The Polycomb group protein CLF emerges as a specific tri-methylase of H3K27 regulating gene expression and development in Physcomitrella patens

Idan Pereman a,1, Assaf Mosquna a,i 1, Aviva Katz a,i 1, Gertrud Wiedemann b, Daniel Lang b, Eva L. Decker b, Yosuke Tamada c,d, Takaaki Ishikawa c, Tomoaki Nishiyama e,f, Mitsuyasu Hasebe c,d,f, Ralf Reski b,g,h,i,1, Nir Ohad a,g,j,i,1

1 These authors have contributed equally to the work.

ARTICLE INFO
Article history:
Received 18 January 2016
Received in revised form 14 April 2016
Accepted 5 May 2016
Available online 11 May 2016

Keywords:
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 have 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.

© 2016 Elsevier B.V. All rights reserved.

1. 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 residues [2]. 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 [34]. 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 (H3K27me3) [23]. 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 (MSH1) [17]; and (iv) a E(z) homolog containing a SET domain (Su(var) E(z)2) 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 PRC2 complex and H3K27 methylation in plants [30,31]. The two methylation of H3K27 [7,32]. Furthermore, human EZH paralogs and the evolutionary ancestral function and origin of the PRC2 complex in flowering plants [33]. 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 also could be observed in the loss-of-function mutants of FIE and CLF [34,35]. Previously, we have shown that CLF and FIE can interact [34]. Moreover, based on their mutant phenotype as well as on complementation assays, it has been suggested that both proteins function as PRC2 proteins [34]. Both ΔPpCLF and ΔppclF mutants in P. patens have 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 programs [29]. PpCLF, being a single SET domain gene, together with the moss unique competence for gene targeting by homologous recombination [36], 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.

2. Results

2.1. H3K27me3 modification is absent in ΔPpCLF and ΔppFIE 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 nucleo-enriched extracts from wild type, ΔPpCLF and ΔPPFIE deletion mutants were subjected to Western blot analysis, using antibodies specific for H3K27me3, H3K27me2 and H3K27me1, respectively. The H3K27me3 mark was completely lost in ΔPpCLF as well as in ΔppFIE 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).

2.2. The loss of H3K27me3 in ΔPpCLF and ΔPpFIE 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 ΔPpFIE plants using a whole-genome microarray. Significantly up-regulated genes in the ΔPpFIE 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 5'-DGE (DRA/ERA/SRA accession nos.:DRR055530-DRR055535 searchable at http://trace.ddbj.nig.ac.jp/dra/index_e.html) of wild type and a ΔPpCLF mutant and H3K27me3 ChiPseq (accession nos.: DRR059713-DRR059717). 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 ΔppFIE [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_66V6.1) [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 ΔPpCLF and ΔPpFIE 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 ΔPpCLF and ΔPpFIE 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).

2.3. 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 Chilorella 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]
**Fig. 1.** Western blot analysis for H3K27 methylation in ΔPpCLF and ΔPpFIE deletion mutants. Wild type, ΔPpCLF and ΔPpFIE nuclear-enriched protein extracts were loaded as four separate groups side by side on a 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 ΔPpCLF and ΔPpFIE 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 ΔPpCLF and ΔPpFIE 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.
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 that 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-1 [39]. Mutation at corresponding amino acid residues in the vSET protein dramatically reduced the methyltransferase activity with the exception of mutating residue 52F-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, designated Control-CLF. To confirm the correct integration of each mutation into PpCLF, a PCR followed by an RFLP analysis was carried out (Supplemental Figs. 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, was shown to only partially reduce the HKTMT activity similar to the vSET F52Y mutation [39], as opposed to line F905A. Thus, these results indicate structural and functional similarities between PpCLF and vSET.

PpCLF binds to PpFIE WD40 protein [34], 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) assay [34], which was used to monitor the interactions between wild

---

**Fig. 3.** Loss of H3K27me3 at PRC2-dependent genes in ΔPpCLF and ΔPpFIE lines Chromatin, purified from four days old protonema of wild type, ΔPpCLF and ΔPpFIE 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.
type PpCLF and PpFIE [34] 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.

2.4. 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 ΔPpCLF and ΔPpFIE 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.

2.5. 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 ΔPpCLF, 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 (ΔPpCLF) resulted in small plants (Fig. 7B–D, F) as compared to wild type (Fig. 7A) with a denser pattern of the protonema (Fig. 7H–J, L) as compared to wild type (Fig. 7 G). Additionally, lines ΔPpCLF, F905A, NH926AA and F964Y developed abnormal buds, forming...
sporophyte-like bodies (Fig. 7N-P, R), as previously described for ΔPpFIE and ppcf-del deletion mutants [34,35]. Line F905Y exhibited only minor developmental defects (Fig. 7E, K and Q), as compared to line F905A (Fig. 7F, L and R).

3. 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 system [41]. 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 to the overall methyl-transferase activity and further determine developmental consequences.

In this study we have shown in-vivo that ΔPpFIE, ΔPpCLF 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 mutant [23]. In addition, it was described that H3K27me2 is partially lost in Arabidopsis lacking the AtFIE PcG protein [25]. Together, these studies indicate a role of the PcG complex in H3K27me2.

Unlike in Arabidopsis our results show that ΔPpCLF and, moreover, point mutations in the PpCLF-SET domain, specifically abolish tri-methylation but not di- nor 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 H3K27me1 [42] 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 homolog [43]. 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 identified [45].

An additional gene which was found in this study is a homolog of AtJAGGED which encodes a zinc finger transcription factor [46]. AtJAGGED was described as a regulator of growth rates in a particular region of the petal while also modulating tissue polarity [47]. 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 AtBEL1 [46]. A homologue of AtBEL1 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 plant land evolution.

4. Material and methods

4.1. Plant material, culture conditions and treatments

The ‘Grandsen 2004’ strain of P. patens which was used for genome sequencing [48] was propagated on BCD and BCDAT media [49] at 25 °C under a 16-hour light and 8-hour dark cycle [50].

4.2. Expression analysis using microarrays

Gene expression in wild type and ΔPpFIE plants was compared using whole genome microarrays (Combimatrix, Mukilteo, WA, USA) based on all gene models v1.2 [48] as described [51]. 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 V1.6 [52]. The array analysis was carried based on the probes as described previously [51]. 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 Milio 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 comparison [53] 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 ppcf-del lines (DNA accession number listed in Supplementary Table 2) and TAP proteins identified previously [37].

4.3. 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;f=;p=;w=;h=;e=;f=;s=;dataset=1

4.4. Constructing moss ΔPpFIE knock out lines

All mutant plants were generated in the background of ‘Grandsen 2004’ strain of P. patens [49] and deposited at the International Moss Stock Centre IMSC (www.moss-stock-center.org; Supplementary Table 2).

The 5′ (805 bp) and 3′ (598 bp) of the untranslated regions (UTR) of PpFIE were amplified using the following primers: PpFIE5′-UTR-Fw, 5′-GAA GCT TGA CTA GAA AAA AAT TGT GAT AGT GTG TTA TGC GAA TG-3′ and PpFIE5′-UTR-Rv, 5′-GAA GCT TCA CGG GAT CCG TCG CAA. PpFIE3′-UTR-Rv, 5′-GCA TGC TGA TCG ATA CCT GCA GCC A-3′ and PpFIE3′-UTR-Rv, 5′-GCA TGC CCG TAC CCT TGA CAC ATG CAT AAC A-3′. All amplified fragments were sub-cloned into the pTiZ 57 vector (Fermentas) and sequenced to ensure their integrity. Following, the 5′ UTR was cloned into the plMBS vector [54] in the HindIII site. The resulting vector was then linearized with SphI and the 3′ UTR of FIE was cloned into the same site resulting in the plMBS ΔPpFIE knock out vector. Proper integration at the 5′ region was validated with primers: PpFIE5′-6-UTR-Fw, 5′-GCA AGC GCA CGG CAG ACT-3′ and Nptii-KO-Rv, 5′-TGC ATG ATC AGG ATG ATC TGC A-3′. Proper integration at the 3′ region was validated with primers: Nptii-KO-Fw, 5′-CTA TCG CCT TCT TCA CGA G-3′ and PpFIE3′-UTR-Rv, 5′-GAA AGG TGG TAC CAT GAA CAT GAT TTG G-3′ (Supplemental Figs. 1A and 3A).

4.5. Constructing moss ΔPpCLF knock out lines

The 5′ (546 bp) and 3′ (561 bp) of CLF untranslated region (UTR) were amplified using the following primers: CLF5′-UTR-Fw, 5′-GAA TTC ATG AAC TGT TTT ACT CAA CCT GTG-3′ and CLF5′-UTR-Rv, 5′-GTA CAC CCA CTT CTT CCT ATC TGC AC-3′. CLF3′-UTR-Fw, 5′-GCA TGC TGT ACT CTA GGC GTT TTC-3′ and CLF3′-UTR-Rv, 5′-GAT ATC ATC CCT TGA TAT GAT GGC AGC GGA-3′. All the 5′UTRs were sub-cloned into the pTiZ 57 vector (Fermentas) and sequenced to ensure their integrity. Following, the 5′ UTR was cloned into the plMBS vector [54] in SalI and EcoRV sites. The resulting vector was then linearized with SphI and the 3′ UTR of CLF was cloned into the same site resulting in the plMBS ΔPpCLF knock out vector. Proper integration at the 5′ region was validated with primers: CLF-KO1-Fw, 5′-GTG ATC TTC GCC GAA ATC GAC GAC-3′ and Nptii-KO-Rv, 5′-TGC ATG ATC AGG AGT ATC TGC G-3′. Proper integration at the 3′ region was validated with primers: Nptii-KO-Fw, 5′-CTA TCG CCT TCT TCA CGA G-3′ and CLF-KO1-Rv, 5′-AGC CAC TGT TTT GCA CAT GAT GG-3′. (Supplemental Figs. 1B and 3B).

4.6. Genomic CLF-SET knock-in vectors

A PpCLF 3′ genomic DNA fragment was amplified with primers: PpCLF 5′Fw, 5′-AAG ATG TTC ATG CCA TCA TCA ATG TAC-3′ and ppcf stop Hill Ry, 5′-AAG ATG TTA ATC AGC AAG TTC CGT TCG ATC-3′. The fragment was then cloned to the pTi 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 sites are listed below, where the restriction site is underlined and the introduced mutation is marked in bold:

4.6.1. Y964F including AfaI restriction site

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

4.6.2. F905Y including DraI restriction site

PpCLF SET-Y Fw, 5'-GAA TTC GTC GTT CCT C-3 PpCLF SET-Y Rv, 5'-GTC GAC TAG TTA CTC GAG AGC AAC TT-3'.

4.6.3. F905A including DraI restriction site

PpCLF SET-A Fw, 5'-GAA TTC GTC GTT CCT CCG CAA TTT AAA TGA TCA TGG AAG CCT C-3' PpCLF SET-A Rv, 5'-GAG GCT TGC CTC ATG ATT TAA ATT GGA CGA CGA ATT C-3'.

4.6.4. NH926AA including NotI restriction site

PpCLF SET-CF Fw, 5'-GCG GAC AAG CTC ATA TTA TAT GGC GGC TCA CCT ACT CCT ACT CCC AAC TGC-3' PpCLF SET-CF Rv, 5'-GCA GGT GGG AGT AGG TGA GGG GCC GCC AAA TTT GAT CTT GGC C-3'.

Each of the mutated PpCLF-SET fragments described above and the 3' UTR of PpCLF (described above in the CLF knock-out paragraph), were cloned into the pMBL5 in the HindIII and Spel sites respectively, creating the pMBL5 PpCLF SET-Y964F, F905Y, F905A, NH926AA vectors.

The corresponding restriction site which was cloned together with each point mutation was used in a RFLP analysis to ensure proper construct integration in each of the above mutant line (Supplement Figs. 2 and 3).

Two independent lines were generated and analyzed for each CLF-SET mutation.

4.7. Validation of PpCLF-SET mutant lines integrity

In order to validate the presence of the correct mutation in each mutated PpCLF gene, a restriction analysis was performed allowing detecting particular restriction products as indicated in Supplemental Fig. 2 which results from silent mutations adjacent to the mutated sites. Transcripts from two independent transgenic lines of each mutant were isolated and following RT-PCR the presence of the correct mutated scripts from two independent transgenic lines of each mutant were isolated which results from silent mutations adjacent to the mutated sites. Transcripts from two independent lines were generated and analyzed for each CLF-SET mutation.

4.8. Construction of CLF-SET knock-in vectors for BiFC assay

PpCLF coding sequence (3003 bp) was amplified in three parts: part A (1−1067) part B (1062−2052) part C (2047−3003). The construction of PpCLF cDNA partition was performed using the following primers:

4.8.1. Part A

PpCLF cDNA-A-Fw 5'-ACT AGT GGA TCA TTG GCG TCC TCC AGC TAC TAC GCC A-3' PpCLF cDNA-A-Rv 5'-TTG CTG CAA GGA GTA GCC CTC CTC TTA A-3'.

4.8.2. Part B

PpCLF cDNA-B-Fw 5'-GCT CAT CCA AGT GAA AGG CAA C-3' PpCLF cDNA-B-Rv 5'-GCA GCA TCT AGT TGC ATT GCC C-3'.

4.8.3. Part C

PpCLF cDNA-C-Fw 5'-AGG GCT GCA AGA AGT GTG CAG-3' PpCLF cDNA-C-Rv 5'-GTC GAC TAG TTA CTC GAG AGC AAC TT-3'.

All amplified fragments were sub-cloned into the pTZ57 vector and sequenced (designated as pTZ-A, pTZ-B and pTZ-C). Fragment B was digested and inserted to pTZ A into the NcoI and Sall sites giving rise to pTZ AB. Then, fragment "C" was inserted into the Bsal and Sall sites of pTZ AB, giving rise to pTZ PpCLF cDNA. In turn, the set domain of PpCLF cDNA was mutated (as described above for the PpCLF knock-in lines) using the "Quickchange" site directed mutagenesis kit (Stratagene) using the following primers where the restriction site is underlined and the introduced mutation is marked in bold:

4.8.4. Y964F including AfaI restriction site

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

4.8.5. F905Y including DraI restriction site

PpCLF SET-Y Fw, 5'-GAA TTC GTC GTT CCT C-3 PpCLF SET-Y Rv, 5'-GTC GAC TAG TTA CTC GAG AGC AAC TT-3'.

4.8.6. F905A including DraI restriction site

PpCLF SET-A Fw, 5'-GAA TTC GTC GTT CCT CCG CAA TTT AAA TGA TCA TGG AAG CCT C-3' PpCLF SET-A Rv, 5'-GAG GCT TGC CTC ATG ATT TAA ATT GGA CGA CGA ATT C-3'.

4.8.7. NH926AA including NotI restriction site

PpCLF SET-CF Fw, 5'-GCG GAC AAG CTC ATA TTA TAT GGC GGC TCA CCT ACT CCT ACT CCC AAC TGC-3' PpCLF SET-CF Rv, 5'-GCA GGT GGG AGT AGG TGA GGG GCC GCC AAA TTT GAT CTT GGC C-3'.

All clones were verified by sequencing. Wild type and mutated PpCLF cDNA were cloned into the pSY735 vector through the SpeI site.

4.9. Monitoring protein–protein interaction via bimolecular fluorescence complementation assay

Protein–protein interactions in planta were examined by bimolecular fluorescence complementation (BiFC) assay [55]. PpHE full-length cDNA was cloned into the pSY 736 vector at the SpeI site, PpHE full-length mutated cDNAs were cloned into the pSY 735 vector which contains the C-terminal (YC) fragment of the YFP protein [55]. Equal concentrations of Agrobacterium tumefaciens strain GV3101/pMP90 containing plasmids of interest were transiently co-expressed in Nicotiana benthamiana leaves via the leaf injection procedure. Image annotation was performed with Zeiss AxioVision, Zeiss CLSM-5 and Adobe Photoshop 7.0 (Mountain View, CA, USA). Negative controls with vectors bearing only YN or YC alone were carried out in order to verify the specificity of the interactions.

4.10. Protoplast isolation and PEG-mediated transformation of P. patens

PEG transformation was performed as described in Nishiyama et al. [54] PHYSOCbase (http://moss.nibb.ac.jp). Six days after regeneration, transformants were selected on BCDAT medium containing 20 mg/l of G418.

4.11. Chromatin immunoprecipitation (ChIP)

The ChIP assays protocol was kindly provided by T. Widiez and performed as previously described in Berr et al. [56] with modifications. Following one week of cultivation, protonema tissue of P. patens was homogenized and re-plated. This tissue was further homogenized twice more in 4 days’ intervals. Samples were harvested and dried completely using Whatman paper, frozen in liquid nitrogen and stored in −80 °C until chromatin isolation. 500 mg of the sample’s tissue was...
ground to fine powder using a mortar and pestle and liquid nitrogen. The powder was placed in a 50 ml tube to which 15 ml of Nuclear Isolation Buffer I was added in room temperature (10 mM HEPES pH 7.5, 0.4 M Sucrose, 5 mM KCl, 5 mM MgCl2, 5 mM EDTA, 0.6% Triton X-100). The following solutions were added just prior to use: β-mercaptoethanol to a final concentration of 0.1%, PMSF to a final concentration of 1 mM, cComplete protease inhibitor (Roche, USA), according to the manufacturer’s recommendations and 0.2 ml formaldehyde 37%. Samples were sonicated (Misonix, sonicator s-4000) at 20% capacity in 6 cycles of 3 s active sonication a 20 s 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 the H3K27me3 regardless of H3 quantity.

Specific primers were used for qPCR following ChIP as follows:

<table>
<thead>
<tr>
<th>Fw</th>
<th>Rv</th>
</tr>
</thead>
<tbody>
<tr>
<td>BELL1 ChIP Fw, 5′-GAG CTC TTA ACC GAG ATT AAC-3′</td>
<td>BELL1 ChIP Rv, 5′-GAG TTA ACC GAG ATT AAC-3′</td>
</tr>
<tr>
<td>BELL2 ChIP Fw, 5′-CGG AAG ATC TCC AGG TCG A-3′</td>
<td>BELL2 ChIP Rv, 5′-TCC ATT TAT ATT ATC CCA GCT ATT GC-3′</td>
</tr>
<tr>
<td>AIB ChIP Fw, 5′-GATGAACATCCAGTATTGTTGG-3′</td>
<td>AIB ChIP Rv, 5′-CCC TGG TCC TTC TGT CTC ACT AA-3′</td>
</tr>
<tr>
<td>PpJAGGED1 ChIP Fw, 5′-CTC TGG CTC TTC TGC TTC T-3′</td>
<td>PpJAGGED1 ChIP Rv, 5′-CTG GTC TGC TTA CTA CTC AG-3′</td>
</tr>
<tr>
<td>Negative control Fw, 5′-CGT TCC GAT TTT TGC-3′</td>
<td>Negative control Rv, 5′-AAA TGT TTT GCC TCC TTA AG-3′</td>
</tr>
</tbody>
</table>

4.12. Real-time qPCR analysis of gene expression

Total RNA was extracted from protomenata using SV-TOTAL RNA Isolation System (Promega) followed by RT-PCR were performed as described [57]. RT-qPCR analysis was performed by ΔΔct 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 [34]. 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. PhHistone3 genes (Pp1s19631_16V1 and Pp1s35_594V61) were used to normalize the values of relative expression by primers PhHistone3100 Fw and PhHistone3143 Rv.

Specific Primers used are as follows:

<table>
<thead>
<tr>
<th>Fw</th>
<th>Rv</th>
</tr>
</thead>
<tbody>
<tr>
<td>BELL1 Fw, 5′-TGC ACC ATT TGG AAG GAA GG-3′</td>
<td>BELL1 Rv, 5′-GCT CAT AGT TGC GGA GAA TT-3′</td>
</tr>
<tr>
<td>BELL2 Fw, 5′-GAC CTC TTA GAG ACG ATT AAC-3′</td>
<td>BELL2 Rv, 5′-AGG TGG CAC AAG ATG ACC AC-3′</td>
</tr>
<tr>
<td>AIB Fw, 5′-TCC ATT TTA CCG GGT TCC AT-3′</td>
<td>AIB Rv, 5′-CCT ATT AAA CAT CGG ATT TCT GAC A-3′</td>
</tr>
<tr>
<td>JAGGED1 Fw, 5′-CAC GAT GCC ATG TCT CGA T-3′</td>
<td>JAGGED1 Rv, 5′-CAC GCC GAG TAT ATT GAA C-3′</td>
</tr>
<tr>
<td>JAGGED2 Fw, 5′-GCC CTT TCT TCT TCT GCC-3′</td>
<td>JAGGED2 Rv, 5′-GTC TTC GTC CCA GTG AGA CTC-3′</td>
</tr>
<tr>
<td>PhHistone3 143 Fw, 5′-TGC AGG TCT CCA GCC GAG-3′</td>
<td>PhHistone3 143 Rv, 5′-GAT CTA GCT ATA AGC CTG ATC TAC CG-3′</td>
</tr>
<tr>
<td>BELL1 Fw, 5′-GAG GAG CAG GGA GAT TC-3′</td>
<td>BELL2 Rv, 5′-GAA CTA GTG AAG GAA GAG-3′</td>
</tr>
</tbody>
</table>

4.13. Western blot

Western blots were performed as described previously [24], with modifications. After SDS PAGE separation, proteins were blotted onto a nitrocellulose transfer membrane (Whatman, Schleicher & Schuell) applying 400 mA for 65 min, at 16 °C. Antibodies against H3K27me1 (07-448), Lot: DAM1598790, H3K27me2 (07-452), Lot: NG1661707, H3K27me3 (07-449), Lot: DAM1662421, were obtained from Millipore. Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagen.2016.05.004.


