction of different modification types affecting the same nucleosome or even the same histone,

the regulatory effect of these modifications on transcription may result in activation or repression.

The Polycomb group (PcG) genes, first described in *Drosophila melanogaster*, encode for proteins which epigenetically regulate developmental programs through gene repression in animals and in plants^{3,4}. At least three distinct multi-subunit Polycomb Repressive Complexes (PRCs) were identified in animals: Polycomb Repressive Complex 2 (PRC2), Polycomb-like PRC2 (Pcl-PRC2) and Polycomb Repressive Complex 1 (PRC1)^{5,6}. In animals, gene expression is regulated by methylation of H3K27 (H3K27me) mediated by PRC2 and related Pcl-PRC2⁷⁻¹¹. In plants, four core protein components of PRC2 were identified *via* genetic and biochemical analyses. In the flowering plant *Arabidopsis thaliana*, these components comprise: (i) a WD-40 protein homologous to the *Drosophila Extra Sex Combs (ESC)*¹², encoded by *Fertilization-Independent Endosperm (FIE)*^{13,14}; (ii) a C2H2 type zinc finger protein homologous to the *Drosophila SU(Z)12*¹⁵, encoded by either *Embryonic Flower 2 (EMF2)*, *Vernalization 2 (VRN2)*¹⁶ or *Fertilization Independent Seed 2 (FIS2)*¹³; (iii) a p55 homolog containing a WD-40 motif designated *Multicopy Suppressor of Ira 1 (MSI1)*¹⁷; and (iv) a E(z) homolog containing a SET domain (Su(var) E(z) Thritorax) domain, encoded by either *Curly Leaf (CLF)*¹⁸, *Medea (MEA)*^{15,19,20} or *Swinger*^{21,22} (SWN). Mutants of either *AtFIE*, *AtCLF*, *AtMEA* or *AtSWN* result in the loss of tri-methylation of H3K27 (H3K27me3)²³⁻²⁷. As the correlation between the genome-wide occurrence of H3K27me and the presence of the E(z) homolog was observed in all eukaryotic organisms tested so far²⁸, it is feasible to assume that PRC2 originated early in the evolution of eukaryotes²⁹.

A growing body of evidence demonstrates the correlation between the PRC2 complex and H3K27 methylation in plants^{30,31}. The two human EZH paralogs and the fly E(z) were annotated to catalyze the methylation of H3K27^{7,32}. Furthermore, *Arabidopsis* orthologs CLF, SWN and MEA are regarded as catalyzers of tri-methylation of H3K27 specifically. However, systematic experimental evidences for this activity including the identity of the functional sites of the methyl-transferase activity of either plant or animal homologs are still lacking. The correlation between PRC2 complex presence, histone methylation, gene expression and development as well as responses to environmental cues is well established for flowering plants³³. However, the evolutionary ancestral function and origin of the PRC2 complex in land plants has not been determined yet. The genome of the moss *P. patens*, a

representative of early-diverged land plants, encodes all four components of the PRC2, including one E(z) ortholog containing SET domain (PpCLF)^{29, 35}. Furthermore, a correlation between gene repression and the presence of the H3K27me3 mark has already been demonstrated in *P. patens*³¹. These genes were mainly addressed as developmental genes as could also be observed in the loss-of-function mutants of FIE and CLF^{34,35}. Previously, we have shown that CLF and FIE can interact³⁴. Moreover, based on their mutant phenotype as well as on complementation assays, it has been suggested that both proteins function as PRC2 proteins³⁴. Both $\Delta PpFIE$ and *ppclf-del* mutants in *P. patens* have a similar phenotype - a deficiency in gametophore development. Strikingly, both mutants develop sporophyte-like structures despite the absence of fertilization^{34,35}. This phenotype can be explained by the known function of PcG proteins in governing differentiation programs²⁹.

PpCLF, being a single SET domain gene, together with the moss's unique competence for gene targeting by homologous recombination³⁶, provides an ideal tool to study the putative role of the PpCLF-SET domain methyltransferase activity in PRC2. Here, we applied *in-vivo* site-directed mutagenesis analysis of the PpCLF SET domain by specifically substituting single amino acids of key residues within the SET domain. Analysis of *P. patens* CLF mutants demonstrated a global loss of tri-, but not mono- or di-, H3K27 methylation. In addition, we observed an up-regulation of transcription factor genes, including members of the Homeobox family. Taken together, these data indicate the enzymatic methyltransferase function of the PpCLF SET domain in the moss PcG complex and its role in regulating the gametophyte-to-sporophyte transition by gene expression regulation.

Results

H3K27me3 modification is absent in $\Delta PpCLF$ and $\Delta PpFIE$ mutants.

To assess the contribution of CLF for the genome-wide levels of H3K27me we first aimed to monitor the overall H3K27 methylation status in *PpCLF* and *PpFIE* deletion mutants. To this end nuclei-enriched extracts from wild type, $\Delta PpCLF$ and $\Delta PpFIE$ deletion mutants were subjected to Western blot analysis, using antibodies specific for H3K27me3, H3K27me2 and H3K27me1, respectively. The H3K27me3 mark was completely lost in $\Delta PpCLF$ as well as in $\Delta PpFIE$ mutants when compared

to wild type, emphasizing the necessity of these proteins for global tri-methylation of H3K27 in moss. These results also highlight the role of PpFIE and PpCLF in mediating H3K27 tri-methylation specifically as H3K27me2 and H3K27me1 modifications remained unchanged (Fig. 1).

The loss of H3K27me3 in $\Delta PpCLF$ and $\Delta PpFIE$ results in the up-regulation of PRC2 target gene expression.

To test for CLF and FIE effects on H3K27me3 levels, we first identified PcG target genes in *P. patens*. To this end, we compared global gene expression in wild type and $\Delta PpFIE$ plants using a whole-genome microarray. Significantly up-regulated genes in the $\Delta PpFIE$ mutant were determined at 95% confidence. Putative PRC2-regulated genes were further compared with the result from a combined set of genes resulting from differential expression analysis by CAPSeq (see DDBJ DRA database: files DRR055530-DRR055535) of a $\Delta PpCLF$ mutant and H3K27me3 ChIPseq (see DDBJ DRA database, files: DRR059713-DRR059717 at http://trace.ddbj.nig.ac.jp/dra/index_e.html). Out of the list of genes overlapping these two sets, we chose to focus on transcription factors as defined previously on a genome-wide scale for *P. patens*³⁷ as these may contribute directly to the developmental phenotypes observed for *ppclf-del* and $\Delta PpFIE$ ^{34,35}.

This procedure resulted in the identification of three homologs of the *Arabidopsis* transcription factors *BELL1* (AT5G41410), *AIB* (AT2G46510) and *JAGGED* (AT1G68480). The homologous genes from *P. patens* were designated *BELL2* (Pp1s220_67V6.1)³⁸, *AIB* (Pp1s57_226V6.1) and *JAGGED1* (Pp1s210_60V6.1), respectively. Searching the *P. patens* genome further, we identified paralogs for *BELL2* and *JAGGED1* and designated them *BELL1* (Pp1s258_6V6.1)³⁸ and *JAGGED2* (Pp1s192_8V6.1), respectively.

Further analysis employing quantitative RT-PCR confirmed that the expression of these five genes is up-regulated in the $\Delta PpCLF$ and $\Delta PpFIE$ mutants (Fig. 2), thus identifying them as PRC2-regulated genes.

To test for a possible correlation between the aberrant expression and the methylation status, we determined H3K27me3 modification levels in these five genes, using a ChIP-qPCR assay. The level of enrichment for H3K27me3 was tested for each gene in wild type and mutant background. For all five genes we found that H3K27me3 was almost abolished in $\Delta PpCLF$ and $\Delta PpFIE$ lines as compared to wild

type (Fig. 3). Thus, for all tested PRC2 target genes a positive correlation exists
between the loss of H3K27me3 mark and up-regulation of their expression (Fig. 3).

PpCLF mediates global H3K27 tri-methylation

To determine whether PpCLF has histone lysine tri-methyltransferase (HKTMT)
activity, we employed an *in-vivo* site-directed mutagenesis approach. For this we
aligned the PpCLF sequence with its orthologs from *Arabidopsis* AtCLF and human
HsEZH2, as well as with *Paramecium bursaria* Chlorella Virus-1 (PBCV-1) vSET,
revealing highly conserved amino acids within the SET domain (Fig. 4).

Specifically, invariant amino acid sites were identified within conserved sites (The
sites names were adapted from Qian *et al.*³⁹ which in turn designated them according
to their role in vSET.): 'multiplicity site', 'cofactor binding site' and 'active site' (Fig.
4). Of note, point mutations in these vSET sites hampered the HKTMT activity to
different degrees^{39,40}.

To test for the contribution of these particular residues in PpCLF enzymatic activity,
transgenic moss plants were generated (two independent lines for each mutation, as
described in the material and methods section), in which these residues were replaced
(highlighted in black in Fig. 4). The lines were designated according to the point
mutation which was introduced: *F905A*, *F905Y*, *NH926AA* and *Y964F*. These
residues were selected based on their conserved position in relation to the well-
defined crystal structure of the vSET protein from PBCV-1³⁹. Mutation at
corresponding amino acid residues in the vSET protein dramatically reduced the
methyltransferase activity with the exception of mutating residue 52 F-to-Y (F52Y)
that showed only a mild effect. In our study, as a control we have generated a
transgenic line containing the wild type *CLF* gene and the same resistance cassette
which was used in all other mutant lines, inserted down-stream to the *CLF* gene,
designated Control-CLF. To confirm the correct integration of each mutation into
PpCLF, a PCR followed by an RFLP analysis was carried out (Supplemental Fig. 2
and 3C).

To determine the effects of the mutations on global H3K27me3 levels, Western blot
analysis was carried out on enriched nuclear protein extracts prepared from wild type

and mutants, using antibodies directed against H3K27me3 (Fig. 5A). In comparison to wild type and the control line containing wild type *PpCLF* and the transgenic cassette (Fig. 5A Control-CLF), all mutant lines, except for line *F905Y*, showed a complete loss of the mark. The mutation of *F905Y*, which differ by a single hydroxyl group, was shown to only partially reduce the HKTMT activity similar to the vSET F52Y mutation³⁹, as opposed to line *F905A* (in which a methyl group is replaced by the alanine residue in the same position). Thus, these results indicate structural and functional similarities between PpCLF and vSET.

PpCLF binds to PpFIE WD40 protein³⁴, which is required for the HKTMT activity of PRC2. To rule out the possibility that the above mentioned SET mutations may affect the ability of CLF to bind to FIE, resulting in an indirect disruption of the HKTMT activity, we tested whether each mutant protein can interact with FIE *in planta*. To this end, we used a BiFC (Bimolecular Fluorescence Complementation) assay³⁴, which was used to monitor the interactions between wild type PpCLF and PpFIE³⁴ *in planta*. Wild type and mutant PpCLF proteins were co-expressed fused to the N-terminal half of the YFP protein (YN) together with PpFIE fused to the C-terminal half of YFP (YC). Reconstitution of YFP fluorescence in the nucleus was observed for wild type PpCLF as well as for all PpCLF mutants (Fig. 5B), indicating that amino acid changes did not hamper FIE-CLF interaction rather specifically disrupted the CLF HKTMT activity.

Point mutations in the CLF-SET domain affect H3K27me3 modification and subsequently expression of target genes

Our ChIP-qPCR and RT-qPCR data obtained from both $\Delta PpCLF$ and $\Delta PpFIE$ lines demonstrated a positive correlation between loss of H3K27me3 and increase in *BELL1*, *BELL2*, *JAGGED1*, *JAGGED2* and *AIB* expression (Fig. 2). If PpCLF is indeed an HKTMT, then one should expect residues within the SET domain, described above, to facilitate tri-methylation of H3K27. To test this directly, we analyzed H3K27me3 levels for the target genes in the background of SET domain point mutation lines applying ChIP-qPCR. For all tested genes, H3K27me3 was found to be significantly reduced in NH926AA, Y964F and F905A CLF-SET mutants (Fig. 6A). H3K27me3 was mildly reduced for genes tested in the F905Y line. This result is in line with the observation that the HKTMT activity for PpCLF harboring this mutation was intact.

To test for a putative correlation between SET domain mutations, loss of H3K27me3 and target gene expression, RT-qPCR analysis was performed in the background of the PpCLF-SET mutants. This analysis showed that all tested genes were up-regulated in NH926AA, Y964F, and F905A (Fig. 6B). Thus, a positive correlation exists between the loss of the H3K27me3 modification and an up-regulation of the target genes, revealing that H3K27 trimethylation by PpCLF is involved in the control of gene expression.

Mutations in the CLF SET domain lead to the development of fertilization-independent sporophyte-like structures.

To determine the relation between the particular residues mutated in the CLF SET domain and their potential effect on moss development, we examined the growth and morphology of the mutant plants, *F905A*, *F905Y*, *NH926AA*, *Y964F* and Δ *PpCLF*, as compared to wild type.

To apply a comparable developmental starting point, protoplasts from wild type and each mutant were generated and cultured for four weeks, allowing the growth of protonema filaments and gametophores. Mutations within each of the methylation multiplicity site, cofactor-binding and active sites as well as the deletion of the complete *CLF* sequence (Δ *PpCLF*) resulted in small plants (Fig. 7 B-D, F) as compared to wild type (Fig. 7 A) with a denser pattern of the protonema (Fig. 7 H-J, L) as compared to wild type (Fig. 7 G). Additionally, lines Δ *PpCLF*, *F905A*, *NH926AA* and *F964Y* developed abnormal buds, forming sporophyte-like bodies (Fig. 7 N-P, R), as previously described for Δ *PpFIE* and *ppclf-del* deletion mutants^{34,35}. Line *F905Y* exhibited only minor developmental defects (Fig. 7 E, K and Q), as compared to line *F905A* (Fig. 7 F, L and R).

Discussion

Studies of the Polycomb complex in plants have demonstrated that it has a pivotal role as a regulator of developmental genes mediated through an epigenetic memory system⁴¹. We have previously shown that the mutation of Polycomb-homologous genes in moss lead to an abnormal transition of the gametophyte to sporophyte phase^{34,35}. It was thus intriguing to test the contribution of key amino acids within the SET domain

of PpCLF to the overall methyl-transferase activity and further determine developmental consequences.

In this study we have shown *in-vivo* that Δ PpFIE, Δ PpCLF and CLF SET domain mutant plants have lost the H3K27me3 mark completely. These results demonstrate the central role of PRC2 in mediating this epigenetic mark, and more specifically the role of the CLF SET domain, in placing this modification. Moreover, our results show a correlation between H3K27me3 levels, expression of PRC2-targeted transcription factor genes and regulation of moss development.

In *Arabidopsis*, representing flowering plants, AtCLF was shown to have specificity for H3K27 tri-methylation. H3K27 di-methylation was also partially decreased in an *AtCLF* mutant²³. In addition, it was described that H3K27me2 is partially lost in *Arabidopsis* lacking the AtFIE PcG protein²⁵. Together, these studies indicate a role of the PcG complex in H3K27me2.

Unlike in *Arabidopsis* our results show that Δ PpCLF and, moreover, point mutations in the PpCLF-SET domain, specifically abolish tri-methylation but not di- or mono-methylation at H3K27. Thus, the function of the PRC2 towards tri-methylation of H3K27 was conserved throughout evolution of land plants, while the ability to di-methylate H3K27, as observed in flowering plants, was probably gained by PRC2 members later in plant evolution. In view of the fact that *CLF* is a single copy gene in *P. patens*, we speculate that additional proteins such as a non-PRC2 SET domain protein may promote di- or mono- methylation of H3K27 in moss. Of note, a PRC2-independent enzyme (ATXR6) has been identified in *Arabidopsis* as a mediator of H3K27me1⁴² with specificity to the histone 3 variant H3.1, hence implying a specific role for ATXR6 in heterochromatin replication regulation. *P. patens* harbors an ATXR6 homolog which is a SET-domain protein (Pp1s403_14V6.1), with 58% sequence identity in the conserved SET domain region. It is possible that the product of this gene exerts a similar function in moss.

In this study we showed that the CLF protein is specifically responsible for H3K27me3. In the absence of an *in vitro* system that monitors HKTMT activity of purified CLF protein, we cannot exclude the possibility that other CLF interactor(s) possess the HKTMT activity. Yet, the biochemical evidence obtained from purified vSET protein demonstrating such activity^{39,40}, together with the full correlation

between the loss of HKTMT activity of vSET and PpCLF mutant plants containing
analogous mutations in the SET domain, strongly suggests that PpCLF is the enzyme
that directly tri-methylates H3K27. The poikilohydric, non-vascular plant *P. patens*
represents an intermediate evolutionary stage between aquatic green algae and
terrestrial vascular plants. In view of the structural and functional similarities
demonstrated here between vSET and PpCLF, it is intriguing to observe the study of
Mujtaba *et al.* which shows how vSET can successfully repress gene expression of
the unicellular eukaryotic green algae *Chlorella*, most likely by mimicking the
endogenous function of a E(z) protein homolog⁴³. Together, these observations
highlight the evolutionary conservation of the E(z) family throughout green plant
evolution.

Overall, our results show a full correlation between the mutations in the CLF SET
domain, in respect to H3K27me3 modification and gene expression, and their effects
on plant development. We have shown a role for PRC2 function in determining tissue
identity^{34,35}. Our current study presents a link between PRC2 and its molecular
function, in mediating tri-methylation of H3K27, and the transition between
gametophytic and sporophytic phases. Thus, these correlations highlight the early
assignment of PRC2 as a regulator of phase transitions during plant development. In
order to test the CLF methylase activity, we have identified novel CLF target genes
that are regulated by H3K27me3. These target genes include orthologs of known
transcription factors, which regulate a variety of developmental processes in
Arabidopsis. Interestingly, among these targets we identified members of the BELL
proteins which are known to exert their function as heterodimers with KNOX
homeodomain proteins. Lee *et al.*⁴⁴ have shown that ectopic expression of a BELL-
related protein in *Chlamydomonas* can initiate the formation of a zygote without
gametes fusion while Sakakibara *et al.*³⁸ have shown that a member of the class 2
KNOX (KNOX2) subfamily takes part in repressing the transition from sporophytic
to gametophytic phases in *P. patens* as the deletion of KNOX2 genes (*mkn1 mkn6*)
resulted in development of gametophytic tissue from a diploid embryo without
meiosis. BELL1 which was up-regulated in the mutants tested in this study was found
to be a regulator of the transition from gametophytic to sporophytic phase, most likely
via the interaction with a KNOX member yet to be identified⁴⁵.

An additional gene which was found in this study is a homolog of *AtJAGGED* which encodes a zinc finger transcription factor⁴⁶. *AtJAGGED* was described as a regulator of growth rates in a particular region of the petal while also modulating tissue polarity⁴⁷. Furthermore, *JAGGED* was identified to serve as a key regulator of the transition from meristem to primordium identity, and also to bind directly and inhibit the expression of *AtBELI*⁴⁶. A homologue of *AtBELI* was found in this study to be targeted by Polycomb. This in turn may indicate an evolutionary conserved regulation network involving these two Polycomb-dependent genes, which in turn interact in order to mediate meristematic-tissue fate identity.

Altogether, by exhibiting an evolutionary conserved role for CLF protein in H3K27 tri-methylation and its role in regulating gene expression, this work further establishes the origins of the Polycomb complex as a pivotal developmental orchestrator that seems to appear early during land plant evolution.

Material and Methods

Plant material, culture conditions and treatments

The 'Gransden 2004' strain of *P. patens* which was used for genome sequencing⁴⁸ was propagated on BCD and BCDAT media⁴⁹ at 25°C under a 16-hour light and 8-hour dark cycle⁵⁰.

Expression analysis using microarrays

Gene expression in wild type and $\Delta PpFIE$ plants was compared using whole genome microarrays (CombiMatrix, Mukileto, WA, USA) based on all gene models v1.2⁴⁸ as described⁵¹. RNA samples were obtained from protonema grown for three days on solid medium and then recovered after scission. This procedure was repeated three times to assure the presence of mainly protonema and as little as possible of young gametophore buds. Total RNA was extracted using the "SV-TOTAL RNA Isolation System" (Promega, Madison, USA). 1.5 µg of total RNA were used for amplification and subsequently 5 µg of amplified RNA labelled with Cyanin-5 according to the manufacturer's instructions (ampULSE, Kreatech, Amsterdam, The

Netherlands). Microarray hybridization was performed following the manufacturer's instructions using 5 µg of fragmented, labelled RNA. For imaging a laser scanner (Genepix 4200A, Molecular Devices, Ismaning / Munich, Germany) was used. The microarray experiments were done in biological triplicates. All analyses were carried out based on the latest version of the moss genome annotation V1.6⁵². The array analysis was carried based on the probes as described previously⁵¹. As the array design was based on a previous version of the genome V1.2, each of the gene-wise probes was mapped to the V1.6. Signals for mapped probes were normalized using quantile normalization as implemented in the Bioconductor R package "preprocessCore" (Benjamin Milo Bolstad. preprocessCore: A collection of pre-processing functions. R package version 1.16.0) and log2 transformed. Based on the normalized signal of negative controls also present on the array, probes with signals below twice the standard deviation of the negative controls were excluded from further analysis. Of the remaining probes, for each transcript the 5 probes with the strongest signals were used as input for a Cyber-T comparison⁵³ of wild type and mutant lines. Transcripts up-regulated in the mutants at 95% confidence (Supplementary table 1) were used to compare to the sets of H3K27me3-CLF-targets determined by the comparison of the expression profiles between the protonema of wild-type and *ppclf-del* lines (DRA accession number listed in Supplementary Table 2) and TAP proteins identified previously³⁷.

Link for genome browser track: "Widiez et al. 2014 histone 3 code MACS peaks":

https://www.cosmoss.org/fgb2/gbrowse/physcome/?start=760300;stop=765799;ref=scaffold_192;width=1024;version=100;flip=0;grid=1;id=2bd55c04ee7b3f7a1521467a90250f29;l=Phypa1_6_coding%1EWidiez_Histone.TRACK%1EA_Gap%1Ecosmos_user_models%1EEvidences_EST%1EEvidences_FLcDNA_gth%1ELTR_angela%1EGap%3Aoverview%1EPHYPAV16%3Aoverview%1EPHYPAV16f%3Aoverview%1EGap%3Aregion%1EPHYPAV16%3Aregion%1EPHYPAV16f%3Aregion;h_feat=pp1s192_94v6.1%40yellow.

Constructing moss $\Delta PpFIE$ knock out lines 391

All mutant plants were generated in the background of 'Gransden 2004' strain of *P. patens*⁴⁹ and deposited at the International Moss Stock Centre IMSC (www.moss-stock-center.org; Supplementary Table 2). 392
393
394

The 5' (805 bp) and 3' (598 bp) of the untranslated regions (UTR) of *PpFIE* were 395
amplified using the following primers: PpFIE5'-UTR-Fw, 5'- GAA GCT TGA CTA 396
GAG CAA AAA AAT TGT GAT AGT GTG T -3' and PpFIE5'-UTR-Rv, 5'- GAA 397
GCT TCA CGG GAT CCG TGC CGA. PpFIE3'-UTR-Fw, 5'-GCA TGC TGA TCG 398
TGG ATA TCT GGA GCC A- 3' and PpFIE3'-UTR-Rv, 5'- GCA TGC CGC TAG 399
GGT CAT AGC CAT ATA AAC A-3'. All amplified fragments were sub-cloned 400
into the pTZ 57 vector (Fermentas) and sequenced to ensure their integrity. 401
Following, the 5' UTR was cloned into the pMBL5 vector⁵⁴ in the *HindIII* site. The 402
resulting vector was then linearized with *SphI* and the 3' UTR of *FIE* was cloned into 403
the same site resulting in the pMBL5 $\Delta PpFIE$ knock out vector. Proper integration at 404
the 5' region was validated with primers: PpFIE5'-6-UTR-Fw, 5'-GCA AGC GCA 405
GGG CAC AGT -3' and Nptii-KO-Rv, 5'-TCG ATC AGG ATG ATC TGG A-3'. 406
Proper integration at the 3' region was validated with primers: Nptii-KO-Fw, 5'-CTA 407
TCG CCT TCT TGA CGA G-3' and PpFIE3'-UTR-Rv, 5'-GAA AGG TGG CAT 408
GTT CAT CAA GAT GTT TG-3 (Supplemental Fig.1A and 3A). 409

Constructing moss $\Delta PpCLF$ knock out lines 410

The 5' (546 bp) and 3' (561 bp) of CLF untranslated region (UTR) were amplified 411
using the following primers: CLF5'-UTR-Fw, 5'-GAA TTC ATG AAC TGT TTG 412
ACT CAA CGT GTG-3' and CLF5'-UTR-Rv, 5'- GTC GAC CCA GTT CTC TTT 413
CCT ATC TGC AG -3' . CLF3'-UTR-Fw, 5'- GCA TGC TGT ACT CGA ATT GGC 414
GGT TAC-3' and CLF3'-UTR-Rv, 5'-GAT ATC GTT TAC TGG ATG TTA TGC 415
GAA TG-3'. The two UTRs were sub-cloned into the pTZ 57 vector (Fermentas) and 416
sequenced to ensure their integrity. Following, the 5' UTR was cloned into the 417
pMBL5 vector⁵⁴ in *SalI* and *EcoRV* sites. The resulting vector was then linearized 418
with *SphI* and the 3' UTR of *CLF* was cloned into the same site resulting in the 419
pMBL5 $\Delta PpCLF$ knock out vector. Proper integration at the 5' region was validated 420
with primers: CLF-KO1-Fw, 5'-GTG ATC CTT GCC GAA ATC TAC G-3' and 421

Nptii-KO-Rv, 5'-TCG ATC AGG ATG ATC TGG A-3'. Proper integration at the 3' region was validated with primers: NptII-KO-Fw, 5'-CTA TCG CCT TCT TGA CGA G-3' and CLF-KO1-Rv, 5'- AGC CAC TGT TTT GGA CAT GG-3'. (Supplemental Fig. 1B and 3B)

Genomic CLF-SET knock-in vectors

A PpCLF 3' genomic DNA fragment was amplified with primers: PpCLF HIII Fw, 5'-AAG CTT ATG CGG CCA TCA ATT CGT ATC -3' and ppCLF stop HIII Rv, 5'-AAG CTT TTA AGC AAC TTT CTG TGC TCG TCC AC-3'. The fragment was then cloned to the pTZ vector followed by sequencing. This fragment was used as a template for the introduction of various point mutations in the SET domain using the "Quickchange" site directed mutagenesis kit (Stratagene) verified by sequencing. In addition, a silent mutation which resulted in a restriction site was included near each site mutated to be used later for identifying proper integration of the desired mutation to the genome. The primers which were used to mutate particular amino acid residues in the SET domain and introduce new restriction site are listed below, where the restriction site is underlined and the introduced mutation is marked in bold:

Y964F including *AfaI* restriction site -

PpCLF SET-Ac Fw, 5'-AGA ACT ATT CTA TGA CTT CCA GTA CGA GCC TGA TAG AGC G-3' PpCLF SET-Ac Rv, 5'- CGC TCT ATC AGG CTC GTA CTG GAA GTC ATA GAA TAG TTC T-3'

F905Y including *DraI* restriction site-

PpCLF SET-Y Fw, 5'-GAA TTC GTC GTT CCT CTA CAA TTT AAA TGA TCA GGC AAG CCT C-3' PpCLF SET-Y Rv, 5'-GAG GCT TGC CTG ATC ATT TAA ATT GTA GAG GAA CGA CGA ATT-3'.

F905A including *DraI* restriction site-

PpCLF SET-A Fw, 5'-GAA TTC GTC GTT CCT CGC CAA TTT AAA TGA TCA GGC AAG CCT -3', PpCLF SET-A Rv, 5'-GAG GCT TGC CTG ATC ATT TAA ATT GGC GAG GAA CGA CGA ATT C-3'.

	454
NH926AA including <i>NotI</i> restriction site-	455
PpCLF SET-CF Fw, 5'-GGC GAC AAG CTC AAA TTT <u>GCG GCC GCC</u> TCA	456
CCT ACT CCT ACT CCC AAC TGC-3' PpCLF SET-CF Rv, 5'- GCA GTT GGG	457
AGT AGG TGA <u>GGC GGC CGC</u> AAA TTT GAG CTT GTC GCC-3'.	458
Each of the mutated PpCLF-SET fragments described above and the 3' UTR of	459
<i>PpCLF</i> (described above in the CLF knock-out paragraph), were cloned into the	460
pMBL5 in the <i>HindIII</i> and <i>SphI</i> sites respectively, creating the pMBL5 PpCLF SET	461
Y964F, F905Y, F905A, NH926AA vectors.	462
The corresponding restriction site which was cloned together with each point	463
mutation was used in a RFLP analysis to ensure proper construct integration in each	464
of the above mutant line (Supplement Fig. 2 and 3)	465
Two independent lines were generated and analyzed for each CLF-SET mutation.	466
	467
Validation of PpCLF-SET mutant lines integrity	468
In order to validate the presence of the correct mutation in each mutated PpCLF gene,	469
a restriction analysis was performed allowing detecting particular restriction products	470
as indicated in Supplemental Fig. 2 which results from silent mutations adjacent to the	471
mutated sites. Transcripts from two independent transgenic lines of each mutant were	472
isolated and following RT-PCR the presence of the correct mutated transcript was	473
verified by the appearance of pattern of specific restriction products. In-addition,	474
qRT-PCR analyses showed that the mutagenesis did not result in mRNA instability	475
(data not shown).	476
	477
Construction of <i>CLF</i>-SET knock-in vectors for BiFC assay	478
<i>PpCLF</i> coding sequence (3003 bp) was amplified in three parts: part A (1-1067) part	479
B (1062-2052) part C (2047-3003). The construction of PpCLF cDNA partition was	480
performed using the following primers:	481
Part A	482
PpCLF cDNA-A-Fw 5'-ACT AGT GGA TCC ATG GCG TCC TCC AGC TAC	483
TAC GCC A-3', PpCLF cDNA-A-Rv 5'- TTG CTG CAA GGA CTA GGG TCT TC	484
-3'.	485
	486

Part B	487
PpCLF cDNA-B-Fw 5'-GTC CAT CCA AGT GAA AGG CAA C-3' PpCLF cDNA-	488
B-Rv 5'-GCA GCA TCT AGT TCC ATT GCG -3'.	489
Part C	490
PpCLF cDNA-C-Fw 5'-AAG GCT GCA AGA CGT GTG CAG -3' PpCLF cDNA-	491
C-Rv 5'-GTC GAC TAG TTA CTC GAG AGC AAC TT-3'.	492
All amplified fragments were sub-cloned into the pTZ 57 vector and sequenced	493
(designated as pTZ-A, pTZ-B and pTZ-C). Fragment B was digested and inserted to	494
pTZ A into the <i>NcoI</i> and <i>SalI</i> sites giving rise to pTZ AB. Then, fragment “C” was	495
inserted into the <i>BalI</i> and <i>SalI</i> sites of pTZ AB, giving rise to pTZ PpCLF cDNA. In	496
turn, the SET domain of <i>PpCLF</i> cDNA was mutated (as described above for the	497
PpCLF knock-in lines) using the “Quickchange” site directed mutagenesis kit	498
(Stratagene) using the following primers where the restriction site is underlined and	499
the introduced mutation is marked in bold:	500
Y964F including <i>AfaI</i> restriction site -	501
PpCLF SET-Ac Fw, 5'-AGA ACT ATT CTA TGA CTT CCA <u>GTA</u> CGA GCC	502
TGA TAG AGC G-3'. PpCLF SET-Ac Rv, 5'- CGC TCT ATC AGG CTC <u>GTA</u>	503
<u>CTG</u> GAA GTC ATA GAA TAG TTC T-3'	504
F905Y including <i>DraI</i> restriction site-	505
PpCLF SET-Y cDNA Fw, 5' –GGG AGA ATT CGT CGT TCC TCT ACA <u>ATT</u>	506
<u>TAA</u> <u>ATG</u> ATC AGT ATG TGT GCT GG AC-3'. PpCLF SET-Y cDNA Rv, 5'-	507
GTC CAG CAC ATA CTG ATC <u>ATT TAA</u> <u>ATT</u> GTA GAG GAA CGA CGA	508
ATT CTC CC -3'.	509
F905A including <i>DraI</i> restriction site-	510
PpCLF SET-A cDNA Fw, 5'-GGG AGA ATT CGT CCG TCG TTC CTC GCC	511
<u>AAT TTA</u> <u>AAT</u> GAT CAG TAT GTG TGC TGG AC -3', PpCLF SET-A cDNA Rv,	512
5'- GTC CAG CAC ATA CTG ATC <u>ATT TAA</u> <u>ATT</u> GGC GAG GAA CGA CGA	513
ATT CTC CC -3'.	514
NH926AA including <i>NotI</i> restriction site-	515
PpCLF SET-CF Fw, 5'-GGC GAC AAG CTC AAA TTT <u>GCG</u> <u>GCC</u> <u>GCC</u> TCA	516
CCT ACT CCT ACT CCC AAC TGC-3' PpCLF SET-CF Rv, 5'- GCA GTT GGG	517
AGT AGG TGA <u>GGC</u> <u>GGC</u> <u>CGC</u> AAA TTT GAG CTT GTC GCC-3'.	518
All clones were verified by sequencing. Wild type and mutated <i>PpCLF</i> cDNA were	519
cloned into the pSY735 vector through the <i>SpeI</i> site.	520

Monitoring protein-protein interaction via bimolecular fluorescence	521
complementation assay	522
Protein-protein interactions <i>in planta</i> were examined by bimolecular fluorescence	523
complementation (BiFC) assay ⁵⁵ . PpFIE full-length cDNA was cloned into the pSY	524
736 vector at the <i>SpeI</i> site, PpCLF full-length mutated cDNAs were cloned into the	525
pSY 735 vector which contains the C-terminal (YC) fragment of the YFP protein ⁵⁵ .	526
Equal concentrations of <i>Agrobacterium tumefaciens</i> strain GV3101/pMp90	527
containing plasmids of interest were transiently co-expressed in <i>Nicotiana</i>	528
<i>benthamiana</i> leaves via the leaf injection procedure. Image annotation was	529
performed with Zeiss AxioVision, Zeiss CLSM-5 and Adobe Photoshop 7.0	530
(Mountain View, CA, USA). Negative controls with vectors bearing only YN or YC	531
alone were carried out in order to verify the specificity of the interactions.	532
	533
Protoplast isolation and PEG-mediated transformation of <i>P. patens</i>	534
PEG transformation was performed as described in Nishiyama <i>et al.</i> ⁵⁴ PHYSCObase	535
(http://moss.nibb.ac.jp). Six days after regeneration, transformants were selected on	536
BCDAT medium containing 20 mg/l of G418.	537
	538
Chromatin Immunoprecipitation (ChIP).	539
	540
The ChIP assays protocol was kindly provided by T. Widiez and performed as	541
previously described in Berr <i>et al.</i> ⁵⁶ with modifications. Following one week of	542
cultivation, protonema tissue of <i>P. patens</i> was homogenized and re-plated. This	543
tissue was further homogenized twice more in 4 days' intervals. Samples were	544
harvested and dried completely using Whatman paper, frozen in liquid nitrogen and	545
stored in -80C° until chromatin isolation. 500 mg of the sample's tissue was ground	546
to fine powder using a mortar and pestle and liquid nitrogen. The powder was placed	547
in a 50 ml tube to which 15 ml of Nuclear Isolation Buffer I was added in room	548
temperature (10 mM HEPES pH 7.5, 0.4 M Sucrose, 5m M KCl, 5 mM MgCl ₂ , 5	549
mM EDTA, 0.6% Triton X-100). The following solutions were added just prior to	550
use: β-mercaptoethanol to a final concentration of 0.1%, PMSF to a final	551
concentration of 1 mM, cOmplete protease inhibitor (Roche, USA), according to the	552
manufactures recommendations and 0.2ml formaldehyde 37%. Samples were	553
sonicated (Misonix, sonicator s-4000) at 20% capacity in 6 cycles of 3 seconds active	554

sonication a 20 seconds break. Chromatin was immunoprecipitated with specific antibodies (H3K27me3 Millipore (07-449), Lot: JBC1854858, Histone H3 abcam (ab1791), Lot: GR10862-1) together with protein A agarose beads (Millipore). DNA was recovered using QIAquick PCR Purification Kit (Qiagen). The resulting ChIP DNA was subjected to qPCR analysis. In order to compare the relative values of H3K27me3 for each sample, the relative value of H3K27me3 was normalized through division by the corresponding H3 value. This normalization assured the comparison of H3K27me3 regardless of H3 quantity. Specific primers were used for qPCR following ChIP as follows: BELL1 ChIP Fw, 5'- GAC CCC ATC TAC TAA GCC TGA A-3'. BELL1 ChIP Rv, 5'- GAG CTT GTA GAG ACA CGA GAT CG-3'. BELL2 ChIP Fw, 5' –CGG AAG CTT TCC AGG TCG A-3'. BELL2 ChIP Rv, 5' –TCC ATT TAT ATT ATC CCA GCT ATT GC-3'. AIB ChIP Fw, 5' – GATCGACATTCCAGGTTTTGA -3'. AIB ChIP Rv, 5'- CCC TGG TTC CTC GTC ACT AA-3'. PpJAGGED1 ChIP Fw, 5'- CTC CTG CTG CTC CTC GTC T-3'. JAGGED1 ChIP Rv, 5'-CGC AGA GGT CAA TCA CTC AG-3'. Negative control Fw, 5'-CGT TGC TCC GAT TTT TGC -3'. Negative control Rv, 5'- AAA TGT TTG GCG TCC TCA AG – 3'.

Real-time qPCR analysis of gene expression

Total RNA was extracted from protonemata using SV-TOTAL RNA Isolation System (Promega) followed by RT-PCR were performed as described⁵⁷. RT-qPCR analysis was performed by $\Delta\Delta C_t$ method of relative quantification with a StepOne Thermal Cycler (Applied Biosystems, Foster City, CA, USA), using SYBR Green to monitor dsDNA synthesis as previously described in³⁴. The amount of cDNA for each gene was quantified using a log-linear regression curve of the threshold cycle and the amount of standard template prepared from a cDNA clone. PpHistone3 genes (Pp1s1963_1V6.1 and Pp1s3_594V6.1) were used to normalize the values of relative expression by primers PpHistone3 100 Fw and PpHistone3 143 Rv.

Specific Primers used are as followed:

BELL1 Fw, 5'-TGG ACC ATT TTG AAG GAA GG-3'. BELL1 Rv, 5'- GCT CAT AGT TGC GGA GGA TT-3'. BELL2 Fw, 5'-GCC AAT TGT TTG TTC AGT AGG TT-3'. BELL2 Rv, 5'-AGA TGG CAC AAC ATG ACC AC-3'. AIB Fw, 5'–TGC

ATT TTA CCT GGG TCC AT-3'. AIB Rv, 5'- CCT AAA CTT GCG ATT CTT 587
GAC A-3'. JAGGED1 Fw, 5'-CAC CAT GCC ATG TCT CGA T-3'. JAGGED1 Rv, 588
5'- AGC GCC GAG GTT ATT GAA C-3'. JAGGED2 Fw, 5'- GCC CTC TCC TCT 589
CTT GCT TC-3'. JAGGED2 Rv, 5'-GTC GTC GGA GTG AGA GTC CA-3'. 590
PpHistone3 100 Fw, 5'- GGA GGA GTG AAG AAG CCA CA-3'. PpHistone3 143 591
Rv, 5'-TGC GAA TCT CCC GAA GAG-3'. Tata Fw 5'-GAT CTA GCT ATA AGC 592
CTG ATC TAC CG-3'. Tata Rv 5'- CAG GAG CAG GGA GAG ATT TG -3'. 593

Western blot 594

Western blots were performed as described previously ²⁴, with modifications. After 595
SDS PAGE separation, proteins were blotted on a nitrocellulose transfer membrane 596
(Whatman, Schleicher & Schuell) applying 400 mA for 65 min, at 16C°. Antibodies 597
against H3K27me1 (07-448), Lot: DAM1598790, H3K27me2 (07-452), Lot: 598
NG1631707, H3K27me3 (07-449), Lot: DAM1662421, were obtained from 599
Millipore. 600

Acknowledgments 603

We would like to thank Daniel Bouyer and Arp Schnittger for their kind support with 604
Western blot analysis and Moran Bodas for technical assistance. We thank Eran 605
Bachrach for critical reading and helpful suggestions. I.P was supported in part by a 606
matching Tel-Aviv University Deans doctoral fellowship and the Manna foundation. 607
This research was supported by the MEXT and JSPS KAKENHI grants to Y.T. and 608
M. H. and JST ERATO grant to M.H., the Excellence Initiative of the German 609
Federal and States Governments to R.R. and the German Israeli Foundation (GIF) 610
Grant No.: 154/2008, granted to N.O. and R.R. 611

Fig. 1. Western blot analysis for H3K27 methylation in $\Delta PpCLF$ and $\Delta PpFIE$ deletion mutants. Wild type, $\Delta PpCLF$ and $\Delta PpFIE$ nuclear-enriched protein extracts were loaded as four separate groups side by side on an polyacrylamide gel. Each group was subsequently analyzed independently with either H3, H3K27me3 (Bouyer *et al.* ²⁴), H3K27me2 or H3K27me1-specific antibodies. Histone H3 served as loading control.

Fig. 2. Up-regulation of putative PRC2 target genes in $\Delta PpCLF$ and $\Delta PpFIE$ lines. RT-qPCR served to analyze the expression of *BELL1*, *BELL2*, *AIB*, *JAGGED1* and *JAGGED2*. The analysis was performed using total mRNA extracted from four days old protonema recovered after dispersion of $\Delta PpCLF$ and $\Delta PpFIE$ lines as described in the material and methods section. PpHistone3 amplicon, using primers PpHistone3 100 Fw and PpHistone3 143 Rv, was used for normalization. RQ- Relative Quantification.

Fig. 3. Loss of H3K27me3 at PRC2-dependent genes in $\Delta PpCLF$ and $\Delta PpFIE$ lines Chromatin, purified from four days old protonema of wild type, $\Delta PpCLF$ and $\Delta PpFIE$ lines, was analyzed via ChIP-qPCR using antibody against H3K27me3 modification to test the methylation status of *BELL1*, *JAGGED1* and *AIB*, and the corresponding paralogs, *BELL2* and *JAGGED2*. As a reference gene we used the *Pp1s192_94V6.1* locus which is not marked by H3K27me3 modification (cosmoss.org genome browser track: “histone 3 code MACS peaks”)³¹. The values of H3K27me3 were normalized to the corresponding values obtained with H3 antibodies as described in the materials and methods section.

Fig. 4. Alignment of the SET domains: *Arabidopsis thaliana* CLF (AT2G23380.1) (At), *Physcomitrella patens* CLF Pp1s100_146V6.1 (Pp), *Drosophila melanogaster* E(z) NP_524021 (Dm), *Homo sapiens* EZH2 CAA64955(Hs) and *Paramecium bursaria* chlorella virus vSET NP_048968 (V). The SET domain multiplicity site is marked with (*), cofactor binding site marked with (**) and histone active site marked with (***). Conserved sites across all species which were mutated are shaded black while the remaining conserved sites are shaded gray.

Fig. 5. Mutations in CLF SET-domain hamper H3K27me3 methylation without affecting CLF – FIE interaction. **A.** Equal amounts of nuclear protein extracts from wild type, control line and each of the CLF-SET mutants were analyzed by Western blot analysis, using H3K27me3 antibodies. Arrow indicates the migration of the H3 protein. **B.** *In planta* BiFC assay reveals that the CLF mutant proteins can interact with their partner FIE. Interaction was tested between YC-PpFIE and the following: **(a)** YN-PpCLF, **(b)** YN-F905A, **(c)** YN-F905Y, **(d)** YN-Y964F, **(e)** YN-NH926AA. **(f)** YN-only serves as a negative control. Localization was determined in leaf epidermis of *Nicotiana benthamiana*. Green reflects protein interaction and red color is the outcome of chloroplasts auto-fluorescence. YFP fluorescence image was obtained from single confocal sections. Magnification 500x.

Fig. 6. H3K27me3 and expression levels of PRC2 dependent genes in $\Delta PpCLF$ mutants. **A.** H3K27me3 status of *BELL1*, *BELL2*, *JAGGED1*, *JAGGED2* and *AIB* analyzed by ChIP-qPCR using α H3K27me3 antibodies. Chromatin was purified from four days old protonema of *CLF* mutated lines: NH926AA, Y964F, F905A and F905Y. As a reference gene we used the *Pp1s192_94V6.1* locus which is not marked by H3K27me3 modification (cosmoss.org genome browser track: “histone 3 code MACS peaks”)³¹. The values of H3K27me3 were normalized to the corresponding values obtained with H3 antibodies as described in the materials and methods section. **B.** Relative expression levels of *BELL1*, *BELL2*, *JAGGED1*, *JAGGED2* and *AIB*, were determined via RT-qPCR on mRNA obtained from four days old protonema as described in the material and methods section.

Figure 7. Growth morphology of transgenic lines which bear point mutations in the CLF-SET domain. Overall plant morphology of wild type, and CLF-SET mutant lines, each originating from a protoplast was monitored after 23 days of growth. Overall plant size of protonema stage (**A-F**). Gross morphology of protonema stage (**G-H**) Close up of gametophores and sporophyte-like structures of CLF-SET mutants (**M-R**). Bar size: A-F 2000 μ M G-L 500 μ M, M-R 100 μ M.

References

1. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature*. 2000;403(6765):41-45.
2. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res*. 2011;21(3):381-395.
3. Kohler C, Villar CB. Programming of gene expression by polycomb group proteins. *Trends Cell Biol*. 2008;18(5):236-243.
4. Schwartz YB, Pirrotta V. Polycomb complexes and epigenetic states. *Curr Opin Cell Biol*. 2008;20(3):266-273.
5. Papp B, Muller J. Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins. *Genes Dev*. 2006;20(15):2041-2054.
6. Muller J, Verrijzer P. Biochemical mechanisms of gene regulation by polycomb group protein complexes. *Curr Opin Genet Dev*. 2009;19(2):150-158.
7. Cao R, Wang L, Wang H, et al. Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science*. 2002;298(5595):1039-1043.

8. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. Drosophila enhancer	687
of zeste/ESC complexes have a histone H3 methyltransferase activity that marks	688
chromosomal polycomb sites. <i>Cell</i> . 2002;111(2):185-196.	689
9. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. Histone	690
methyltransferase activity associated with a human multiprotein complex containing	691
the enhancer of zeste protein. <i>Genes Dev</i> . 2002;16(22):2893-2905.	692
10. Muller J, Hart CM, Francis NJ, et al. Histone methyltransferase activity of a	693
drosophila polycomb group repressor complex. <i>Cell</i> . 2002;111(2):197-208.	694
11. Nekrasov M, Klymenko T, Fraterman S, et al. Pcl-PRC2 is needed to generate	695
high levels of H3-K27 trimethylation at polycomb target genes. <i>EMBO J</i> .	696
2007;26(18):4078-4088.	697
12. Tie F, Stratton CA, Kurzhals RL, Harte PJ. The N terminus of drosophila ESC	698
binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3	699
lysine 27. <i>Mol Cell Biol</i> . 2007;27(6):2014-2026.	700
13. Chaudhury AM, Ming L, Miller C, Craig S, Dennis ES, Peacock WJ.	701
Fertilization-independent seed development in arabidopsis thaliana. <i>Proc Natl Acad</i>	702
<i>Sci U S A</i> . 1997;94(8):4223-4228.	703
14. Ohad N, Margossian L, Hsu YC, Williams C, Repetti P, Fischer RL. A mutation	704
that allows endosperm development without fertilization. <i>Proc Natl Acad Sci U S A</i> .	705
1996;93(11):5319-5324.	706

15. Luo M, Bilodeau P, Koltunow A, Dennis ES, Peacock WJ, Chaudhury AM. Genes	707
controlling fertilization-independent seed development in arabidopsis thaliana. <i>Proc</i>	708
<i>Natl Acad Sci U S A</i> . 1999;96(1):296-301.	709
16. Gendall AR, Levy YY, Wilson A, Dean C. The VERNALIZATION 2 gene	710
mediates the epigenetic regulation of vernalization in arabidopsis. <i>Cell</i> .	711
2001;107(4):525-535.	712
17. Kohler C, Hennig L, Bouveret R, Gheyselinck J, Grossniklaus U, Gruissem W.	713
Arabidopsis MSI1 is a component of the MEA/FIE polycomb group complex and	714
required for seed development. <i>EMBO J</i> . 2003;22(18):4804-4814.	715
18. Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz EM, Coupland G. A	716
polycomb-group gene regulates homeotic gene expression in arabidopsis. <i>Nature</i> .	717
1997;386(6620):44-51.	718
19. Grossniklaus U, Vielle-Calzada JP, Hoepfner MA, Gagliano WB. Maternal	719
control of embryogenesis by MEDEA, a polycomb group gene in arabidopsis.	720
<i>Science</i> . 1998;280(5362):446-450.	721
20. Kiyosue T, Ohad N, Yadegari R, et al. Control of fertilization-independent	722
endosperm development by the MEDEA polycomb gene in arabidopsis. <i>Proc Natl</i>	723
<i>Acad Sci U S A</i> . 1999;96(7):4186-4191.	724
21. Chanvivattana Y, Bishopp A, Schubert D, et al. Interaction of polycomb-group	725
proteins controlling flowering in arabidopsis. <i>Development</i> . 2004;131(21):5263-5276.	726

22. Wang D, Tyson MD, Jackson SS, Yadegari R. Partially redundant functions of	727
two SET-domain polycomb-group proteins in controlling initiation of seed	728
development in arabidopsis. <i>Proc Natl Acad Sci U S A</i> . 2006;103(35):13244-13249.	729
23. Schubert D, Primavesi L, Bishopp A, et al. Silencing by plant polycomb-group	730
genes requires dispersed trimethylation of histone H3 at lysine 27. <i>EMBO J</i> .	731
2006;25(19):4638-4649.	732
24. Bouyer D, Roudier F, Heese M, et al. Polycomb repressive complex 2 controls the	733
embryo-to-seedling phase transition. <i>PLoS Genet</i> . 2011;7(3):e1002014.	734
25. Lindroth AM, Shultis D, Jasencakova Z, et al. Dual histone H3 methylation marks	735
at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. <i>EMBO J</i> .	736
2004;23(21):4286-4296.	737
26. Makarevich G, Leroy O, Akinci U, et al. Different polycomb group complexes	738
regulate common target genes in arabidopsis. <i>EMBO Rep</i> . 2006;7(9):947-952.	739
27. Jullien PE, Katz A, Oliva M, Ohad N, Berger F. Polycomb group complexes self-	740
regulate imprinting of the polycomb group gene MEDEA in arabidopsis. <i>Curr Biol</i> .	741
2006;16(5):486-492.	742
28. Shaver S, Casas-Mollano JA, Cerny RL, Cerutti H. Origin of the polycomb	743
repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga	744
chlamydomonas. <i>Epigenetics</i> . 2010;5(4):301-312.	745
29. Butenko Y, Ohad N. Polycomb-group mediated epigenetic mechanisms through	746
plant evolution. <i>Biochim Biophys Acta</i> . 2011;1809(8):395-406.	747

30. Schmitges FW, Prusty AB, Faty M, et al. Histone methylation by PRC2 is inhibited by active chromatin marks. <i>Mol Cell</i> . 2011;42(3):330-341.	748 749
31. Widiez T, Symeonidi A, Luo C, Lam E, Lawton M, Rensing SA. The chromatin landscape of the moss <i>Physcomitrella patens</i> and its dynamics during development and drought stress. <i>Plant J</i> . 2014;79(1):67-81.	750 751 752
32. Shen X, Liu Y, Hsu YJ, et al. EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. <i>Mol Cell</i> . 2008;32(4):491-502.	753 754 755
33. Zheng B, Chen X. Dynamics of histone H3 lysine 27 trimethylation in plant development. <i>Curr Opin Plant Biol</i> . 2011;14(2):123-129.	756 757
34. Mosquna A, Katz A, Decker EL, Rensing SA, Reski R, Ohad N. Regulation of stem cell maintenance by the polycomb protein FIE has been conserved during land plant evolution. <i>Development</i> . 2009;136(14):2433-2444.	758 759 760
35. Okano Y, Aono N, Hiwatashi Y, et al. A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. <i>Proc Natl Acad Sci U S A</i> . 2009;106(38):16321-16326.	761 762 763
36. Strepp R, Scholz S, Kruse S, Speth V, Reski R. Plant nuclear gene knockout reveals a role in plastid division for the homolog of the bacterial cell division protein FtsZ, an ancestral tubulin. <i>Proc Natl Acad Sci U S A</i> . 1998;95(8):4368-4373.	764 765 766
37. Lang D, Weiche B, Timmerhaus G, et al. Genome-wide phylogenetic comparative analysis of plant transcriptional regulation: A timeline of loss, gain, expansion, and correlation with complexity. <i>Genome Biol Evol</i> . 2010;2:488-503.	767 768 769

38. Sakakibara K, Ando S, Yip HK, et al. KNOX2 genes regulate the haploid-to-	770
diploid morphological transition in land plants. <i>Science</i> . 2013;339(6123):1067-1070.	771
39. Qian C, Wang X, Manzur K, et al. Structural insights of the specificity and	772
catalysis of a viral histone H3 lysine 27 methyltransferase. <i>J Mol Biol</i> .	773
2006;359(1):86-96.	774
40. Manzur KL, Farooq A, Zeng L, et al. A dimeric viral SET domain	775
methyltransferase specific to Lys27 of histone H3. <i>Nat Struct Biol</i> . 2003;10(3):187-	776
196.	777
41. Hennig L, Derkacheva M. Diversity of polycomb group complexes in plants:	778
Same rules, different players? <i>Trends Genet</i> . 2009;25(9):414-423.	779
42. Jacob Y, Feng S, LeBlanc CA, et al. ATXR5 and ATXR6 are H3K27	780
monomethyltransferases required for chromatin structure and gene silencing. <i>Nat</i>	781
<i>Struct Mol Biol</i> . 2009;16(7):763-768.	782
43. Mujtaba S, Manzur KL, Gurnon JR, Kang M, Van Etten JL, Zhou MM.	783
Epigenetic transcriptional repression of cellular genes by a viral SET protein. <i>Nat Cell</i>	784
<i>Biol</i> . 2008;10(9):1114-1122.	785
44. Lee JH, Lin H, Joo S, Goodenough U. Early sexual origins of homeoprotein	786
heterodimerization and evolution of the plant KNOX/BELL family. <i>Cell</i> .	787
2008;133(5):829-840.	788
45. Horst N, Katz A, Pereman I, Decker E, Ohad N, Reski R. A single homeobox	789
gene triggers phase transition, embryogenesis and asexual reproduction. <i>Nature</i>	790
<i>Plants</i> . 01/2016; 2(2):15209. DOI: 10.1038/nplants.2015.209	791

46. Schiessl K, Kausika S, Southam P, Bush M, Sablowski R. JAGGED controls	792
growth anisotropy and coordination between cell size and cell cycle during plant	793
organogenesis. <i>Curr Biol.</i> 2012;22(19):1739-1746.	794
47. Sauret-Gueto S, Schiessl K, Bangham A, Sablowski R, Coen E. JAGGED	795
controls arabidopsis petal growth and shape by interacting with a divergent polarity	796
field. <i>PLoS Biol.</i> 2013;11(4):e1001550.	797
48. Rensing SA, Lang D, Zimmer AD, et al. The physcomitrella genome reveals	798
evolutionary insights into the conquest of land by plants. <i>Science.</i> 2008;319(5859):64-	799
69.	800
49. Ashton NW, Cove DJ. The isolation and preliminary characterisation of	801
auxotrophic and analogue resistant mutants of the moss, physcomitrella patens.	802
<i>Molecular and General Genetics.</i> 1977;154(1):87-95.	803
50. Frank W, Decker EL, Reski R. Molecular tools to study physcomitrella patens.	804
<i>Plant Biol (Stuttg).</i> 2005;7(3):220-227.	805
51. Wolf L, Rizzini L, Stracke R, Ulm R, Rensing SA. The molecular and	806
physiological responses of physcomitrella patens to ultraviolet-B radiation. <i>Plant</i>	807
<i>Physiol.</i> 2010;153(3):1123-1134.	808
52. Zimmer AD, Lang D, Buchta K, et al. Reannotation and extended community	809
resources for the genome of the non-seed plant physcomitrella patens provide insights	810
into the evolution of plant gene structures and functions. <i>BMC Genomics.</i>	811
2013;14:498.	812

53. Baldi P, Long AD. A bayesian framework for the analysis of microarray	813
expression data: Regularized t -test and statistical inferences of gene changes.	814
<i>Bioinformatics</i> . 2001;17(6):509-519.	815
54. Nishiyama T, Hiwatashi Y, Sakakibara I, Kato M, Hasebe M. Tagged mutagenesis	816
and gene-trap in the moss, <i>physcomitrella patens</i> by shuttle mutagenesis. <i>DNA Res</i> .	817
2000;7(1):9-17.	818
55. Bracha-Drori K, Shichrur K, Katz A, et al. Detection of protein-protein	819
interactions in plants using bimolecular fluorescence complementation. <i>Plant J</i> .	820
2004;40(3):419-427.	821
56. Berr A, McCallum EJ, Alioua A, Heintz D, Heitz T, Shen WH. Arabidopsis	822
histone methyltransferase SET DOMAIN GROUP8 mediates induction of the	823
jasmonate/ethylene pathway genes in plant defense response to necrotrophic fungi.	824
<i>Plant Physiol</i> . 2010;154(3):1403-1414.	825
57. Katz A, Oliva M, Mosquna A, Hakim O, Ohad N. FIE and CURLY LEAF	826
polycomb proteins interact in the regulation of homeobox gene expression during	827
sporophyte development. <i>Plant J</i> . 2004;37(5):707-719.	828
	829

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

A.

B.

Figure 6A

A.

Figure 6B

B.

Figure 7