Plant Mol Biol
DOI 10.1007/s11103-015-0328-8

DNA METHYLTRANSFERASE 1 is involved in ^{m}CG and ^{m}CCG DNA methylation and is essential for sporophyte development in *Physcomitrella patens*

Rafael Yaari[1] · Chen Noy-Malka[1] · Gertrud Wiedemann[2] · Nitzan Auerbach Gershovitz[1] · Ralf Reski[2,3,4] · Aviva Katz[1] · Nir Ohad[1,4]

Received: 18 February 2015 / Accepted: 29 April 2015

Abstract DNA methylation has a crucial role in plant development regulating gene expression and silencing of transposable elements. Maintenance DNA methylation in plants occurs at symmetrical ^{m}CG and ^{m}CHG contexts (^{m} = methylated) and is maintained by DNA METHYLTRANSFERASE 1 (MET1) and CHROMOMETHYLASE (CMT) DNA methyltransferase protein families, respectively. While angiosperm genomes encode for several members of MET1 and CMT families, the moss *Physcomitrella patens*, serving as a model for early divergent land plants, carries a single member of each family. To determine the function of *P. patens* PpMET we generated *ΔPpmet* deletion mutant which lost ^{m}CG and unexpectedly ^{m}CCG methylation at loci tested. In order to evaluate the extent of ^{m}CCG methylation by MET1, we reexamined the *Arabidopsis thaliana Atmet1* mutant methylome and found a similar pattern of methylation loss, suggesting that maintenance of DNA methylation by MET1 is conserved through land plant evolution. While *ΔPpmet* displayed no phenotypic alterations during its gametophytic phase, it failed to develop sporophytes, indicating that PpMET plays a role in gametogenesis or early sporophyte development. Expression array analysis revealed that the deletion of PpMET resulted in upregulation of two genes and multiple repetitive sequences. In parallel, expression analysis of the previously reported *ΔPpcmt* mutant showed that lack of PpCMT triggers overexpression of genes. This overexpression combined with loss of ^{m}CHG and its pleiotropic phenotype, implies that PpCMT has an essential evolutionary conserved role in the epigenetic control of gene expression. Collectively, our results suggest functional conservation of MET1 and CMT families during land plant evolution. A model describing the relationship between MET1 and CMT in CCG methylation is presented.

Keywords DNA METHYLTRANSFERASE 1 MET1 · Chromomethylase CMT · CCG CG CHG maintenance DNA methylation · Epigenetic regulation of development · Quantification of gene copy number · *Physcomitrella patens* phase transition

Abbreviations

CMT	CHROMOMETHYLASE
DRM	DOMAINS REARRANGED METHYLTRANSFERASE
MET1	DNA METHYLTRANSFERASE

Rafael Yaari and Chen Noy-Malka have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s11103-015-0328-8) contains supplementary material, which is available to authorized users.

Transgenic lines described in this study were deposited in the International Moss Stock Center (http://www.moss-stock-center.org/) with the accessions IMSC 40758 (*ΔPpmet* 5), IMSC 40759 (*ΔPpmet* 227) and IMSC 40760 (*ΔPpmet* 262).

✉ Nir Ohad
niro@tauex.tau.ac.il

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

[2] Plant Biotechnology, Faculty of Biology, University of Freiburg, 79104 Freiburg, Germany

[3] BIOSS – Centre for Biological Signalling Studies, 79104 Freiburg, Germany

[4] FRIAS – Freiburg Institute for Advances Studies, 79104 Freiburg, Germany

Published online: 06 May 2015

Introduction

DNA methylation is a heritable epigenetic modification found in most eukaryotic organisms (Goll and Bestor 2005; Feng et al. 2010; Law and Jacobsen 2010; Zemach et al. 2010; Cedar and Bergman 2012). In plants, DNA methylation is classified by sequence context into three groups: ^{m}CG (^{m} = methylated), ^{m}CHG and ^{m}CHH (where H is A, T or G) (Feng et al. 2010; Law and Jacobsen 2010; Zemach et al. 2010). Asymmetric ^{m}CHH methylation is accomplished by the de-novo DOMAINS REARRANGED METHYLTRANSFERASE (DRM) protein family via the RNA-directed DNA methylation pathway (Cao et al. 2000; Cao and Jacobsen 2002; Law and Jacobsen 2010) and by CHROMOMETHYLASE 2 (CMT2) (Stroud et al. 2013a; Zemach et al. 2013). Symmetric ^{m}CG and ^{m}CHG methylation is preserved following DNA replication by maintenance DNA methyltransferases using the hemimethylated strand of the mother cell as a template (Goll and Bestor 2005; Law and Jacobsen 2010). ^{m}CG is maintained by the DNA METHYLTRANSFERASE 1 (MET1) protein family, while ^{m}CHG methylation, which is unique to land plants, is maintained by the CHROMOMETHYLASE 3 (CMT3) protein family (Law and Jacobsen 2010). In *Arabidopsis thaliana*, DNA METHYLTRANSFERASE 1 (AtMET1), one of four MET1 homologs, methylates mainly CG sites, but also CHG and CHH sites (Genger et al. 1999; Pavlopoulou and Kossida 2007; Cokus et al. 2008; Stroud et al. 2013b). ^{m}CG methylation mainly controls silencing of DNA repetitive elements but also affects gene expression (Zhang et al. 2006; Lister et al. 2008). *Atmet1* mutants exhibit pleiotropic morphological defects at the sporophytic stage, including delayed flowering, abnormal embryos and seed abortion, resulting from DNA hypomethylation in gametes (Vongs et al. 1993; Kankel et al. 2003; Saze et al. 2003; Jullien et al. 2006; Xiao et al. 2006; Mathieu et al. 2007). Furthermore, DNA methylation mediated by AtMET1 was shown to take part in genomic imprinting (Jullien et al. 2006, et al. 2008; Gehring 2013). In *Oryza sativa* L., disruption of *OsMET1-2*, one of two *MET1* homologs, led to loss of most ^{m}CG methylation which resulted in altered expression of genes and repeats (Hu et al. 2014). *OsMET1-2*$^{-/-}$ displayed abnormal seeds and an arrest of growth after germination (Hu et al. 2014; Yamauchi et al. 2014). While the MET1 protein family is conserved in most eukaryotes, the CMT protein family is unique to land plants, methylating CHG and CHH contexts (Henikoff and Comai 1998; Pavlopoulou and Kossida 2007; Noy-Malka et al. 2014). A single CMT3 homolog of the moss *Physcomitrella patens*, a representative of early land plants, affects ^{m}CHG and mildly ^{m}CHH (Noy-Malka et al. 2014). *ΔPpcmt* deletion mutant shows pleiotropic phenotypes during the gametophytic stage, including abnormal protonema and terminal stunted gametophores (Dangwal et al. 2014; Noy-Malka et al. 2014). In angiosperms, the CMT protein family has diverged into two subfamilies, CMT2 and CMT3 (Zemach et al. 2013). *Arabidopsis* AtCMT2 mainly affects ^{m}CHH, while AtCMT3 mainly affects ^{m}CHG (Bartee et al. 2001; Lindroth et al. 2001; Cokus et al. 2008; Lister et al. 2008; Stroud et al. 2013a; Zemach et al. 2013). No morphological defects were detected in either of the single mutants, although altered expression of genes and DNA repetitive elements was reported for *Atcmt3* mutant (Lindroth et al. 2001; Stroud et al. 2013a). Further, CMT3 homologs, Zea methyltransferase 2 (ZMET2) in *Zea mays* and NbCMT3 in *Nicotiana benthamiana*, also affect ^{m}CHG methylation (Papa et al. 2001; Hou et al. 2014). In *Arabidopsis*, multiple DNA methyltransferases have distinct as well as overlapping activities at all three sequence contexts (Cokus et al. 2008; Stroud et al. 2013a, b). *Arabidopsis Atcmt2 Atcmt3* and *Atmet1 Atcmt3* double mutants showed a wider reduction of ^{m}CHG and ^{m}CHH methylation than each of the single mutants (Cokus et al. 2008; Stroud et al. 2013a). ^{m}CHH methylation is virtually eliminated in *Atdrm1 Atdrm2 Atcmt2* triple mutant while ^{m}CHG methylation is lost only in the quadruple *Atdrm1 Atdrm2 Atcmt2 Atcmt3* mutant, indicating biochemical redundancy between the different *Arabidopsis* DNA methyltransferases (Zhang et al. 2006; Cokus et al. 2008; Lister et al. 2008; Stroud et al. 2013a).

DNMT1, the mammalian homolog of AtMET1, was shown to bind UHRF1, a specific ^{m}CG binding protein, during DNA replication (Bostick et al. 2007; Sharif et al. 2007). Thus, in mammals it was suggested that DNMT1 is recruited to hemimethylated CG sites via an adaptor protein allowing the faithful maintenance of CG methylation. In *Arabidopsis*, three UHRF1 homologs were detected: Variance In Methylation (VIM) 1, 2 and 3 (Woo et al. 2007). Indeed, CG methylation was abolished while CHG and CHH methylation decreased in *Arabidopsis Atvim1/2/3* triple mutant similarly to the *Atmet1* mutant (Stroud et al. 2013b). This suggests that AtMET1 is targeted in *Arabidopsis* via VIM binding to hemimethylated CG sites. Two pathways were shown to control DNA methylation in *Arabidopsis*: the RNA dependent DNA Methylation (RdDM) pathway (Law and Jacobsen 2010) and DDM1-mediated methylation pathway (Zemach et al. 2013). Based on mutant analysis is was shown that AtMet1 affects DNA methylation in both pathways mainly in CG context as well as CHG and CHH to a lesser extent (Cokus et al. 2008; Stroud et al. 2013b; Zemach et al. 2013). Analysis of *Arabidopsis Atmet1* mutant methylome showed complete loss of DNA methylation in all sequence contexts in four specific loci (Watson et al. 2014). Thus, AtMET1 may participate in coordinating DNA methylation in all sequence contexts in some loci while maintaining CG methylation genome wide.

DNA methylation in land plants is located mainly in repeat regions in all sequence contexts (Feng et al. 2010; Zemach et al. 2010). It was argued that DNA methylation in repetitive regions may serve as a defense mechanism by the host against genomic parasites, repressing the expression of transposable elements, as reviewed by (Kim and Zilberman 2014). In early divergent land plants such as *P. patens* and *Selaginella moellendorffii*, the genic regions are methylated upstream and downstream but not in the coding region (Zemach et al. 2010). Angiosperms however, present a high degree of ^{m}CG methylation also in coding regions with a sharp decline of DNA methylation levels towards the transcriptional start and termination sites. This decline is in correlation with increased gene expression suggesting a regulatory role for DNA methylation in controlling expression (Zemach et al. 2010).

P. patens represents an early step in land plant evolution allowing an insight into the evolution of epigenetic mechanisms in plants (Cove 2005; Rensing et al. 2008; Mosquna et al. 2009; Prigge and Bezanilla 2010; Malik et al. 2012; Dangwal et al. 2014; Noy-Malka et al. 2014). Biochemical redundancy within and between the different DNA methyltransferases families limits the ability to investigate maintenance DNA methylation in angiosperms. *P. patens* has a single member for each of the MET1 and CMT families (Malik et al. 2012), allowing us to dissect the specific roles of each family. Here, we report a functional analysis of the maintenance DNA methyltransferases PpMET and PpCMT in *P. patens*. Furthermore, the dominant gametophytic phase of mosses allows studying the role of maintenance methylation through gametophyte development, which is small and short-lived in angiosperms. Taking advantage of highly efficient homologous recombination in *P. patens* we generated *ΔPpmet* deletion mutant by gene targeting. Our study indicates evolutionary conservation of MET1 in ^{m}CG and ^{m}CCG methylation, repeat silencing and sporophyte development. Analysis of gene expression in *ΔPpcmt* mutant revealed positive correlation between loss of ^{m}CHG methylation, gene expression and altered development of the gametophyte.

Materials and methods

Plant material, culture conditions and treatments

All mutant plants were generated in the background of 'Gransden 2004' strain of *P. patens* (Ashton and Cove 1977; Rensing et al. 2008) which is deposited at the International Moss Stock Centre IMSC, and were propagated on BCD or BCDAT media (Nishiyama et al. 2000) at 25 °C under a 16-h light and 8-h dark regime (Frank et al. 2005). For induction of gametangia, plants were transferred to 16 °C under a 8 h of light and 16 h of dark regime according to (Hohe et al. 2002).

Generation of *ΔPpmet* deletion plants

ΔPpmet deletion mutant plants were generated by replacing the *PpMET* genomic coding region with either a hygromycin resistance cassette (*hptII*) or a G418 resistance cassette (*nptII*). Genomic fragments corresponding to the 5′ and 3′ flanking regions (Fig. S1a) were amplified using the following primers: 5′ fragment—HindIII Met1 5′ Fw and Met1 5′ HindIII Rv; 3′ fragment—SphI Met1 3′ Fw and Met1 3′ SphI Rv (Table S1), then cloned into the pTZ57 vector (Fermentas, Lithuania) and sequenced to ensure their integrity. Subsequently, the 5′ and 3′ fragments were subcloned into either the pMBL5 vector (GenBank: DQ228130.1) or the pMBL5 Nos Hyg vector, in which the *nptII* cassette of the pMBL5 vector was replaced by Nos following an *hptII* cassette subcloned from the pMHubi vector (Bezanilla et al. 2003). Protoplasts were subjected to PEG-mediated transformation as described (Nishiyama et al. 2000) using 15 µg of linearized plasmid. Six days after regeneration, transformants were selected on BCDAT medium containing 25 µg/ml of hygromycin (Sigma-Aldrich, USA) or 25 µg/ml of G418 (Merck, USA). Following transformation, selection and regeneration of protoplasts, hygromycin or G418 resistant plants were screened by tissue PCR to verify correct integration of the construct into the genome, by amplifying the junction regions between the insert and *PpMET* locus at both the 5′ and 3′ ends (Fig. S1b, c). Transgene copy number was determined by real-time PCR (Fig. S2) as described (Noy-Malka et al. 2014) and loss of *PpMET* transcripts was confirmed by RT-PCR analysis in three independent mutant deletion lines (Fig. S1d).

Screening for *ΔPpmet* plants by tissue PCR

A pinch of protonema tissue from each plant was immersed in 20 µl Biolong PCR buffer (Biolabs, Israel), frozen and thawed three times in liquid nitrogen followed by incubation at 68 °C for 10 min. 2 µl of each sample were used as a template for PCR reaction (Biolong enzyme, Biolabs, Israel) following the manufacturer instructions. Verification of 5′ genomic integration (Fig. S1b, c) was conducted using the following primers: MET1 KO 5′ screen Fw and 35S Rv. Verification of 3′ genomic integration (Fig. S1b, c) for the *nptII* selection cassette was conducted using primers: 35S-Ter-R-Fw and MET1 KO 3′ screen Rv; or pMHubi Fw and MET1 KO 3′ screen Rv for *hpt* selection cassette (primers listed in Table S1). PCR cycling conditions: 94 °C for 2 min, and 38 cycles of 94 °C for 15 s, 60 °C for 15 s and 72 °C for 2 min. All primers were synthesized by Hylabs (Israel).

DNA extraction

The genomic DNA samples used for the quantitative qPCR analyses and bisulfite sequencing were extracted from 120 mg of 7 days old protonema grown on BCD medium and purified using a DNeasy plant mini kit (Qiagen, Germany) following the manufacturer's instructions. DNA integrity was analyzed by 1 % agarose gel electrophoresis in 0.5 × TBE with ethidium bromide staining. 0.5 µl was used per qPCR reaction and 10 µl per bisulfite reaction.

Rt-pcr

Total RNA was extracted using SV Total RNA Isolation System (Promega, USA) from 7 days old protonema grown on BCD medium. The cDNA was synthesized from 2 µg of total RNA using SuperScript® III First-Strand Synthesis System (Life Technologies, USA) with the oligo-T 20 primer. RT-PCR was performed as described (Katz et al. 2004), using the following primers: PpMET—(PpMET 3124 Fw and PpMET 10299 Rv); rRNA—(rRNA 3402 Fw and rRNA 3403 Rv) (Table S1) using the following cycling conditions: 94 °C for 2 min, and cycles of 94 °C for 15 s, 57 °C for 15 s and 72 °C for 60 s; 40 cycles of amplification were used for PpMET and 20 cycles for rRNA.

RT-qPCR

qPCR was conducted as described before (Noy-Malka et al. 2014). The following primers were used to detect expression of R1 and R2 ORFs using cDNA generated from wild type (WT) and *ΔPpmet* line 5: R1- Sca 256 RT Fw, Sca 256 RT Rv, Sca 256 RTM Fw and Sca 256 RTM Rv; R2-Sca 332 RT Fw and Sca 332 RT Rv (Table S1).

The following primers were used to validate expression levels of four PpCMT regulated genes using cDNA generated from WT or *ΔPpcmt* line 281 (Noy-Malka et al. 2014): Pp1s10_78V6 Fw and Pp1s10_78V6 Rv; Pp1s311_8V6 Fw and Pp1s311_8V6 Rv; Pp1s39_36V6 Fw and Pp1s39_36V6 Rv; Pp1s71_52V6 Fw and Pp1s71_52V6 Rv (Table S1).

The following primers were used to validate expression levels of twelve PpMET regulated ORFs using cDNA generated from WT or *ΔPpmet* lines 5, 227 and 262: 333 and 334; 335 and 336; 339 and 340; 341 and 342; 345 and 346; 347 and 348; 349 and 350; 355 and 356; Pp1s102_90V6.1 fw and Pp1s102_90V6.1 rev; Pp1s56_169V6.1 fw and Pp1s56_169V6.1 rev; Pp1s104_50V6.1 fw and Pp1s104_50V6.1 rev; Pp1s424_17V6.1 fw and Pp1s424_17V6.1 rev (Table S1).

The following primers were used to amplify normalizing or validation accuracy genes: PpTATA-binding protein 2 (Pp1s246_34V6.1)—TATA binding p Left and TATA binding p Right; PpHistone3 (Pp1s1963_1V6.1 and Pp1s3_594V6.1)—PpHistone3 157 Fw and PpHistone3 211 Rv or PpHistone3 100 Fw and PpHistone3 143 Rv; and PpCLF (Pp1s100_146V6.1)—PpCLF mix 8 7739 Fw and PpCLF mix 8 7804 Rv (Table S1).

Quantification of genomic copy number by qPCR

Quantification of transgene copy number was performed as described before (Noy-Malka et al. 2014). *ΔPpmet* lines 5, 227 and 262 transgene copy number was determined by amplifying the 5′ fragment used for homologues recombination compared to its amplification in WT which bears only one copy. PpMET 5′ KO 325 Fw and PpMET 3′ KO 390 Rv primers (Table S1) were used for amplification of the PpMET 5′ fragment. PpCLF gene (Pp1s100_146V6.1) was used for normalization using primers: PpCLF mix 8 7739 Fw and PpCLF mix 8 7804 Rv (Table S1). As an internal control for quantification, an additional set of primers, PpFIE—245 Left and PpFIE—213 Right primers (Table S1), amplifying a region of the PpFIE gene (Pp1s535_6V6.1), were used. This internal control would be expected to give the relative ratio showing a single copy sequence.

Bisulfite assay and methylation sensitive DNA restriction followed by qPCR

Bisulfite conversion and methylation sensitive DNA restriction followed by qPCR were performed as described before (Noy-Malka et al. 2014).

Genomic BS-seq analysis

BS-seq data adopted from (Zemach et al. 2013) of *Arabidopsis* WT and *Atmet1-6* roots was analyzed for ^{m}CHG DNA methylation levels in sequences (^{m}CAG, ^{m}CTG and ^{m}CCG). Only sites with at least five reads (C+T>4) were included in the analysis.

Microscopy

Plant morphology along the *P. patens* life cycle was observed using stereomicroscope Stemi SV11 Apo (Carl Zeiss AG, Germany), SZRX-ILLB2-200 (Olympus, Japan) and Axioplan2 (Zeiss) microscopes equipped with Olympus DP71 and Coolpix P5100 cameras (Nikon, Japan).

Protoplast preparation for characterization

A single protoplast served as a starting point for phenotypic characterization. To this end, 7 days old protonema tissue, grown on BCDAT medium was incubated for

30 min in room temperature with 0.125 mg Driselase (Kyowa Hakko Kogyo Co., Ltd, Japan) in 1 ml 8 % mannitol (Sigma-Aldrich) for protoplast isolation. Next, the protoplast solution was filtrated with a sterile 40 μM mesh, washed twice with 1 ml 8 % mannitol, resuspended with 1 ml PLM (Nishiyama et al. 2000) and incubated overnight in 22 °C at dark. Next day, the protoplasts were resuspended with 5 ml modified liquid PRM/T (Nishiyama et al. 2000) containing 7 % mannitol instead of 8 % and no agar, and incubated for 7 days under normal growth conditions with gentle shaking. Finally, 1 ml of the protoplast solution was relocated to solid BCD medium plates.

Gene expression array

Total RNA was extracted from 120 mg fresh weight material from 7 days old protonema grown on BCDAT medium from WT, *ΔPpcmt* line 281 (Noy-Malka et al. 2014) and *ΔPpmet* line 5 using the SV Total RNA Isolation System (Promega) according to the manufacturer's protocol. Next, residual DNA was removed by DNA-free™ DNA Removal Kit (AM1906, Ambion Life technologies) according to the manufacturer's protocol. 200 ng of total RNA was reverse transcribed and amplified with the TransPlex Whole Transcriptome Amplification (WTA) Kit (Sigma-Aldrich, USA). The design of the array NimbleGen_Ppat_SR_exp_HX12: NimbleGen 12 × 135 k chip (Roche, Switzerland) is based on the V1.6 gene models (Zimmer et al. 2013) and represents 32,275 transcripts with an average of four 60 mer probes per gene. 1 μg of cDNA was labeled with Cy3 according to the NimbleGen One-Color DNA Labeling Kit (Roche). The results were deposited in ArrayExpress (www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-3431/) including protocols of microarray hybridization using 4 μg of labelled cDNA, washing, imaging and data processing.

Microarray expression values were analyzed as previously described (Wolf et al. 2010; Hiss et al. 2014; Beike et al. 2015) using the Expressionist Analyst 7.5.7 (Genedata, Switzerland). Differentially regulated genes were detected using the Bayesian regularized unpaired CyberT test (Baldi and Long 2001) with Benjamini-Hochberg false discovery rate correction ($q \leq 0.005$).

Putative protein domains were detected using the NCBI conserved domains database (Marchler-Bauer et al. 2011, 2013, 2015). Homology analyses were conducted using NCBI BLAST+ version 2.2.29 locally (Camacho et al. 2009). Sequences were considered similar when blast results showed identity of at least 75 nucleotides in a minimal alignment of 75–100 nucleotides. Genomic loci separated by <1000 bp were integrated.

Results

PpMET is involved in ^{m}CG and ^{m}CCG methylation

The *P. patens* genome encodes for a single MET1 homolog designated PpMET (Malik et al. 2012). To examine the role of PpMET in DNA methylation, we have generated three independent *ΔPpmet* deletion mutant plants in which the entire coding region was deleted (Figure S1, S2). DNA methylation levels of *ΔPpmet* plants were tested at four, non-repetitive highly methylated genomic loci designated R1 to R4, described previously (Noy-Malka et al. 2014). R1 and R2 loci are enriched for ^{m}CG and ^{m}CHG contexts, and located within predicted ORFs. R3 and R4 loci are enriched for ^{m}CHH context and located in intergenic regions.

Bisulfite analysis showed almost complete loss of ^{m}CG at R1 and R2 loci in *ΔPpmet* (Fig. 1; Table S2), demonstrating the role of PpMET in ^{m}CG. The CHG context represents three DNA sequences: CAG, CTG and CCG. Interestingly, while DNA methylation was almost unaffected at ^{m}CAG and ^{m}CTG contexts, it was completely lost at ^{m}CCG context (Fig. 1; Table S2). To note, the single sites ^{m}CAG and ^{m}CTG at R3 and a single ^{m}CAG site at R4 preserved methylation in *ΔPpmet*, while the two ^{m}CG sites at the R4 locus lost methylation (there are no CG sites at R3) similarly to the specific loss of ^{m}CG DNA methylation in R1 and R2 (Table S2).

Similarly, ^{m}CCG methylation loss was previously described in centromeric repeats of *Arabidopsis Atmet1* antisense mutant, via genomic restriction (Finnegan et al. 1996). In view of the above, we re-examined the distribution of methylation in CAG, CTG and CCG contexts based on genomic BS-seq data reported for *Arabidopsis* WT and *AtMet1* mutant (Zemach et al. 2013). Global ^{m}CHG methylation is reduced by 33 % in *AtMet1* mutant (Cokus et al. 2008). Our analysis revealed that the ^{m}CHG methylation loss was specific to the ^{m}CCG context (Fig. 2), while ^{m}CAG and ^{m}CTG methylation levels were similar to those of WT in this dataset from *Arabidopsis*. These results are in agreement with the loss of ^{m}CCG methylation identified in *ΔPpmet* in this study and with the loss of ^{m}CCG methylation in centromeric repeats of *Arabidopsis met1* antisense mutant described previously (Finnegan et al. 1996), indicating that in addition to their role in CG methylation PpMET as well as AtMet1 take part in methylating ^{m}CCG.

CHH methylation was mildly reduced in *ΔPpmet* from 55 and 41 to 31 and 29 % at R1 and R2, respectively (Fig. 1; Table S2), however, it was unaffected in the R3 and R4 loci, which are enriched for ^{m}CHH (Fig. 3; Table S2). This indicates that PpMET is not involved in methylation at loci which are highly methylated in CHH context.

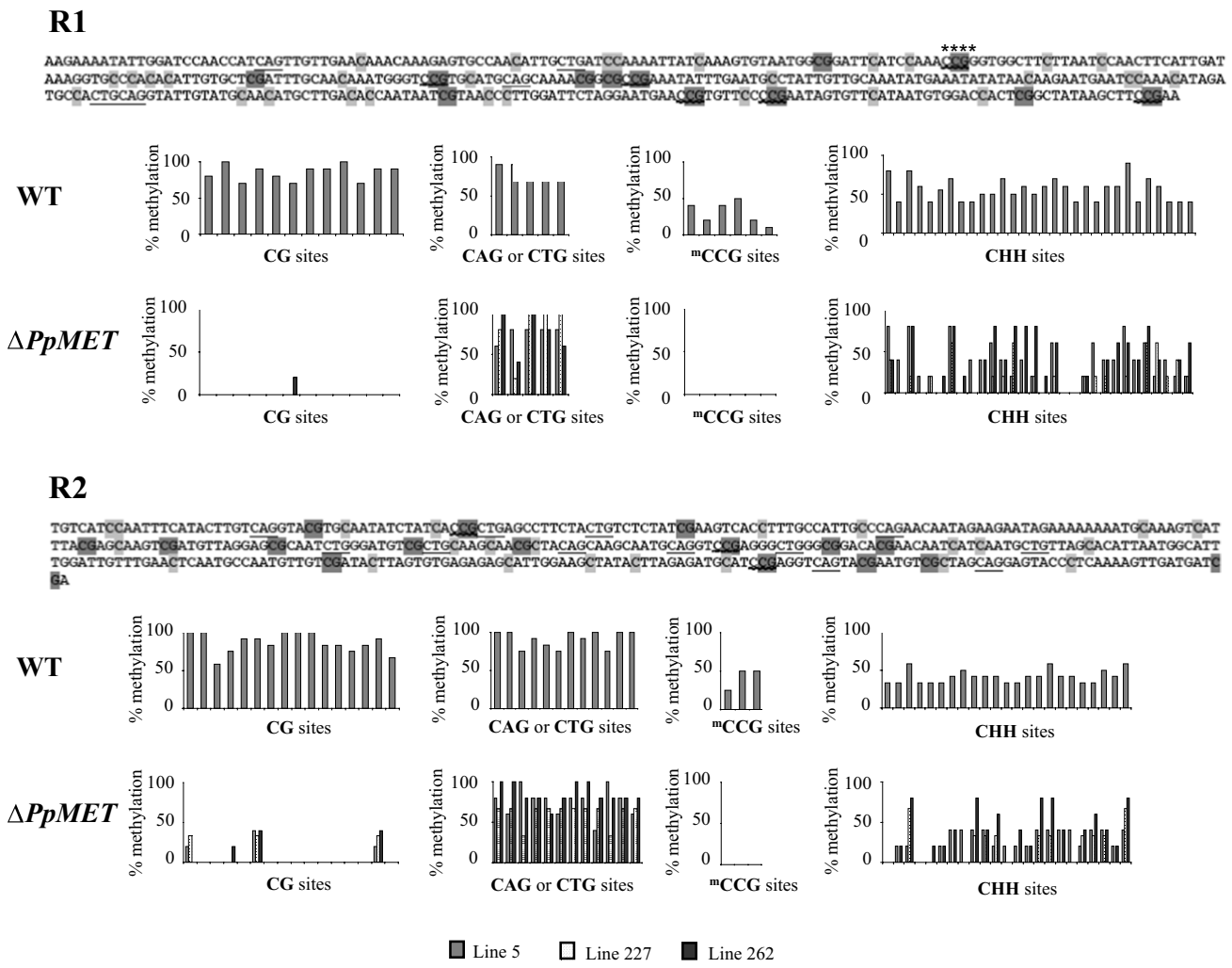

Fig. 1 DNA methylation levels in WT *P. patens* and *ΔPpmet* deletion lines at the R1 and R2 loci. The methylation status of three independent *ΔPpmet* deletion lines (lines 5, 227 and 262) were compared to WT by bisulfide sequencing. Through the sequences, CG sites are marked by dark grey, CAG and CTG sites are underlined and CCG sites are underlined by a waved line. CHH bar charts include only the sites which were over 30 % methylated in WT and are marked by light grey in the sequences. The bar charts are separated into CG, CAG and CTG, ^{m}CCG or CHH-context groups for each locus. ^{m}CCG context refers to methylation of the first C residue in the CCG trinucleotide. The bar charts indicate the methylation levels for individual C residues (*X-axis*) in the order they appear in the sequence. The methylation levels are calculated from 100 % methylation potential (if all reads were methylated). DNA methylation is completely lost at CG and CCG sites in *ΔPpmet* but unaffected at CAG and CTG sites

To further establish the role of PpMET in DNA methylation, we employed a complementary tissue-wide approach comparing DNA methylation levels of WT and *ΔPpmet*. DNA methylation was evaluated using methylation-sensitive DNA restriction assay followed by qPCR at a single CCGG genomic site within the R1 locus (Noy-Malka et al. 2014). *HpaII* and *MspI* restriction enzymes recognize the same CCGG sequence but show different sensitivity to methylation. *MspI* is blocked when the first cytosine is methylated (^{m}CCGG), while *HpaII* is unable to restrict when the second cytosine is methylated (C^{m}CGG). Genomic DNA extracted from WT or *ΔPpmet* protonema tissue was incubated with either *HpaII* or *MspI*. Next, the extent of methylation was evaluated via qPCR using primers flanking the CCGG site at the R1 locus in comparison to untreated genomic DNA. While in WT almost all DNA molecules were protected from *HpaII* restriction (Fig. 4), 84 % of *ΔPpmet* DNA molecules were digested, indicating almost a complete loss of ^{m}CG at this site in *ΔPpmet*. Restriction with *MspI* showed that 67 % of the DNA molecules were protected in WT (Fig. 4), while only 4 % were protected in *ΔPpmet* plants, indicating that ^{m}CCG methylation levels at this site were severely affected by the absence of PpMET. Similar results were obtained using genomic DNA extracted from gametophore tissue of WT and *ΔPpmet* (Figure S3). Collectively, our results indicate

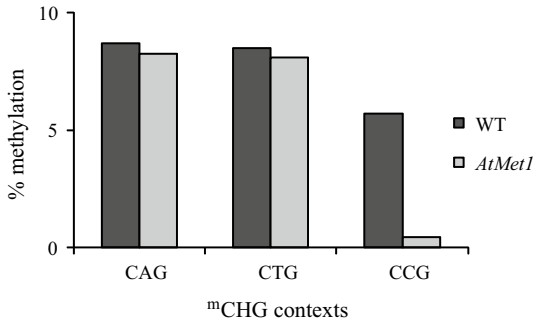

Fig. 2 Genomic CHG DNA methylation levels at CAG, CTG and CCG contexts of WT *Arabidopsis* versus *Atmet1* mutant. Genomic BS-seq data adopted from (Zemach et al. 2013) of WT and *Atmet1-6* roots was analyzed for methylation levels at CAG, CTG and CCG contexts. ^{m}CCG but not ^{m}CAG or ^{m}CTG was eliminated in the *Atmet1* mutant. The methylation percentage is considered as 100 %, if all cytosines were methylated

that PpMET is involved in methylation of both C^{m}CG and ^{m}CCG contexts.

ΔPpmet plants fail to form sporophytes

To study the role of PpMET during *P. patens* development, morphological analysis of *ΔPpmet* plants was conducted throughout their lifecycle and compared to WT (Fig. 5). All three deletion lines showed similar morphology. *ΔPpmet* development during the gametophyte vegetative phase was similar to WT, including protonema and buds (Fig. 5a, c vs. b, d), as well as gametophores (Fig. 5e vs. f). Turning into the reproductive phase, male and female sex organs were indistinguishable from WT (Fig. 5g, i vs. h, j). However, no sporophytes were observed in *ΔPpmet* plants subjected to fertilization-inductive conditions in three different experiments, even after 6 months (Fig. 5k vs. l). This suggests that PpMET is not involved in differentiation of protonema, bud initiation and gametophore development but has an essential role in either gamete development, fertilization or sporophyte development.

Loss of PpMET affects expression of multiple repetitive sequences

To test whether loss of DNA methylation has an effect on gene expression genome-wide, we have used the NimbleGen_Ppat_SR_exp_HX12 microarray comparing WT, *ΔPpmet* and *ΔPpcmt* expression patterns in 7 days old protonema tissue. The array provided coverage for all gene models predicted in the annotation version 1.6 (Zimmer et al. 2013). For *ΔPpmet*, the analysis revealed 79 significantly upregulated putative ORFs and one significantly downregulated with no homology to known genes (Table S3). In general, the length of these ORFs ranged from 150 to 600 bp and did not contain functional protein domains. Our analysis identified that these ORFs have multiple similar sequences at different loci in the genome. For 8 ORFs less than 10 similar loci were identified while the remaining 72 sequences have between 17 and 237 copies in the genome, some of which reside in close proximity to each other but were annotated as separate gene models and represented as such on the array (Table S3). 73 out of the 79 upregulated ORFs clustered by sequence homology into six different groups (designed groups A–F) (Table S3). To note, the probes used in our array do not necessarily discriminate between the expression of different ORF's within the same homology group due to sequence similarity. To validate the array results, RT-qPCR analysis was performed on a subset of targets (Figure S4), where seven out of eight targets tested were indeed upregulated in the mutant. These results indicate a role for PpMET in regulating the expression of a subset of repetitive sequences in *P. patens* genome at the protonema stage.

To test whether the increase in expression of these genomic repetitive sequences in *ΔPpmet* is a result of DNA methylation loss, we examined the published methylome landscape (Zemach et al. 2010) within these regions of WT. Furthermore, we examined the methylome landscape of the homologous sequences of the identified PpMET regulated repetitive sequences (Table S4) as the expression readout may arise from distinct genomic loci which have similar sequences but are not annotated as gene models and are therefore not represented on the array. Our analysis showed that PpMET regulated repetitive sequences and their homologs are highly methylated in WT, characterized by 60–80 % ^{m}CG, 60–80 % ^{m}CHG, and 30–40 % ^{m}CHH and by a decline of methylation in the regions flanking these sequences (Fig. 6c). This pattern of methylation is characteristic for DNA repeats in *P. patens* as described by (Zemach et al. 2010).

Four ORFs regulated by PpMET, Pp1s56_169V6.1, Pp1s102_90V6.1, Pp1s104_50V6.1 and Pp1s424_17V6.1 are characterized by up to 8 genomic copies and long transcripts (Table S3). For Pp1s102_90V6.1 and Pp1s104_50V6.1 protein domains are predicted, lipase and peptidase, respectively, indicating a functional role of these gene products. RT-qPCR on three independent *ΔPpmet* mutant lines confirmed the array results for two of the upregulated genes Pp1s104_50V6.1 and Pp1s424_17V6.1 (Fig. 6a). Interestingly, the methylome landscape of these two genes in WT, based on (Zemach et al. 2010), show high levels of DNA methylation upstream to their coding region (Figure S5), which may indicate a role for ^{m}CG and ^{m}CCG methylation in regulating their expression.

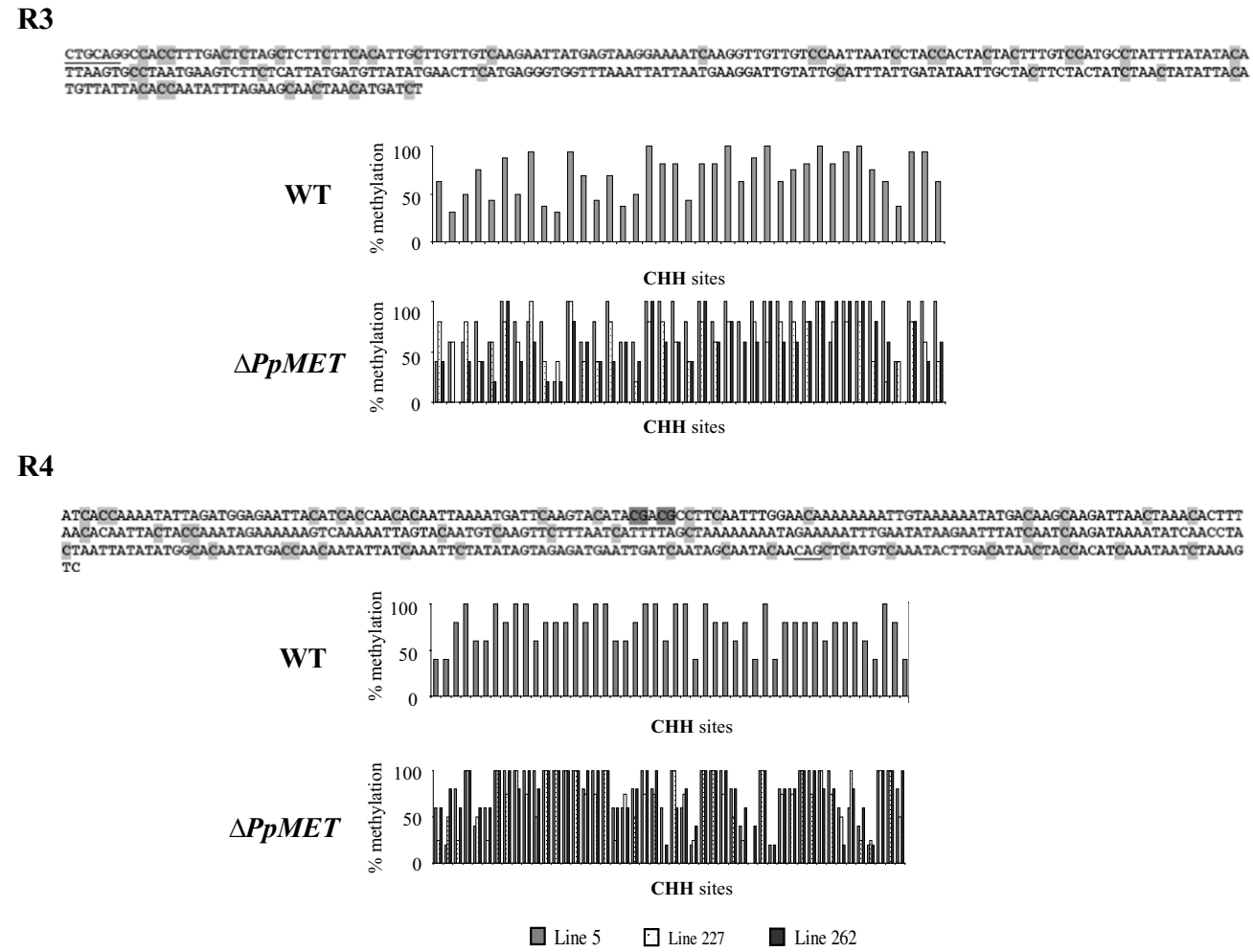

Fig. 3 CHH DNA methylation levels in WT *P. patens* and *ΔPpmet* deletion lines at the R3 and R4 loci. The methylation status of three independent *ΔPpmet* deletion lines (lines 5, 227 and 262) were compared to WT. Through the sequences, CG sites are marked by dark grey and CAG/CTG sites are underlined. CHH bar charts include only the sites which were over 30 % methylated in WT and are marked by light grey in the sequence. These sites are presented in the bar charts, indicating the methylation levels for individual C residues (*X-axis*), in the order they appear in the sequence. The methylation levels are calculated from 100 % methylation potential (if all reads were methylated). CHH methylation levels were similar to WT in all three *ΔPpmet* lines

PpCMT regulates gene expression in protonema tissue

Our microarray analysis for *ΔPpcmt* revealed 35 upregulated and 6 downregulated genes (Table S5). As there was no overlap between PpCMT and PpMET regulated ORFs we infer that distinct mechanisms affect their expression. For half of the 41 target genes putative protein domains associated with distinct cellular functions were assigned (Table S5). RT-qPCR analysis validated the enhanced expression of three out of four *ΔPpcmt* regulated genes tested (Fig. 6b). Analysis of the methylation pattern in PpCMT regulated genes based on the published *P. patens* methylome (Zemach et al. 2010) revealed absence of DNA methylation within the coding regions and increased levels (up to 35 % ^{m}CG and ^{m}CHG, up to 20 % ^{m}CHH) of methylation up- or down-stream of the coding regions (Fig. 6d). This is in accordance with the prevailing DNA methylation landscape of genes in *P. patens* (Zemach et al. 2010). To note, the array used in this study did not include probes for a previously reported PpCMT regulated gene, Pp1s234_91V6.2 (Dangwal et al. 2014), due to differences in gene annotation between versions 1.2 (Rensing et al. 2008) and 1.6 (Zimmer et al. 2013). The downregulation of Pp1s234_91V6.2 was verified by qRT-PCR in *ΔPpcmt* (Figure S6).

The microarray analysis of *ΔPpcmt* additionally identified two upregulated repetitive sequences (Table S5) which have tens of thousands of similar sequences in the genome. Our analysis show that PpCMT regulated repetitive sequences and their homologs are highly methylated

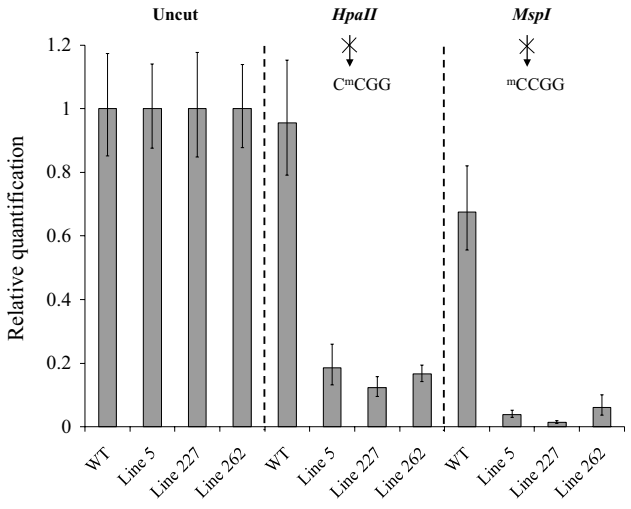

Fig. 4 DNA methylation levels at a single site within the R1 locus in WT *P. patens* and *ΔPpmet*. Relative quantification (RQ) of DNA methylation levels at a specific site (indicated in the R1 sequence by *asterisks* in Fig. 1) was performed by methylation sensitive restriction assay using *MspI* (sensitive to ^{m}CHG methylation) or *HpaII* (sensitive to ^{m}CG methylation), followed by qPCR. *X marked on the arrows* indicates inability of the restriction enzyme to cleave when the site is methylated. RQ of WT and three *ΔPpmet* lines 5, 227 and 262 was calculated relatively to uncut DNA which served as a reference in each genetic background. Results were normalized to the *PpHistone3* gene. *Error bars* indicate standard deviation of three technical replicates. Results indicated almost a complete loss of methylation at both ^{m}CG and ^{m}CCG contexts at this site in *ΔPpmet* mutant

in WT (Fig. 6e), characterized by 60–80 % ^{m}CG, 60–80 % ^{m}CHG, and 30–40 % ^{m}CHH and by a decline of methylation in the regions flanking these sequences. This landscape is similar to the DNA methylation landscape characteristic to genomic repeats described by (Zemach et al. 2010). These results may indicate an additional role for PpCMT in repressing specific sets of repetitive sequences.

Discussion

Methylation of cytosines is an epigenetic modification found in many eukaryote species, including plants and animals, which affects gene expression and transposon silencing (Goll and Bestor 2005; Zhang et al. 2006; Lister et al. 2008; Feng et al. 2010; Law and Jacobsen 2010; Zemach et al. 2010; Cedar and Bergman 2012; Stroud et al. 2013a). Maintenance of symmetrical DNA methylation in plants occurs in both CG and CHG sequence contexts (Law and Jacobsen 2010). In *Arabidopsis*, AtMET1 and AtCMT3 are required for the maintenance of ^{m}CG and ^{m}CHG DNA methylation, respectively (Cokus et al. 2008; Lister et al. 2008). In an effort to understand the evolutionary functions of maintenance DNA methylation in land plants, we studied mutants in which the single homologs of CMT and MET1 in *P. patens* were deleted. We previously reported loss of ^{m}CHG methylation in four loci and a severe gametophore stunted phenotype in *ΔPpcmt* deletion plants (Noy-Malka et al. 2014). In this study, we generated *ΔPpmet* deletion plants which resulted in elimination of ^{m}CG DNA methylation at the same loci (Figs. 1, 3, 4) and in sterility, as the mutants were unable to produce sporophytes (Fig. 5). The specific context methylation loss of *ΔPpcmt* and *ΔPpmet* in the four genomic loci tested may indicate a global role for PpCMT and PpMET in CHG and CG methylation in *P. patens*, respectively. Thus, we suggest an evolutionary conservation of the biochemical activity of MET1 and CMT protein families in land plants.

Furthermore, our study revealed a specific methylation loss at ^{m}CCG sites, but not ^{m}CAG nor ^{m}CTG, in four loci analyzed in *ΔPpmet* (Figs. 1, 4). This is in agreement with DNA methylation loss detected in *Arabidopsis Atmet1* antisense mutant centromeric repeats (Finnegan et al. 1996). Our re-analysis of the *Arabidopsis Atmet1* published methylome showed a genome wide specific methylation loss at ^{m}CCG context (Fig. 2). Taken together, these results indicate that MET1 family affects methylation of ^{m}CCG sites in both, *Physcomitrella* and *Arabidopsis.* Specific loss of ^{m}CHG, including methylation at ^{m}CCG sites, was shown also in *Arabidopsis Atcmt3* and *P. patens ΔPpcmt* (Cokus et al. 2008; Lister et al. 2008; Noy-Malka et al. 2014). We suggest that only the CMT family is involved in ^{m}CCG methylation but is dependent on the complementary hemimethylated ^{m}CGG site which is methylated first by a MET1 family member (Fig. 7). CCG and CGG sequence contexts, although complementary to each other, are not symmetric with regards to DNA methylation. While the CCG context contains two cytosines in the CG and CHG contexts, the CGG context contains only one cytosine in the CG context. The loss of ^{m}CCG methylation in *cmt* mutants indicates that CMT family, but not MET1 family proteins, methylate ^{m}CCG sites. In agreement, methylomes of eukaryotes reported so far, which lack CMT homologs also lack CHG methylation including ^{m}CCG context, thus supporting the role of CMTs in ^{m}CCG methylation (Noy-Malka et al. 2014). The above results indicate that MET1 is not involved in ^{m}CCG methylation, yet *met1* mutants showed loss of ^{m}CCG methylation. This discrepancy can be explained by the dependency of ^{m}CCG methylation by CMT on the complementary hemimethylated ^{m}CGG site which in WT is methylated by MET1 (Fig. 7). The specific loss of ^{m}CCG methylation but not ^{m}CAG nor ^{m}CTG in *met1* mutants (Figs. 1, 2) suggests that CMT targeting is not affected but rather its efficiency in ^{m}CCG methylation. In support of the above, in vitro assays showed that methylation of unmethylated substrate by CMTs is less efficient than hemmimethylated substrate (Pradhan and Adams 1995; Du et al. 2012). These results allow proposing a mechanistic explanation

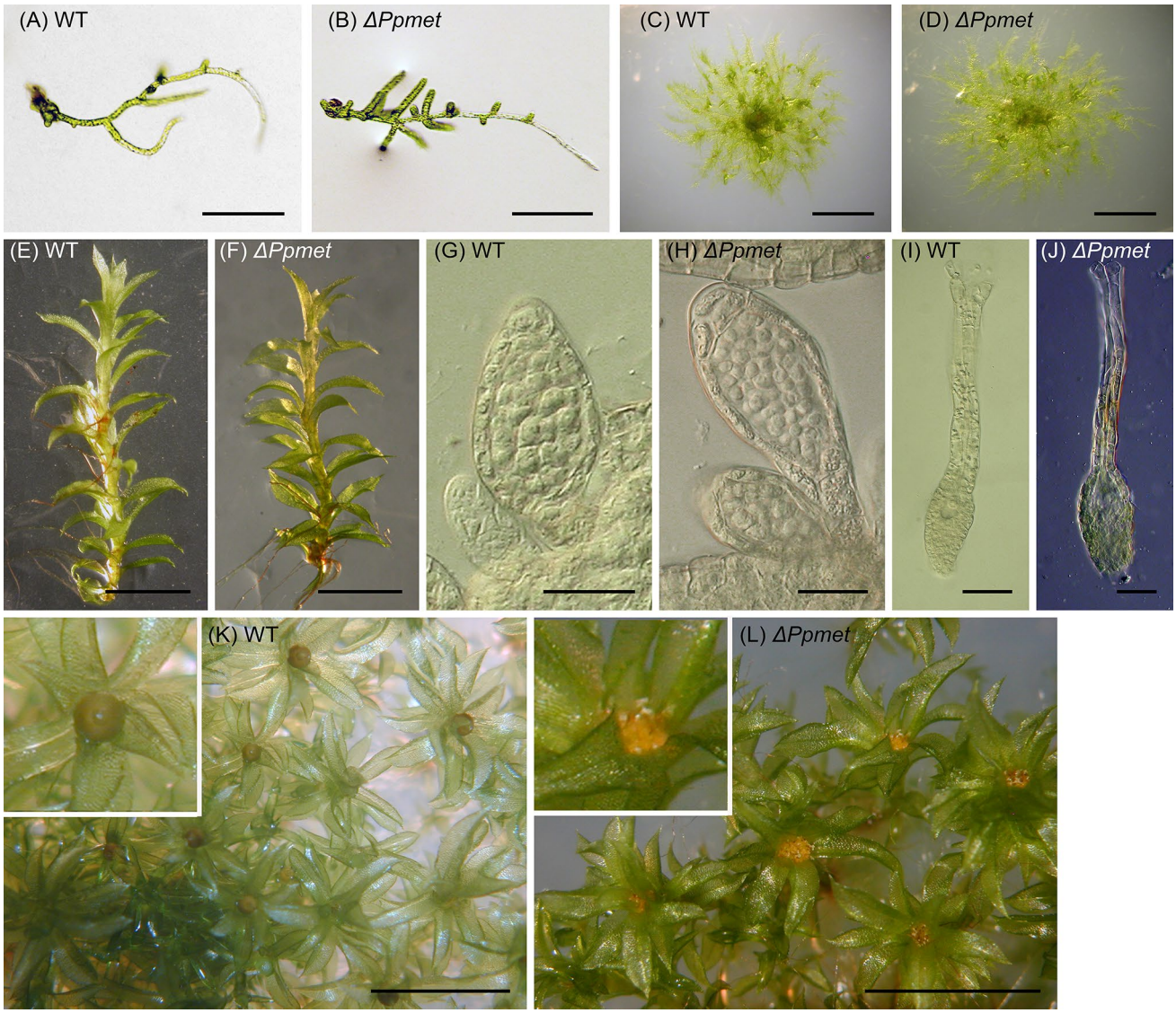

Fig. 5 Morphological analysis of *ΔPpmet* deletion mutant as compared to WT *P. patens*. **a** WT and **b** *ΔPpmet* protonema were grown for 5 days after protoplast plating. **c** WT and **d** *ΔPpmet* 2 weeks old plants after protoplast plating. **e** WT and **f** *ΔPpmet* mature gametophores. **g** WT and **h** *ΔPpmet* antheridia. **i** WT and **j** *ΔPpmet* archegonia. **k** WT gametophore apexes bearing sporophytes (*insertion*) and **l** *ΔPpmet* gametophore apexes 8 weeks after gametangia induction bearing only gametangia and no sporophytes (insertion). Plants were grown on solid BCD medium. *Scale bars* **g**, **h** = 25 μm; **i**, **j** = 50 μm; **a**, **b** = 200 μm; **e**, **f** = 1 mm; **c**, **d**, **k**, **l** = 5 mm

for ^{m}CCG methylation loss in both, *cmt* and *met*1 mutants, although only CMT is involved in ^{m}CCG methylation (Fig. 7).

Upregulation of genes in *ΔPpcmt*, as identified through the expression array (Table S5), suggests that regulation of expression presumably mediated by CHG methylation is a molecular mechanism which existed already at the onset of early land plants. This is consistent with *P. patens* being the earliest diverged plant in which a CMT gene was identified (Noy-Malka et al. 2014) and its emergence may have contributed to the regulation of developmental processes leading to a more complex structural morphology of plants. In addition, two highly repetitive sequences were also identified to be regulated by PpCMT, which may indicate a role for PpCMT in silencing of repetitive sequences as well. Similarly, *Arabidopsis* AtCMT3 was shown to take part in regulating the expression of genes and repetitive sequences (Stroud et al. 2013a), which indicates conservation of CMT role in expression regulation during land plant evolution. *ΔPpcmt* phenotype (Noy-Malka et al. 2014) together with miss-expression of genes indicate that PpCMT is involved in regulating growth and development of the gametophytic stage.

Based on the upregulation of a limited number of genomic repetitive sequences in *ΔPpmet* (Table S3)

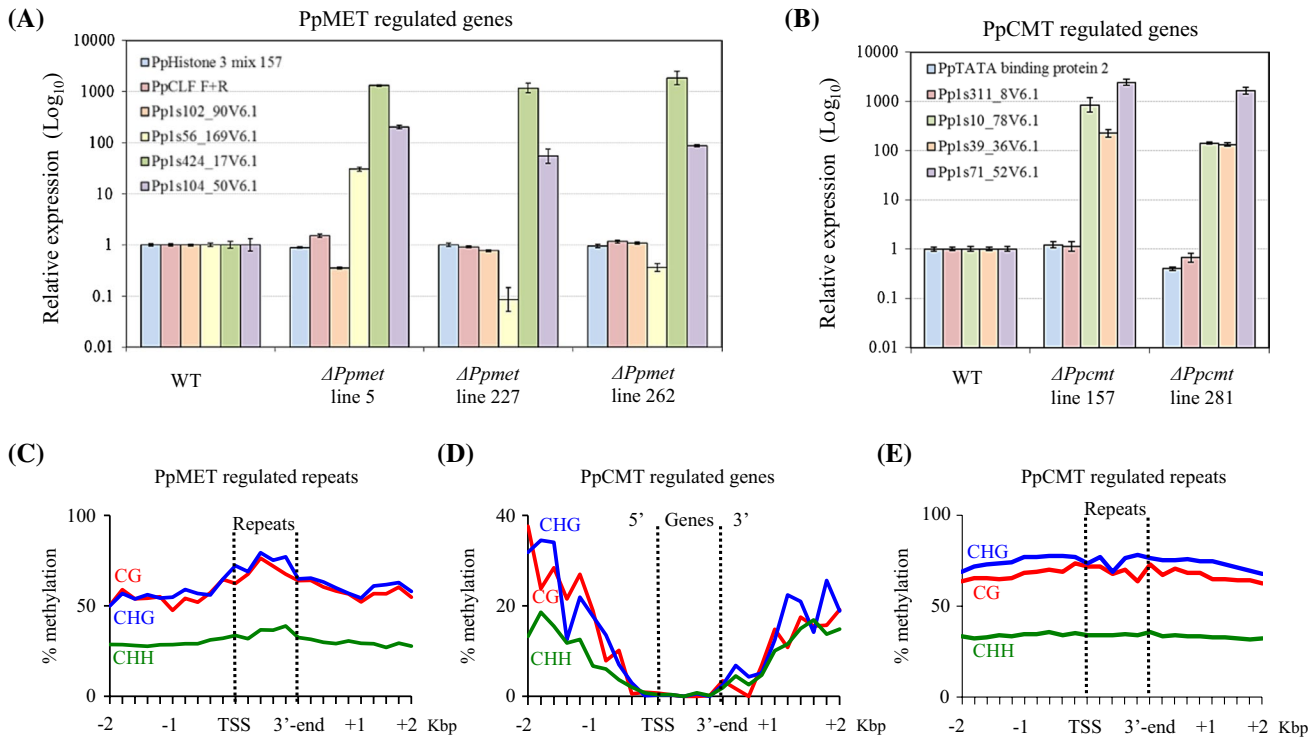

Fig. 6 Expression analyses and DNA methylation landscape of PpCMT and PpMET regulated genes and repetitive sequences. **a** RT-qPCR analysis of four PpMET regulated genes (Table S3) in three *ΔPpmet* deletion lines (5, 227 and 262). *PpTATA binding protein 2* (primers listed in Table S1) was used to normalize the expression levels. *PpHistone3* and *PpCLF* genes were used to validate quantification accuracy. Relative expression values are presented on a logarithmic scale. *Errors bars* indicate standard deviation of three technical replicates. **b** RT-qPCR analysis of four PpCMT up-regulated genes (Table S5) in *Ppcmt* lines 157 and 281 and WT. The expression levels were normalized to the expression levels of *PpHistone3* (primers listed in Table S1). *PpTATA binding protein 2* gene was used to validate quantification accuracy. Relative expression values are presented on a logarithmic scale. *Errors bars* indicate standard deviation of three technical replicates. **c–e** Average of DNA methylation levels in 200 bp windows adopted from (Zemach et al. 2010) plotted continuously for the first 1000 bp of ORF sequences and 2000 bp upstream (5′) or downstream (3′) away from the ORF. Dashed lines represent ORF borders starting at the transcription start site (TSS) and ending at the 3′-end (which represents the end of the first 1000 bp of ORF sequences). The data is plotted by sequence context ^{m}CG, ^{m}CHG and ^{m}CHH represented by red, blue and green lines, respectively. **c** DNA methylation landscape at 318 loci similar to PpMET regulated repeats (Table S4). **d** 41 PpCMT regulated genes (Table S5). **e** Tens of thousands loci similar to the two PpCMT regulated repetitive sequences (Table S5)

present on the array, we suggest that PpMET is involved in silencing of such repetitive genomic loci. The extent by which PpMET affects the expression of repeats in the genome of *P. patens* could be wider than those we identified, as the expression array used in this study contained probes for only some repetitive sequences. Upregulation of expression from repetitive genomic loci in *ΔPpmet* is similar to that reported for *Arabidopsis* AtMET1 (Saze et al. 2003; Zhang et al. 2006; Mathieu et al. 2007; Lister et al. 2008) and for *O. sativa* L. OsMET1-2 (Hu et al. 2014; Yamauchi et al. 2014), suggesting an evolutionary conservation of MET protein family in silencing DNA repeats in land plants. Gene expression analysis of *ΔPpmet* protonema and subsequent RT-qPCR allowed to identify only two upregulated genes (Fig. 6a) suggesting that ^{m}CG methylation loss in *ΔPpmet* as detected by four genomic loci is not a robust mechanism for governing expression of protein-coding genes at the gamethopytic vegetative stage in *P patens*. However, the lack of sporophyte development in *ΔPpmet* (Fig. 5) indicates impairment of either gamete development, fertilization or early embryo development. In view of the correlation between loss of CG methylation in *ΔPpmet* protonema tissue and the upregulation of particular genes, we suggest that loss of ^{m}CG methylation in *ΔPpmet* can affect gene expression which may explain the lack of sporophyte development in *ΔPpmet*. Although one cannot rule out the possibility that ^{m}CG methylation is indispensable for early sporophyte development, it is more likely a result of ^{m}CG methylation loss in the *ΔPpmet* mutant during gamete formation, as shown for angiosperms. In *Arabidopsis*, lack of AtMET1 in the gametophyte led to epigenetic alterations during gamete formation resulting in impaired sporophytes presenting pleiotropic and stochastic phenotypes including late

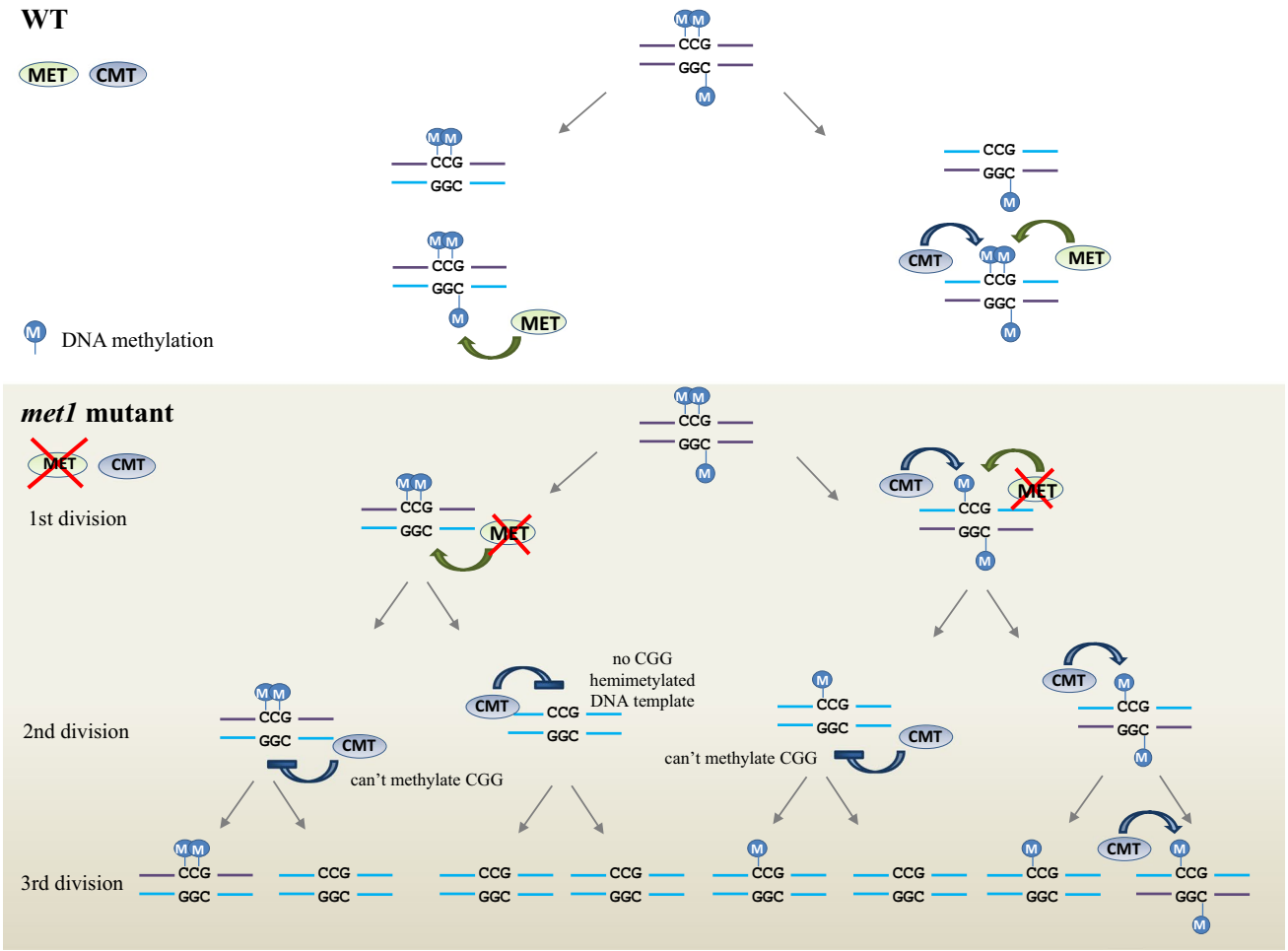

Fig. 7 A schematic model representing suggested events which allow maintenance of $^{m}C^{m}CG$ methylation by MET1 and CMT. In WT, maintenance of $^{m}C^{m}CG$ methylation is mediated by MET1 methylating $C^{m}CG$ and ^{m}CGG context and CMT methylating ^{m}CCG context using the original hemimethylated ^{m}CGG as a template. In the absence of MET1, ^{m}CG methylation is lost and cannot be methylated by CMT. Upon DNA replication, newly synthesized CCG sequences, complementary to unmethylated CGG sites, may not be methylated by CMT as such sites are not hemimethylated thus cannot serve as template. In *met1* mutants, eventually ^{m}CCG methylation is lost through DNA replication cycles passively

flowering, partial fertility and abnormal seeds (Saze et al. 2003; Jullien et al. 2006; Mathieu et al. 2007). The late flowering phenotype was attributed to ectopic expression of a single gene, *AtFWA*, as a result of ^{m}CG methylation loss (Saze et al. 2003). *O. sativa* L. *OsMET1-2*$^{-/-}$ mutant plants also showed abnormal seeds, which could not germinate under regular growth conditions (Hu et al. 2014; Yamauchi et al. 2014). Thus, it is possible that MET1 function in maintaining the epigenetic state during gametogenesis is essential for sporophyte development.

The limited effect in *ΔPpmet* on overall gene expression (Table S3) and development of the gametophyte (Fig. 5) could be explained by several mechanisms. One possibility is a functional overlap between CG and CHG methylation. In *P. patens*, the distribution of CG and CHG methylation marks highly coincide along the genome (Zemach et al. 2010). It is likely that when PpMET function is compromised in *ΔPpmet*, the remaining CHG methylation mediated by PpCMT could be sufficient to maintain the regulatory function of such mark in the same loci. Indeed, severe morphological and molecular phenotypes were observed in *Arabidopsis Atmet1 Atcmt3* double mutant, in contrast to each of the single mutants phenotypes (Zhang et al. 2006; Cokus et al. 2008). Furthermore, CHG hypermethylation was reported in *Arabidopsis Atmet1* mutants (Cokus et al. 2008; Lister et al. 2008). Similarly, in *ΔPpmet* hypermethylation of CHG context may further compensate for the absence of CG methylation thus maintaining regulation of gene expression. The loci we have tested are hypermethylated in WT at CHG context and therefore no additional methylation can be detected in *ΔPpmet* at these loci (Fig. 1). This proposed mechanism may explain our failure

to generate *ΔPpmet ΔPpcmt* double deletion mutants by reciprocal transformations in the background of each of the single mutants in 14 independent transformation experiments, supporting a possible functional overlap between CG and CHG methylation.

A second possible mechanism explaining why *ΔPpmet* gametophyte development was unaffected while sporophyte development was arrested is a crosstalk between histone modifications and DNA methylation. An emerging concept is that epigenetic gene repression is established by histone modifications and is reinforced by DNA methylation (Cedar and Bergman 2012; Kawashima and Berger 2014; She and Baroux 2014). Thus, loss of DNA methylation may not lead to alteration of gene expression as the heterochromatin state may be maintained by existing histone modifications present at the same loci. *ΔPpmet* mutant was generated through transformation of *P. patens* WT protoplasts, cells in which the chromatin landscape was already established. Therefore, we speculate that loss of DNA methylation in *ΔPpmet* does not affect epigenetic gene silencing in protonema. However, during gametogenesis somatic cells undergo epigenetic reprogramming, a process which requires remodeling of the chromatin landscape including DNA methylation, histone modifications and histone variants (Cedar and Bergman 2012; Kawashima and Berger 2014; She and Baroux 2014). Thus, in the reproductive phase CG methylation may be required for the correct establishment of epigenetic reprogramming which may in turn explain the inability of *ΔPpmet* to produce sporophytes.

In angiosperms, generating loss of function homozygous mutants requires crossing of heterozygous mutant plants, a process which entails chromatin remodeling during gametogenesis (She and Baroux 2014). In *P. patens*, the ability to generate mutants directly in somatic haploid tissue allows studying protein function prior to epigenetic reprogramming taking place during the reproductive phase. Thus, *P. patens* serves as a model organism which can be used to dissect between somatic and germ line epigenetic effects.

Acknowledgments C. N. M and R. Y were supported in part by a matching Tel-Aviv University Deans doctoral fellowship and the Manna foundation. This research was supported by the Israeli Science Foundation Grant #767/09, and by the Israel Korea Program #3-824 financed by the Ministry of Science and Technology, both granted to N.O. Additional support from the German-Israeli Foundation for Scientific Research and Development (GIF I-832-130.12/2004 and I-1008-154.13-2008 to N. O. and R. R.) and by the Excellence Initiative of the German Federal and State Governments (EXC294 to R. R.) is gratefully acknowledged.

Conflict of interest The authors declare that they have no conflict of interest.

References

Ashton NW, Cove DJ (1977) The isolation and preliminary characterisation of auxotrophic and analogue resistant mutants of the moss *Physcomitrella patens*. Mol Gen Genet 154:87–95

Baldi P, Long AD (2001) A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. Bioinformatics 17:509–519

Bartee L, Malagnac F, Bender J (2001) Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. Genes Dev 15:1753–1758

Beike AK et al (2015) Insights from the cold transcriptome of *Physcomitrella patens*: global specialization pattern of conserved transcriptional regulators and identification of orphan genes involved in cold acclimation. New Phytol 205:869–881

Bezanilla M, Pan A, Quatrano RS (2003) RNA interference in the moss *Physcomitrella patens*. Plant Physiol 133:470–474

Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. Science 317:1760–1764

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinform 10:421

Cao X, Jacobsen SE (2002) Role of the arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. Curr Biol 12:1138–1144

Cao X, Springer NM, Muszynski MG, Phillips RL, Kaeppler S, Jacobsen SE (2000) Conserved plant genes with similarity to mammalian de novo DNA methyltransferases. Proc Natl Acad Sci USA 97:4979–4984

Cedar H, Bergman Y (2012) Programming of DNA methylation patterns. Annu Rev Biochem 81:97–117

Cokus SJ et al (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452:215–219

Cove D (2005) The moss *Physcomitrella patens*. Annu Rev Genet 39:339–358

Dangwal M, Kapoor S, Kapoor M (2014) The PpCMT chromomethylase affects cell growth and interacts with the homolog of LIKE HETEROCHROMATIN PROTEIN 1 in the moss *Physcomitrella patens*. Plant J 77:589–603

Du J et al (2012) Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell 151:167–180

Feng S et al (2010) Conservation and divergence of methylation patterning in plants and animals. Proc Natl Acad Sci USA 107:8689–8694

Finnegan EJ, Peacock WJ, Dennis ES (1996) Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. Proc Natl Acad Sci USA 93:8449–8454

Frank W, Decker EL, Reski R (2005) Molecular tools to study *Physcomitrella patens*. Plant Biol (Stuttg) 7:220–227

Gehring M (2013) Genomic imprinting: insights from plants. Annu Rev Genet 47:187–208

Genger RK, Kovac KA, Dennis ES, Peacock WJ, Finnegan EJ (1999) Multiple DNA methyltransferase genes in *Arabidopsis thaliana*. Plant Mol Biol 41:269–278

Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. Annu Rev Biochem 74:481–514

Henikoff S, Comai L (1998) A DNA methyltransferase homolog with a chromodomain exists in multiple polymorphic forms in Arabidopsis. Genetics 149:307–318

Hiss M et al (2014) Large-scale gene expression profiling data for the model moss *Physcomitrella patens* aid understanding of

developmental progression, culture and stress conditions. Plant J 79:530–539

Hohe A, Rensing SA, Mildner M, Lang D, Reski R (2002) Day length and temperature strongly influence sexual reproduction and expression of a novel MADS-box gene in the moss *Physcomitrella patens*. Plant Biol 4:595–762

Hou PQ et al (2014) Functional characterization of *Nicotiana benthamiana* chromomethylase 3 in developmental programs by virus-induced gene silencing. Physiol Plant 150:119–132

Hu L et al (2014) Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. Proc Natl Acad Sci USA 111:10642–10647

Jullien PE, Kinoshita T, Ohad N, Berger F (2006) Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. Plant Cell 18:1360–1372

Jullien PE, Mosquna A, Ingouff M, Sakata T, Ohad N, Berger F (2008) Retinoblastoma and its binding partner MSI1 control imprinting in Arabidopsis. PLoS Biol 6:e194

Kankel MW et al (2003) Arabidopsis MET1 cytosine methyltransferase mutants. Genetics 163:1109–1122

Katz A, Oliva M, Mosquna A, Hakim O, Ohad N (2004) FIE and CURLY LEAF polycomb proteins interact in the regulation of homeobox gene expression during sporophyte development. Plant J 37:707–719

Kawashima T, Berger F (2014) Epigenetic reprogramming in plant sexual reproduction. Nat Rev Genet 15:613–624

Kim MY, Zilberman D (2014) DNA methylation as a system of plant genomic immunity. Trends Plant Sci 19:320–326

Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 11:204–220

Lindroth AM, Cao X, Jackson JP, Zilberman D, McCallum CM, Henikoff S, Jacobsen SE (2001) Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science 292:2077–2080

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell 133:523–536

Malik G, Dangwal M, Kapoor S, Kapoor M (2012) Role of DNA methylation in growth and differentiation in *Physcomitrella patens* and characterization of cytosine DNA methyltransferases. FEBS J 279:4081–4094

Marchler-Bauer A et al (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res 39:D225–D229

Marchler-Bauer A et al (2013) CDD: conserved domains and protein three-dimensional structure. Nucleic Acids Res 41:D348–D352

Marchler-Bauer A et al (2015) CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D226

Mathieu O, Reinders J, Caikovski M, Smathajitt C, Paszkowski J (2007) Transgenerational stability of the Arabidopsis epigenome is coordinated by CG methylation. Cell 130:851–862

Mosquna A, Katz A, Decker EL, Rensing SA, Reski R, Ohad N (2009) Regulation of stem cell maintenance by the Polycomb protein FIE has been conserved during land plant evolution. Development 136:2433–2444

Nishiyama T, Hiwatashi Y, Sakakibara I, Kato M, Hasebe M (2000) Tagged mutagenesis and gene-trap in the moss *Physcomitrella patens* by shuttle mutagenesis. DNA Res 7:9–17

Noy-Malka C, Yaari R, Itzhaki R, Mosquna A, Auerbach Gershovitz N, Katz A, Ohad N (2014) A single CMT methyltransferase homolog is involved in CHG DNA methylation and development of *Physcomitrella patens*. Plant Mol Biol 84:719–735

Papa CM, Springer NM, Muszynski MG, Meeley R, Kaeppler SM (2001) Maize chromomethylase Zea methyltransferase2 is required for CpNpG methylation. Plant Cell 13:1919–1928

Pavlopoulou A, Kossida S (2007) Plant cytosine-5 DNA methyltransferases: structure, function, and molecular evolution. Genomics 90:530–541

Pradhan S, Adams RL (1995) Distinct CG and CNG DNA methyltransferases in *Pisum sativum*. Plant J 7:471–481

Prigge MJ, Bezanilla M (2010) Evolutionary crossroads in developmental biology: *Physcomitrella patens*. Development 137:3535–3543

Rensing SA et al (2008) The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. Science 319:64–69

Saze H, Mittelsten Scheid O, Paszkowski J (2003) Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. Nat Genet 34:65–69

Sharif J et al (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature 450:908–912

She W, Baroux C (2014) Chromatin dynamics during plant sexual reproduction. Front Plant Sci 5:354

Stroud H et al (2013a) Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. Nat Struct Mol Biol 21:64–72

Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE (2013b) Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. Cell 152:352–364

Vongs A, Kakutani T, Martienssen RA, Richards EJ (1993) *Arabidopsis thaliana* DNA methylation mutants. Science 260:1926–1928

Watson M, Hawkes E, Meyer P (2014) Transmission of epi-alleles with MET1-dependent dense methylation in *Arabidopsis thaliana*. PLoS One 9:e105338

Wolf L, Rizzini L, Stracke R, Ulm R, Rensing SA (2010) The molecular and physiological responses of *Physcomitrella patens* to ultraviolet-B radiation. Plant Physiol 153:1123–1134

Woo HR, Pontes O, Pikaard CS, Richards EJ (2007) VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. Genes Dev 21:267–277

Xiao W, Custard KD, Brown RC, Lemmon BE, Harada JJ, Goldberg RB, Fischer RL (2006) DNA methylation is critical for Arabidopsis embryogenesis and seed viability. Plant Cell 18:805–814

Yamauchi T, Johzuka-Hisatomi Y, Terada R, Nakamura I, Iida S (2014) The MET1b gene encoding a maintenance DNA methyltransferase is indispensable for normal development in rice. Plant Mol Biol 85:219–232

Zemach A, McDaniel IE, Silva P, Zilberman D (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 328:916–919

Zemach A et al (2013) The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell 153:193–205

Zhang X et al (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. Cell 126:1189–1201

Zimmer AD et al (2013) Reannotation and extended community resources for the genome of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions. BMC Genom 14:498