Partial redundancy buffers deleterious effects of mutating *DNA methyltransferase 1-1* (*MET1-1*) in polyploid wheat

Samuel Burrows[1]*, Delfi Dorussen[1]*, Joseph Crudgington[1], Giorgia Di Santolo[1], James Simmonds[1], Marco Catoni[2], Philippa Borrill[1†]

[1] Department of Crop Genetics, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, United Kingdom.

[2] School of Biosciences, University of Birmingham, Birmingham, B15 2TT, United Kingdom.

* These authors contributed equally

[†] Corresponding author, email: philippa.borrill@jic.ac.uk

Keywords: DNA methylation, DNA methyltransferases, Wheat (*Triticum aestivum*), polyploidy, epibreeding

Short title: Partial redundancy of *MET1-1* in wheat

The author responsible for distribution of materials integral to the findings presented in this article is: Philippa Borrill (philippa.borrill@jic.ac.uk).

Abstract

DNA methylation is conserved across biological kingdoms, playing important roles in gene expression, transposable element silencing and genome stability. Altering DNA methylation could generate additional phenotypic variation for crop breeding, however the lethality of epigenetic mutants in crop species has hindered its investigation. Here, we exploit partial redundancy between homoeologs in polyploid wheat to generate viable mutants in the DNA *methyltransferase 1-1* (*MET1-1*) gene with altered methylation profiles. In both *Triticum turgidum* (tetraploid wheat) and *Triticum aestivum* (hexaploid wheat) we identified clear segregation distortions of higher-order mutants (5/6 and 6/6 mutant *met1-1* copies in hexaploid and 3/4 and 4/4 copies in tetraploid) when genotyping segregating seeds and seedlings, which we attribute to reduced transmission of null mutant gametes. We found that the reduced transmission occurred from both the maternal and paternal gametes, however, we did not detect any

deleterious effects on pollen development. The loss of four or more functional copies of *MET1-1* results in decreased CG methylation in hexaploid wheat. Changes to gene expression increase stepwise with the number of mutant alleles suggesting a dosage dependent effect. Finally, we identify heritable changes to flowering and awn phenotypes which segregate independently of *MET1-1*. Together our results demonstrate that polyploidy can be leveraged to generate quantitative changes to CG methylation without the lethal consequences observed in other crops, opening the potential to exploit novel epialleles in plant breeding.

Introduction

DNA methylation on cytosine bases is an epigenetic modification which is conserved across biological kingdoms and plays important roles in gene expression, transposable element silencing and genome stability. However, differences occur between kingdoms, with cytosine methylation mainly occurring in the CG context in mammals, whereas in plants it can occur in any context (CG, CHG and CHH, where H is A, T, or C) (Feng *et al.*, 2010). Nonetheless, CG methylation is the most prevalent context in plant species examined to date (Niederhuth *et al.*, 2016; Domb *et al.*, 2020; Li *et al.*, 2023). Methylation is established and maintained by multiple enzymes which function in specific or overlapping contexts. CG methylation is maintained during cell division by the methyltransferase 1 (MET1) which recognises hemi-methylated CG sites following DNA replication and methylates the opposite strand (Kankel *et al.*, 2003; Zhang *et al.*, 2018).

Disruption of the *MET1* gene in Arabidopsis (*Arabidopsis thaliana*), tomato (*Solanum lycopersicum*) and rice (*Oryza sativa*) has revealed major differences in the phenotypic impact and severity of null mutations between species. In Arabidopsis *met1* mutants, including complete and partial loss-of-function alleles, have defects in multiple developmental stages including dwarfism, delayed flowering, abnormal embryo development and seed abortion (Kankel *et al.*, 2003; Saze *et al.*, 2003; Xiao *et al.*, 2006; Srikant *et al.*, 2022). Null *met1* alleles have more severe consequences with segregation distortion resulting in an underrepresentation of homozygous mutants, which are only occasionally viable to maturity (Saze *et al.*, 2003; Srikant *et al.*, 2022). In contrast, in rice, null homozygous mutants were formed at the expected ratio

from heterozygous parents but could only grow on specific media and underwent necrotic death at seedling stage (Hu *et al.*, 2014). In tomato, homozygous null mutants could not be recovered from bi-allelic CRISPR mutants, although only 90 plants were examined so it is possible very rare homozygous mutants may be viable (Yang *et al.*, 2019).

Differences in the segregation and viability of homozygous null mutants may be due to variation in epigenetic backup systems: in Arabidopsis these backup systems operate through changes to the RNA-directed DNA methylation (RdDM) process, changes to expression of DNA demethylases and progressive H3K9 remethylation of heterochromatin (Mathieu *et al.*, 2007). Stochasticity in these processes explains the rarity of homozygous null mutants (Mathieu *et al.*, 2007). In rice a duplicate copy of *MET1* has been proposed to provide a back-up mechanism for essential CG methylation, although the lethality of homozygous mutants at seedling stage indicates incomplete compensation (Hu *et al.*, 2014; Yamauchi *et al.*, 2014).

The species examined to date also vary regarding the degree to which CG methylation is lost in *met1* mutants, and compensatory changes to other methylation contexts. Whole-genome bisulfite sequencing of null homozygous mutants revealed a stronger reduction in CG methylation in Arabidopsis than in rice (98.3% vs 75.7%) (Cokus *et al.*, 2008; Hu *et al.*, 2014). In tomato only 25% of CG methylation was lost in a CRISPR mutant but there were multiple *met1* alleles detected, suggesting chimerism within the plant and therefore incomplete penetrance of the phenotype. Heterozygous mutants in tomato and rice maintained similar CG methylation to wild type plants (Hu *et al.*, 2014; Yang *et al.*, 2019), indicating only one allele was sufficient to compensate, whereas effects in Arabidopsis were variable between genotypes: heterozygous mutants were similar to wild type in some accessions, while CG methylation was reduced up to 64% in others (Srikant *et al.*, 2022). In rice homozygous mutants, hypomethylation in CHG and CHH contexts was also observed (Hu *et al.*, 2014), whereas hypermethylation was observed in CHG contexts in euchromatin in Arabidopsis and in CHH contexts in tomato (Lister *et al.*, 2008; Yang *et al.*, 2019).

It is possible that despite retaining more CG methylation in *met1* mutants, the larger genome size in tomato (~950 Mb) and rice (~430 Mb) than in Arabidopsis (~135 Mb), is associated with the stronger developmental problems observed in null mutants in these species. Higher proportions of transposable elements in larger genomes may cause stronger genetic and developmental

defects in *met1* mutants through their remobilisation (Mirouze & Vitte, 2014). This hypothesis is supported by the lack of null mutants in *MET1* orthologs identified in crop species with large genomes including in a barley TILLING population (genome size of ~5Gb) (Schreiber *et al.*, 2019) and in maize TILLING and transposon populations (genome size of ~2.4 Gb) (Li *et al.*, 2014).

Changes to methylation status can result in the production of stable epialleles which cause transcriptional and phenotypic alterations across generations. Epialleles affecting diverse traits have been identified, including alterations in floral morphology, fruit ripening, plant height and climate adaptation in a range of plant species (Cubas *et al.*, 1999; Manning *et al.*, 2006; Miura *et al.*, 2009; He *et al.*, 2018; Srikant & Tri Wibowo, 2021). The deliberate induction of stable epialleles has been demonstrated through epiRIL populations generated genetically by the use of *met1* and *ddm1* (*Decreased DNA Methylation 1*) mutants in Arabidopsis (Johannes *et al.*, 2009; Reinders *et al.*, 2009; Zhang *et al.*, 2013; Catoni & Cortijo, 2018). The potential to exploit novel epialleles in plant breeding has been proposed, but the production of epiRIL populations in crop species as a foundation to generate novel epialleles has been limited by the lethality of mutants (Springer & Schmitz, 2017).

Here we explore the use of a polyploid crop species to generate quantitative changes to CG methylation by knocking out differing numbers of *met1* alleles, without causing lethal consequences, aiming to find a balance between altering epigenetic status while retaining fertility and viability. We use polyploid wheat as a model system due to the availability of sequenced mutant populations which can be readily applied to breeding worldwide without legislative complications surrounding gene editing and high-quality genome sequences allowing accurate analysis of epigenomic consequences (Krasileva *et al.*, 2017; IWGSC *et al.*, 2018). We generate *met1* mutants in tetraploid (*Triticum turgidum*) and hexaploid (*Triticum aestivum*) wheat with up to three null alleles (out of four) in tetraploid wheat, and five null alleles (out of six) in hexaploid wheat. We examine changes to global DNA methylation, transcriptional responses and effects on fertility and plant growth linked to mutant allele copy number and identify stably inherited phenotypic alterations. Our findings provide insights into using polyploidy to generate quantitative changes to CG methylation for epigenetic breeding in crops.

Results

Null *met1-1* mutants cannot be recovered in polyploid wheat

METHYLTRANSFERASE 1 (*MET1*) genes were previously identified in hexaploid wheat (*Triticum aestivum*) on the group 2, 5, and 7 chromosomes, with the group 5 and 7 *MET1* homologs predicted to be pseudogenes (Thomas *et al.*, 2014). While the group 5 and group 7 *MET1* genes are orthologous to *MET1a* in rice, the group 2 *MET1* genes are orthologous to *MET1b* in rice (Thomas *et al.*, 2014) (Figure 1A). *OsMET1b* is also known as *OsMET1-2*, and homozygous null mutants in this gene were found to be seedling lethal in rice (Hu *et al.*, 2014). To determine which *MET1* homologs are functional in wheat, we examined expression of the *MET1* genes using transcriptomic data from a developmental time course in the Chinese Spring cultivar (Choulet *et al.*, 2014). While the group 2 genes (*TraesCS2A02G235900*, *TraesCS2B02G260800*, *TraesCS2D02G241800*) are expressed in numerous above-ground tissues (Figure 1B), expression of the group 5 (*TraesCS5A02G443200*, *TraesCS5B02G447800*, *TraesCS5D02G450700*) and group 7 (*TraesCS7A02G531500*, *TraesCS7B02G448900*, *TraesCS7D02G519000*) genes was virtually undetectable in most of these tissues (Supplementary Figure 1). In addition, expression of the group 2 genes is highest in actively dividing tissues such as the shoot apical meristem (SAM), in line with the role of MET1 in maintaining CG methylation across cell divisions (Zhang *et al.*, 2018). We therefore concentrated our analysis on the group 2 genes which will be referred to as *MET1-A1*, *MET1-B1*, and *MET1-D1* in accordance with the gene nomenclature guidelines for wheat (Boden *et al.*, 2023) and a previous study by Li *et al.* (2021).

We selected lines from the wheat EMS-mutagenised TILLING population with premature termination codon mutations in the three *MET1-1* homoeologs, predicted to disrupt the cytosine methyltransferase domain (Figure 1C). Two independent segregating F_2 populations were produced in hexaploid wheat from which, in theory, all combinations of loss-of-function *met1-1* mutants could be recovered (Figure 1D, Supplementary Figure 2). However, genotyping of the two F_2 populations (3,068 individuals from the C0465xC0884 population, and 328 individuals from the C0451xC2028xC1292 population) revealed that we were unable to recover null *met1-1* mutants with homozygous loss-of-function mutations in all three homoeologs. Moreover, we

only found one plant with five non-functional *MET1-1* alleles which was capable of germination, although its growth was severely stunted and the plant died before progressing to tillering (Zadoks growth stage 20) (Figure 1E). The genotype ratios observed in the C0465xC0884 population were significantly different from the ratios expected from Mendelian segregation of the *MET1-1* alleles (Figure 1F, χ^2 $p=1.63x10^{-103}$), as mutants with four or more non-functional *MET1-1* alleles were underrepresented in the population. This deviation from the expected Mendelian segregation ratio was also observed in the C0451xC2028xC2192 population (Supplementary Table 1, χ^2 $p=5.05x10^{-12}$). We therefore concluded that higher-order *met1-1* mutants have reduced viability in hexaploid wheat.

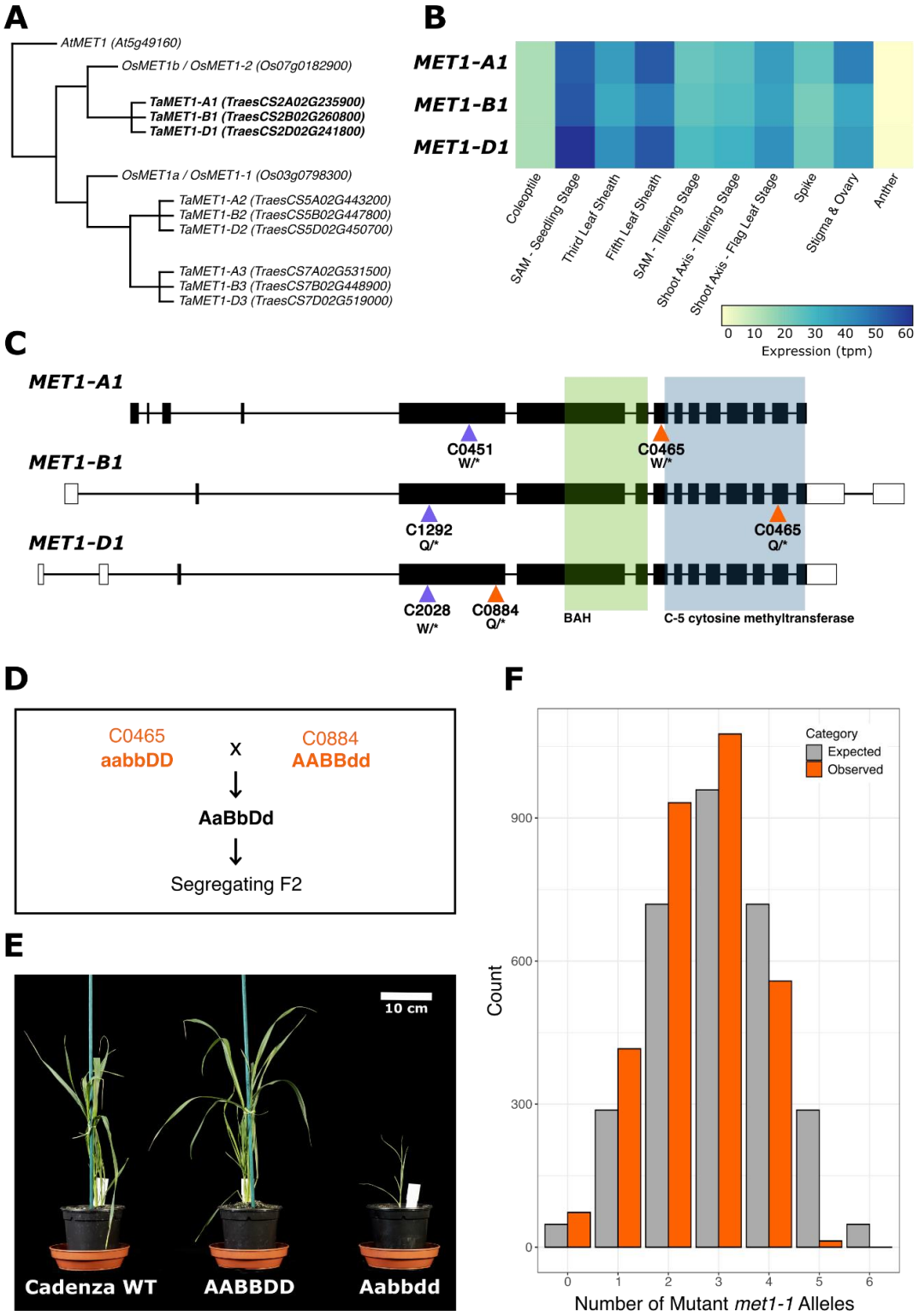


Figure 1. Null *met1-1* mutants cannot be produced by crossing TILLING lines with disruptions in the three *MET1-1* homoeologs. A) Evolutionary relationships between *MET1* homologs in

Arabidopsis (*AtMET1*), rice (*OsMET1a* and *OsMET1b*), and wheat (homoeolog groups *TaMET1-1*, *TaMET1-2*, and *TaMET1-3*). B) Expression of the three *MET1-1* homoeologs (*TraesCS2A02G235900*, *TraesCS2B02G260800*, *TraesCS2D02G241800*) in above-ground tissues across development in the Chinese Spring cultivar. SAM = shoot apical meristem. C) Structure of the three *MET1-1* homoeologs. Filled rectangles represent exons, empty rectangles represent untranslated regions, and lines represent introns. Triangles indicate the positions of the premature termination codon mutations in the TILLING lines used to produce the two hexaploid F_2 populations (C0465xC0884, orange, and C0451xC2028xC1292, purple). Green box signifies Bromo adjacent homology domain (BAH) and blue box signifies C-5 cytosine methyltransferase domain. D) Crossing diagram showing the structure of the C0456xC0884 population. E) Cadenza WT, WT segregant (AABBDD), and Aabbdd individuals from the C0465xC0884 segregating F_2 population. F) Expected (grey) and observed (orange) number of individuals with each possible number of mutant *met1-1* alleles in C0465xC0884 segregating F_2 population.

Reduced viability of null *met1-1* gametes hinders recovery of higher-order *met1-1* mutants

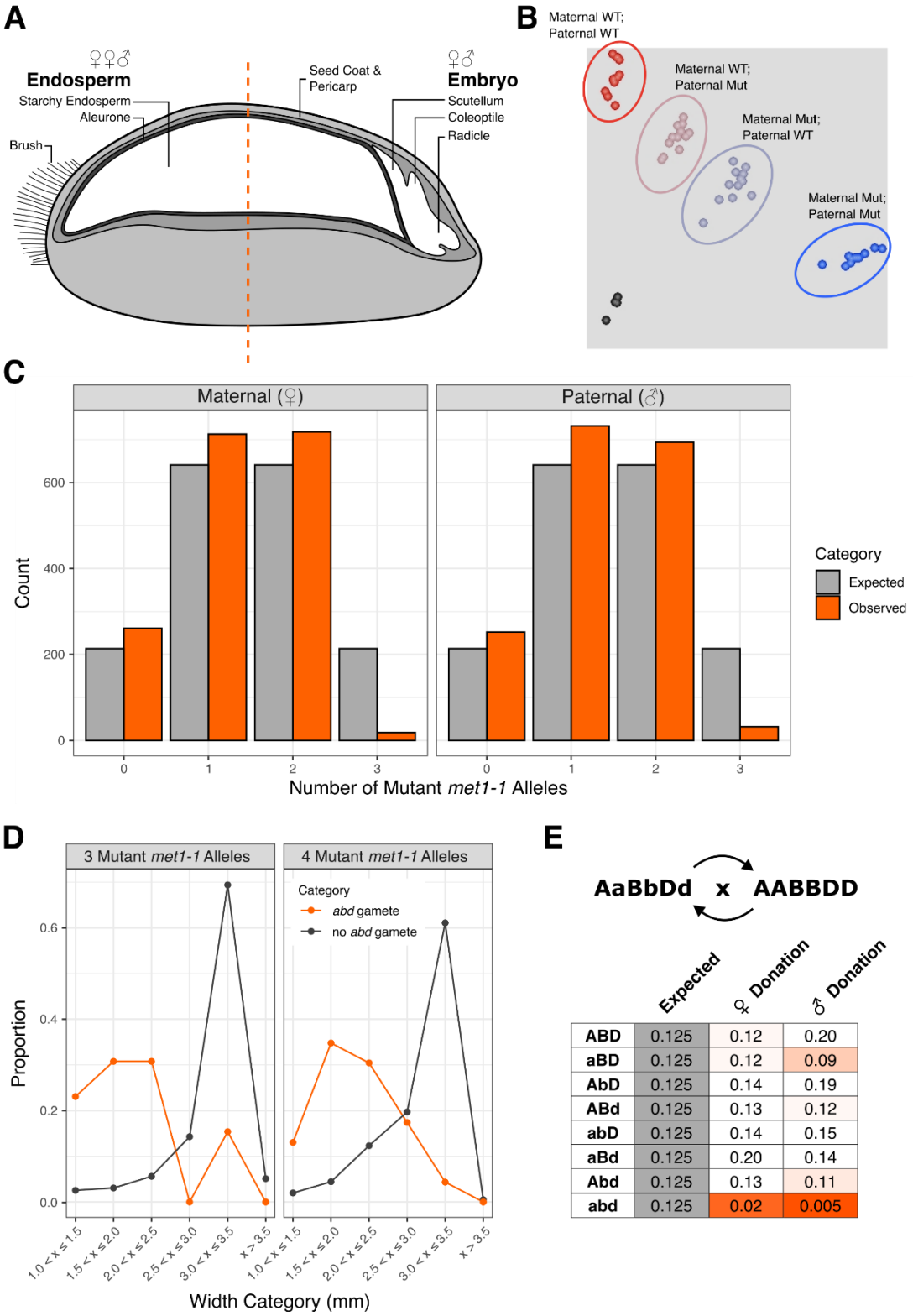
We hypothesized that the absence of functional alleles of *MET1-1* within a gamete might cause the segregation distortion observed. We developed a grain genotyping method using the endosperm-containing half of the grain (Figure 2A) which contains maternal to paternal DNA at a 2:1 ratio (the endosperm originates from the fertilisation of two female central cell nuclei by a single male sperm cell). This allowed us to determine which parent donated the mutant alleles to the offspring (grain) using KASP genotyping by separating the heterozygous genotypes into two clusters (Figure 2B). We confirmed the accuracy of this method using reciprocal crosses for an independent mutation in the semi-dwarfing gene *Rht13* (Borrill *et al.*, 2022) (Figure 2B). Moreover, we were able to grow seedlings from the embryo-containing halves of the grain, allowing us to genotype the leaf tissue and confirm that it matched the grain genotype (Supplementary Table 2). We genotyped 1,710 grains from a segregating hexaploid *MET1-1* Cadenza F_2 population (C0465xC0884) using this method and found that far fewer grains with three mutant alleles of *met1-1* inherited together from one parent were detected than expected (generalised linear model with log link, $p<2E-16$; Figure 2C). We detected fewer grains which inherited three mutant alleles from the maternal side than from the paternal side (18 vs 32 grains). We did not observe a reduction in the number of grains which inherited one or two mutant alleles from the same parent (Figure 2C).

We confirmed these observations in an independent tetraploid wheat population, where we combined premature truncation mutations in the A and B subgenomes by crossing (Supplementary Figure 3). In a segregating F_4 population with 332 grains, we observed a significant segregation distortion (χ^2 p= 3.39×10^{-30}). We were not able to identify any homozygous null mutants (aabb), despite 21 grains expected in a mendelian ratio. We found only 11 grains with three mutant alleles compared to 83 grains expected (Supplementary Table 3). Similar to hexaploid wheat, tetraploid wheat gametes with a complete knock-out of *MET1-1* (i.e. ab gametes) were strongly underrepresented (generalised linear model with log link, $p<2 \times 10^{-16}$), with reduced transmission from both the maternal and paternal (Supplementary Table 3). Consistent with hexaploid wheat, in tetraploid wheat we detected fewer grains which inherited two mutant alleles from the maternal side than from the paternal side (1 vs 18 grains; Supplementary Table 3).

In hexaploid wheat, we found the distribution of genotypes observed when genotyping the grains to be significantly different from the distribution observed when genotyping seedling leaf tissue (χ^2 p=0.00058), with a greater proportion of individuals with five non-functional *MET1-1* alleles observed in the grain population (12 out of 1,710 in the grain population [0.702%] compared to 1 out of 1,352 in the seedling population [0.074%]). This suggests that higher-order *met1-1* mutants may be formed as grains, but not be capable of germination. Previous research characterising *met1* mutants in rice showed that homozygous mutants were associated with abnormal, shrivelled seeds with reduced endosperm and embryo (Hu *et al.*, 2014; Yamauchi *et al.*, 2014). We therefore hypothesised *met1-1* mutants in wheat inheriting three non-functional alleles from the same parent may be associated with a reduction in grain size. The inheritance of three non-functional alleles from the same parent was found to have a significant effect on grain size (ordered logistic regression model, $p<0.05$). We found that, for grains with a total of three mutant *met1-1* alleles, those inheriting all three mutant alleles from one parent were smaller than those inheriting the mutant alleles separately from both parents (mode 1.5 - 2.0/2.0 - 2.5 mm wide vs 3.0 - 3.5 mm wide; Figure 2D). The same pattern was observed for individuals with a total of four mutant *met1-1* alleles; those inheriting three mutant alleles from one parent were smaller than those inheriting two mutant alleles from each parent (mode 1.5 - 2.0 mm wide vs 3.0 - 3.5 mm wide; Figure 2D).

To further understand why the recovery of *met1-1* mutants formed from at least one gamete with no functional copies of *MET1-1* is so low, we investigated whether pollen development was affected in the *met1-1* mutants. We measured the number of pollen grains per anther, pollen diameter, pollen viability, and the proportion of pollen grains that were non- or mononucleated for plants from the segregating F_2 population, including those that should be capable of producing gametes with no functional copies of *MET1-1* (AaBbdd, AabbDd, and aaBbDd). However, no significant differences (at α=0.05) were detected between those with an AaBbdd, AabbDd, or aaBbDd genotype and the WT segregants (AABBDD) for any of the aforementioned phenotypes (Table 1).

To further confirm whether the inheritance of three mutant *met1-1* alleles from the same parent reduces viability, we performed reciprocal crosses between triple heterozygous (AaBbDd) *met1-1* mutants and Cadenza WT. This allowed us to determine the number of progeny formed from the inheritance of each gamete type (ABD, aBD, AbD, ABd, abD, aBd, Abd, and abd) from the AaBbDd parent. A total of 454 progeny were genotyped, 235 with AaBbDd as the female parent and 219 with AaBbDd as the male parent. Again, we found that plants inheriting three non-functional *MET1-1* alleles from one parent were significantly underrepresented compared to the expected Mendelian ratio (χ^2 $p<0.001$ for both maternal and paternal donation; Figure 2E). The proportion of abd gametes was lower from the paternal side of the reciprocal crosses than the maternal side, which was the opposite to the trend observed in the selfed progeny of a triple heterozygous plant (Figure 2C).



	Expected	♀ Donation	♂ Donation
ABD	0.125	0.12	0.20
aBD	0.125	0.12	0.09
AbD	0.125	0.14	0.19
ABd	0.125	0.13	0.12
abD	0.125	0.14	0.15
aBd	0.125	0.20	0.14
Abd	0.125	0.13	0.11
abd	0.125	0.02	0.005

Figure 2. Inheritance of three mutant alleles *met1-1* from the female or male parent (i.e. abd gamete) is rare relative to inheritance of zero, one or two mutant alleles. A) Wheat grains divided

transversely to separate two halves of the grain, with one half containing the embryo for future germination and the other half containing largely endosperm for DNA extraction and genotyping. B) KASP genotyping for *Rht13* mutants (*Rht-B13b*) from reciprocal crosses between mutant and wild type parents with known genotypes, showing homozygous wild type (WT) (red cluster), homozygous mutant (blue cluster), heterozygous clusters with a maternally inherited WT allele and paternally inherited mutant allele (brown), a paternally inherited WT allele and maternally inherited mutant allele (grey) and no template control (black). C) Expected (grey) and observed (orange) number of individuals inheriting zero, one, two, or three mutant *met1-1* alleles from the maternal or paternal gamete from a self-pollinated AaBbDd plant. D) The proportion of grains in each width category, for grains with a total of three or four mutant *met1-1* alleles, formed from an abd gamete (orange) or not (grey). E) Expected and observed proportion of individuals formed from each type of gamete when inherited maternally (♀ donation) or paternally (♂ donation) from an AaBbDd parent when crossed with AABBDD. Orange shades indicate proportions below the expected proportion of 0.125.

Table 1. Pollen phenotypes are not significantly different between *met1-1* mutants and AABBDD plants. The estimated marginal means for each phenotype are shown for each genotype category. Shared letters show genotypes that are not significantly different from one another (FDR adjusted p<0.05).

Genotype Category	Pollen Grains per Anther	Modal Pollen Diameter (µm)	Non-Viable Pollen Grains (%)	Non- and Mononucleated Pollen Grains (%)
Cadenza WT	2,459 (a)	49.8 (a)	1.05 (a)	8.41 (a)
AABBDD	1,829 (bc)	46.2 (b)	2.62 (b)	12.45 (ab)
aaBBDD/AAbbDD/AABBdd	2,085 (ab)	46.8 (b)	1.96 (b)	-
aabbDD/aaBBdd/AAbbdd	1,623 (c)	45.3 (b)	1.60 (ab)	11.40 (a)
AaBbdd/AabbDd/aaBbDd	1,780 (bc)	45.8 (b)	2.10 (b)	19.56 (b)
N per genotype category	10-15	10-16	16-38	4-15

Loss of four or more functional copies of *MET1-1* results in decreased CG methylation

We carried out whole genome bisulfite sequencing to determine whether a stepwise reduction in the number of functional *MET1-1* alleles was associated with a proportional change in DNA methylation. We found evidence for partial functional redundancy between the *MET1-1* homoeologs; genome-wide CG methylation in the single mutants (aaBBDD, AAbbDD, and AABBdd) was between 87.0% and 87.1% which was very similar to the WT segregant

(AABBDD, 87.3% CG methylation) (Figure 3A). Genome-wide CG methylation was reduced in the double mutants, ranging from 74.0% in the aabbDD mutant, to 79.3% in the aaBBdd mutant, and 81.2% in the AAbbdd mutant (Figure 3A). The lowest level of CG methylation was observed in the Aabbdd mutant, with 72.4% (Figure 3A). While methylation in the CG context was reduced in the Aabbdd mutant, it had the highest CHG (57.4%) and CHH (4.1%) methylation out of all the lines tested (Figure 3A). This suggests that CHG and CHH methylation may be increased to compensate for the loss in CG methylation.

To explore the impact of these global changes in methylation, we examined how methylation was altered across genes and transposable elements (TEs) using chromosome 5D as an example. We found that the average CG methylation was decreased across genes and TEs relative to the WT segregant (AABBDD) in both the double mutant (AAbbdd) and the Aabbdd mutant (Fig 3B and C). For both sequence features, the greatest decrease was observed for the Aabbdd mutant, in line with the global changes in methylation. No obvious difference was observed in CHG methylation across genes or TEs, while CHH methylation was highest in the Aabbdd mutant across both genes and TEs (Supplementary Figure 4).

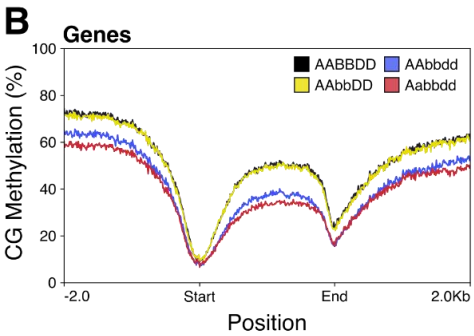
Spatial correlations between the different methylation contexts can also be observed when looking at the distribution of cytosine methylation across each chromosome. Along chromosome 5D, CG methylation is reduced in the Aabbdd mutant compared to the WT segregant, single and double mutants in the R1, R2a, C, and R2b regions. In the R3 region, however, CG methylation in the Aabbdd mutant is comparable to the AAbbdd double mutant (Figure 3D). In this region, there is also an increase in CHG methylation in the Aabbdd mutant, such that it is higher than any of the other lines (Figure 3E), and an increase in CHH methylation (Figure 3F). There is a clear transition point between states in the Aabbdd mutant across all three methylation contexts (Supplementary Figure 5). This could be due to targeting of the *de novo* DNA methylation machinery to the R3 region of chromosome 5D in the Aabbdd mutant, resulting in increased DNA methylation in all sequence contexts. The positive correlation between CG and CHG/CHH methylation is also observed in specific regions of another ten chromosomes, affecting all possible chromosome regions not only R3, and across the whole of chromosomes 3A, 4D and 6B (Supplementary Figures 6-12). These region-specific increases in CG, CHG and CHH

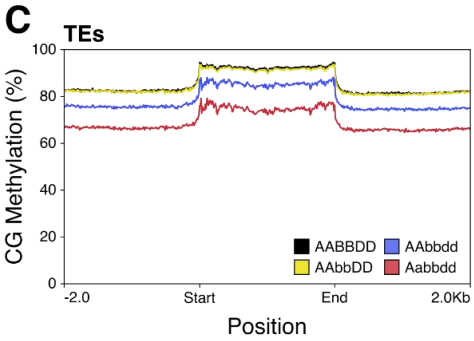
methylation were only observed in the Aabbdd mutant, and not in any of the single or double mutants, as shown for chromosome 5D (Supplementary Figure 13).

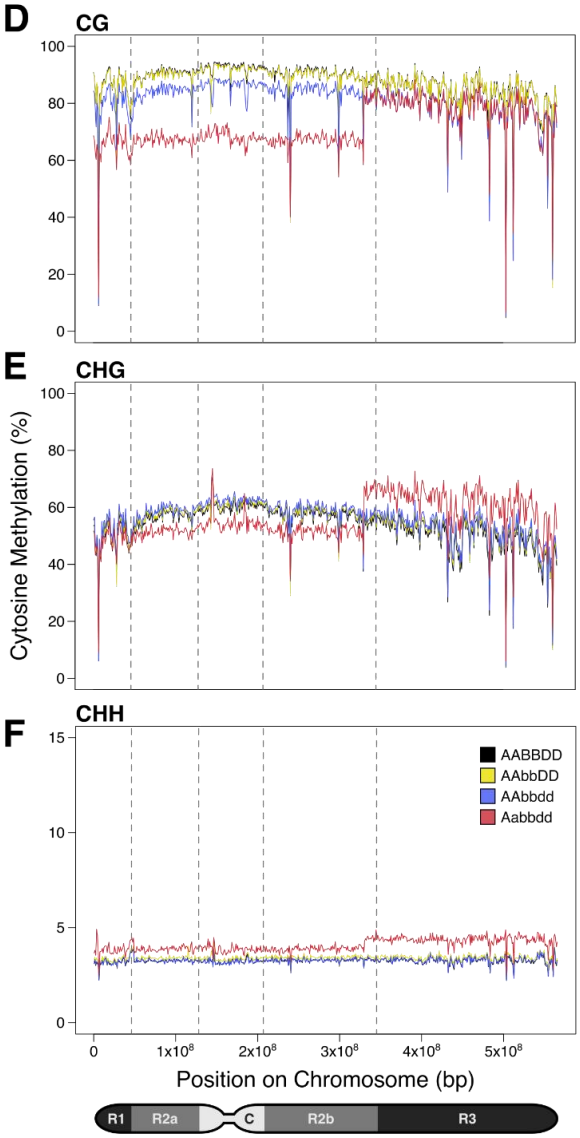
To further understand the spatial nature of changes in methylation in the *met1-1* mutants, we identified 100 bp differentially methylation regions (DMRs) in the CG context in each of the mutants relative to the WT segregant. Substantially more DMRs were detected in the Aabbdd mutant (1,156,551 DMRs), compared to between 98,626 and 282,113 DMRs in the double mutants, and between 10,616 and 11,803 DMRs in the single mutants (Figure 3G). The vast majority of the DMRs were hypo-methylated in the *met1-1* mutants – only 0.27% of the DMRs in the Aabbdd mutant were hyper-methylated (Figure 3G). We calculated the density of the hypo-methylated DMRs in each of the chromosome regions and observed that, in the double mutants, the greatest density of DMRs is found in the distal regions R1 and R3, and the lowest in the centromeric region C (Figure 3H). However, in the Aabbdd mutant the centromeric region has the highest density of DMRs (Figure 3H). In all of the *met1-1* mutants, the density of hyper-methylated DMRs is lowest in the proximal and centromeric regions, and highest in the distal regions (Supplementary Figure 14). As the proximal and centromeric regions have the highest transposable element (TE) density (Daron *et al.*, 2014; Wicker *et al.*, 2018), loss of CG methylation in the centromere could cause significant activation of TEs. This suggests that the severe developmental phenotype observed in the Aabbdd mutant could be due to greater loss of CG methylation from the centromeric regions and subsequent activation of TEs.

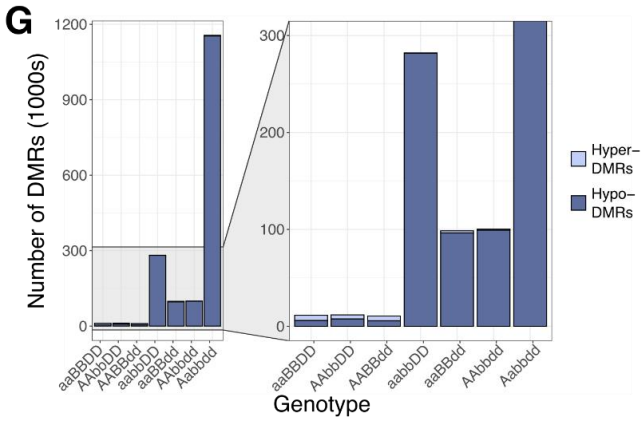
A

	Cytosine Methylation (%)		
	CG	**CHG**	**CHH**
Cad WT	88.7	57.1	3.4
AABBDD	87.3	54.4	3.3
aaBBDD	87.0	56.3	3.5
AAbbDD	87.0	55.4	3.5
AABBdd	87.1	56.4	3.6
aabbDD	74.0	55.7	3.2
aaBBdd	79.3	56.3	3.5
AAbbdd	81.2	56.6	3.3
Aabbdd	72.4	57.4	4.1









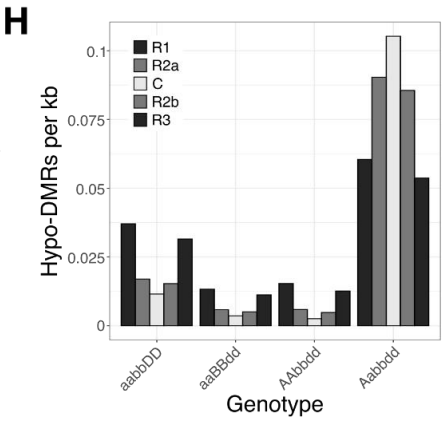


Figure 3: Reduced number of functional copies of *MET1-1* results in loss of CG methylation and the formation of differentially methylated regions (DMRs). A) Genome-wide percentage of cytosines methylated in the CG, CHG, and CHH contexts. B-C) Percentage CG methylation

across genes (B) and transposable elements (TEs) (C) on chromosome 5D, including 2 kb up and downstream of the feature. The WT segregant (AABBDD, black), AAbbDD single mutant (yellow), AAbbdd double mutant (blue), and Aabbdd mutant (red) are shown. D-F) Percentage of methylated cytosines in the CG (D), CHG (E), and CHH (F) contexts across chromosome 5D, calculated as an average in 1 Mb bins across the chromosome. The WT segregant (AABBDD, black), AAbbDD single mutant (yellow), AAbbdd double mutant (blue), and Aabbdd mutant (red) are shown. Partitions between the chromosome regions are indicated by the dashed lines. G) Number of differentially methylated regions (DMRs) in 1000s relative to AABBDD in each genotype. Both hypo-methylated (dark dark) and hyper-methylated (light blue) regions are shown. H) Number of hypo-methylated DMRs per kb in each of the chromosome regions – the distal regions R1 and R3 are shown in dark grey, the proximal regions R2a and R2b in medium grey, and the centromeric region C in light grey.

Reduced function of *MET1-1* results in stepwise changes in gene expression

To determine the effect of these changes in methylation, we conducted RNA-sequencing and differential expression analysis for the WT segregant, each of the single and double mutants, and the Aabbdd mutant. The number of differentially expressed genes (DEGs) compared to the WT segregant (FDR adjusted $p<0.01$, fold change ≥ 2) was calculated for each of the mutant classes (Figure 4A). We found that the Aabbdd mutant had the largest number of DEGs (5,904), while the double mutants had an average of 79 DEGs, and the single mutants had an average of 9 DEGs. Generally, the number of up-regulated DEGs was higher than the number of down-regulated, except for the Aabbdd mutant, in which 3,900 genes were down-regulated and 2,004 genes were up-regulated (Figure 4A). These changes in gene expression are in line with the global changes in CG methylation, with those mutants maintaining the highest level of CG methylation relative to the WT segregant having the fewest DEGs, and vice versa (Figure 4B).

To further examine the transcriptional response in the *met1-1* mutants, we identified genes that were differentially expressed in both the Aabbdd mutant and at least one of the double mutants. A total of 86 DEGs were identified in this way, 76 of which were up-regulated, and 10 of which were down-regulated. Several of these DEGs have putative functions relating to DNA methylation, other epigenetic marks, and gene silencing (Figure 4C), suggesting that reduced function of *MET1-1* initiates active changes in transcriptional regulation. The gene *TraesCS6B02G315200* is up-regulated in all of the double mutants and the Aabbdd mutant (adjusted $p<0.01$) (Figure 4C), and is annotated in the Panther Classification System as RNA Polymerase IV and V Subunit 4 (PTHR21297:SF2). The homoeologous gene

TraesCS6D02G266900 is also up-regulated in all of the double mutants and the Aabbdd mutant, while its other homoeolog *TraesCS6A02G286200* and tandem duplicated gene copy *TraesCS6B02G15300*, are up-regulated only in the aabbDD and Aabbdd mutants. A gene with the same Panther classification in Arabidopsis, *NRPD4* (*AT4G15950*), is required for RNA-directed DNA methylation and loss-of-function of this gene results in loss of DNA methylation in all sequence contexts (He *et al.*, 2009). We also found that the gene *TraesCS2D02G114600* was up-regulated in all of the double *met1-1* mutants (Figure 4C). This gene encodes a small RNA 2'-O-methyltransferase annotated as HEN1 in the Panther Classification System (PTHR21404). In Arabidopsis, *HEN1* has also been implicated in RNA-directed DNA methylation as methylation by HEN1 stabilises small RNAs produced by RNA Polymerase IV/V (Erdmann & Picard, 2020). Knock-out mutations in *HEN1* were found to relieve silencing conferred by DNA methylation (He *et al.*, 2009). Furthermore, in both the aabbDD and Aabbdd mutants, *TraesCS3B02G023200*, a DEMETER (DME)-like DNA glycosylase, is significantly down-regulated (Figure 4C). In Arabidopsis, loss-of-function mutations in the DME-like DNA glycosylases *DML2* and *DML3* result in increased cytosine methylation in regions which have low to intermediate methylation levels in wild type (Ortega-Galisteo *et al.*, 2008). Overall, this suggests that reduced function of *MET1-1* is buffered by up-regulation of genes that facilitate DNA methylation and down-regulation of genes that remove DNA methylation.

We also observed up-regulation of *TraesCS2B02G503200* in the Aabbdd, aabbDD and AAbbdd mutants (adjusted $p<0.01$) and the aaBBdd mutants (adjusted $p<0.05$) (Figure 4C). This gene is predicted to contain a histone-lysine N-methyltransferase SUVR4/SUVR1/SUVR2 domain (IPR025776). This suggests that reduced function of *MET1-1* may promote changes in histone methylation that could further alter transcription.

To gain additional insight into the biological processes that are affected by the loss-of-function mutations in *MET1-1*, we performed gene ontology (GO) enrichment analysis for the DEGs in the Aabbdd mutant (Figure 4D and E). We found that the up-regulated genes were enriched for GO terms related to translation (such as ribosome biogenesis and translational elongation) and stress responses (such as response to high light intensity, hydrogen peroxide, and heat) (Figure 4D). The down-regulated genes were enriched in GO terms related to photosynthesis and other metabolic processes (such as Photosystem II assembly, pentose-phosphate shunt, and starch

biosynthesis) (Figure 4E). These results indicate that multiple core pathways are dysregulated in the Aabbdd mutant which likely contributed to its death before tillering.

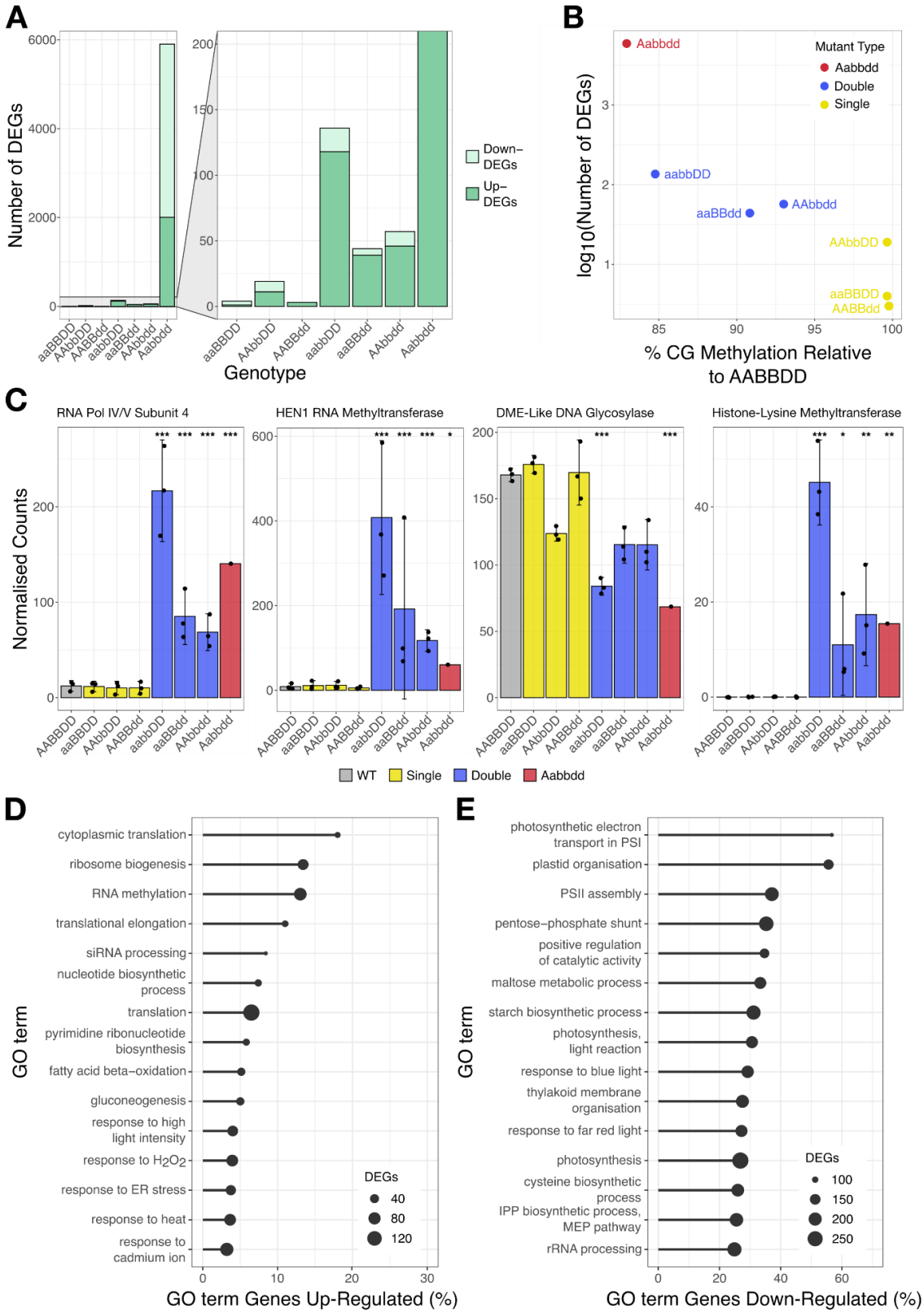


Figure 4. Reduced number of functional copies of *MET1-1* results in transcriptional reprogramming. A) Number of differentially expressed genes (DEGs) relative to AABBDD in

each genotype. Both down-regulated (pale green) and up-regulated (dark green) genes are shown. B) The $\log_{10}$(Number of DEGs) against the percentage CG methylation relative to AABBDD, for each of the single (yellow) and double (blue) mutants and the Aabbdd (red) mutant. C) Normalised transcript counts in each genotype for genes putatively encoding an RNA Polymerase IV/V subunit (*TraesCS6B02G315200*), a HEN1 small RNA 2'-O-methyltransferase (*TraesCS2D02G114600*), a DEMETER (DME)-like DNA glycosylase (*TraesCS3B02G023200*), and a histone-lysine methyltransferase (*TraesCS3B02G503200*). Error bars represent the 95% confidence interval, estimated as 1.96 x standard error. Asterisks show adjusted p-values for differential expression relative to AABBDD; * $p<0.05$, ** $p<0.01$, *** $p<0.001$. D-E) Significantly enriched GO terms (adjusted $p<0.05$) amongst the significantly (adjusted $p<0.01$) up-regulated (D) and down-regulated (E) DEGs in the Aabbdd mutant. The 15 most significant GO terms encompassing at least 15 DEGs are shown in each case.

Next, we investigated whether there was overlap between the CG-DMRs identified in the *met1-1* mutants (Figure 3) and genes and TEs. We found that up to 35.7% of high-confidence genes (IWGSC *et al.*, 2018) overlapped with CG-DMRs in the *met1-1* mutants (either within the gene body, or within 1 kb up- or downstream) (Figure 5A). Although the Aabbdd mutant had between 4.1x and 11.7x as many CG-DMRs as the double mutants (Figure 3G), the percentage of genes intersected by at least one CG-DMR (26.0%) was within the range observed in the double mutants (between 17.9% and 35.7%). In line with the overall distribution of DMRs, the majority of DMRs intersecting genes were hypo-methylated in the *met1-1* mutants (Figure 5A).

In contrast to the DMRs intersecting genes, we observed a stark difference between the percentage of TEs overlapped by at least one CG-DMR in the Aabbdd mutant and the double mutants (Figure 5B). In the Aabbdd mutant, 20.1% of non-fragmented TEs (Wicker *et al.*, 2018) were intersected by at least one DMR, compared to between 0.95% and 4.7% in the double mutants. This is in line with the greater density of hypo-DMRs in the transposon-dense centromeric region in the Aabbdd mutant (Figure 3H). 87.0% of the hypomethylated TEs in the Aabbdd mutant were long terminal repeat retrotransposons (LTR-RTs), compared to between 56.5% and 72.5% in the double mutants (Supplementary Table 4).

To examine whether changes in methylation status may be directly related to changes in gene expression, we quantified the percentage of DEGs which overlapped with at least one CG-DMR. For this analysis, we focused on the double and Aabbdd mutants, as only a small number of DEGs were identified in the single mutants (Figure 4A). We found that between 32.6% and

51.5% of DEGs were associated with a DMR (Figure 5C). For each mutant, the proportion of DEGs associated with a DMR is greater than the proportion of genes associated with a DMR, suggesting that changes in DNA methylation may be directly responsible for the change in expression of a subset of DEGs. We next examined the relationship between methylation status and the direction of differential expression. In the double mutants, we found that, on average, 79.7% of hypomethylated DEGs increased in expression relative to the WT segregant (Figure 5D). However, in the Aabbdd mutant, this trend was reversed, with only 26.9% of hypomethylated DEGs being up-regulated (Figure 5D). This could reflect the overall greater proportion of down-regulated genes identified in the Aabbdd mutant compared to the double mutants (Figure 4A).

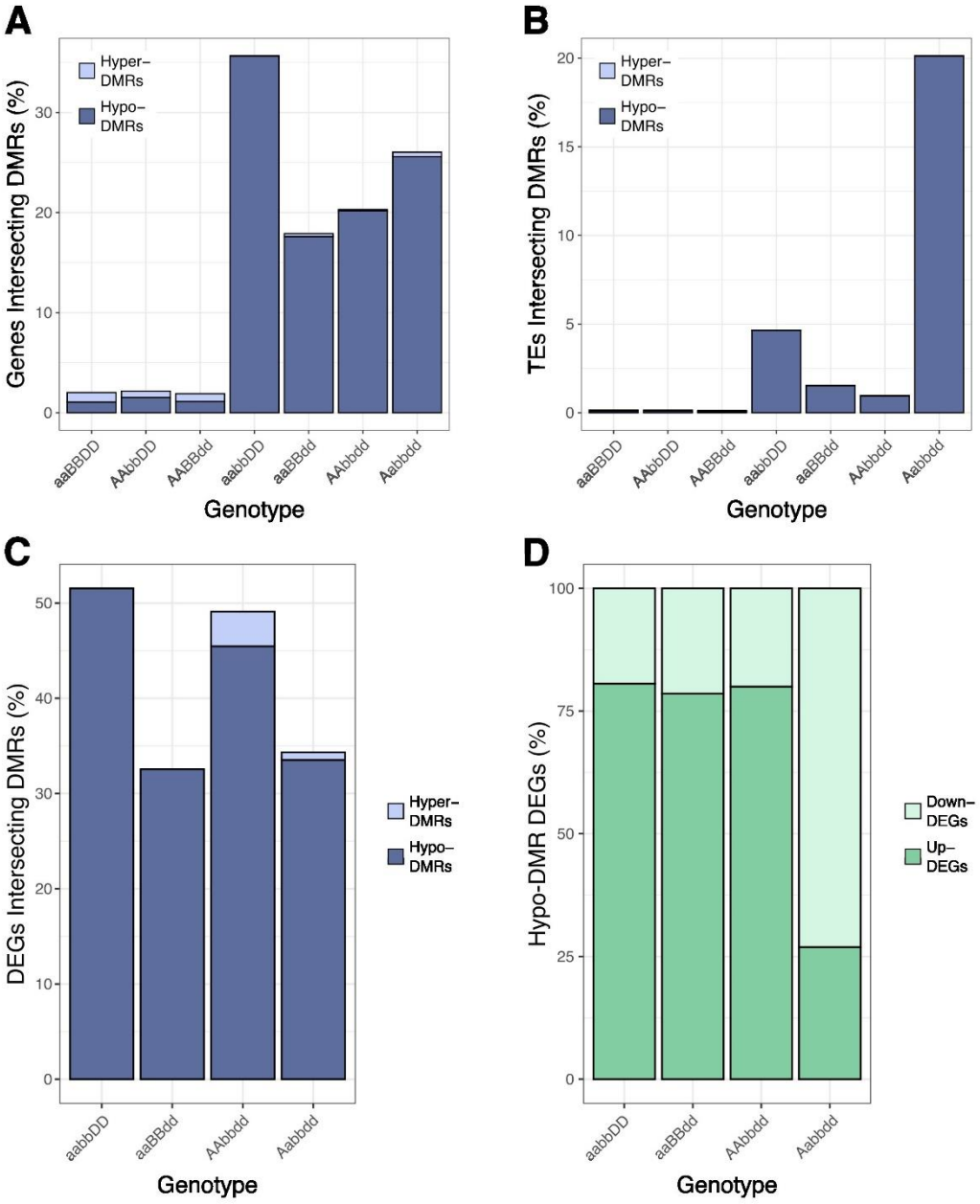


Figure 5. Differentially methylated regions (DMRs) are found associated with genes and transposable elements (TEs). A-B) The percentage of genes (A) or transposable elements (B) which intersect at least one DMR in each of the *met1-1* mutants. Hyper-DMRs are shown in light blue and hypo-DMRs are shown in dark blue. C) The percentage of differentially expressed genes (DEGs) which intersect at least one DMR in the double and Aabbdd *met1-1* mutants. Hyper-DMRs are shown in light blue and hypo-DMRs are shown in dark blue. D) The percentage of DEGs associated with a hypo-DMR which are down-regulated (light green) or up-regulated (dark green) in the double and Aabbdd *met1-1* mutants.

Heritable phenotypic changes were identified in the *MET1-1* segregating F_2 population

When growing the C0451xC1292xC2028 F_2 population (Supplementary Figure 2) we identified an individual plant, with an AaBbdd genotype, that flowered 23 days later than any other plant in the population. This plant showed a striking phenotype – the plant was taller, had lax ears (thin ears composed of many spikelets) and took longer to reach maturity compared to Cadenza WT and sibling plants. Offspring from this late-flowering individual segregated into normal types which matured similar to Cadenza WT and had normal ears, and late-flowering types which had lax ears (Figure 6A and B). We grouped the offspring into normal (<10 days relative to earliest flowering plant) and late (≥10 days relative to earliest flowering plant) flowering categories, although all offspring flowered later than Cadenza WT (Figure 6C). We observed a difference between median flowering date for the late and normal flowering plants of 27 days. The numbers of segregant plants in the late and normal flowering categories did not significantly differ from a 3:1 ratio ($\chi^2 = 0.214$, p-value = 0.644), suggesting that a single dominant gene controls the phenotype. The late flowering phenotype was independent of *MET1-1* genotype (Figure 6D). The mean height of late flowering F_3 plants was 21.1 cm taller than Cadenza WT (FDR adjusted p-value = 2.5×10^{-14}) and 26.2 cm taller than normal flowering F_3 plants (FDR adjusted p-value = 8.2×10^{-15}; Figure 6A, E). Although visually the late flowering segregants frequently appeared to have longer spikes, there was no significant difference in spike length between the late and normal flowering segregants (FDR adjusted p-value = 0.22, Figure 6F). Despite having a similar spike length, late flowering segregant plants had a higher number of fully formed spikelets per spike than normal flowering segregants (mean 25.3 vs 18.6; FDR adjusted p-value = 5.5×10^{-9}, Figure 6B and G). These data show that the observed late flowering phenotype is heritable and co-segregates with increased plant height and increased fully formed spikelets per spike. Moreover, the phenotype is independent of *MET1-1* genotype, suggesting that it is controlled by a single stable allele generated in the mutant background.

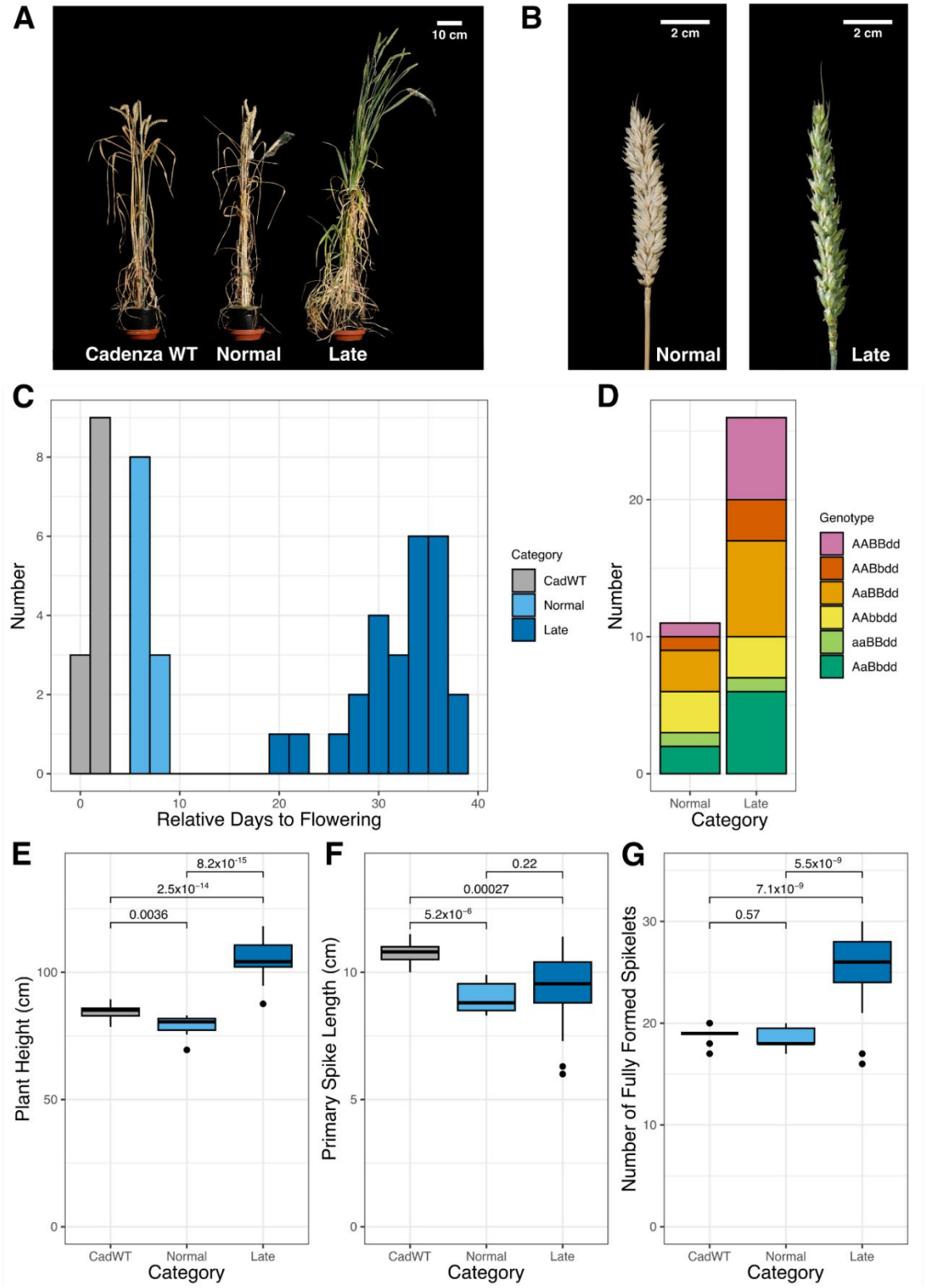


Figure 6. The late flowering phenotype in the C0451xC1292xC2028 *MET1-1* population was independent of *MET1-1* genotype and was accompanied by increased plant height and number of fully formed spikelets. A) Phenotypes of Cadenza WT and normal and late flowering F_3

segregant plants. B) Normal and late flowering segregant plants showing the increased number of fully formed spikelets in the late flowering plant. C) Histogram showing the days to flowering (full emergence of primary spike) of the different flowering time categories relative to the first plant. D) *MET1-1* genotypes of late and normal flowering plants. Boxplots of plant height (E), primary spike length (F) and number of fully formed spikelets (G) for each of the phenotype categories. For (E) to (G), the means were compared by a t-test and the False Discovery Rate (FDR) adjusted p-value calculated is shown.

A distinct second phenotype was observed in the same segregating F_2 population (C0451xC1292xC2028): a single plant, with an AAbbdd genotype, that exhibited an awned phenotype. As Cadenza WT and each of the TILLING lines used to generate the population were awnless this observation was unexpected. To investigate the phenotype further we generated a segregating population by back crossing the awned plant with Cadenza WT. We grew the F_1 generation, which showed an intermediate (awnletted) phenotype, and allowed the plant to self-pollinate. We harvested and sowed the segregating F_2 seed and grew 215 plants to maturity to score the awn phenotype. We found that the segregation ratio of awnless:awnletted:awned individuals did not significantly deviate from a single co-dominant gene 1:2:1 Mendelian ratio ($\chi^2 = 1.13$, p-value = 0.568) (Figure 7A and B). The awnletted individual had longer awns at the top of the ear which was reminiscent of the effect of the *Tipped1* (*B1*) awn suppressor gene (Huang *et al.*, 2020) and we hypothesised that this was the causal gene controlling the phenotype. To test this hypothesis, we used published primer sequences to screen *B1* and the surrounding region in our awned population (Figure 7C) (Huang *et al.*, 2020). Initially we attempted to amplify the entire *B1* gene by PCR and sequence the product for polymorphisms, however, no product could be amplified for any awned plants (Figure 7D). This result suggested that there was either a polymorphism affecting primer binding, a large insertion, or a deletion of the *B1* gene preventing amplification. To investigate the wider *B1* region we used a panel of flanking KASP markers to amplify the region surrounding *B1*, spanning a total distance of 28.5 Mbp. We found that all six of the KASP markers flanking *B1* amplified for the awnless and awnletted plants, but no amplification was achieved for the awned plants (Figure 7C). Together, these results indicate that the presence of awns in our population is due to a heritable deletion >28.5Mbp which includes *B1*.

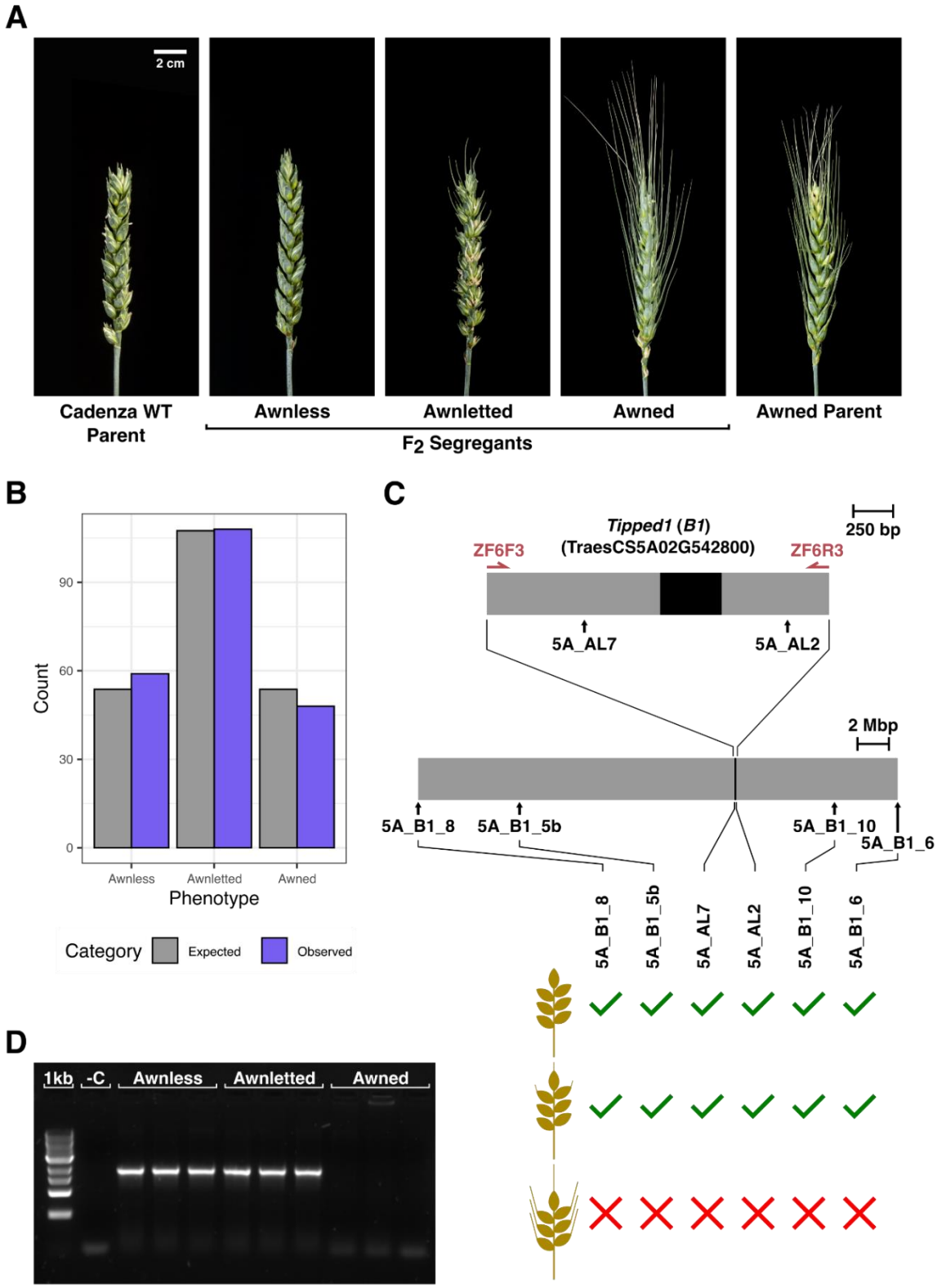


Figure 7. The observed awned phenotype in the C0451xC1292xC2028 *MET1-1* population followed a single gene semi dominant mendelian inheritance pattern, which was caused by a

deletion encompassing *Tipped1* (*B1*). A) Phenotypes of the awnless Cadenza WT parent and spontaneous awned individual which were crossed together to generate the F_1 generation. The F_1 was self-pollinated to generate the segregating F_2 population which displayed awnless, awnletted and awned phenotypes. B) Expected (grey) and observed (purple) number of individuals which were awnless, awnletted and awned in the segregating F_2 population. C) Diagram of the *B1* gene (black) and surrounding region (grey) with KASP marker binding sites (black arrows) and PCR primer binding sites (maroon half arrows). Schematic summarising successful (green tick) and unsuccessful (red cross) amplification of KASP marker sets for each phenotype class of plant (awned, awnletted and awnless). D) Agarose gel image showing PCR product of the ZF6F3 and ZF6R3 primer set (expected product ≈2kb) from 3 individual F_2 segregant plants of each phenotype class and 1kb DNA ladder (1kb) and no DNA (-C) controls.

Discussion

Complete loss of *MET1-1* is lethal in wheat, but single and double mutants can be recovered

Despite genotyping over 3,000 progeny from two independent segregating *MET1-1* TILLING populations, we were unable to recover any null *met1-1* mutants in hexaploid wheat, in which all three homoeologs are interrupted by a premature termination codon mutation (Figure 1F, Supplementary Table 1). Similarly, we were unable to recover any null *met1-1* mutants in tetraploid wheat, in which both the A and B homoeologs are interrupted by a premature termination codon mutation (Supplementary Table 3). This indicates that complete loss of *MET1-1* is lethal in wheat, consistent with other monocotyledonous crop species with large genomes such as maize and barley (Li *et al.*, 2014; Schreiber *et al.*, 2019).

However, unlike maize and barley (both diploids), wheat is a polyploid, allowing intermediate mutants to be recovered. We were able to recover all possible homozygous single and double *met1-1* mutants from the segregating *MET1-1* populations (Figure 1F, Supplementary Table 1) suggesting that the function of *MET1-1* homoeologs in wheat is mostly redundant. We also recovered a single Aabbdd mutant, in which five out of six copies of *MET1-1* are non-functional (Figure 1E). While the single mutants had no decrease in CG methylation compared to the WT segregant, the double and Aabbdd mutants showed altered patterns of CG methylation and gene expression (Figure 3, Figure 4). The double mutant classes showed variability in the degree of

CG methylation loss and changes to gene expression, with the aabbDD mutants showing the greatest change compared to the AAbbdd and aaBBdd double mutants (Figure 3, Figure 4). This suggests that *MET1-D1* is less effective at compensating for the loss of the other two homoeologs than *MET1-A1* or *MET1-B1*. Furthermore, the double mutants did not show any growth defects, despite the aabbDD mutants having a similar degree of demethylation to the Aabbdd mutant (74.0% CG methylation in the aabbDD mutant compared to 72.4% CG methylation in the Aabbdd mutant).The proportion of genes associated with at least one differentially methylated region (DMR) was comparable between the double mutants (average of 24.6% of genes) and the Aabbdd mutant (26.0% of genes). Together these features suggests that double mutants with four mutant alleles may allow the generation of epialleles which could be utilised in epigenetic breeding (Figure 8). Double mutants and the Aabbdd mutant also had significant numbers of DMRs in centromeric and pericentromeric regions (Figure 3). These DMRs may increase variation in gene expression in these regions which could be useful for plant breeding because these regions suffer from low recombination which limits alterations to gene expression through genetic changes.

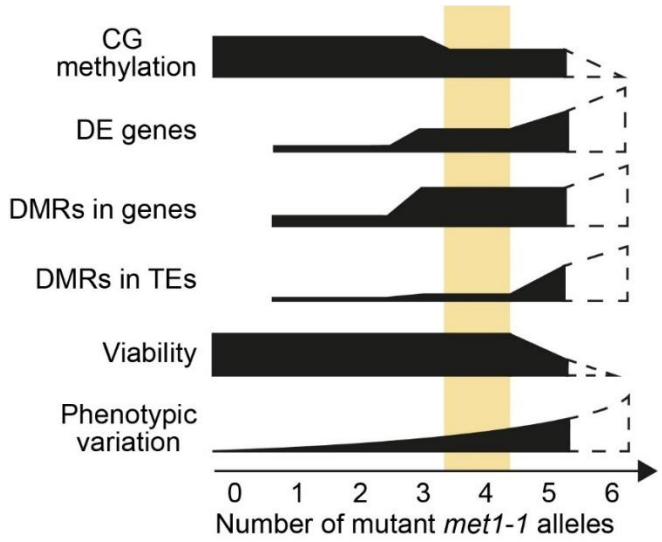


Figure 8. Increasing the number of mutant *met1-1* alleles in hexaploid wheat has quantitative yet variable effects on CG methylation, differentially expressed (DE) genes, differentially methylated regions (DMRs) in genes and in transposable elements (TEs), viability and phenotypic variation. Dashed regions indicate hypothetical data for a complete null mutant. The region highlighted in yellow represents an optimum where CG methylation is substantially reduced and viability is retained, made possible by partial redundancy between homoeologs.

Indeed, we observed novel, heritable phenotypic variation in one of the segregating *MET1-1* populations (Figure 6, Figure 7). These phenotypes (delayed flowering and the development of awns) were both identified in plants that had four out of six non-functional copies of *MET1-1*, indicating that a null *met1-1* mutant is not necessary to generate phenotypic variation. Loss of DNA methylation can generate novel variation in multiple ways. Firstly, epialleles may be generated resulting in altered patterns of gene expression. Secondly, hypomethylation of transposable elements (TEs) is associated with their transcription and remobilisation, leading to abnormal developmental phenotypes (Miura *et al.*, 2001; Mirouze *et al.*, 2009; Reinders *et al.*, 2009; Tsukahara *et al.*, 2009). We found that one fifth of TEs were associated with at least one DMR in the Aabbdd mutant, potentially resulting in their activation. This may explain the stunted growth of the Aabbdd mutant, and the severely reduced viability of higher-order *met1-1* mutants. The genome instability caused by loss of DNA methylation can also cause structural variation (Zhang *et al.*, 2023), and we found that the development of awns in one of the *met1-1* mutants was due to a deletion greater than 28.5 Mb (Figure 7). A similar large deletion was found in wheat treated with the demethylation agent zebularine – this deletion encompassed *FT-B1* and was associated with an increase in spikelet number and a delay in flowering (Finnegan *et al.*, 2019).

Lethal effect of null *met1-1* mutants associated with mutant (abd) gametes

We observed a clear segregation distortion when genotyping selfed seed from a triple heterozygous (AaBbDd) *MET1-1* parent. There was a clear underrepresentation of higher-level mutants (5/6 and 6/6 non-functional copies) when genotyping the leaves (Figure 1F). By genotyping the grain, we attributed this distortion to a reduced transmission of gametes with no functional copy of *MET1-1* (abd gametes) when donated maternally or paternally (Figure 2C). The reduced transmission through both male and female gametes was confirmed by reciprocal crosses to wild type (Figure 2E). In the reciprocal crosses, abd gametes transmitted paternally were rarer than abd gametes transmitted maternally, whereas the opposite trend was observed using selfed heterozygous (AaBbDd) mutants. This discrepancy may be explained by post-fertilisation effects of *met1-1* being made more severe in a scenario where mutant alleles may be present in both gametes, as for the selfed heterozygous plants. This is consistent with work in

Arabidopsis that found that selfed heterozygous *met1* mutants had higher rates of embryo abortion than wild type x homozygous *met1* mutant reciprocal crosses (Xiao *et al.*, 2006).

In Arabidopsis, transmission of mutant *met1* alleles has also been shown to be distorted in the gametes. However, in contrast to our results where both maternal and paternal transmission was affected, only the paternal transmission of *met1* was reduced to 36.8% of the expected rate in Arabidopsis while maternal transmission did not differ from expected (Liang *et al.*, 2022). Liang *et al.* (2022) demonstrated that this single parental bias in *met1* transmission was, in part, due to impaired pollen development resulting in abnormal numbers of nuclei in the pollen grains. In our work we found no change to any pollen phenotype measured (pollen grains per anther, modal pollen grain diameter, non-viable pollen grains and non- or mono-nucleated pollen) in the *met1-1* mutant plants compared to WT segregant plants, even in genotypes which we expect to produce *met1-1* abd pollen at a proportion of 25%. This suggests that in hexaploid wheat the reduction in CG methylation in pollen does not affect the formation of phenotypically normal pollen and instead the barrier to the production of null mutants may occur during fertilisation or embryogenesis.

Further differences in the effects of parent on *met1* transmission have been found in Arabidopsis. Maternal donation of mutant *met1* has been shown to produce larger seeds, while paternal donation produced smaller seeds (Xiao *et al.*, 2006). This opposing parental effect was not observed in wheat, though we did find that transmission of mutant *met1-1* copies through abd gametes were associated with smaller seeds irrespective of parent of origin. In rice homozygous *met1* mutants were accompanied by a small and shrivelled seed phenotype and the effect was more extreme than we observed in wheat (Hu *et al.*, 2014; Yamauchi *et al.*, 2014). Alterations to seed size across species may be a result of a dysregulation of gene expression in *met1* mutants following fertilisation, where normally core imprinted genes regulate endosperm proliferation. This has been demonstrated in Arabidopsis where *MET1* is required to silence the paternal allele *FIS2* to ensure normal endosperm development (Jullien *et al.*, 2006).

Reduced function of *MET1-1* is buffered by activation of other epigenetic pathways

We found that loss of five out of six functional copies of *MET1-1* resulted in a significant decrease in CG methylation, but an increase in CHG and CHH methylation (Figure 3),

suggesting that other methylation pathways are activated upon reduction of *MET1-1* activity. A similar increase in CHG and CHH methylation was observed in tomato *met1* mutants (Yang *et al.*, 2019). In rice, on the other hand, increases in CHG and CHH methylation were observed in the *MET1/met1* heterozygote, but not the homozygous *met1* mutant (Hu *et al.*, 2014). In Arabidopsis, CHG methylation was found to increase in *met1* mutants, while CHH methylation was decreased (Srikant *et al.*, 2022). In wheat, we found that these increases in CHG and CHH methylation are found in extremely well-defined regions where CG methylation is maintained to at least the level of the AAbbdd double mutant (Figure 3, Supplementary Figures 5-12). This could be due to increased activity of the *de novo* methylation machinery in these regions. Future studies will be required to determine whether these putative *de novo* methylated regions are stochastic or reproducible and whether they are a feature only present in species with large genomes and thus have been undetectable until now.

In plants, *de novo* DNA methylation is facilitated by the RNA-directed DNA methylation (RdDM) pathway and acts to methylate cytosines in all sequence contexts via the DRM2 methyltransferase (Zhang *et al.*, 2018).The compensatory increase in CHG and CHH methylation we observed may be promoted by transcriptional up-regulation of genes involved in the RdDM pathway. We observed increased expression of multiple homoeologs encoding RNA Polymerase IV/V Subunit 4 and increased expression of the HEN1 RNA methyltransferase in the double and Aabbdd *met1-1* mutants (Figure 4C). These proteins are both involved in the generation of the short interfering RNAs (siRNAs) which direct the RdDM machinery to the loci to be methylated and both are required for RdDM to function in Arabidopsis (He *et al.*, 2009). We also observed down-regulation of a DME-like DNA glycosylase in the mutants with the largest reduction in CG methylation (aabbDD and Aabbdd) (Figure 4C). Transcriptional down-regulation of genes encoding DNA glycosylases has also been observed in *met1* mutants in Arabidopsis and rice (Mathieu *et al.*, 2007; Hu *et al.*, 2014). In Arabidopsis, transcriptional down-regulation of the *ROS1* DNA glycosylase is caused by loss of methylation at a 39 bp region in its promoter known as the methylation monitoring sequence (MEMS) (Lei *et al.*, 2015). We found that the DME-like DNA glycosylase down-regulated in the aabbDD and Aabbdd *met1-1* mutants overlapped a hypo-DMR in all of the double and Aabbdd mutants. This suggests that a similar MEMS mechanism may regulate the expression of this DNA glycosylase, allowing the plant to sense changes in the DNA methylation status and compensate for them.

Compensation for a reduction in CG methylation can also occur through changes in histone modifications. In Arabidopsis, alternations to the histone epigenetic landscape occurs in stages – in the first generation of homozygous *met1* mutants H3K9 dimethylation is decreased, while in later generations H3K9 dimethylation is increased above WT levels (Tariq *et al.*, 2003; Mathieu *et al.*, 2007). We observed transcriptional up-regulation of a SUVR4/SUVR1/SUVR2 histone lysine methyltransferase in the double and Aabbdd *met1-1* mutants (Figure 4C). SUVR4/SUVR1/SUVR2 histone lysine methyltransferases have a preference for H3K9 methylation (Thorstensen *et al.*, 2006), suggesting that this epigenetic mark may also be altered upon reduction of CG methylation in wheat.

In summary, we leveraged partial redundancy between homoeologs in polyploid wheat to recover *met1-1* mutants. Similar to other plants with large genomes, complete null mutants were lethal. However, wheat plants with four null alleles (i.e. retaining two functional alleles) develop normally, have substantially reduced CG methylation and occasionally exhibit novel heritable phenotypes. Generating partial mutants may also be effective in other polyploid species to manipulate CG methylation without lethal consequences, opening new avenues to understand the role of CG methylation in polyploid species and generate epialleles for crop breeding.

Materials and Methods

Plant materials and growth

Exome capture sequenced EMS mutagenized populations of hexaploid bread wheat (*Triticum aestivum* cv. Cadenza) and tetraploid pasta wheat (*Triticum turgidum* cv. Kronos) were screened for mutations in *MET1-1* (*TraesCS2A02G235900*, *TraesCS2B02G260800* and *TraesCS2D02G241800*) (Krasileva *et al.*, 2017). We selected 5 Cadenza lines and 2 Kronos lines, all containing mutations which generated premature stop codons (Figure 1C). These were predicted to disrupt the MET1-1 protein and eliminate C5 Cytosine methyltransferase domain activity.

Two independent *MET1-1* populations were developed in Cadenza. The first population was developed by crossing the double *met1-1* mutant line Cadenza0465 with Cadenza0884 to generate a triple heterozygous F_1. This was then self-pollinated to produce a segregating F_2

population (Fig 1D). This population is referred to as C0465xC0884. The second was developed by initially crossing Cadenza0451 with Cadenza2028 and Cadenza2028 with Cadenza1292. The resulting F_1 progeny were crossed together to produce a triple heterozygous F_1 which was self-pollinated to generate a segregating F_2 population (Supplementary Figure 2). This population is referred to as C0451xC2028xC1292.

When growing segregating F_2 plants from C0451xC2028xC1292 we observed an individual plant that was awned. To investigate this unexpected phenotype further we developed a segregating population by performing a backcross to Cadenza WT. The resulting BC_1F_1 plants were grown and allowed to self-pollinate. We harvested the segregating BC_1F_2 seeds and grew these for phenotyping and further investigation of the awned trait.

When growing C0451xC2028xC1292 we also observed a plant which was delayed in its flowering. To investigate this phenotype further we allowed the plant to self-pollinate and grew the F_3 seeds for phenotyping.

The Kronos *MET1-1* population was produced by crossing Kronos3085 with Kronos0809, giving double heterozygous (AaBb) F_1 progeny. These were then self-pollinated to generate a segregating F_2 population. From this segregating F_2 population the seeds from double heterozygous plants were harvested and sown to give a segregating F_3 generation. F_4 seeds from double heterozygous F_3 plants were harvested and used for seed genotyping (Supplementary Figure 3).

Seeds were germinated in petri dishes at 4°C for 48 hours, followed by room temperature for 48 hours before sowing into 96-cell trays containing John Innes F_2 starter + Grit (90% peat, 10% grit, 4kg m^{-3} dolomitic limestone, 1.2kg m^{-3} osmocote start). Once plants were leaf sampled and genotyped, selected plants were potted on into 1L pots containing John Innes Cereal Mix (65% peat, 25% loam, 10% grit, 3kg m^{-3} dolomitic limestone, 1.3kg m^{-3} pg mix, 3kg m^{-3} osmocote exact). Plants were grown in the John Innes Centre glasshouses (Norwich, UK) with supplementary lighting as required for minimum 16h light and supplementary heating to 20°C day and 12°C night.

Genotyping

Leaf

Once plants reached the second leaf stage (Zadoks growth stage 12), 2cm leaf samples were harvested into 96-well deep well DNA plates and genomic DNA was extracted following the protocol from www.wheat-training.com adapted from Pallotta *et al.* (2003). KASP markers were used to genotype plant populations (Supplementary Table 5). Genotyping was carried out in 384 well format with 5µl final reaction volume (2.5µl DNA ≈10ng/µl, 2.43µl PACE mix and 0.07µl primer mix). Standard cycling conditions for hot start PACE genotyping were used for 30-40 cycles (94°C for 15:00, 10 cycles of 94°C 0:20, 65°C-57°C 1:00 dropping 0.8°C per cycle, 30-40 cycles of 94°C 0:20, 57°C 1:00, ending with 10°C hold). Data analyses were carried out manually using Kluster-Caller (v3.4.1.36). In the case of the KASP primer mix 'Cad1292' we found that it could not distinguish between heterozygous and homozygous mutant clusters. Consequently, we developed a CAPS marker using standard PCR primers (Cad1292_F and Cad1292_R) which was amplified using standard Taq DNA Polymerase (New England Biolabs, 94°C for 2:00, 40 cycles of 94°C for 0:15, 50°C for 0:25, 72°C for 1:00 ending with 2:30 final extension at 72°C). The resulting product was digested by the restriction enzyme PstI-HF (New England Biolabs, 37°C, 1 hour). The banding pattern was visualised by running the product through a 96 well agarose gel where homozygous *MET1-1* plants produced 3 bands (94bp, 157bp, 249bp) and homozygous *met1-1* mutants produced 2 bands (94bp and 406bp).

Grain

Grain genotyping was used to determine the parental donation of *MET1-1* alleles. Wheat grains were surface sterilised and while working in a flow hood they were cut transversely. The half of the seed containing the embryo was germinated and the half of the grain containing only endosperm was used for genotyping. Full details can be found at dx.doi.org/10.17504/protocols.io.6qpvr84xplmk/v1. Parental donation of the *met1-1* allele was determined using KASP genotyping, where the usual single heterozygous cluster was split into two heterozygous clusters corresponding to whether the mutant *met1-1* allele donation was from the maternal or the paternal parent. The method was validated using reciprocal crosses between *Rht-B13b* (semi-dwarf) and wild type plants using KASP markers from (Borrill *et al.*, 2022).

Phenotyping

Grain size measurements

To investigate the relationship between seed size and *met1-1* inheritance we graded grains into size categories using 3D printed sieves with apertures ranging from 1.0mm – 3.0mm increasing by 0.5mm increments The grain sieve design is available at https://www.printables.com/model/890045-grain-sieve. Following seed size grading grain genotyping was carried out. The effect of total number of mutant *met1-1* alleles and inheritance of an *abd* gamete on grain size was tested using an ordered logistic regression model with the MASS package in R (v4.2.2).

Pollen Assays

Plants used for the pollen assays were grown in a randomised block design. The total number of pollen grains per anther and pollen grain diameter were analysed using the Multisizer 4e Coulter Counter (Beckman Coulter, California, USA). For each plant, three non-extruded, mature anthers were collected from three florets on the secondary spike. The anthers were stored in 70% ethanol at 4°C. Pollen grains were released and analysed as previously described (Alabdullah *et al.*, 2021). Two measurements were performed for each floret sample; one measuring all pollen grains in a total volume of 2 mL, and another measuring up to 10,000 particles. The first was used to estimate the total number of pollen grains per anther. For each plant, the mean number of pollen grains per anther from the three florets was calculated. The second measurement was used to determine the modal pollen grain diameter. The modal pollen grain diameter class was identified for each floret, and the mean diameter was calculated from the three florets per plant. To remove the effect of debris, only particles with diameter between 30 and 60 µm were considered for this analysis.

Pollen viability was analysed using the Ampha Z32 Pollen Analyzer (Amphasys, Lucerne, Switzerland). For each plant, 10 non-extruded, mature anthers were collected from the primary spike and ruptured to release the pollen grains into AmphaFluid9 buffer (Amphasys). Up to 10,000 pollen grains were measured per sample; samples with fewer than 2,000 total pollen grains were excluded from the downstream analysis. To identify the cluster corresponding to non-viable pollen grains, anthers were collected from Cadenza WT plants and heated to 65°C to kill the pollen grains. Separately, we agitated a sample of AmphaFluid9 buffer to identify the

cluster corresponding to bubbles. The number of pollen grains in the clusters corresponding to the viable and non-viable pollen grains were used to calculate the percentage of non-viable pollen grains in each sample. Linear mixed models were used to calculate the estimated marginal means for pollen number per anther, pollen size, and pollen viability, per genotype category, with block as a random variable (phenotype ~ genotype category + (1|block)). Pairwise significant differences were assigned using a false discovery rate multiple testing correction. Statistical analysis was carried out using the lme4 and emmeans packages in R (v4.2.2).

To determine the number of nuclei per pollen grain, pollen grains were stained with DAPI (4',6-diamidine-2-phenylindole). Four non-extruded, mature anthers from the primary spike were fixed in Carnoy's solution. After a minimum incubation time of 48 hours, the anthers were removed from the Carnoy's solution and stained with DAPI solution (0.001 mg/mL DAPI, 1% (v/v) Triton X-100). The sample was sonicated to release the pollen grains and incubated in the dark for 5 minutes. Samples were imaged with an Axio Zoom v16 II stereomicroscope (Zeiss, Oberkochen, Germany) with an ORCA-Flash4.0 Digital CMOS camera (Hamamatsu Photonics, Shizuoko, Japan). Six non-overlapping images were collected per sample. The number of non-, mono-, bi-, and tri-nucleated pollen grains were counted manually for each sample. Pollen grains with an ambiguous number of nuclei were not counted. As fewer replicates were used for DAPI staining, estimated marginal means were calculated irrespective of blocks and significant differences were assigned by a Wilcoxon signed-rank test with false discovery rate multiple testing correction. The percentage of pollen grains with abnormal numbers of nuclei were log transformed before the test was applied.

Flowering Time

The seeds from the self-pollinated late flowering individual from the C0451xC2028xC1292 population were sown and genotyped with the C0451 and C2028 KASP primers and the C1292 CAPS primers (Supplementary Table 5). The date of complete emergence of the primary spike (Zadok's growth stage Z59) was recorded and the number of days to flowering (relative to the first plant to reach Z59) was calculated. Plant height and the length of the primary spike were measured. The number of fully formed spikelets on the primary spike were counted.

Awn Development

215 segregating BC_1F_2 plants were grown and scored for the presence of awns based on the method previously described (Huang *et al.*, 2020). Fully awned individuals had awns >1cm throughout the length of the spike whereas awnletted plants had awns <1cm in the mid and basal portions of the spike while the length of awns at the apex reached up to 3cm. We scored any plants that had awns <1cm across the entire spike as awnless (Figure 7). We investigated the *Tipped1* (*B1*) gene (*TraesCS5A02G542800*), a characterised awn suppression gene, using published PCR and KASP markers (Huang *et al.*, 2020).

Reciprocal crossing

To test the transmissibility of *met1-1* mutant alleles in each of the parental gametes in hexaploid wheat we made reciprocal crosses between Cadenza WT and triple heterozygous plants (AaBbDd) from the C0465xC0884 population. The seeds of these crosses were harvested and grown. Once the seedlings had reached the two-leaf stage (Zadoks growth stage Z12) leaf samples were taken, DNA was extracted and KASP genotyping was carried out.

Whole Genome Bisulfite Sequencing

We collected leaf samples from the fourth leaf at Zadok's growth stage Z15.6/23, or from the sixth leaf at Z16 for the Aabbdd mutant. The leaf samples were frozen in liquid nitrogen and stored at -70°C. DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) from three biological replicates per genotype, excluding the Aabbdd genotype, of which only one plant was recovered. DNA concentration was estimated for each sample using an Invitrogen Qubit 4 fluorometer (Thermo Fisher Scientific, Massachusetts, USA) and the three replicates per genotype were pooled in equimolar ratios. Library preparation, bisulfite conversion, and sequencing by Illumina NovaSeq 6000 were performed by BMKgene (Münster, Germany), to obtain an average of 2.62 billion 150 bp paired-end reads per sample. Bisulfite conversion rates were calculated by BMKgene using the Bismark software (Krueger & Andrews, 2011), and rates between 99.52% and 99.70% were obtained. Fastp (Chen *et al.*, 2018) was used to remove adapter sequences, low quality reads, and polyG sequences. The Bismark Bisulfite Mapper (Krueger & Andrews, 2011) was used to process the sequencing data. Reads were

aligned to a bisulfite-converted version of the IWGSC RefSeq v1.0 reference genome (IWGSC *et al.*, 2018). The average mapping efficiency was 40.49%, resulting in average coverage of 11X. Methylation calls in all sequence contexts (CpG, CHG, CHH) were extracted. Average cytosine methylation across genes and transposable elements was plotted using the computeMatrix and plotProfile functions from deepTools (Ramírez *et al.*, 2016). The lengths of all genes and transposable elements were scaled to 2,000 bp. Only cytosines with at least 4 reads were included in the analysis. To make these plots, GTF files for high confidence genes and transposable elements were generated using gffread (Pertea & Pertea, 2020) from previously published gff3 files (Catoni *et al.*, 2018). Methylation profiles across chromosomes were calculated using the computeMethylationProfile function from DMRCaller (Catoni *et al.*, 2018), using bins of 1 Mb. Partitions between the chromosome segments (R1, R2a, C, R2b, and R3) were assigned based on the segments defined in (IWGSC *et al.*, 2018), available at https://opendata.earlham.ac.uk/wheat/under_license/toronto/Ramirez-Gonzalez_etal_2018-06025-Transcriptome-Landscape/data/TablesForExploration/. Differentially methylated regions (DMRs) in the CG context were called for each chromosome separately using the computeDMRs function from DMRCaller (Catoni *et al.*, 2018), specifying the method as 'bins', binSize as 100 bp, pValueThreshold as 0.01, minCytosinesCount as 10, minProportionDifference as 0.4, minGap as 0, and minReadsPerCytosine as 4. DMRs associated with genomic features were identified using bedtools intersect (Quinlan & Hall, 2010) to find overlapping genomic coordinates between previously annotated high-confidence genes (including 1 kb up- and downstream of the coding sequence) (IWGSC *et al.*, 2018) and transposable elements (Wicker *et al.*, 2018).

RNA-Sequencing

We collected leaf samples from the fourth leaf at Zadok's growth stage Z15.6/23, or from the sixth leaf at Z16 for the Aabbdd mutant. The leaf samples were frozen in liquid nitrogen and stored at -70°C. RNA was extracted using trizol-chloroform and the RNA Clean & Concentrator Kit (Zymo Research). RNA was extracted from three biological replicates per genotype, excluding the Aabbdd genotype. Library preparation and sequencing by Illumina NovaSeq 6000 was performed by BMKgene (Münster, Germany), to obtain 150 bp paired-end reads. Fastp (Chen *et al.*, 2018) was used to remove adapter sequences, low quality reads, and polyG

sequences. Reads were pseudoaligned to the IWGSC RefSeq v1.1 reference transcriptome (IWGSC *et al.*, 2018) and quantified using kallisto (Bray *et al.*, 2016). All downstream analysis was carried out in R (v4.3.2). The output from kallisto was imported into R using tximport (Soneson *et al.*, 2015). Differentially expressed genes between genotypes were identified using DESeq2, using an adjusted p-value threshold of <0.01 and ≥ 2 fold change (Love *et al.*, 2014). Enriched gene ontology (GO) terms were identified using the Goseq package in R (Young *et al.*, 2010) and GO term-annotated wheat transcript IDs from (Andleeb *et al.*, 2023).

Gene expression data across multiple tissues was downloaded as tpm values from https://bar.utoronto.ca/efp_wheat/cgi-bin/efpWeb.cgi (Choulet *et al.*, 2014; Borrill *et al.*, 2016; Ramírez-González *et al.*, 2018).

Accession numbers

Raw reads from whole genome bisulfite sequencing have been deposited in the European Nucletoide Archive under project PRJEB77426 and reads from RNA-sequencing have been deposited in project PRJEB77425.

All scripts used for whole genome bisulfite sequencing and RNA-sequencing analyses can be found on GitHub at https://github.com/Borrill-Lab/Wheat_MET1/ .

Acknowledgements

We thank Abdulkader Alabdullah for advice on pollen phenotyping, Marek Glombik for advice on genomic analysis and James Brown for advice on statistical analysis of grain genotyping data. We thank Phil Robinson from Scientific Photography at JIC for taking high-quality images of plants and spikes. This research was supported by the NBI Research Computing group through HPC resources.

Author contributions

MC and PB conceived the study. SB, DD and PB designed the research. JS and SB generated the mutant populations. SB and DD performed most experiments and statistical analysis. DD carried out the data analysis for the RNA-seq and whole genome bisulfite sequencing. JC assisted with genotyping and phenotyping of awn development. GDS helped with genotyping and carried out

pollen phenotyping experiments. DD analysed the pollen phenotyping data. SB, DD and PB created the figures. SB, DD and PB wrote the manuscript with input from all authors. All authors have read and approved the manuscript.

Funding

This work was supported by the UK Biotechnology and Biological Science Research Council (BBSRC) through grant BB/T013524/2 and the Institute Strategic Programmes Delivering Sustainable Wheat (DSW) (BB/X011003/1) and Building Robustness in Crops (BRiC) (BB/X01102X/1). DD was funded through a BBSRC Doctoral Training Partnership studentship (BB/T008717/1). GDS was funded by a scholarship from Sant'Anna School of Advanced Studies, Pisa.

References

Alabdullah AK, Moore G, Martín AC. 2021. A duplicated copy of the meiotic gene *ZIP4* preserves up to 50% pollen viability and grain number in polyploid wheat. *Biology-Basel* **10**: 290.

Andleeb T, Knight E, Borrill P. 2023. Wheat *NAM* genes regulate the majority of early monocarpic senescence transcriptional changes including nitrogen remobilization genes. *G3 Genes|Genomes|Genetics* **13**: jkac275.

Boden SA, McIntosh RA, Uauy C, Krattinger SG, Dubcovsky J, Rogers WJ, Xia XC, Badaeva ED, Bentley AR, Brown-Guedira G, et al. 2023. Updated guidelines for gene nomenclature in wheat. *Theoretical and Applied Genetics* **136**: 72.

Borrill P, Mago R, Xu T, Ford B, Williams SJ, Derkx A, Bovill WD, Hyles J, Bhatt D, Xia X, et al. 2022. An autoactive *NB-LRR* gene causes *Rht13* dwarfism in wheat. *Proceedings of the National Academy of Sciences* **119**: e2209875119.

Borrill P, Ramirez-Gonzalez R, Uauy C. 2016. expVIP: a Customizable RNA-seq data analysis and visualization platform *Plant Physiology* **170**: 2172-2186.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **34**: 525-527.

Catoni M, Cortijo S 2018. Chapter Four - EpiRILs: Lessons From Arabidopsis. In: Mirouze M, Bucher E, Gallusci P eds. *Advances in Botanical Research*: Academic Press, 87-116.

Catoni M, Tsang JM, Greco AP, Zabet NR. 2018. DMRcaller: a versatile R/Bioconductor package for detection and visualization of differentially methylated regions in CpG and non-CpG contexts. *Nucleic Acids Research* **46**: e114.

Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: i884-i890.

Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, et al. 2014. Structural and functional partitioning of bread wheat chromosome 3B. *Science* **345**: 1249721.

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. 2008. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* **452**: 215-219.

Cubas P, Vincent C, Coen E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401**: 157-161.

Daron J, Glover N, Pingault L, Theil S, Jamilloux V, Paux E, Barbe V, Mangenot S, Alberti A, Wincker P, et al. 2014. Organization and evolution of transposable elements along the bread wheat chromosome 3B. *Genome Biology* **15**: 546.

Domb K, Katz A, Harris KD, Yaari R, Kaisler E, Nguyen VH, Hong UVT, Griess O, Heskiau KG, Ohad N, et al. 2020. DNA methylation mutants in *Physcomitrella patens* elucidate individual roles of CG and non-CG methylation in genome regulation. *Proceedings of the National Academy of Sciences* **117**: 33700-33710.

Erdmann RM, Picard CL. 2020. RNA-directed DNA Methylation. *PLOS Genetics* **16**: e1009034.

Feng S, Cokus SJ, Zhang X, Chen PY, Bostick M, Goll MG, Hetzel J, Jain J, Strauss SH, Halpern ME, et al. 2010. Conservation and divergence of methylation patterning in plants and animals. *Proceedings of the National Academy of Sciences* **107**: 8689-8694.

Finnegan EJ, Ford B, Wallace X, Pettolino F, Griffin PT, Schmitz RJ, Zhang P, Barrero JM, Hayden MJ, Boden SA, et al. 2019. Zebularine treatment is associated with deletion of *FT-B1* leading to an increase in spikelet number in bread wheat *Plant Cell and Environment* **41**: 1346–1360.

He L, Wu W, Zinta G, Yang L, Wang D, Liu R, Zhang H, Zheng Z, Huang H, Zhang Q, et al. 2018. A naturally occurring epiallele associates with leaf senescence and local climate adaptation in *Arabidopsis* accessions. *Nature Communications* **9**: 460.

He XJ, Hsu YF, Pontes O, Zhu JH, Lu J, Bressan RA, Pikaard C, Wang CS, Zhu JK. 2009. NRPD4, a protein related to the RPB4 subunit of RNA polymerase II, is a component of RNA polymerases IV and V and is required for RNA-directed DNA methylation. *Genes & Development* **23**: 318-330.

Hu L, Li N, Xu C, Zhong S, Lin X, Yang J, Zhou T, Yuliang A, Wu Y, Chen Y-R, et al. 2014. Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proceedings of the National Academy of Sciences* **111**: 10642-10647.

Huang D, Zheng Q, Melchkart T, Bekkaoui Y, Konkin DJ, Kagale S, Martucci M, You FM, Clarke M, Adamski NM. 2020. Dominant inhibition of awn development by a putative zinc-finger transcriptional repressor expressed at the B1 locus in wheat. *New Phytologist* **225**: 340-355.

IWGSC, Appels R, Eversole K, Stein N, Feuillet C, Keller B, Rogers J, Pozniak CJ, Choulet F, Distelfeld A, et al. 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**: eaar7191.

Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuisson J, Heredia F, Audigier P, et al. 2009. Assessing the impact of transgenerational epigenetic variation on complex traits. *PLOS Genetics* **5**: e1000530.

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

Kankel MW, Ramsey DE, Stokes TL, Flowers SK, Haag JR, Jeddeloh JA, Riddle NC, Verbsky ML, Richards EJ. 2003. Arabidopsis *MET1* cytosine methyltransferase mutants. *Genetics* **163**: 1109-1122.

Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L, Simmonds J, Ramirez-Gonzalez RH, Wang X, Borrill P, et al. 2017. Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences* **114**: E913-E921.

Krueger F, Andrews SR. 2011. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**: 1571-1572.

Lei MG, Zhang HM, Julian R, Tang K, Xie SJ, Zhu JK. 2015. Regulatory link between DNA methylation and active demethylation in. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 3553-3557.

Li D, Feng BE, Liu YJ, Gong J, Tang YM, Zhang LP, Pang BS, Sun RW, Zhang FT, Chen ZB, et al. 2021. Genome-wide identification and transcriptional characterization of DNA

methyltransferases conferring temperature-sensitive male sterility in wheat. *BMC Genomics* **22**: 310

Li J, Han F, Yuan T, Li W, Li Y, Wu HX, Wei H, Niu S. 2023. The methylation landscape of giga-genome and the epigenetic timer of age in Chinese pine. *Nature Communications* **14**: 1947.

Li Q, Eichten SR, Hermanson PJ, Zaunbrecher VM, Song J, Wendt J, Rosenbaum H, Madzima TF, Sloan AE, Huang J, et al. 2014. Genetic perturbation of the maize methylome. *The Plant Cell* **26**: 4602-4616.

Liang W, Li J, Sun L, Liu Y, Lan Z, Qian W. 2022. Deciphering the synergistic and redundant roles of CG and non-CG DNA methylation in plant development and transposable element silencing. *New Phytologist* **233**: 722-737.

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.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Manning K, Tör M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB. 2006. A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. *Nature Genetics* **38**: 948-952.

Mathieu O, Reinders J, Čaikovski M, Smathajitt C, Paszkowski J. 2007. Transgenerational stability of the *Arabidopsis* epigenome is coordinated by CG methylation. *Cell* **130**: 851-862.

Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, Cao J, Weigel D, Paszkowski J, Mathieu O. 2009. Selective epigenetic control of retrotransposition in. *Nature* **461**: 427-U130.

Mirouze M, Vitte C. 2014. Transposable elements, a treasure trove to decipher epigenetic variation: insights from Arabidopsis and crop epigenomes. *Journal of Experimental Botany* **65**: 2801-2812.

Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T. 2001. Mobilization of transposons by a mutation abolishing full DNA methylation in. *Nature* **411**: 212-214.

Miura K, Agetsuma M, Kitano H, Yoshimura A, Matsuoka M, Jacobsen SE, Ashikari M. 2009. A metastable *DWARF1* epigenetic mutant affecting plant stature in rice. *Proceedings of the National Academy of Sciences* **106**: 11218-11223.

Niederhuth CE, Bewick AJ, Ji L, Alabady MS, Kim KD, Li Q, Rohr NA, Rambani A, Burke JM, Udall JA, et al. 2016. Widespread natural variation of DNA methylation within angiosperms. *Genome Biology* **17**: 194.

Ortega-Galisteo AP, Morales-Ruiz T, Ariza RR, Roldán-Arjona T. 2008. Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Molecular Biology* **67**: 671-681.

Pallotta M, Warner P, Fox R, Kuchel H, Jefferies S, Langridge P 2003. Marker assisted wheat breeding in the southern region of Australia. *Proceedings of the Tenth International Wheat Genetics Symposium*: Istituto Sperimentale per la Cerealicoltura Puglia, Italy. 789-791.

Pertea G, Pertea M. 2020. GFF Utilities: GffRead and GffCompare. *F1000Res* **9**.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.

Ramírez-González R, Borrill P, Lang D, Harrington S, Brinton J, Venturini L, Davey M, Jacobs J, Van Ex F, Pasha A. 2018. The transcriptional landscape of polyploid wheat. *Science* **361**: eaar6089.

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* **44**: W160-W165.

Reinders J, Wulff BB, Mirouze M, Marí-Ordóñez A, Dapp M, Rozhon W, Bucher E, Theiler G, Paszkowski J. 2009. Compromised stability of DNA methylation and transposon immobilization in mosaic *Arabidopsis* epigenomes. *Genes & Development* **23**: 939-950.

Saze H, Scheid OM, Paszkowski J. 2003. Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. *Nature Genetics* **34**: 65-69.

Schreiber M, Barakate A, Uzrek N, Macaulay M, Sourdille A, Morris J, Hedley PE, Ramsay L, Waugh R. 2019. A highly mutagenised barley (cv. Golden Promise) TILLING population coupled with strategies for screening-by-sequencing. *Plant Methods* **15**: 99.

Soneson C, Love MI, Robinson MD. 2015. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* **4**: 1521.

Springer NM, Schmitz RJ. 2017. Exploiting induced and natural epigenetic variation for crop improvement. *Nature Reviews Genetics* **18**: 563-575.

Srikant T, Tri Wibowo A. 2021. The underlying nature of epigenetic variation: origin, establishment, and regulatory function of plant epialleles. *International Journal of Molecular Sciences* **22**: 8618.

Srikant T, Yuan W, Berendzen KW, Contreras-Garrido A, Drost H-G, Schwab R, Weigel D. 2022. Canalization of genome-wide transcriptional activity in *Arabidopsis thaliana* accessions by MET1-dependent CG methylation. *Genome Biology* **23**: 263.

Tariq M, Saze H, Probst AV, Lichota J, Habu Y, Paszkowski J. 2003. Erasure of CpG methylation in *Arabidopsis* alters patterns of histone H3 methylation in heterochromatin. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 8823-8827.

Thomas M, Pingault L, Poulet A, Duarte J, Throude M, Faure S, Pichon JP, Paux E, Probst AV, Tatout C. 2014. Evolutionary history of Methyltransferase 1 genes in hexaploid wheat. *BMC Genomics* **15**: 922.

Thorstensen T, Fischer A, Sandvik SV, Johnsen SS, Grini PE, Reuter G, Aalen RB. 2006. The SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. *Nucleic Acids Research* **34**: 5461-5470.

Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. 2009. Bursts of retrotransposition reproduced in. *Nature* **461**: 423-U125.

Wicker T, Gundlach H, Spannagl M, Uauy C, Borrill P, Ramirez-Gonzalez RH, De Oliveira R, International Wheat Genome Sequencing C, Mayer KFX, Paux E, et al. 2018. Impact of transposable elements on genome structure and evolution in bread wheat. *Genome Biology* **19**: 103.

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. *The 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 Molecular Biology* **85**: 219-232.

Yang Y, Tang K, Datsenka TU, Liu W, Lv S, Lang Z, Wang X, Gao J, Wang W, Nie W, et al. 2019. Critical function of DNA methyltransferase 1 in tomato development and regulation of the DNA methylome and transcriptome. *Journal of Integrative Plant Biology* **61**: 1224-1242.

Young MD, Wakefield MJ, Smyth GK, Oshlack A. 2010. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology* **11**: R14.

Zhang H, Lang Z, Zhu J-K. 2018. Dynamics and function of DNA methylation in plants. *Nature Reviews Molecular Cell Biology* **19**: 489-506.

Zhang P, Mbodj A, Soundiramourtty A, Llauro C, Ghesquière A, Ingouff M, Keith Slotkin R, Pontvianne F, Catoni M, Mirouze M. 2023. Extrachromosomal circular DNA and structural variants highlight genome instability in Arabidopsis epigenetic mutants. *Nature Communications* **14**: 5236.

Zhang Y-Y, Fischer M, Colot V, Bossdorf O. 2013. Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytologist* **197**: 314-322.