

# UCLA

## UCLA Previously Published Works

### Title

MTHFD1 controls DNA methylation in Arabidopsis

### Permalink

<https://escholarship.org/uc/item/0bn043tc>

### Journal

Nature Communications, 7(1)

### ISSN

2041-1723

### Authors

Groth, Martin  
Moissiard, Guillaume  
Wirtz, Markus  
et al.

### Publication Date

2016

### DOI

10.1038/ncomms11640

Peer reviewed

ARTICLE

Received 26 Oct 2015 | Accepted 14 Apr 2016 | Published 13 Jun 2016

DOI: 10.1038/ncomms11640

OPEN

# MTHFD1 controls DNA methylation in *Arabidopsis*

Martin Groth<sup>1</sup>, Guillaume Moissiard<sup>1</sup>, Markus Wirtz<sup>2</sup>, Haifeng Wang<sup>3</sup>, Carolina Garcia-Salinas<sup>4</sup>, Perla A. Ramos-Parra<sup>4</sup>, Sylvain Bischof<sup>1</sup>, Suhua Feng<sup>1,5,6</sup>, Shawn J. Cokus<sup>1</sup>, Amala John<sup>1</sup>, Danielle C. Smith<sup>1</sup>, Jixian Zhai<sup>1</sup>, Christopher J. Hale<sup>1</sup>, Jeff A. Long<sup>1</sup>, Ruediger Hell<sup>2</sup>, Rocío I. Díaz de la Garza<sup>4</sup> & Steven E. Jacobsen<sup>1,5,6</sup>

DNA methylation is an epigenetic mechanism that has important functions in transcriptional silencing and is associated with repressive histone methylation (H3K9me). To further investigate silencing mechanisms, we screened a mutagenized *Arabidopsis thaliana* population for expression of *SDCpro-GFP*, redundantly controlled by DNA methyltransferases DRM2 and CMT3. Here, we identify the hypomorphic mutant *mthfd1-1*, carrying a mutation (R175Q) in the cytoplasmic bifunctional methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase (MTHFD1). Decreased levels of oxidized tetrahydrofolates in *mthfd1-1* and lethality of loss-of-function demonstrate the essential enzymatic role of MTHFD1 in *Arabidopsis*. Accumulation of homocysteine and S-adenosylhomocysteine, genome-wide DNA hypomethylation, loss of H3K9me and transposon derepression indicate that S-adenosylmethionine-dependent transmethylation is inhibited in *mthfd1-1*. Comparative analysis of DNA methylation revealed that the CMT3 and CMT2 pathways involving positive feedback with H3K9me are mostly affected. Our work highlights the sensitivity of epigenetic networks to one-carbon metabolism due to their common S-adenosylmethionine-dependent transmethylation and has implications for human MTHFD1-associated diseases.

<sup>1</sup>Department of Molecular, Cell, and Developmental Biology, University of California Los Angeles, Los Angeles, California 90095, USA. <sup>2</sup>Centre for Organismal Studies, University of Heidelberg, Heidelberg 69120, Germany. <sup>3</sup>Basic Forestry and Proteomics Research Center, Haixia Institute of Science and Technology (HIST), Fujian Province Key Laboratory of Plant Virology, Institute of Plant Virology, Fujian Agriculture and Forestry University, Fuzhou, Fujian 350002, China. <sup>4</sup>Tecnologico de Monterrey, Campus Monterrey, Ave. Eugenio Garza Sada 2501, Monterrey 64849, México. <sup>5</sup>Eli & Edythe Broad Center of Regenerative Medicine & Stem Cell Research, University of California Los Angeles, Los Angeles, California 90095, USA. <sup>6</sup>Howard Hughes Medical Institute, University of California Los Angeles, Los Angeles, California 90095, USA. Correspondence and requests for materials should be addressed to M.G. (email: grothmn@gmail.com) or to S.E.J. (email: jacobsen@ucla.edu).

**D**NA methylation serves as a defense mechanism against transposable elements (TEs) and other types of repetitive DNA that can harm the genome of the organism they inhabit. DNA methylation promotes the packaging of DNA into so-called heterochromatin as to enforce a silent state, for example, by making the DNA inaccessible to transcription activators. Because of its transgenerational stability, DNA methylation is the prime example of an epigenetic mechanism. The stability is provided by feedback loops within, as well as crosstalk between, different methylation pathways, as established in *Arabidopsis thaliana* (*Arabidopsis*), where DNA methylation occurs at CG, CHG and CHH (H = A, T or C)<sup>1</sup>. In contrast to CHG and CHH methylation, which are exclusively involved in heterochromatin formation and transcriptional gene silencing (TGS), CG methylation also occurs over gene bodies.

DNA methylation is generated by the activity of DNA methyltransferases (DNMTs), which enzymatically transfer a methyl group from S-adenosyl methionine (SAM) to cytosine. In *Arabidopsis*, DOMAINS REARRANGED METHYLTRANSFERASE2 (DRM2) catalyses *de novo* methylation in all sequence contexts<sup>2</sup>. Once established, DNA methylation at symmetric CG and CHG sequences is maintained by DNA METHYLTRANSFERASE1 (MET1) and CHROMOMETHYLASE3 (CMT3), respectively. Maintenance of CG methylation is based on the recognition of hemimethylated signatures after semi-conservative DNA replication. Analogous to the recruitment of the mammalian maintenance methylase DNMT1 through UHRF1, members of the VARIANT IN METHYLATION (VIM) family bind to hemimethylated DNA with their SET and RING-associated (SRA) domains and are required for CG methylation by MET1 (refs 1,3). In contrast, CHG methylation is maintained by a reinforcing loop between non-CG methylation and methylation of lysine 9 of histone H3 (H3K9me), which involves the histone methyltransferases KRYPTONITE/SUVH4, 5 and 6. These preferentially bind methylated non-CG sequences via their SRA domains and modify the wrapped nucleosome with H3K9me (ref. 4). In turn, CMT3 binds H3K9me through its chromo and BAH domains and catalyses the remethylation of CHG sites during replication<sup>5</sup>. Similarly, maintenance of CHH methylation by CMT2 also depends on SUVH4/5/6-mediated H3K9me (ref. 6).

While CMT2 and 3 mostly target transposons in the pericentromeric heterochromatin, DRM2 is mainly required for maintenance of CHH methylation and TGS in the chromosomal arms<sup>7</sup>. Targeting of DRM2 is mediated by the concerted action of short transcripts that are processed into 24 nt small-interfering RNAs and complementary long noncoding transcripts produced by the plant-specific RNA polymerase complexes Pol IV and Pol V, respectively. In the canonical RNA-directed DNA methylation pathway (RdDM), 24 nt RNAs are incorporated into ARGONAUT4 (AGO4) in order to match the RNA-induced silencing complexes with Pol V transcripts<sup>8</sup>. Subsequently, DRM2 is recruited to the target CHH sites by direct interaction with AGO4 (ref. 9). Recruitment of Pol IV to heterochromatin is also dependent on H3K9me interaction via SAWADEE HOMEODOMAIN HOMOLOGUE 1 (SHH1) (ref. 10), whereas Pol V is recruited via the non-catalytic SUV39 homologues SUVH2 and 9 that bind to methylated DNA via their SRA domains<sup>11</sup>.

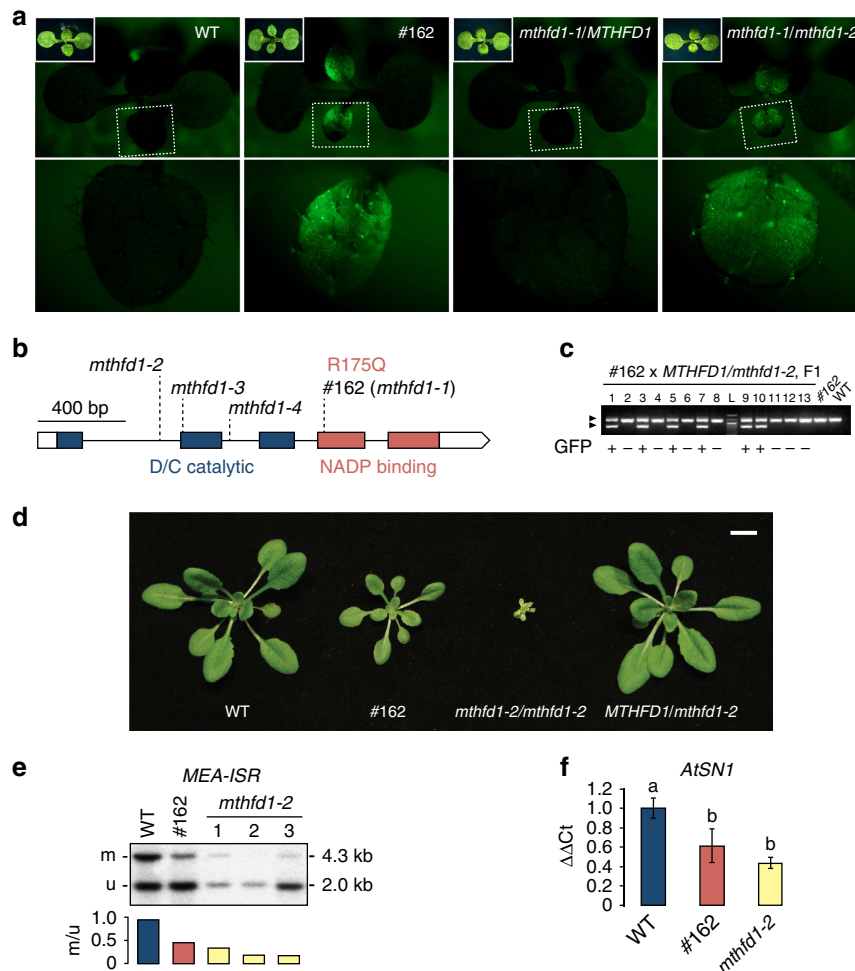
TGS is often reinforced by the synergistic action of different DNA methylation pathways. This is exemplified by the SUPPRESSOR OF DRM2 CMT3 (*SDC*) locus, which is redundantly silenced by the CMT3- and DRM2-mediated methylation of tandem repeats in the promoter region<sup>12</sup>. Hence, *SDC* is ectopically expressed in *drm2 cmt3* double mutants, but repressed during most of development in the single mutants, making *SDC* a

powerful genetic marker of simultaneous impairment of CHG and CHH methylation pathways. We generated stable transgenic lines carrying an *SDC<sub>pro</sub>-GFP* fusion construct in wild-type (WT) and *cmt3* genetic background and screened M2 populations for EMS-mutants that express GFP. The identification of *microrchidia 1* (*morc1*) and *morc6* mutants from this screen was published previously<sup>13</sup>. Here, we identified a mutant from the WT background that carries a missense mutation in the *Arabidopsis* METHYLENETETRAHYDROFOLATE DEHYDROGENASE/METHENYLTETRAHYDROFOLATE CYCLOHYDROLASE 1 (*MTHFD1*) gene. The mutation disrupts folate metabolism and leads to accumulation of homocysteine (Hcy), a hallmark of an impaired methionine (Met) cycle, whose main function is to produce SAM for transmethylation reactions and recycle the byproduct S-adenosyl homocysteine (SAH)<sup>14</sup>. Genome-wide loss of CHG and CHH methylation, reduced H3K9me and derepression of TEs in the *mthfd1-1* mutant indicate that the Met cycle constitutes an ‘Achilles heel’ of the feedback mechanisms between DNA and histone methylation.

## Results

**The R175Q mutation in MTHFD1 leads to *SDC<sub>pro</sub>-GFP* expression.** Mutant #162 was identified by screening M2 seedlings of an EMS-mutagenized population of *Arabidopsis* that carried an *SDC<sub>pro</sub>-GFP* insertion event in WT background (herein after referred to as WT; Col refers to non-transgenic WT) for individuals that showed GFP fluorescence (Fig. 1a). Using deep sequencing of bulked GFP-positive F2 progeny of mutant #162 crossed with a WT plant of ecotype Landsberg *erecta* (*Ler*), we confined the target region containing the causative mutation to the north end of chromosome 3 (Supplementary Fig. 1). To identify the causative mutation, mutant #162 was crossed to WT and the co-segregation of candidate EMS mutations in GFP-positive F2 progeny was analysed using dCAPS markers (Supplementary Fig. 2). A guanine to adenine transition in a gene (At3g12290) encoding a putative methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase (*MTHFD1*) showed 100% co-segregation in 98 GFP-positive F2 individuals (Supplementary Fig. 2). The mutant phenotype segregated as a recessive monogenic trait (103 GFP-positive versus 368 GFP-negative,  $\chi^2$  *P* value (3:1) = 0.12). The mutation, herein after named *mthfd1-1*, leads to a predicted substitution of a conserved arginine by glutamine at residue 175 (R175Q) (Fig. 1b). To confirm that *mthfd1-1* caused the expression of *SDC<sub>pro</sub>-GFP*, a #162 M3 mutant (*mthfd1-1/mthfd1-1*) was crossed with a heterozygous plant containing a transfer DNA (T-DNA) insertion allele, *MTHFD1/mthfd1-2* (Fig. 1b). GFP expression in F1 progeny co-segregated with the *mthfd1-2* allele (Fig. 1a,c), confirming that *mthfd1-1* caused the expression of *SDC<sub>pro</sub>-GFP* in #162.

GFP-positive plants originating from mutant #162 had pale leaves, reduced seed set, were smaller and developed more slowly compared with WT but did not display other morphological defects (Fig. 1d and Supplementary Fig. 3a). In contrast, *mthfd1-2* homozygous mutants showed severe developmental defects, including dwarfism, pale, shortened leaves, reduced apical dominance, delayed flowering, prolonged vegetative phase and infertility (Fig. 1d and Supplementary Fig. 3a,b). Moreover, on average we only retrieved one viable homozygous *mthfd1-2* mutant out of 18 seeds from a heterozygous parent, indicating that more than 75% of homozygous *mthfd1-2* mutants died prematurely (Supplementary Table 1). Viability seemed to be affected during or after germination, because siliques from WT and heterozygous *MTHFD1/mthfd1-2* did not show differences in ovule and seed development (Supplementary Fig. 3a). We did not



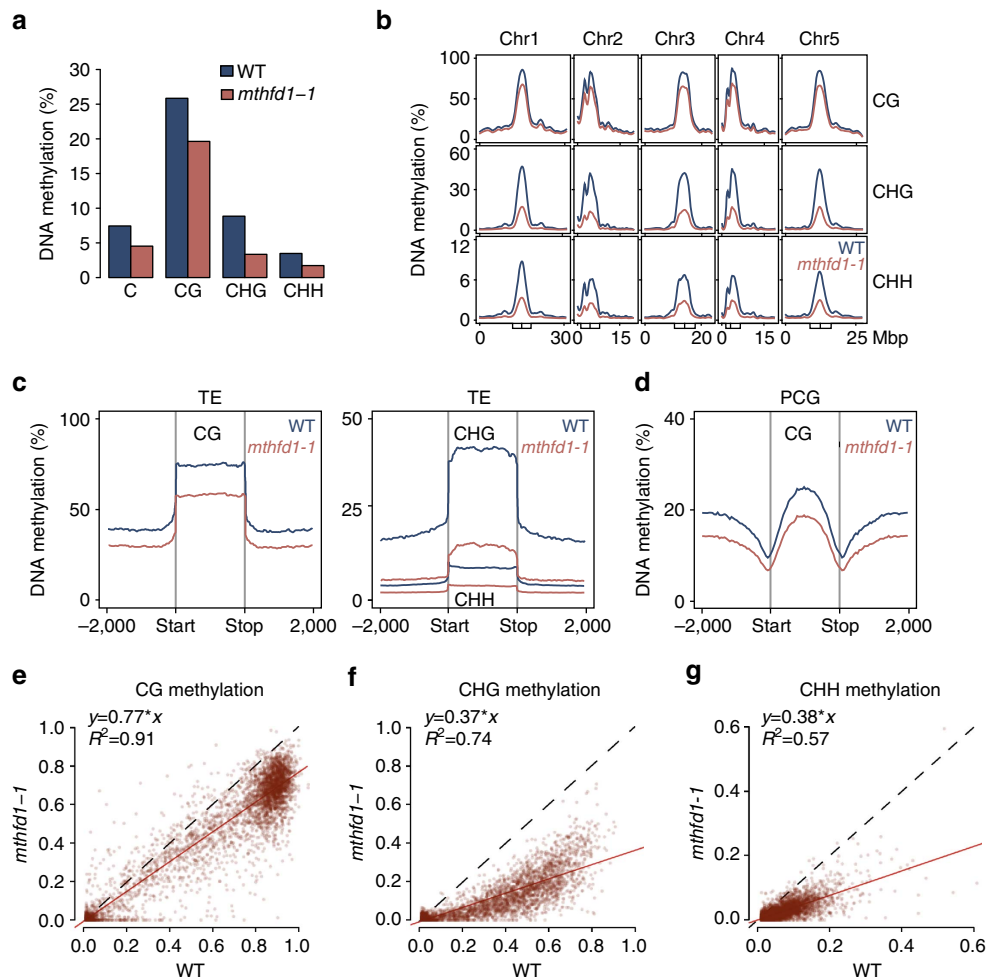
**Figure 1 | SDCpro-GFP expression and DNA demethylation caused by R175Q mutation in MTHFD1.** (a) GFP fluorescence micrographs of WT, #162 M2, MTHFD1/*mthfd1-2* F1 and #162/*mthfd1-2* F1 seedlings. F1 are progeny of #162 M2 x MTHFD1/*mthfd1-2*. Dashed boxes indicate magnified areas shown in lower panels. Inlets show bright-field images. (b) Gene structure, positions of mutations and conserved domains of MTHFD1. The EMS mutation in #162 lead to a R175Q substitution of a conserved residue required for NADP binding<sup>28</sup>. (c) PCR-based genotype analysis of 13 F1 seedlings and two control samples. Arrowheads mark bands corresponding to WT/*mthfd1-1* (upper) and *mthfd1-2* (lower). The *mthfd1-2* allele co-segregates with GFP fluorescence in F1 (+: present, -: absent). L, ladder. (d) Habit of different genotype plants 20 days after germination. Scale bar, 10 mm. (e) DNA blot analysis of non-CG methylation at the MEA-ISR locus. Genomic DNA was digested with methylation-sensitive MspI; upper and lower bands correspond to methylated (m) and unmethylated (u) fragments, respectively. Ratios of band intensities for each lane are shown under the gel image. (f) Levels of non-CG methylation at the AtSN1 locus by quantitative chop PCR analysis of genomic DNA after digestion with methylation-sensitive HaeIII relative to undigested DNA. Mean values  $\pm$  s.d. ( $n = 3$ ). Different letters above bars indicate significant differences between pairwise comparisons by Student's *t*-test ( $P < 0.05$ ).

retrieve any homozygous mutants from the T-DNA alleles *mthfd1-3* or *mthfd1-4* (Fig. 1b). These results led to the conclusion that MTHFD1 is an essential gene in *Arabidopsis* and that the T-DNA insertion in the first intron in *mthfd1-2* does not completely abolish the function of MTHFD1. Moreover, the results indicate that R175Q in *mthfd1-1* partially impairs gene function, causing a hypomorphic phenotype that does not affect viability.

To test if DNA methylation is altered in *mthfd1* mutants, we analysed the MEDEA INTERGENIC SUBTELOMERIC REPEAT (MEA-ISR) locus using a well-established Southern blot assay<sup>15</sup>. In WT, approximately half of the MEA-ISR alleles are methylated and restriction digest with methylation-sensitive MspI produces two fragments of similar abundance. Both #162 and *mthfd1-2* mutants showed a reduction of the methylated band, and the loss of methylation was stronger in *mthfd1-2* than in #162 (Fig. 1e). We also analysed DNA methylation at *Arabidopsis thaliana* SHORT INTERSPERSED ELEMENT 1 (AtSN1) by quantitative PCR following methylation-sensitive restriction digestion and

observed similar reductions in DNA methylation (Fig. 1f). Furthermore, bisulfite (BS)-PCR analysis of the levels of methylation at CG, CHG and CHH sites in the tandem repeat region of the transgenic SDC promoter showed a decrease in CHG methylation of the transgenic SDC in #162 mutants compared with the WT reference (Supplementary Fig. 4). In summary, the DNA methylation assays confirmed that MTHFD1 is required for DNA methylation at different loci in different sequence contexts, and the R175Q amino-acid substitution leads to reduced DNA methylation. Because of the limiting amounts of tissue available from *mthfd1-2* mutants, only #162 mutants (referred to as *mthfd1-1* herein after) were analysed subsequently.

**MTHFD1 is required for epigenetic silencing.** To get a general view of the DNA methylation defects in *mthfd1-1*, we analysed genome-wide DNA methylation at single-nucleotide resolution by BS-seq. The average global DNA methylation was reduced by  $\sim 40\%$  relative to WT (Fig. 2a). The strongest effect was observed



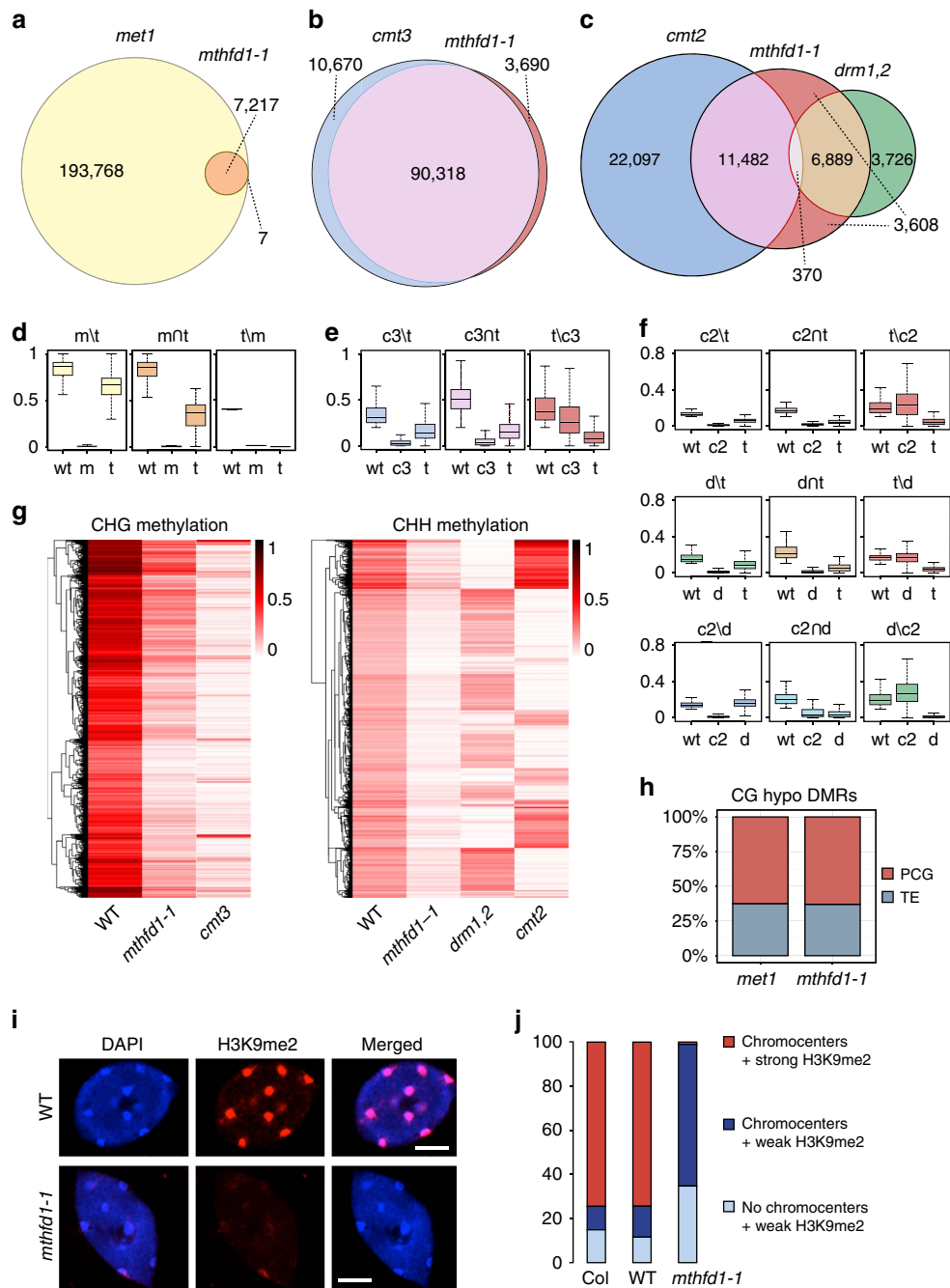
**Figure 2 | DNA methylation is globally decreased in *mthfd1-1* mutants.** (a) Average genome-wide DNA methylation for all Cs and in individual sequence contexts (H = C, A or T). (b) Chromosomal distribution of fractional DNA methylation in individual sequence contexts. Boxes and vertical lines inside boxes mark pericentromeric regions and centromeres, respectively. (c,d) Average distribution of DNA methylation over TEs (c) and PCGs (d) and the flanking 2,000 bp in individual sequence contexts. (e–g) Comparison of DNA methylation levels in 5,000 random 100 bp bins with WT methylation levels > 0.01 in CG (e), CHG (f) and CHH (g) contexts. Red line: linear regression between *mthfd1-1* and WT levels; corresponding coefficients are shown in top left corners. Dashed: identity line.

in the CHG context, which lost 62% of DNA methylation relative to WT, followed by CHH and CG methylation with 50% and 24% decreases, respectively (Fig. 2a). DNA methylation in all sequence contexts was mostly decreased over the TE-rich pericentromeric regions, which contain most of the DNA methylation along the chromosomes (Fig. 2b). Accordingly, average methylation levels over TEs were strongly decreased in *mthfd1-1* compared with WT, especially in CHG and CHH contexts (Fig. 2c). DNA methylation was also moderately decreased over protein-coding genes (PCGs) (Fig. 2d), indicating that MTHFD1 is not only required for repressive DNA methylation at TEs but also for efficient gene body CG methylation, although to a lesser degree, compared with non-CG methylation. The different effects on DNA methylation in the different sequence contexts was also apparent when comparing WT and *mthfd1-1* CG, CHG and CHH methylation levels in randomly selected 100 bp windows of the genome with methylation thresholds > 1% (in order to exclude unmethylated bins during randomization) (Fig. 2e–g). This comparison additionally shows a positive correlation of DNA methylation levels between WT and *mthfd1-1* in all three sequence contexts, which indicates that DNA methylation was decreased uniformly across the genome (Fig. 2e–g), as opposed to the DNA methylation patterns in loss-of-function DNMT

mutants, which show either nearly complete loss across the sample pool (Supplementary Fig. 5a,b), or at a subset of regions (Supplementary Fig. 5c).

To define the effect of the *mthfd1-1* mutation on the different DNA methylation pathways, we calculated differentially methylated regions (DMRs) that had decreased DNA methylation in *mthfd1-1* or the DNMT mutants compared with the WT reference (hypo-DMRs). The comparison of hypo-DMRs clearly showed that CMT3-dependent CHG methylation was the most affected, followed by CMT2- and DRM1,2-dependent CHH methylation (DRM1 is a lowly expressed paralog of DRM2 (ref. 16)), and finally MET1-dependent CG methylation that was the least affected in *mthfd1-1* (Fig. 3a–c). It is noteworthy that *mthfd1-1* DNA methylation levels were also decreased in regions that were only defined as DMRs in the DNMT mutants, but not in *mthfd1-1* (Fig. 3d–f). Moreover, all subsets of *mthfd1-1* hypo-DMRs showed residual DNA methylation, which supports a uniform genome-wide decrease in *mthfd1-1* (Fig. 3d–f). Heat maps of hierarchically clustered CHG and CHH hypo-DMRs further illustrated that DNA methylation levels in *mthfd1-1* are evenly decreased at moderate degrees and thus generally proportional to WT levels (Fig. 3g). The equal ratio of TE- versus PCG-overlapping CG hypo-DMRs in *met1* and *mthfd1-1* is also





**Figure 3 | *mthfd1-1* mostly interferes with non-CG and H3K9 methylation.** (a–c) Overlap between hypo-DMRs of different mutants in CG (a), CHG (b) and CHH (c) contexts. (d–f) CG (d), CHG (e) and CHH (f) methylation levels in DMR fractions corresponding to (a–c), respectively. w = wild-type, m = *met1*, t = *mthfd1-1*, c3 = *cmt3*, c2 = *cmt2*, d = *drm1,2*. Box plot (herein and after): horizontal line, median; edges of boxes, 25th (bottom) and 75th (top) percentiles; error bars, minimum and maximum points within  $1.5 \times$  interquartile range. (g) Heat map of DNA methylation levels in *mthfd1-1* hypo-DMRs (rows) clustered by methylation levels. (h) Overlap of *met1* or *mthfd1-1* CG hypo-DMRs with PCGs or TEs. (i) Fluorescence micrographs of representative nuclei from WT and *mthfd1-1*. DNA was stained with DAPI and H3K9me2 was immunostained using Alexa Fluor 647 as secondary antibody. Scale bar, 5  $\mu$ m. (j) Number of nuclei classified by DAPI staining and H3K9me2 immunofluorescence.

in agreement with a uniform decrease in genomic DNA methylation in *mthfd1-1* (Fig. 3h). Accordingly, the small subset of *met1* CG hypo-DMRs shared by *mthfd1-1* does not seem to represent site-specific MTHFD1 function, but is more likely due to stringent criteria for DMR calling, which were not met by the majority of sites in *mthfd1-1* despite reduced CG methylation levels (Fig. 3a,d). In the CHH context, RdDM and the CMT2 pathway were both affected by *mthfd1-1* (Fig. 3c,g). The chromosomal distribution of *mthfd1-1* CHH hypo-DMRs, with

high densities inside and at the peripheries of the pericentromeric regions, and lower densities in the chromosomal arms, reflects the overlaps with the alternative CHH methylation pathways (Supplementary Fig. 5d). In summary, the DNA methylation analysis of *mthfd1-1* revealed a rather uniform genome-wide decrease, which is most pronounced in the CHG and CHH context.

DNA methylation by CMT3, CMT2 and RdDM is functionally linked to histone methylation, because CMT2 and CMT3 directly

bind H3K9me in a feedback loop with the histone methyltransferases KYP/SUVH4, SUVH5 and SUVH6 (refs 5,6), and Pol IV is recruited to chromatin by K3K9me-binding SHH1 (ref. 10). To test if H3K9me is also affected in *methfd1-1*, we analysed H3K9 dimethylation (H3K9me<sub>2</sub>) by immunofluorescence. Nuclei of *methfd1-1* mutants showed a strong decrease in H3K9me<sub>2</sub>, but the majority of the nuclei still contained DNA-dense chromocenters visualized by DAPI staining (Fig. 3i,j). Therefore, the strong decrease in CHG and CHH methylation in contrast to the small decrease in CG methylation is likely explained by the combined effect of impaired DNA and H3K9 methylation.

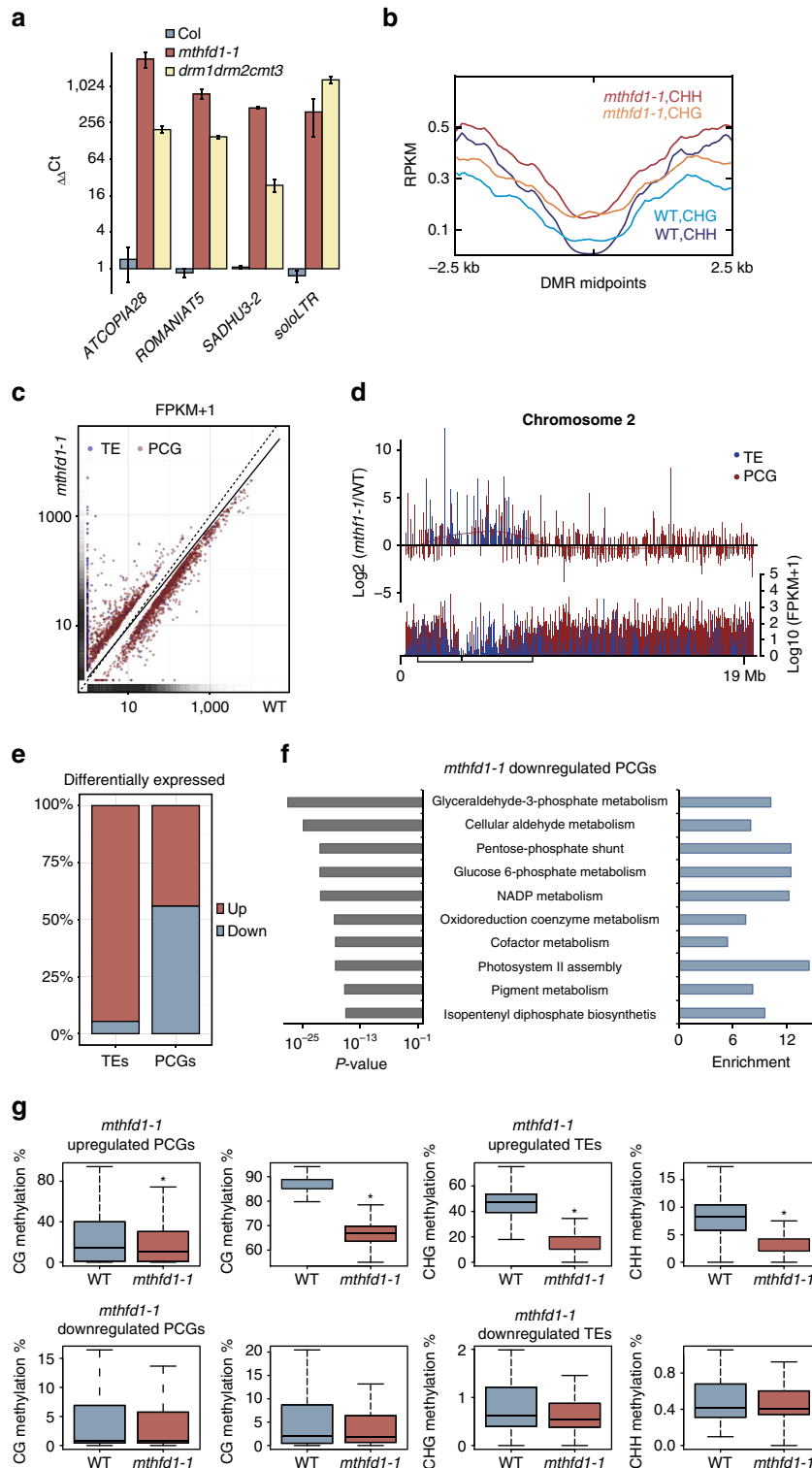
To test if loss of DNA methylation in *methfd1-1* led to transcriptional derepression, we analysed the expression levels of different retrotransposons that were previously identified as upregulated in *drm1,2 cmt3* triple mutants<sup>13</sup>. Quantitative reverse transcription (RT)-PCR showed that these TEs are also strongly induced in *methfd1-1* (Fig. 4a and Supplementary Data 1). Transcriptome analysis of *methfd1-1* and WT by RNA-seq confirmed that average transcript levels over CHH and CHG hypo-DMRs are higher in *methfd1-1* than in WT (Fig. 4b). Many transcripts from TEs that were silenced in WT were highly abundant in *methfd1-1*, whereas transcriptional differences of PCGs were more even and showed a slight tendency towards higher transcript levels in WT (Fig. 4c–e). Correspondingly, pericentromeric regions showed many differentially upregulated TEs and—to a lesser degree—PCGs in *methfd1-1*, whereas chromosome arms contained approximately equal distributions of up- and downregulated PCGs (Fig. 4d,e). The differentially upregulated TEs in *methfd1-1* belonged to class I, as well as class II transposons. Among the differentially upregulated TEs, members of the LTR/Gypsy family were overrepresented, and members of the RC/Helitron family were under-represented compared with the genomic distribution (Supplementary Fig. 6). In summary, the transcriptional analyses have shown that the loss of DNA methylation in *methfd1-1* led to derepression of transposons and genes (Fig. 4g), predominantly in the pericentromeric region (Fig. 4d). The observed transcriptional changes in the chromosome arms seem to be mainly a consequence of pleiotropic effects of impaired MTHFD1 function. This is supported by an analysis of GO terms annotated to genes that are significantly downregulated in *methfd1-1* compared with WT (Fig. 4f and Supplementary Data 2). The 10 statistically most significantly enriched biological processes indicate that MTHFD1 serves important functions in sugar metabolism, isoprenoid synthesis, redox homeostasis and photosynthesis. Since *methfd1-1*-downregulated genes did not show a significant loss of DNA methylation, the overall effects on transcript abundance caused by decreased gene body methylation in *methfd1-1* are likely to be negligible (Fig. 4g).

#### ***methfd1-1* mutants show accumulation of S-adenosylhomocysteine.**

*Arabidopsis* MTHFD1 contains two highly conserved protein domains, a catalytic domain in the N-terminal half, and a NAD(P<sup>+</sup>)-binding domain of the Rossmann fold superfamily in the C-terminal half (Fig. 1b). Therefore MTHFD1 is probably required for the interconversion of tetrahydrofolate (THF) species in one-carbon metabolism of *Arabidopsis*. Members of the bifunctional enzyme family catalyse the reversible interconversion of 5,10-methylenetetrahydrofolate (5,10-CH<sub>2</sub>-THF) to 5,10-methenyltetrahydrofolate (5,10-CH=THF) (NADP<sup>+</sup>-dependent dehydrogenase activity) and further to 10-formyltetrahydrofolate (10-CHO-THF) (cyclohydrolase activity) (Fig. 5) (ref. 17). These enzymatic activities have previously been detected in plant extracts<sup>18</sup>. The *Arabidopsis* genome encodes four homologues, the mitochondrial MTHFD2/FOLD1 (AT2G38660), FOLD3

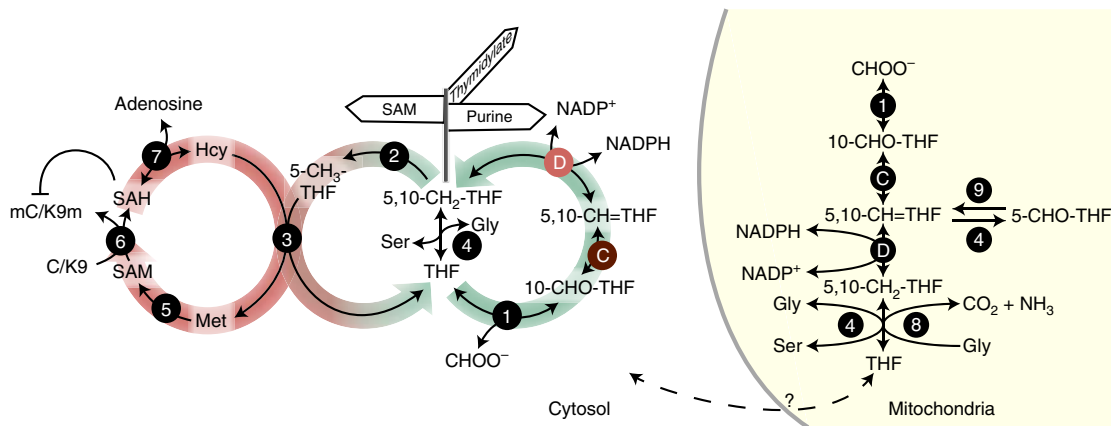
(At4g00600) and FOLD4 (At4g00620), which are putatively plastidic, and MTHFD1/FOLD2, which lacks an N-terminal targeting peptide and is presumably localized in the cytoplasm (Supplementary Fig. 7) (refs 19–21). We confirmed the subcellular localization of the latter two homologues by expression and *in vivo* imaging of the full-length fusion proteins MTHFD1-YPET-3xFLAG, MTHFD1\_R175Q-YPET-3xFLAG and FOLD4-YPET-3xFLAG in *Nicotiana benthamiana* (Fig. 6a,b and Supplementary Fig. 8), showing that MTHFD1 is predominantly in the cytoplasm. The Met cycle is exclusively localized in the cytoplasm and is required for the synthesis of SAM, which serves as methyl donor for many transmethylation reactions, including those catalysed by histone and DNMTs (Fig. 5). During transmethylation SAM is converted to SAH, which is further processed into adenosine and Hcy by the SAH hydrolase (SAHH) (Fig. 5). Hcy is recycled to Met by 5-CH<sub>3</sub>-THF-dependent transmethylation activity of methionine synthase and can serve for a new round of SAM synthesis (Fig. 5) (refs 14,22). To test if the DNA methylation defects in *methfd1-1* are caused by changes in the Met cycle, we analysed the levels of SAM, SAH and Hcy, as well as cysteine in *methfd1-1*, WT and Col leaves. SAM and SAH were both significantly increased in *methfd1-1*, but the stronger increase in SAH levels led to an overall decrease of the methylation index (MI = SAM/SAH) (Fig. 7a–c). MI is an important measure of the organismal methylation status, because SAH is a strong competitive inhibitor of SAM-dependent transmethylation<sup>23</sup>. Because of the low intracellular concentration and high affinity for methyltransferases, it has been suggested that even small changes in the MI can lead to a reduction in transmethylation activity<sup>24</sup>. Therefore, it is likely that the decreased MI in *methfd1-1* leads to decreased activities of DNA and histone methyltransferases, as reflected by the observed DNA and histone methylation defects. This is further supported by a 12-fold increase in Hcy (and an associated increase in cysteine levels) in *methfd1-1* (Fig. 7d), because Hcy accumulation leads to inhibition of SAH hydrolysis and consequently to a lower MI<sup>22,25</sup>.

**Regulation of folate homeostasis by MTHFD1.** Analysis of folate metabolites in leaves of *methfd1-1* and WT showed that total folate content did not differ significantly, but *methfd1-1* had an 8.8-fold increase in THF + 5,10-CH<sub>2</sub>-THF (these two compounds cannot be distinguished by the analysis). On the other hand, the levels of 5,10-CH=THF, which also include the 10-CHO-THF pool, were reduced by 30% and 5-CHO-THF levels were also ~33% lower (Fig. 7e). Unexpectedly, there was no significant difference in 5-CH<sub>3</sub>-THF contents, which is the product of 5,10-CH<sub>2</sub>-THF reductase that serves as co-substrate for the methylation of Hcy to Met and constitutes the most abundant active THF species<sup>26</sup>. The low levels of the oxidized folates, 10-CHO-THF + 5,10-CH=THF and 5-CHO-THF, suggest an impairment in the dehydrogenase and cyclohydrolase activities towards the formation of 10-CHO-THF. On the basis of the crystal structure and site-directed mutagenesis of human MTHFD1, the conserved arginine that is mutated in *methfd1-1* (R175Q), forms a hydrogen bond with NADP<sup>+</sup> and is required for dehydrogenase, but not cyclohydrolase activity<sup>27,28</sup>. Therefore, we anticipate that the R175Q mutation has a similar effect in *methfd1-1* mutants. Accordingly, the accumulation of the reduced forms THF + 5,10-CH<sub>2</sub>-THF in *methfd1-1* might reflect reduced conversion of 5,10-CH<sub>2</sub>-THF to 5,10-CH=THF, which is in agreement with the homology-based prediction that the R175Q mutation affects NADP<sup>+</sup> binding and dehydrogenase activity<sup>27,28</sup>. The decreased pool of oxidized THFs further suggests that one-carbon supply from formate via FTHFS and reverse cyclohydrolase by MTHFD1 is not sufficient to

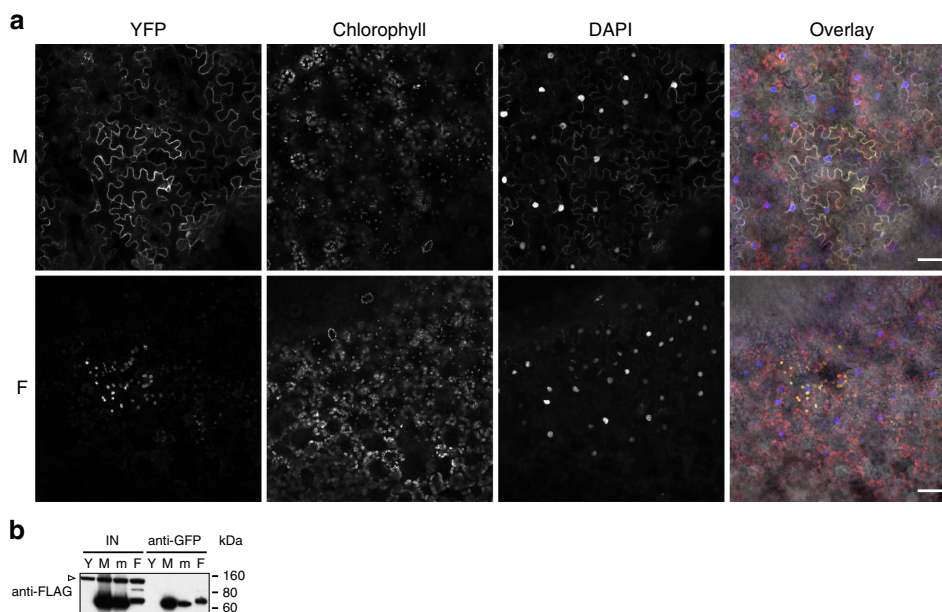


**Figure 4 | *mthfd1-1* mutants show loss of TE silencing and pleiotropic transcriptional deregulation of PCGs.** (a) Quantitative RT-PCR analysis of four exemplary TEs. Transcript levels are normalized to *ACTIN7* and relative to WT, and mean values  $\pm$  s.e.m. ( $n=3$ ) are shown. (b) Average distribution of normalized RNA-seq reads (RPKM) over *mthfd1-1* hypo-DMRs in CHG or CHH context. X axis indicates distance from the DMR midpoints. (c) Scatter plot showing normalized transcript levels (FPKM + 1) of differentially expressed PCGs and TEs in *mthfd1-1* versus WT. Solid line: linear regression through all TEs and PCGs. Dashed: identity line. Marginal density plots: distribution of all TEs and PCGs in WT (x axis) and *mthfd1-1* (y-axis). (d) Distribution of average normalized transcript levels (FPKM + 1) of all TEs and PCGs from *mthfd1-1* and WT (lower panel) and fold change in normalized transcript levels of differentially expressed PCGs and TEs (upper panel) along chromosome 2. Box and vertical line inside box mark pericentromeric region and centromere, respectively. (e) Fraction of TEs and PCGs significantly up- or downregulated in *mthfd1-1*. (f) Enrichment score and statistical significance of GO processes annotated to *mthfd1-1*-downregulated PCGs. (g) Methylation levels over *mthfd1-1* differentially expressed PCGs and TEs in different sequence contexts.





**Figure 5 | Schematic representation of plant SAM and folate metabolism in the cytosol and mitochondria.** One-carbon enters the cytoplasmic folate cycle (green) either through formyltetrahydrofolate synthetase (1) or SHMT (4); MTHFD1 reversibly interconverts 10-CHO-THF to 5,10-CH<sub>2</sub>-THF by cyclohydrolase (C) and NADP<sup>+</sup>-dependent dehydrogenase (D) activity. 5,10-CH<sub>2</sub>-THF serves for thymidylate synthesis or is converted by methylenetetrahydrofolate reductase (2) to 5-CH<sub>3</sub>-THF, which enters the Met cycle (red) and serves for Hcy remethylation to Met by methionine synthase (3) (ref. 17). SAM synthetase (5) converts Met to SAM, which is further converted to SAH (6) during methylation of cytosines, H3K9 and so on. SAH is a competitive inhibitor of methyltransferases (6) and is recycled to Hcy by SAH hydrolase (7) (ref. 23). In mitochondria, one-carbon is transferred to THF during the oxidation of Gly by the glycine decarboxylase complex (8), but surplus of Gly due to photorespiration leads to consumption of one-carbon by SHMT during serine production<sup>30</sup>. 5-CHO-THF, a byproduct of SHMT, is metabolized by mitochondrial 5-formyltetrahydrofolate cycloligase in order to re-enter the folate cycle (9) (ref. 59). Shuttle of THF between mitochondria and the cytosol has been described in other organisms, but remains uncharacterized in plants.

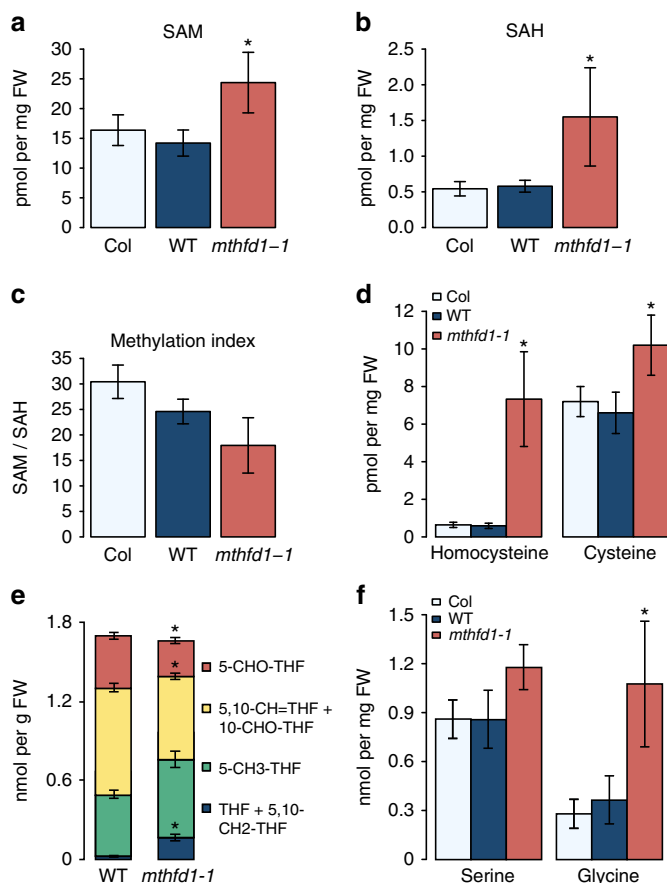


**Figure 6 | Cytoplasmic localization of MTHFD1-YPET-3xFLAG.** (a) Confocal micrographs of MTHFD1-YPET-3xFLAG (M) and FOLD4-YPET-3xFLAG (F) transiently expressed in *N. benthamiana*. Excitation ( $\lambda$ , nm)/filters ( $\lambda$ , nm): YFP = 514/519–559, chlorophyll = 488/630–730, DAPI = 405/409–530, and fluorescence overlay with bright field. Scale bars, 50  $\mu$ m. (b) Western blot using anti-FLAG antibody against anti-GFP-immunopurified extracts from *N. benthamiana* (IN) transiently expressing free YFP (Y), MTHFD1-YPET-3xFLAG (M), MTHFD1\_R175Q-YPET-3xFLAG (m), or FOLD4-YPET-3xFLAG (F). Arrowhead indicates unspecific binding of anti-FLAG (shown as loading reference).

compensate for reduced MTHFD1 dehydrogenase activity (Fig. 5). This is in accordance with metabolic analyses showing relatively low one-carbon flow from formate to serine (Ser)<sup>29</sup>.

Ser can serve as a one-carbon source through the reversible enzymatic activity of serine hydroxymethyltransferase (SHMT), which converts Ser and THF to 5,10-CH<sub>2</sub>-THF and glycine (Gly) (Fig. 5). In addition, the Gly decarboxylase complex converts Gly and THF to 5,10-CH<sub>2</sub>-THF, carbon dioxide and ammonia during photorespiration, which in turn can lead to Ser synthesis by

SHMT in the mitochondria<sup>30</sup> (Fig. 5). We analysed amino-acid levels in rosette leaves and found a threefold increase of Gly levels in *mthfd1-1* compared with WT and Col controls, whereas Ser levels were only slightly increased in *mthfd1-1* (Fig. 7f). In accordance with our folate analysis, the lower Ser/Gly ratio in *mthfd1-1* might be due to increased SHMT activity towards 5,10-CH<sub>2</sub>-THF formation<sup>19</sup>. Furthermore, *mthfd1-1* mutants showed a 5.4-fold and 1.9-fold increase in proline and Met levels, respectively (Supplementary Fig. 9). The increase in Met might be due to increased *de novo* synthesis (which occurs



**Figure 7 | *mthfd1-1* mutants show impaired one-carbon cycle.**

(a–f) Steady-state levels of SAM (a), SAH (b), Methylation Index (MI) (c), selected thiols (d), folates (e) and selected amino acids (f) in leaves of Col, WT and the *mthfd1-1* mutant. Data represent means  $\pm$  SD. Asterisks indicate significant differences determined by Student's *t*-test ( $P < 0.05$ ,  $n \geq 3$ ).

exclusively in plastids) in response to impaired Met recycling, whereas proline accumulation is typically observed as a stress response<sup>31</sup>.

We sought to manipulate one-carbon metabolism in *mthfd1-1* by exogenous application of metabolites involved in the folate and Met cycle and monitor root growth and DNA methylation effects. Previously it has been shown that growth phenotypes, DNA methylation and epigenetic silencing defects caused by decreased activated THF pools because of impaired plastidic folylpolyglutamate synthetase (FPGS1), or chemical inhibition of folate synthesis by sulfamethazine (SMZ) were complemented by exogenous application of 5-CHO-THF or Met<sup>32–35</sup>. Application of 5-CHO-THF, as well as 5-CH3-THF, tests for defects in *mthfd1-1* caused by reduced folate availability before the flow of one-carbon into the Met cycle, whereas application of Met tests for defects in the Met cycle (Fig. 5), and SMZ was used as a control. In contrast to previously described complementation of *fpgs1* mutants, 5-CHO-THF strongly inhibited root growth of *mthfd1-1* seedlings without showing an inhibitory effect on WT seedlings (Fig. 8a). Although analysis of global DNA methylation levels revealed that these are largely independent of the observed root growth responses, average *mthfd1-1* CHG methylation over the previously defined CHG hypo-DMRs significantly increased upon Met application. On the other hand, 5-CHO-THF, as well as 5-CH3-THF, did not rescue the DNA methylation defects in *mthfd1-1* (Fig. 8b). These results indicate that, in contrast to the *fpgs1* mutants and SMZ inhibition<sup>34,35</sup>, the DNA methylation

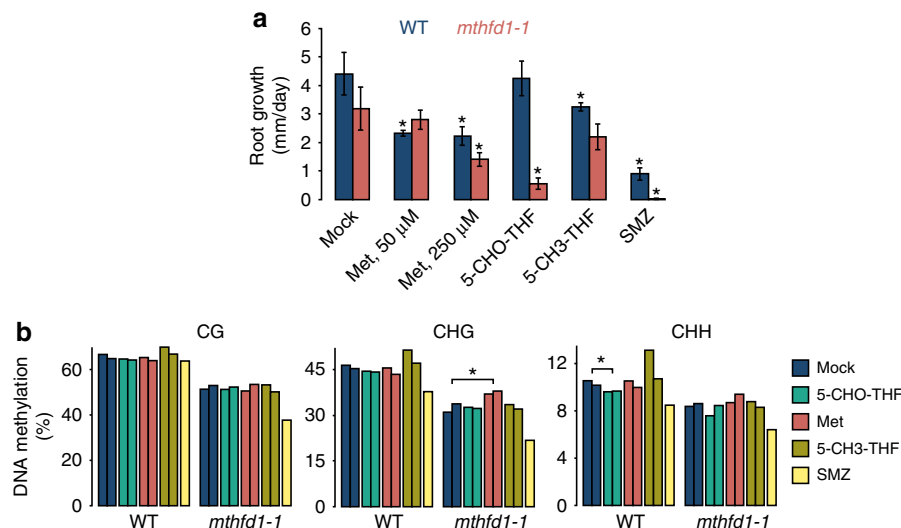
defects in *mthfd1-1* are probably not the mere result of diminished folate pools, but rather point towards an inhibition of methionine synthase.

## Discussion

Methylation patterns in the *Arabidopsis* genome are remarkably stable not only from one generation to the next but also at evolutionary timescales<sup>36–39</sup>. Comparative genomics and genome-wide association studies have linked DNA methylation and phenotypic variation in Brassicaceae to genetic polymorphisms in the DNA methylation machinery<sup>40,41</sup>, and consequently support an adaptive role of spontaneous epigenetic changes. For example, two independent studies have revealed that different alleles of *CMT2* and the concomitant differences in CHH methylation are associated with climate adaptation<sup>40,42</sup>. Here we have identified an EMS-induced polymorphism in the essential folate metabolic enzyme MTHFD1 from *Arabidopsis*, which causes a strong, genome-wide decrease in DNA methylation. This finding highlights that DNA methylation patterns in *Arabidopsis* not only depend on the pathways and catalytic activities of the DNMTs but also on the metabolic network that regulates the availability of the methyl donor SAM and the adequate functioning of the activated methyl cycle. It is therefore conceivable that regulatory mechanisms have evolved, which connect nutritional changes to epigenetic gene regulation by DNA and histone methylation. Although direct examples in plants are still lacking, it has been shown that a folate-rich diet in mice leads to changes in coat colour of the offspring that is caused by altered expression of the *agouti* gene due to increased DNA methylation of a transposon in the *agouti* locus (*A<sup>v</sup>*)<sup>43</sup>. This finding illustrates an example of how such regulatory mechanisms could work.

The EMS-allele *mthfd1-1* was identified through a genetic screen for mutants that simultaneously affect CHG and CHH methylation. Correspondingly, our genome-wide BS-seq analysis of *mthfd1-1* mutants revealed extensive hypomethylation in CHG and CHH sequence contexts. In contrast, loss of CG methylation was comparatively low. We therefore reason that the feedback regulation between CHG/CHH and H3K9 methylation is particularly prone to changes in one-carbon metabolism, because transmethylation by DNA and histone methyltransferases are both SAM-dependent. Accordingly, *mthfd1-1*, as well as previous analyses of *fpgs1* mutants and plants treated with SMZ<sup>34,35</sup>, showed reduced H3K9me2. Because of the mechanistic interdependence of non-CG and H3K9 methylation<sup>5</sup>, it is difficult to tell whether histone or DNA methylation is more directly affected by impaired one-carbon metabolism. The predominant effect on CHG methylation might suggest that loss of H3K9me2 is the primary defect. However, the fact that CG methylation is also decreased in *mthfd1-1*, a type of methylation that is not linked with H3K9 methylation, suggests a general inhibition of the enzymatic activity of different methyltransferases, including MET1.

Our transcriptome analysis demonstrated that loss of DNA methylation in *mthfd1-1* mutants leads to derepression of TEs and a generally higher abundance of transcripts from the pericentromeric heterochromatin compared with WT. With respect to the almost equal numbers of up- and downregulated PCGs, loss of gene body methylation can only account for some of the observed changes, whereas the majority of differential gene expression is probably caused by pleiotropic effects of impaired MTHFD1 function. Secondary to its role in one-carbon metabolism, a shortfall of NADP<sup>+</sup> conversion by MTHFD1 is expected to severely disturb the redox state<sup>44</sup>. Accordingly, GO term analysis revealed a significant enrichment of genes involved in cell redox homeostasis, oxidative pentose phosphate pathway and glycolysis (Fig. 4f and Supplementary Data 2) (ref. 45).



**Figure 8 | *mthfd1-1* mutants are hypersensitive to exogenous 5-CHO-THF but tolerant to exogenous methionine.** (a) Root growth of seedlings on solid media containing mock, Met, 0.5 mM 5-CHO-THF, 0.5 mM 5-CH3-THF or 5 μM SMZ. Mean values ± s.d. ( $n=3$ ) are shown. (b) Average CG, CHG and CHH methylation levels at previously defined *mthfd1-1* CG, CHG and CHH hypo-DMRs, respectively, in seedlings grown for 14 days on solid media containing mock, 0.1 mM Met, 0.5 mM 5-CHO-THF, 0.5 mM 5-CH3-THF or 5 μM SMZ. Two biological replicates are shown, except for SMZ. \* indicates significant difference between mock and chemical treatment ( $P < 0.05$ , Student's *t*-test).

MTHFD1 related proteins in different species have mono-, bi- or trifunctional enzymatic activity. Yeast and mammalian cytosolic homologues, known as C1-THF synthases, are trifunctional and reversibly catalyse the stepwise oxidation from 5,10-CH2-THF to 10-CHO-THF, which serve for thymidylate/pantothenate and *de novo* purine/*N*-formylmethionine synthesis, respectively, and the conversion of 10-CHO-THF to THF and formate (reverse FTHFS activity) (Fig. 5) (refs 14,46,47). Bifunctional forms, which are found in certain bacteria and in plants, lack the FTHFS activity<sup>18,48,49</sup>. Moreover, methylenetetrahydrofolate reductase converts 5,10-CH2-THF to 5-CH3-THF and thereby directs activated methyl towards SAM (Fig. 5). As such, the reversible enzymatic activity of MTHFD1 channels one-carbon into different pathways and acts as a crucial regulatory hub (Fig. 5). Correspondingly, functional mutations, such as in *mthfd1-2*, have severe pleiotropic effects and are mostly lethal. This is in contrast to the subtle morphological defects generally observed in epigenetic *Arabidopsis* mutants, for example, *drm1 drm2 cmt3* triple knockout mutants<sup>50</sup>, and denotes that inhibition of MTHFD1 leads to pleiotropic morphological defects that are independent of its impact on DNA and histone methylation. The essential nature of MTHFD1 further indicates that the additional three *Arabidopsis* MTHFD homologues have plastid- and mitochondrion-specific functions that cannot compensate for a loss of cytoplasmic MTHFD1 function.

Because of its role in nucleotide biosynthesis and DNA methylation, folate metabolism is of central relevance in cancer research, as exemplified by the therapeutic use of antifolates<sup>51</sup>. Polymorphisms in human MTHFD1 C1-THF synthase have been associated with cancers, as well as neural tube defects and other illnesses<sup>51</sup>. Interestingly, one polymorphism (R173C) resides in the same conserved residue that is mutated in the EMS-allele *mthfd1-1* and was linked to severe combined immunodeficiency, megaloblastic anaemia and altered Met metabolism, including Hcy accumulation<sup>52</sup>. Analyses of fibroblasts harbouring this mutation showed signs of DNA damage and uracil misincorporation into DNA due to impaired *de novo* thymidylate synthesis<sup>52</sup>. Interestingly, the mutation had the strongest impact on one-carbon flow towards the Met cycle<sup>52</sup>. Impaired dehydrogenase activity was partially compensated by

increased SHMT activity, as well as increased salvage thymidylate synthesis, whereas *de novo* purine synthesis was not affected<sup>52</sup>. The study did not include DNA methylation analyses, but given the conserved function of MTHFD1 and shared effects on Hcy remethylation, our results predict that DNA methylation is affected by the MTHFD1 R173C mutation and might be involved in certain types of severe combined immunodeficiency and megaloblastic anaemia. In reverse, analogous redirection of one-carbon flow towards nucleotide synthesis at the expense of Hcy remethylation is a possible explanation of the defects observed in *mthfd1-1* and was also suggested to occur upon methotrexate-induced THF depletion in *Arabidopsis*, based on folate measurements and transcriptional analyses<sup>53</sup>.

Hcy accumulation and decreased MI, as observed in *mthfd1-1* and previous studies<sup>32–35</sup>, are hallmarks of impaired Hcy remethylation due to impaired folate metabolism. Increased Hcy levels lead to decreased SAH hydrolase activity and accumulation of SAH, which competitively inhibits SAM-dependent transmethylation, including DNA and histone methylation<sup>54,55</sup>.

Decreased flux of one-carbon towards the remethylation of Hcy should intuitively lead to decreased levels of Met, yet we observed increased cellular Met in *mthfd1-1*. However, primary metabolites, in particular the sulfur amino acids Cys and Met, are often controlled by multiple layers of regulatory circuits, as exemplified by the *sir1-1* mutant, which also show an increased Met steady level despite a 20-fold decreased flux of sulfur through the assimilatory sulfate reduction pathway<sup>56</sup>. This increase was the result of decreased flux into Met sinks, due to an attenuation of translation and growth in *sir1-1* (ref. 56). To that effect, it is conceivable that Met levels were increased in the *mthfd1* mutant because histone and DNA methylation (two major one-carbon sinks) and growth were decreased. Moreover, increased Met levels might also be due to increased Met *de novo* synthesis. Accordingly, the transcriptome analysis of *mthfd1-1* shows a 4.3-fold increase in transcripts corresponding to *MRU1* (At5g35490), which is also upregulated in the Met over-accumulating mutant *mtol-1* (ref. 57).

In the cases of chemically inhibited THF synthesis or impaired THF polyglutamylation, exogenous application of 5-CHO-THF,



which is readily assimilated and metabolized to active THF forms in *Arabidopsis*<sup>53,58</sup>, successfully reversed the feedback inhibition of transmethylation and TGS<sup>34,35</sup>. Interestingly, 5-CHO-THF feeding to *mthfd1-1* mutants did not complement the DNA methylation defect and had a strong adverse effect on root growth. This hypersensitivity could be attributable to an inhibitory effect of 5-CHO-THF on SHMT<sup>59</sup>. As in R173C fibroblasts<sup>52</sup>, it is likely that the supply of 5,10-CH<sub>2</sub>-THF in *mthfd1-1* mutants depends on cytosolic SHMT. Although cell compartmentalization demands a cautious interpretation of the metabolic profiles, the observed decrease in 5-CHO-THF levels by 33% in *mthfd1-1* versus WT might have led to an increase in SHMT activity. Since the reaction equilibrium catalysed by SHMT favours Gly production<sup>60</sup>, increased SHMT activity might have contributed to the threefold increase in steady-state Gly levels observed in *mthfd1-1* versus WT. An inhibition of SHMT by exogenously applied 5-CHO-THF would accordingly cut off the cytosolic 5,10-CH<sub>2</sub>-THF supply and explain the enhanced root growth defect in *mthfd1-1* mutants. It is noteworthy that we did not observe an enhanced DNA methylation defect upon 5-CHO-THF feeding, which suggests that even under normal growth conditions SHMT-dependent 5,10-CH<sub>2</sub>-THF production is unable to perpetuate the Met cycle in *mthfd1-1*. This is further supported by the lack of phenotypic rescue of *mthfd1-1* by exogenous 5-CH<sub>3</sub>-THF. On the other hand, exogenous Met partially restored global CHG methylation, which together with the folate quantifications and feeding experiments suggests that transmethylation in *mthfd1-1* is impaired due to an inhibition of Hcy remethylation, as opposed to limited availability of folate intermediates<sup>34,35</sup> or inhibition of SAHH<sup>54</sup>. As such, the described DNA methylation and gene regulatory defects in *mthfd1-1* highlight a central regulatory role of MTHFD1 in one-carbon distribution towards different cell physiological processes.

## Methods

**Plant material.** All plants used in this study were of the Columbia-0 ecotype. T-DNA insertion mutants *mthfd1-2* (WiscDsLox244C04), *mthfd1-3* (SALK\_015165) and *mthfd1-4* (SALK\_039538) were obtained from the *Arabidopsis* Biological Research Center (Ohio State University). Genotypes were analysed by PCR using primers listed in Supplementary Table 2. The triple mutant *drm1 drm2 cmt3* and the WT transgenic line carrying the *SDC<sub>pro</sub>-GFP* fusion construct were published previously<sup>4,13</sup>. *mthfd1-1* mutants have been backcrossed with WT plants carrying *SDC<sub>pro</sub>-GFP*. Plants were grown in the greenhouse at long day light cycles, unless stated differently.

**Genetics screening and mapping analyses.** WT seeds (2,000) were suspended in 0.3% EMS solution for 13 h with rotation, washed with water and planted on soil. Approximately 1,000 independent M2 populations were collected and screened for GFP fluorescence using a Leica MZ16F Fluorescence Stereomicroscope equipped with a GFP Plus filter. Pictures were taken with a DFC300 FX digital camera. For mapping and identification of EMS mutations, mutant #162 was crossed with WT *Ler* and 10-days-old F2 seedlings grown on media containing 1 × Murashige and Skoog basal salt mixture (MP) and 20 μg ml<sup>-1</sup> glufosinate ammonium (Sigma) were analysed for GFP expression. Genomic DNA was isolated from pooled tissue of 50 GFP-positive F2 mutants and analysed by whole-genome re-sequencing for co-segregating single-nucleotide polymorphisms between Col and *Ler*<sup>13</sup>. Primer sequences of CAPS markers for co-segregation analyses are shown in Supplementary Table 2.

**Local DNA methylation analyses.** Genomic DNA was isolated from aerial tissue of 4–5-weeks-old plants. The *MEA-ISR* probe for DNA blot analysis was amplified using primers JP980 and JP981 (Supplementary Table 2) (ref. 15). Vertically uncropped images of all blots and gels shown in this study are provided in Supplementary Fig. 10. Chop-PCR analysis of *AtSN1* was performed by real-time PCR using primers JP6349 and JP6350 (Supplementary Table 2) (ref. 61). For DNA methylation analysis of the transgenic *SDC* promoter, DNA was BS converted using EZ DNA Methylation Gold kit (Zymo Research) and PCR amplified using primers listed in Supplementary Table 2. PCR fragments were cloned into pCR2.1-TOPO (Thermo Fisher Scientific), and 20 clones per genotype were sequenced.

**Whole-genome bisulfite sequencing.** Genomic DNA was extracted from rosette leaves of 3-weeks-old plants using DNeasy Plant Mini Kit (Qiagen) and fragmented into 200 bp average size with a Covaris S2 sonicator. Next, fragmented DNA was end repaired, adenylated and ligated with TruSeq DNA LT adapters (Illumina) using NEBNext DNA library prep reagent set (NEB). Subsequently, BS conversion was performed with CpGenome DNA modification kit (Millipore). Libraries were amplified using PCR primer cocktail (Illumina) and Pfu Turbo Cx hotstart DNA polymerase (Agilent). Sequencing was performed on a HiSeq 2000 platform at 50 bp length. Identical reads were removed and unique reads were aligned to the *Arabidopsis* reference genome (TAIR10) using BSMAP 2.87 (ref. 62). Read statistics are listed in Supplementary Table 3. Data for mutants other than the *mthfd1-1* were obtained from GSE39901 (ref. 63). Methylation levels were calculated as #C/(#C + #T). DMRs were defined by dividing the genome into 100 bp bins and comparing mutants and WT by the number of methylated and unmethylated Cs with at least four Cs covered using Fisher's exact test and cutoffs of Benjamini–Hochberg corrected false discovery rate < 0.01. Moreover, absolute methylation difference of each bin had to be at least 0.4, 0.2 and 0.1 for CG, CHG and CHH, respectively. Heat maps of DMRs were generated by 'pheatmap' package in R software and clusters were grouped by the complete linkage method with Euclidean distance measurement. Venn diagrams were generated by calculating overlaps of 100 bp DMRs.

**RNA analyses.** Total RNA was isolated with TRIzol (Thermo Fisher Scientific) from 0.1 g of rosette leaves from 3-weeks-old plants. For real-time RT-PCR analysis, 2 μg of DNase I-treated total RNA were reverse-transcribed with SuperScript III (Thermo Fisher Scientific) and cDNA was amplified at target loci (primers listed in Supplementary Table 2) using iQ SYBR Green Supermix (Bio-Rad) and a Mx3005P qPCR system (Agilent Technologies).

For RNA-seq analysis, unstranded libraries from poly-A-tailed RNA were generated according to the manufacturer's instructions (Illumina TruSeq) and sequenced with the HiSeq 2,000 platform at 50 bp length. Reads were mapped to the TAIR10 genome with TopHat2 (ref. 64) using default settings, except that intron length was set to 40–5,000. Read statistics are listed in Supplementary Table 3. Fragments per kilobase of exon per million fragments mapped (FPKM) values and differential gene expression were analysed with Cufflinks<sup>65</sup> using default settings, except that maximum intron length was set to 5,000 and the *-u* option was used. The reference annotation for Cufflinks analysis was downloaded from TAIR and combined genes, including pseudogenes and TE genes, and TEs. GO term enrichment in genes that were significantly down regulated in *mthfd1-1* compared with WT by at least twofold was analysed with GOrilla<sup>66</sup>, using all *Arabidopsis* PCGs as background list.

**Immunofluorescence analysis.** Nuclei from rosette leaves of 3-weeks-old plants were immunostained with anti-H3K9me2 primary (Abcam ab1220, 5 μg ml<sup>-1</sup>) and Alexa Fluor 647-conjugated anti-mouse IgG secondary (Thermo Fisher Scientific A-31571, 10 μg ml<sup>-1</sup>) antibodies, and counterstained with DAPI (1 μg ml<sup>-1</sup>) (ref. 13). Stained nuclei were imaged with a LSM 710 confocal microscope (Zeiss), with a C-Apochromat × 40/1.2 W Corr M27 objective and detection at λ (nm) = 410–504 (DAPI) and λ (nm) = 653–680 (Alexa Fluor 647).

**Subcellular localization.** To generate C-terminally tagged translational fusion proteins MTHFD1-YPET-3xFLAG, MTHFD1\_R175Q-YPET-3xFLAG and FOLD4-YPET-3xFLAG, genomic DNA from Col and *mthfd1-1* was amplified with primer pairs JP14184/5 and JP14190/1 (Supplementary Table 2), spanning the entire ORF (excluding Stop) and 1147 and 866 bp 5' of the ORF of *MTHFD1* and *FOLD4*, respectively. The amplified products were digested with XhoI & SpeI or Sall & SpeI and ligated with the plasmid pBJ36 (ref. 67), which has been linearized with XhoI & XbaI and contained an insertion of *YPET-3xFLAG* on the 3'-side of the XbaI site. NotI fragments from the resulting plasmids were inserted into the NotI site of the binary vector pMLBART<sup>67</sup>. Overnight cultures of transformed *Agrobacterium tumefaciens* strain ASE were adjusted to OD<sub>600</sub> = 0.3 and coinfiltrated with p19 into *N. benthamiana* leaves<sup>68</sup>. Leave discs were imaged 4 days after infiltration with a LSM 710 confocal microscope (Zeiss), using a Plan-Apochromat × 20/0.8 M27 objective and sequential scanning at excitation/detection λ (nm) = 514/519–559 (YFP), 488/630–730 (chlorophyll) and 405/409–530 (DAPI).

**Immunopurification and western blot analysis.** Agro-infiltrated *N. benthamiana* leaves (0.5 g) were ground in liquid nitrogen, and ground tissue was resuspended in 3 ml of IP buffer (50 mM Tris pH7.6, 150 mM NaCl, 5 mM MgCl<sub>2</sub>, 5% (vol/vol) glycerol, 1% Tergitol (Type NP-40, Sigma), 2.8 mM β-mercaptoethanol, 1 μg ml<sup>-1</sup> pepstatin, 1 mM PMSF and 1 × protease inhibitor mixture tablet (Roche, 14696200)). Cleared lysates were incubated with 4 μl of anti-GFP antibody (A-11122, Molecular Probes), followed by 50 μl of Dynabeads Protein G (Thermo Fisher Scientific) at 4 °C for 1 h each. Western blotting was performed with anti-FLAG M2-Peroxidase (horseradish peroxidase) antibody (A8592-1MG, Sigma, 1:7,500 dilution).

**Quantification of metabolites.** Thiols, amino acids and adenosine nucleotides were extracted with 0.5 ml of 0.1 M hydrochloric acid from 0.1 g of in liquid nitrogen grinded rosette leaves from 4-weeks-old plants ( $n=6$ ) grown at long day light and 21 °C. Amino acids and thiol were labelled with AccQ-Tag (Waters) and monobromobimane (Callbiochem), respectively, and quantified after separation by reverse phase chromatography<sup>69</sup>. SAM and SAH were converted by chloroacetaldehyde treatment to their fluorescent etheno-derivates and quantified according to Burstenbinder *et al.*<sup>70</sup> after separation on a Gemini-NX C18 column (150 × 3 mm, 5 µm, 110 Å, Phenomenex, Germany) connected to a Waters 600 HPLC system with a flow rate of 1 ml min<sup>-1</sup> using the following gradient: 5 min 100% buffer A (50 mM tri-sodium phosphate decahydrate, 10 mM sodium 1-heptane sulfonate, 4% acetonitrile, pH 3.2); linear gradient for 15 min to 15% buffer B (pure acetonitrile); 7 min linear gradient to 90% buffer B; and 3 min 90% buffer B followed by re-equilibration of the column in 100% buffer A for 20 min.

**Folate analysis.** *Arabidopsis* rosette leaves (~0.15 g) were pulverized in a mortar with addition of liquid N<sub>2</sub> and homogenized with 10 ml folate extraction buffer (50 mM HEPES, 50 mM CHES, 10 mM β-mercaptoethanol, 2% Na-ascorbate (p/v), pH 7.9). The extracts were deglutamylated with a recombinant conjugase from *Arabidopsis* (100 µg AtGGH2 g<sup>-1</sup> sample) for 1 h at 37 °C. Foliates were purified by affinity chromatography using folate-binding columns. Purified folates were separated by liquid chromatography (Agilent Technologies, Santa Clara, CA, USA) using a Prodigy ODS(2) column (150 × 3.2 mm; 5 µm particle size) (Phenomenex, Torrance, CA, USA) with a 33 min nonlinear gradient of phase A (28 mM K<sub>2</sub>HPO<sub>4</sub>, 59 mM H<sub>3</sub>PO<sub>4</sub>) and phase B (75% phase A, 25% acetonitrile): 10% B (0–2 min); 10–20% B (2–4 min); 20–47% B (4–20 min); 47–80% B (20–25 min); 100% B (25–30 min); and 10% B (30–33 min) with a 1 ml min<sup>-1</sup> flow. Folate derivatives were detected by a four-channel electrochemical detector (CoulArray Model 5600A, ESA, Massachusetts, USA) with potentials set at 100, 200, 300 and 400 mV. THF, 5-methyl-THF, 5,10-methenyl-THF and 5-formyl-THF were quantified using calibration curves made with standards obtained from Schircks (Schircks Laboratories, Buechstrasse, Jona Switzerland). Because of the acidic pH of the mobile phase, in these analyses, THF represents THF + 5,10-methylene-THF and 5,10-CH = THF comprises 5,10-CH = THF + 10-CHO-THF<sup>33</sup>.

**Root growth assays and global DNA methylation analyses.** Seeds were germinated on Phyto agar (RPI Corp.) containing 1 × Murashige and Skoog Basalt Salt Mixture (MP) and 500 µM (6R,S)-5-CHO-5,6,7,8-THF calcium salt (Schircks Laboratories); 50, 100 or 250 µM L-methionine (SIGMA); 5 µM SMZ (SIGMA); 500 µM (6R,S)-5-CH<sub>3</sub>-5,6,7,8-THF calcium salt (Schircks Laboratories), which has been re-applied directly to the roots every 3 days due to its instability; or mock. Seedlings were grown vertically at 16 h light/ 8 h dark cycles and 22 °C. Root lengths were measured at 6, 8, 10, 12 and 14 days after germination and growth rates were calculated by linear regression from at least 10 seedlings per genotype per replicate. For measurement of global DNA methylation, seedlings were pooled per genotype and isolated genomic DNA was analysed by BS-seq as described above, except that EZ DNA Methylation-Lightning kit (Zymo Research) was used for BS conversion. DNA methylation levels were calculated as #C/(#C + #T) at CG, CHG and CHH sites and averaged over previously defined *mtfhd1-1* DMRs. Read statistics are listed in Supplementary Table 3.

**Data availability statement.** Primary high-throughput sequencing data that support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) with the accession code GSE77966 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE77966>).

Secondary high-throughput sequencing data that support the findings of this study are available in the Gene Expression Omnibus (GEO) with the accession code GSE39901 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39901>).

The authors declare that all other relevant data supporting the findings of this study and computer code are available within the article and its Supplementary Information files or on request.

## References

- Law, J. A. & Jacobsen, S. E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 (2010).
- Cao, X. & Jacobsen, S. E. Role of the *Arabidopsis* DRM methyltransferases in *de novo* DNA methylation and gene silencing. *Curr. Biol.* **12**, 1138–1144 (2002).
- Woo, H. R., Dittmer, T. A. & Richards, E. J. Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in *Arabidopsis*. *PLoS Genet.* **4**, e1000156 (2008).
- Johnson, L. M. *et al.* The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr. Biol.* **17**, 379–384 (2007).
- Du, J. *et al.* Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell* **151**, 167–180 (2012).
- Stroud, H. *et al.* Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. *Nat. Struct. Mol. Biol.* **21**, 64–72 (2014).
- Zemach, A. *et al.* The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205 (2013).
- Matzke, M. A. & Moshier, R. A. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* **15**, 394–408 (2014).
- Zhong, X. *et al.* Molecular mechanism of action of plant DRM *de novo* DNA methyltransferases. *Cell* **157**, 1050–1060 (2014).
- Law, J. A. *et al.* Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* **498**, 385–389 (2013).
- Johnson, L. M. *et al.* SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature* **507**, 124–128 (2014).
- Henderson, I. R. & Jacobsen, S. E. Tandem repeats upstream of the *Arabidopsis* endogene SDC recruit non-CG DNA methylation and initiate siRNA spreading. *Genes Dev.* **22**, 1597–1606 (2008).
- Moissiard, G. *et al.* MORC family ATPases required for heterochromatin condensation and gene silencing. *Science* **336**, 1448–1451 (2012).
- Hanson, A. D. & Roje, S. One-carbon metabolism in higher plants. *Annu. Rev. Plant. Physiol. Mol. Biol.* **52**, 119–137 (2001).
- Cao, X. & Jacobsen, S. E. Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. *Proc. Natl Acad. Sci. USA* **99**, 16491–16498 (2002).
- Cao, X. *et al.* Conserved plant genes with similarity to mammalian *de novo* DNA methyltransferases. *Proc. Natl Acad. Sci. USA* **97**, 4979–4984 (2000).
- Cossins, E. A. & Chen, L. Foliates and one-carbon metabolism in plants and fungi. *Phytochemistry* **45**, 437–452 (1997).
- Kirk, C. D., Chen, L., Imeson, H. C. & Cossins, E. A. A 5,10-methylenetetrahydrofolate dehydrogenase: 5,10-methylenetetrahydrofolate cyclohydrolase protein from *Pisum sativum*. *Phytochemistry* **39**, 1309–1317 (1995).
- Collakova, E. *et al.* *Arabidopsis* 10-formyl tetrahydrofolate deformylases are essential for photorespiration. *Plant Cell* **20**, 1818–1832 (2008).
- Ito, J. *et al.* Analysis of the *Arabidopsis* cytosolic proteome highlights subcellular partitioning of central plant metabolism. *J. Proteome. Res.* **10**, 1571–1582 (2011).
- Zybailov, B. *et al.* Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE* **3**, e1994 (2008).
- Sauter, M., Moffatt, B., Saechao, M. C., Hell, R. & Wirtz, M. Methionine salvage and S-adenosylmethionine: essential links between sulfur, ethylene and polyamine biosynthesis. *Biochem. J.* **451**, 145–154 (2013).
- De La Haba, G. & Cantoni, G. L. The enzymatic synthesis of S-adenosyl-L-homocysteine from adenosine and homocysteine. *J. Biol. Chem.* **234**, 603–608 (1959).
- Moffatt, B. A. *et al.* Adenosine kinase deficiency is associated with developmental abnormalities and reduced transmethylation. *Plant. Physiol.* **128**, 812–821 (2002).
- Guranowski, A. & Pawelkiewicz, J. Adenosylhomocysteinase from yellow lupin seeds. Purification and properties. *Eur. J. Biochem.* **80**, 517–523 (1977).
- Roje, S. *et al.* Isolation, characterization, and functional expression of cDNAs encoding NADH-dependent methylenetetrahydrofolate reductase from higher plants. *J. Biol. Chem.* **274**, 36089–36096 (1999).
- Allaire, M., Li, Y., MacKenzie, R. E. & Cygler, M. The 3-D structure of a folate-dependent dehydrogenase/cyclohydrolase bifunctional enzyme at 1.5 Å resolution. *Structure* **6**, 173–182 (1998).
- Pawelek, P. D., Allaire, M., Cygler, M. & MacKenzie, R. E. Channeling efficiency in the bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase domain: the effects of site-directed mutagenesis of NADP binding residues. *Biochim. Biophys. Acta* **1479**, 59–68 (2000).
- Li, R., Moore, M. & King, J. Investigating the regulation of one-carbon metabolism in *Arabidopsis thaliana*. *Plant Cell Physiol.* **44**, 233–241 (2003).
- Prabhu, V., Chatson, K. B., Abrams, G. D. & King, J. 13C nuclear magnetic resonance detection of interactions of serine hydroxymethyltransferase with C1-tetrahydrofolate synthase and glycine decarboxylase complex activities in *Arabidopsis*. *Plant Physiol.* **112**, 207–216 (1996).
- Hare, P. D. & Cress, W. A. Metabolic implications of stress-induced proline accumulation in plants. *Plant Growth Regul.* **21**, 79–102 (1997).
- Mehrshahi, P. *et al.* Functional analysis of folate polyglutamylation and its essential role in plant metabolism and development. *Plant J.* **64**, 267–279 (2010).
- Srivastava, A. C. *et al.* The folylpolyglutamate synthetase plastidial isoform is required for postembryonic root development in *Arabidopsis*. *Plant. Physiol.* **155**, 1237–1251 (2011).
- Zhang, H. *et al.* Sulfamethazine suppresses epigenetic silencing in *Arabidopsis* by impairing folate synthesis. *Plant Cell* **24**, 1230–1241 (2012).
- Zhou, H. R. *et al.* Folate polyglutamylation is involved in chromatin silencing by maintaining global DNA methylation and histone H3K9 dimethylation in *Arabidopsis*. *Plant Cell* **25**, 2545–2559 (2013).



36. Becker, C. *et al.* Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480**, 245–249 (2011).
37. Haggmann, J. *et al.* Century-scale methylome stability in a recently diverged *Arabidopsis thaliana* lineage. *PLoS Genet.* **11**, e1004920 (2015).
38. Schmitz, R. J. *et al.* Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334**, 369–373 (2011).
39. van der Graaf, A. *et al.* Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *Proc. Natl Acad. Sci. USA* **112**, 6676–6681 (2015).
40. Dubin, M. J. *et al.* DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation. *Elife* **4**, e05255 (2015).
41. Willing, E.-M. *et al.* Genome expansion of *Arabidopsis alpina* linked with retrotransposition and reduced symmetric DNA methylation. *Nat. Plants* **1**, 14023 (2015).
42. Shen, X. *et al.* Natural CMT2 variation is associated with genome-wide methylation changes and temperature seasonality. *PLoS Genet* **10**, e1004842 (2014).
43. Waterland, R. A. & Jirtle, R. L. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.* **23**, 5293–5300 (2003).
44. Fan, J. *et al.* Quantitative flux analysis reveals folate-dependent NADPH production. *Nature* **510**, 298–302 (2014).
45. Van Wilder, V. *et al.* C1 metabolism and chlorophyll synthesis: the Mg-protoporphyrin IX methyltransferase activity is dependent on the folate status. *New Phytol.* **182**, 137–145 (2009).
46. Hum, D. W., Bell, A. W., Rozen, R. & MacKenzie, R. E. Primary structure of a human trifunctional enzyme. Isolation of a cDNA encoding methylenetetrahydrofolate dehydrogenase-methylenetetrahydrofolate cyclohydrolase-formyltetrahydrofolate synthetase. *J. Biol. Chem.* **263**, 15946–15950 (1988).
47. Paukert, J. L., Williams, G. R. & Rabinowitz, J. C. Formyl-methylenetetrahydrofolate synthetase (combined); correlation of enzymic activities with limited proteolytic degradation of the protein from yeast. *Biochem. Biophys. Res. Commun.* **77**, 147–154 (1977).
48. Dev, I. K. & Harvey, R. J. A complex of N5,N10-methylenetetrahydrofolate dehydrogenase and N5,N10-methylenetetrahydrofolate cyclohydrolase in *Escherichia coli*. Purification, subunit structure, and allosteric inhibition by N10-formyltetrahydrofolate. *J. Biol. Chem.* **253**, 4245–4253 (1978).
49. Ljungdahl, L. G., O'Brien, W. E., Moore, M. R. & Liu, M. T. Methylenetetrahydrofolate dehydrogenase from *Clostridium formicoaceticum* and methylenetetrahydrofolate dehydrogenase, methylenetetrahydrofolate cyclohydrolase (combined) from *Clostridium thermoaceticum*. *Methods Enzymol.* **66**, 599–609 (1980).
50. Chan, S. W. *et al.* RNAi, DRD1, and histone methylation actively target developmentally important non-CG DNA methylation in *Arabidopsis*. *PLoS Genet.* **2**, e83 (2006).
51. Krajcinovic, M. MTHFD1 gene: role in disease susceptibility and pharmacogenetics. *Pharmacogenomics* **9**, 829–832 (2008).
52. Field, M. S., Kamynina, E., Watkins, D., Rosenblatt, D. S. & Stover, P. J. Human mutations in methylenetetrahydrofolate dehydrogenase 1 impair nuclear *de novo* thymidylate biosynthesis. *Proc. Natl Acad. Sci. USA* **112**, 400–405 (2015).
53. Loizeau, K. *et al.* A genome-wide and metabolic analysis determined the adaptive response of *Arabidopsis* cells to folate depletion induced by methotrexate. *Plant Physiol.* **148**, 2083–2095 (2008).
54. Rocha, P. S. *et al.* The *Arabidopsis* HOMOLOGY-DEPENDENT GENE SILENCING1 gene codes for an S-adenosyl-L-homocysteine hydrolase required for DNA methylation-dependent gene silencing. *Plant Cell* **17**, 404–417 (2005).
55. Tehlivets, O., Malanovic, N., Visram, M., Pavkov-Keller, T. & Keller, W. S-adenosyl-L-homocysteine hydrolase and methylation disorders: yeast as a model system. *Biochim. Biophys. Acta* **1832**, 204–215 (2013).
56. Khan, M. S. *et al.* Sulfite reductase defines a newly discovered bottleneck for assimilatory sulfate reduction and is essential for growth and development in *Arabidopsis thaliana*. *Plant Cell* **22**, 1216–1231 (2010).
57. Goto, D. B. & Naito, S. AtMRD1 and AtMRU1, two novel genes with altered mRNA levels in the methionine over-accumulating mto1-1 mutant of *Arabidopsis thaliana*. *Plant. Cell. Physiol.* **43**, 923–931 (2002).
58. Goyer, A. *et al.* 5-Formyltetrahydrofolate is an inhibitory but well tolerated metabolite in *Arabidopsis* leaves. *J. Biol. Chem.* **280**, 26137–26142 (2005).
59. Roje, S., Janave, M. T., Ziemak, M. J. & Hanson, A. D. Cloning and characterization of mitochondrial 5-formyltetrahydrofolate cycloligase from higher plants. *J. Biol. Chem.* **277**, 42748–42754 (2002).
60. Besson, V., Rebeille, F., Neuburger, M., Douce, R. & Cossins, E. A. Effects of tetrahydrofolate polyglutamates on the kinetic parameters of serine hydroxymethyltransferase and glycine decarboxylase from pea leaf mitochondria. *Biochem. J.* **292**, 425–430 (1993).
61. Deleris, A. *et al.* Involvement of a Jumonji-C domain-containing histone demethylase in DRM2-mediated maintenance of DNA methylation. *EMBO Rep.* **11**, 950–955 (2010).
62. Xi, Y. & Li, W. BSMAP: whole genome bisulfite sequence MAPPING program. *BMC Bioinformatics* **10**, 232 (2009).
63. Stroud, H., Greenberg, M. V., Feng, S., Bernatavichute, Y. V. & Jacobsen, S. E. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* **152**, 352–364 (2013).
64. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
65. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
66. Eden, E., Navon, R., Steinfeld, L., Lipson, D. & Yakhini, Z. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
67. Eshed, Y., Baum, S. F., Perea, J. V. & Bowman, J. L. Establishment of polarity in lateral organs of plants. *Curr. Biol.* **11**, 1251–1260 (2001).
68. Jarsch, I. K. & Ott, T. Quantitative image analysis of membrane microdomains labelled by fluorescently tagged proteins in *Arabidopsis thaliana* and *Nicotiana benthamiana*. *Bio-protocol* **5**, e1497 (2015).
69. Wirtz, M., Droux, M. & Hell, R. O-acetylserine (thiol) lyase: an enigmatic enzyme of plant cysteine biosynthesis revisited in *Arabidopsis thaliana*. *J. Exp. Bot.* **55**, 1785–1798 (2004).
70. Burstenbinder, K., Rzewuski, G., Wirtz, M., Hell, R. & Sauter, M. The role of methionine recycling for ethylene synthesis in *Arabidopsis*. *Plant J.* **49**, 238–249 (2007).

## Acknowledgements

We thank Mahnaz Akhavan for her support in high-throughput sequencing at the University of California, Los Angeles (UCLA) Broad Stem Cell Research Center BioSequencing Core Facility. Work in the Jacobsen laboratory was supported by NIH grant GM60398. M.G. was supported by an EMBO ALTF 986-2011 Long-term Fellowship. Work in the Diaz de la Garza laboratory was supported by CONACYT grant CB243058. C.G.-S. was supported by CONACYT doctoral fellowship 372615. S.F. was a Special Fellow of the Leukemia & Lymphoma Society. S.B. is supported by a postdoctoral fellowship of the Swiss National Science Foundation. J.A.L. is supported by NIH grant GM072764. The work in the Hell/Wirtz laboratory was supported by the German research foundation (grants HE 1848/16-1 and WI 3560/2-1) and the Metabolomics Core Technology Platform of the Excellence cluster 'CellNetworks' (University of Heidelberg). S.E.J. is an Investigator of the Howard Hughes Medical Institute.

## Author contributions

G.M. performed genetic screening. M.G. performed co-segregation analyses, complementation crosses and analysed *mthfd1* T-DNA insertion mutants. M.G., G.M., A.J. and S.F. generated sequencing libraries. H.W., M.G., S.J.C., J.Z. and C.J.H. performed bioinformatics analyses. M.G. performed immunofluorescence analysis. J.L., S.B. and M.G. performed the subcellular localization analysis. M.W. analysed thiols, amino acids and adenosine nucleotides. C.G.-S., P.A.R.-P. and R.I.D.I.G. analysed folates. A.J. and D.C.S. performed growth assays. M.G. and S.E.J. wrote the manuscript. R.H., R.I.D.I.G. and S.E.J. coordinated research.

## Additional information

**Accession codes:** All generated high-throughput sequencing data are available at NCBI's Gene Expression Omnibus (GEO) and are accessible via GEO Series accession number GSE77966.

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no conflict of interest.

**Reprints and permission** information is available online at <http://npng.nature.com/reprintsandpermissions/>

**How to cite this article:** Groth, M. *et al.* MTHFD1 controls DNA methylation in *Arabidopsis*. *Nat. Commun.* **7**:11640 doi: 10.1038/ncomms11640 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>