Enzymatic DNA demethylation enables to the formation of stable epimutations in tomato

Javier Antunez-Sanchez[1*], Julia Engelhorn[1,a], Sara Lopez-Gomollon[2], Peter Meyer[3] and Jose Gutierrez-Marcos[1*]

[1] School of Life Sciences, University of Warwick, Coventry, United Kingdom.

[2]School of Biosciences, University of Kent, Canterbury, UK

[3]Centre for Plant Sciences, University of Leeds, Leeds, UK

[a]Current address: DIADE (Diversity Adaptation and Development of Plants laboratory), IRD, CIRAD, University Montpellier, 911 avenue Agropolis, 34 394 Montpellier, France.

***Correspondence:** J.Antunez@warwick.ac.uk and j.f.gutierrez-marcos@warwick.ac.uk

Keywords: Demethylation, TET3, epiallele, variation.

Running title: TET3-mediated demethylation in tomato.

Abstract

DNA methylation is a critical epigenetic modification in plants that regulates gene expression, silences transposable elements (TEs), and supports proper development. Traditionally, heritable epimutations in plants have been generated using genetic mutants or chemical inhibitors, but these approaches often lack precision or stability. In this study, we investigated the effects of globally altering DNA methylation in tomato, a species with a large, TE-rich genome, through the ectopic expression of the catalytic domain of the human DNA demethylase TEN-ELEVEN TRANSLOCATION3 (hTET3cd). We found that TET3-mediated demethylation induced stable hypomethylation at CG and CHG sites and that these changes were inherited across multiple generations, including in non-transgenic siblings. Interestingly, demethylation in heterochromatic pericentromeric regions was often accompanied by gains in CHH methylation, suggesting the compensatory activation of the RNA-directed DNA methylation (RdDM) pathway. Differentially methylated region (DMR) analysis revealed that CG and CHG methylation loss was widespread, while CHH DMRs showed complex patterns of gain and loss, particularly near gene-rich regions and transposable elements enriched for 24-nt small RNAs. Transcriptomic analyses showed distinct gene expression profiles in both TET3 and non-transgenic progeny, with altered expression of TEs and associated genes. These

findings demonstrate that enzymatic manipulation of the methylome via hTET3cd can generate stable, heritable epigenetic variation, and highlight the dynamic interplay between targeted DNA demethylation and endogenous mechanisms that act to restore epigenetic homeostasis.

Introduction

DNA methylation, which involves the addition of a methyl group to cytosine residues, is one of the most studied epigenetic modifications in plants. It occurs in three different sequence contexts: CG, CHG, and CHH (where H = A, T, or C), and can be established, maintained, or removed by distinct enzymatic pathways targeting each context (Law and Jacobsen 2010, Erdmann and Picard 2020). The establishment of DNA methylation in plants, termed *de novo* methylation, is primarily mediated by the RNA-directed DNA methylation (RdDM) pathway. In this mechanism, 24-nucleotide small interfering RNAs (siRNAs), generated by RNA polymerase IV (Pol IV) and processed by Dicer-like 3 (DCL3), guide the DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2) to specific genomic loci (Matzke and Mosher 2014, Erdmann and Picard 2020). Simultaneously, a long non-coding RNA (lncRNA) produced by RNA polymerase V (Pol V) forms a scaffold for ARGONAUTE4 (AGO4) and other accessory factors, ultimately recruiting DRM2 to catalyse DNA methylation (Stroud, Do et al. 2014). Although RdDM can initiate methylation in previously unmethylated regions, its primary role is to reinforce silencing at loci already marked by DNA methylation and repressive chromatin modifications (Zhang, Lang et al. 2018). The maintenance of DNA methylation in plants is achieved through different pathways for each sequence context. In the CG context, METHYLTRANSFERASE 1 (MET1) faithfully copies symmetrical methylation patterns from the parental to the newly synthesized DNA strand during replication (Finnegan, Peacock et al. 1996). In the CHG context, CHROMOMETHYLASE 3 (CMT3) cooperates with H3K9me2-marked chromatin and SU(VAR)3–9 HOMOLOG (SUVH) proteins to propagate methylation marks (Bartee, Malagnac et al. 2001, Stroud, Do et al. 2014). Meanwhile, CHH methylation relies on a combination of RdDM and chromomethylases like CHROMOMETHYLASE 2 (CMT2), often assisted by the chromatin remodeler DECREASED DNA METHYLATION 1 (DDM1) (Zhong, Du et al. 2014, Zhang, Lang et al. 2018). Active demethylation also plays an essential role in the dynamic regulation of the plant methylome. REPRESSOR OF SILENCING 1 (ROS1), DEMETER (DME), and DEMETER-LIKE (DML) DNA glycosylases/lyases remove methylated cytosines, leaving an abasic site repaired via base excision repair (Gong, Morales-Ruiz et al. 2002, Penterman, Zilberman et al. 2007, Williams, Bechen et al. 2022). While both plants and animals perform active demethylation, the enzymatic mechanisms are different. In mammals, Ten-Eleven Translocation (TET) enzymes oxidize 5mC into intermediates (5hmC, 5fC, 5caC), which are then excised by DNA glycosylases (Wu and Zhang 2017, Zhang, Zhang et al. 2023). By contrast, plants lack true TET homologs and rely directly on bifunctional glycosylases that recognize and excise 5mC (Zhu 2009).

DNA methylation patterns can be altered in plants through various means. Mutant lines defective in major methyltransferases (e.g., *met1*, *cmt3*) or demethylases (e.g., *ros1*) often display striking phenotypic changes, revealing the significance of maintaining methylation levels for development and stress responses (Zhang, Lang et al. 2018). Moreover, chemical inhibitors of DNA methyltransferases like 5-azacytidine (azaC) and 5-azadeoxycytidine (azadC) can be used to induce global demethylation. These nucleoside analogues become incorporated into DNA and bind covalently to DNA methyltransferases, preventing the enzymes from methylating cytosines during DNA replication (Christman 2002). While effective at reducing overall methylation levels, these treatments can be quite aggressive, affecting many genomic regions simultaneously and sometimes

causing developmental abnormalities or toxicity (Christman 2002, Akimoto, Katakami et al. 2007). Tissue de-differentiation and regeneration, a common strategy used for plant clonal propagation and genetic modification, can induce partial or widespread methylation reprogramming, an effect sometimes referred to as somaclonal variation (Ghosh, Igamberdiev et al. 2021). Moreover, somatic embryogenesis induced by embryonic transcription factors can reset certain epigenetic marks during regeneration, leading to stable epigenetic variants in the regenerated plants (Wibowo, Becker et al. 2018, Wibowo, Antunez-Sanchez et al. 2022). Environmental factors offer yet another route for manipulating the methylation landscape of plants. Abiotic stresses such as heat, drought, or salinity, as well as biotic interactions like pathogens, can trigger locus-specific gains or losses in methylation in plants (Dowen, Pelizzola et al. 2012, Wibowo, Becker et al. 2016). Intriguingly, these stress-induced epigenetic modifications persist beyond the initial generation, providing an additional mechanism for acclimation or adaptation.

Heritable changes in DNA methylation that do not involve alterations in the underlying DNA sequence are often termed “epimutations”. These can arise spontaneously or be deliberately induced through the methods described above. Epimutations can activate or silence transposons, modify gene regulatory regions, or otherwise shape gene expression profiles in ways that lead to phenotypic outcomes (Saze, Tsugane et al. 2012). For example, changes in promoter methylation in plants can accelerate or delay flowering (Choi, Gehring et al. 2002, Kankel, Ramsey et al. 2003), alter root architecture (Henderson and Jacobsen 2007), or enhance stress tolerance (Dowen, Pelizzola et al. 2012). Unlike mammals, which often undergo extensive waves of DNA demethylation in the germline or early embryos, plants largely retain their epigenetic marks during reproduction. Because the plant germline is set aside late in development and does not typically undergo wholesale erasure of DNA methylation, epimutations that arise in somatic tissues can be transmitted to subsequent generations (Henderson and Jacobsen 2007, Gutierrez-Marcos and Dickinson 2012, Heard and Martienssen 2014). Although DNA demethylation is common in companion cells surrounding the developing plant germline, methylation is not fully reset in the egg or sperm (Calarco, Borges et al. 2012). Consequently, epigenetic states can be transmitted to embryos and inherited by the progeny. This partial reprogramming contrasts with mammalian development, in which global demethylation waves erase most epigenetic marks in primordial germ cells and zygotes (Hackett and Azim Surani 2013, Seisenberger, Peat et al. 2013, Leeke, Varsally et al. 2025). Because methylation patterns in plants can persist across generations, these epimutations represent a potential resource for plant breeding, enabling the selection of new phenotypes without modifying the primary DNA sequence (Tonosaki, Fujimoto et al. 2022).

Several studies have demonstrated the feasibility of strategies for the precise manipulation of the methylome both in animals and in plants. Genome-editing platforms like TALENs, Zinc fingers and CRISPR/Cas9 can be repurposed for epigenome editing by fusing catalytically inactive Cas9 (dCas9), Zinc fingers or TAL effector domains to methylation-modifying enzymes. For instance, dCas9-TET fusions have been employed in mammalian cells to demethylate specific loci (Amabile, Migliara et al. 2016, Liu, Wu et al. 2016). It has been shown in Arabidopsis that a dCas9-TET1 fusion can erase DNA methylation at targeted sites, leading to altered gene expression (Gallego-Bartolomé, Gardiner et al. 2018). Such locus-specific manipulation of methylation holds promise for generating stable yet precise epimutations, expanding the toolkit available for crop improvement.

In this study, we demonstrate that ectopic expression of the human TET3 catalytic domain in tomato induces widespread loss of DNA methylation, primarily at CG and CHG sites. These hypomethylated marks were largely retained across multiple generations, even in the absence of the transgene, and were accompanied by transcriptional reprogramming. These results indicate

that enzymatic demethylation in plants can generate stable, meiotically heritable epimutations, offering a strategy to induce novel phenotypic variation without altering the underlying DNA sequence.

Results

To assess the impact that global hypomethylation has in tomato we generated transgenic lines expressing the C-terminal catalytic domain (residues 859 to 1795) of the human methylcytosine dioxygenase Ten Eleven Translocation 3 protein (hTET3cd) (Hollwey, Out et al. 2017) (**Figure 1 Supplement 1A-B**). As previous work in *Arabidopsis* showed that the expression of human TET proteins can have varied effects on the methylome (Ji, Jordan et al. 2018), we selected for analysis two independent (TET3) transgenic lines with alike hTET3cd expression, normal growth and fertility. One of the TET3 lines (A) was propagated by self-pollination over three generations, after selection for presence of the transgene (**Figure 1A**). A second TET3 line (B) was self-pollinated and segregating plants were selected for the presence (TET3) or absence (non-transgenic, NT) of the transgene, and propagated by self-pollination (**Figure 1A**). We used a modified whole-genome bisulfite sequencing (WGBS) methodology (see methods) to profile DNA methylation levels in TET3 and NT progenies (**Figure 1 Supplement 1C**). This analysis revealed a significant decrease in CG and CHG methylation in all TET3 plants, with the most significant methylation decrease observed in one of them (A). Notably, DNA hypomethylation was not fully restored in subsequent generations of TET3 nor NT sibling plants (**Figure 1B and Figure 1 Supplement 2**). We found that CG and CHG hypomethylation was similar in TET3 and NT sibling plants, suggesting that TET3-directed demethylation can be stably inherited. In contrast, CHH sites became hypermethylated in T1 TET3 and NT sibling plants and returned to WT levels in subsequent generations indicating that the loss of CHH methylation can be rapidly restored in absence of demethylase activity (**Figure 1B**).

To investigate the possible differences between transgenic lines and generations, we performed a Principal Component Analysis (PCA) of the average %mCG. This analysis showed a clear separation between WT and TET3 or NT sibling plants (**Figure 1C and Figure 1 Supplement 3**), which correlated well with the global level of methylation (**Figure 1B**). The second Principal Component separated the two TET3 lines, suggesting that hTET3cd activity mediate stochastic patterns of demethylation of the tomato genome. We next examined the distribution of these methylation losses across chromosomes and found that the biggest changes occurred in the heterochromatic pericentromeric regions, with the most significant losses in CG and CHG methylation and the greatest gains in CHH methylation (**Figure 1D and Figure 1 Supplement 4A**). When we split the genome into euchromatin and heterochromatin regions (**see Methods**) and calculated the average methylation for each compartment, we observed larger losses of CG methylation in both chromatin compartments, but CHG demethylation was primarily found in heterochromatin (**Figure 1 Supplement 4B**).

To further investigate this observation, we split the genome into euchromatin and heterochromatin and calculated the average methylation levels for each compartment. Heterochromatin accumulated substantial losses in CG and CHG methylation across both TET3 lines and across several generations, likely due to its higher baseline methylation levels. In contrast, euchromatin exhibited consistent reductions in CG methylation, but CHG methylation remained largely unchanged in most plants, except for a notable reduction in the T1 generation of line B (**Figure 1 Supplement 4B**). These findings suggest that hTET3cd can targets CG methylation ubiquitously, while its effect on CHG methylation is more dependent on chromatin context. Interestingly, non-transgenic siblings exhibited similar methylation patterns to their transgenic counterparts, indicating that TET3-induced demethylation can be inherited in the absence of TET3 activity.

Additionally, we observed gains in CHH methylation, particularly in the first transgenic generation, which returned to near WT levels in subsequent generations in the absence of TET3 (**Figure 1B-D and Figure 1 Supplement 4B**). The increase in CHH methylation in TET3 plants reveals the complex interplay between asymmetric methylation and potential compensatory mechanisms.

Collectively, our analysis reveals that the ectopic expression of hTET3cd results in global changes in DNA methylation, predominantly establishing CG and CHG hypomethylation. Notably, the most pronounced demethylation effects were observed in heterochromatic regions, while euchromatin regions appear to be more resistant. Interestingly, non-transgenic siblings exhibited levels of demethylation comparable to their transgenic counterparts, indicating that the induced methylation changes are inherited over several meiotic cycles.

To get a deeper understanding of the changes mediated by TET3 we carried out an in-depth methylome analysis using two TET3 (A.T1_4) and NT sibling (A.T1_17) lines (**See methods**). At the whole genome level, both showed reduced methylation in euchromatin across all methylation contexts (**Figure 2A and Figure 2 Supplement 1A**). However, NT siblings exhibited greater demethylation in heterochromatin compared to TET3 (**Figure 2A and Figure 2 Supplement 1A**). We further examined methylation levels in different chromatin compartments, separating euchromatin and heterochromatin. Both plants exhibited general demethylation in euchromatin, while in heterochromatin, the non-transgenic sibling had a stronger demethylation effect compared to the TET3 positive line (**Figure 2 Supplement 1B**). We computed methylation statistics to identify differentially methylated regions (DMRs) between the TET3 and the WT control. We first examined read coverage per cytosine and the spatial correlation of methylation (**Figure 2 Supplement 2A-B**) and determined the most suitable parameters for the quantification of DMRs (**Figure 2 Supplement 2C-D; see Methods**). We then computed DMRs for the whole genome and found that loss of DNA methylation was the predominant pattern across all contexts (**Figure 2B**). However, in the TET3 line, a large number of regions were hypermethylated in the CHH context. We further analysed the differentially methylated regions (DMRs) determine the strength and number of DMRs in both TET3 and NT sibling lines across the three methylation contexts (CG, CHG, CHH) (**Figure 2C**). The analysis revealed that the NT siblings exhibited a stronger and more abundant set of DMRs than TET3 plants. DMR analysis revealed distinct methylation patterns between both genotypes (**Figure 2D**). In the CG context, most gain and loss DMRs were largely common in both lines, with a large number of DMRs showing stronger demethylation. CHG DMRs resembled CG DMRs but were exclusively found in NT siblings (**Figure 2 Supplement 3**). The pattern of CHH DMRs was more complex, as the majority of gain DMRs in TET3 were not affected in NT siblings, and vice versa, while loss DMRs showed a mixture of common and unique demethylation (**Figure 2D and Figure 2 Supplement 4**).

To gain a deeper understanding of the methylation changes found in the TET3 and NT siblings we looked at the distribution of DMRs across different chromosomes. This analysis revealed that CG and CHG demethylation, both in TET3 and NT siblings, affected different chromosomes and was independent of the type of chromatin (**Figure 3A**). However, CHH DMRs exhibited stronger methylation differences, both gain and loss, in chromosomal arms (**Figure 3A**). As chromosome arms are rich in genes, we looked at the correlation between gene density and DMR density, and found a strong correlation with genes at gain DMRs in CG and gain or loss DMRs in CHH (**Figure 3B** and **Figure 3 Supplement 1**). Further analysis revealed that loss DMRs in CG and CHG contexts were distal from genes, whereas gain DMRs and CHH DMRs were closer to genes (**Figure 3 Supplement 2**). Notably, the identified CHH DMRs, both gains and losses, were enriched at upstream and downstream gene flanking regions, but depleted at gene bodies (**Figure 3C**). In contrast, CG DMRs

were evenly distributed across of genes. We then quantified the distribution of DMRs at different genomic features, including genes, intergenic regions, and transposable element (TE) (**Figure 3D**). We found that CHH DMRs were predominantly genic and associated with DNA transposons, which contrasted with CG loss DMRs that were largely intergenic and enriched at Gypsy retrotransposons. Since RNA-directed DNA methylation (RdDM) plays a pivotal role in CHH methylation (Law and Jacobsen 2010), we examined whether these regions were associated with the production of 24 nt small RNAs (sRNAs). We found that mCHH DMRs strongly methylated in TET3 plants showed low levels of 24 nt sRNAs in WT plants (**Figure 3E**). In contrast, regions that showed strong mCHH loss DMRs accumulated 24 nt sRNAs in the WT plants (**Figure 3F**). Collectively, these data indicates that TET3-induced methylation changes accumulate primarily at upstream genic regions, affect primarily CHH methylation, are associated with transposable elements and that accumulate non-coding siRNAs.

Given the observed changes in methylation at TE-rich regions, we measured the expression of different transposon families in TET3 and NT siblings. We found several differentially expressed TEs (DETEs), primarily belonging to DNA transposon and LTR retrotransposon misregulated both in TET3 and NT sibling plants (**Figure 4A and Figure 4 Supplement 1A and C**). Further analysis revealed that in both genotypes downregulated DETEs were strongly repressed but upregulated DETEs were transcriptionally active in WT plants (**Figure 4B and Figure 4 Supplement 1B**). Notably, we found that upregulated DETEs were primarily high-copy LTR retrotransposons (**Figure 4C**), suggesting that this class is particularly susceptible to TET3-mediated demethylation. We investigated the spatial distribution of the identified DETEs and found that upregulated DETEs were more concentrated in heterochromatic genomic regions, while downregulated DETEs were evenly distributed across the genome (**Figure 4D**). This localisation pattern mirrors the methylation dynamics observed in demethylated plants, reinforcing the link between methylation and transcriptional activity of transposons. However, when we looked in more detail at the relationship between methylation changes and transposon expression, we found that demethylation of discrete genomic regions resulted in transcriptional misregulation of nearby transposons (**Figure 4E**).

We then investigated the impact that TET3-directed demethylation has on gene expression. We found a clear separation between different genotypes, with profound differences between TET3 and NT sibling plants (**Figure 5A**). We identified a notable number of differentially expressed genes (DEGs): TET3 plants had 279 downregulated and 110 upregulated genes compared to WT, NT sibling plants had 256 downregulated and 829 upregulated genes compared to WT and 468 downregulated and 1559 upregulated genes between TET3 and NT sibling plants (**Figure 5B and Figure 5 Supplement 1**). Hierarchical clustering analysis of DEGs from different comparisons confirmed that the transcriptome of TET3 and NT sibling plants were different (**Figure 5C**). Gene Ontology (GO) and KEGG pathway enrichment analyses showed that in TET3 plants, downregulated genes were significantly enriched in cell wall function and alkaloid biosynthesis, where in NT siblings upregulated genes were associated with energy metabolism and photosynthetic activity (**Figure 5D**). To assess if the methylation changes correlated with gene expression changes, we measured the intersection between DEGs and DMRs. We found that DEGs from both TET3 and NT sibling plants had a significant correlation with DMRs across various methylation contexts (**Figure 5E**). Notably, upregulated genes in TET3 plants were associated with CHH loss and CG gain DMRs, while downregulated genes in NT sibling plants showed a strong enrichment for CHH loss DMRs (**Figure 5E and Figure 5 Supplement 2**). Collectively, these findings suggest that TET3-mediated methylation changes are somatically and meiotically stable, and result in notable changes in gene expression.

Discussion

We set to investigate the consequences of enzymatic demethylating the tomato genome using the mammalian TET3 DNA demethylase. We found that enzymatic DNA demethylation mediated by TET3 results in somatically and meiotically stable epimutations. Previous studies in *Arabidopsis*, which has a small and gene-rich genome, have shown that mammalian TET1 enzymes can catalyse the removal of DNA methylation resulting in gene expression changes (Hollwey, Watson et al. 2016, Gallego-Bartolomé, Gardiner et al. 2018, Ji, Jordan et al. 2018). Our study shows that the TET3-mediated demethylation of tomato, a species with a larger and transposon-rich genome, results in global methylation changes that are stable that are associated novel transcriptional states. These findings open the door to harnessing the activity of different DNA demethylases in plants, as stable epimutations could help to generate phenotypic diversity while avoiding direct edits to the underlying DNA sequence (Tonosaki, Fujimoto et al. 2022, Zhang and Zhu 2025). A particular aspect of the TET3-mediated demethylation in tomato, that contrast with Arabidopsis, is its partial or "mosaic" nature across different genomic contexts, with CG and CHG sites generally losing methylation but certain CHH regions showing notable gains in methylation. One possible explanation for the observed hypermethylation could be that demethylation of pericentromeric or transposon-dense regions triggers compensatory RNA-directed DNA methylation (RdDM) (Law and Jacobsen 2010, Erdmann and Picard 2020). In plants with complex genomes, as it is the case in tomato, such local expansions of CHH methylation (often referred to as "CHH islands") are thought to serve as protective boundaries, preventing transcriptional reactivation of transposable elements (Zemach, Kim et al. 2013, Gouil and Baulcombe 2016). Moreover, differences in TET3 expression could result in the demethylation of different regions of the genome and as such some hypermethylated CHH regions accumulate around specific TEs or gene bodies, while others appear stochastically, but ultimately producing specific patterns across different lines.

Molecular analyses revealed that numerous TEs exhibited altered DNA methylation in TET3 and NT siblings, often accompanied by shifts in expression. Interestingly, while certain classes of LTR retrotransposons showed transient upregulation, we also detected cases of TE downregulation—possibly reflecting secondary chromatin effects or local transcriptional repression that follows initial demethylation. The modest scale of TE methylation dynamics mediated by TET3 in tomato contrast with the strong TE demethylation found in TET1 in *Arabidopsis* (Ji, Jordan et al. 2018) may be linked to the stronger epigenetic surveillance machinery operating in tomato to control TE activity (Wang and Baulcombe 2020). Similar partial reprogramming effects on TEs has been reported in epimutant lines generated with tomato *met1* mutants (Yang, Tang et al. 2019). In addition to transposon-related changes, we also found in NT sibling large transcriptomic changes in the expression of genes implicated in photosynthesis, which may be attributed to a reduction in stomata function. In *Arabidopsis*, defects in active DNA demethylation or related chromatin-remodelling factors have been linked to altered stomatal patterns, including overproduction or mis-localization of stomatal lineage cells (Yamamuro, Miki et al. 2014, Wang, Xue et al. 2016). Whether a similar epigenetic pathway governs stomatal development in tomato remain to be tested.

Notably, we found that TET3-mediated demethylation can persist across multiple generations in NT siblings. This stable inheritance highlights the absent a resetting mechanism for DNA methylation during gametogenesis and fertilisation, enabling newly formed epialleles to be stable (Henderson and Jacobsen 2007, Tonosaki, Fujimoto et al. 2022, Hollwey, Briffa et al. 2023). The current view is that DNA methylation mediated by mammalian demethylases do not undergoes genome-wide reprogramming early in primordial germ cells and early embryo development, thus enabling the transgenerational inheritance of most epimutations (Hollwey, Briffa et al. 2023). The stability of

newly formed epialleles results in novel phenotypic traits in plants, as it has been also reported in studies with epigenetic recombinant inbred lines (epiRILs) (Johannes, Porcher et al. 2009, Reinders, Wulff et al. 2009), or in response to environmental stress or during developmental reprogramming using zygotic factors on exogenous phytohormones (Wibowo, Becker et al. 2018, Wibowo, Antunez-Sanchez et al. 2022).

In *Arabidopsis*, the ectopic expression of TET1 causes widespread DNA hypomethylation and in the activation of transposable element (Gallego-Bartolomé, Gardiner et al. 2018, Ji, Jordan et al. 2018), However, we found that TET3 produces moderate decreases in DNA methylation in tomato, which may be attributed to the often lethal or severely deleterious effect of strong demethylation arising from methylation defective mutants (Yang, Tang et al. 2019, Wang, Baulcombe et al. 2020) and the strong demethylation defences operating in complex genomes rich in mobile transposable elements (Gouil and Baulcombe 2016, Yang, Tang et al. 2019). In a similar vein, targeted TET1-mediated demethylation in *Arabidopsis* could completely demethylate the *FWA* promoter but only partially demethylate the *CACTA1* transposon, which regained partial methylation once TET1 was removed (Gallego-Bartolomé, Gardiner et al. 2018). Although TET1 and TET3 both belong to the same family of 5-methylcytosine dioxygenases, but they exhibit distinct catalytic oxidation efficiencies (Ito, Shen et al. 2011), which could contribute to the moderate demethylation generated in tomato. Moreover, it is also possible that a transient 5-hydroxymethylcytosine (5-hmC) formation directed by TET3 goes undetected in standard bisulfite data. However, a previous study have shown that *Arabidopsis* TET1 lines do not accumulate stable 5-hmC methylation pointing to the presence of a rapid base-excision repair and cytosine demethylation (Ji, Jordan et al. 2018).

Collectively, our findings highlight the complexity of TET-mediated epimutagenesis in plants with large and transposon-rich genomes. Our data reveals that exogenous demethylases can produce heritable epigenetic variation, which in large extent is modulated by the activity of the RdDM pathway that controls transposable elements. Further work should be directed to examine the stability of TET-induced epialleles during gametogenesis and across multiple cycles of sexual reproduction.

Methods

Plant material, growth conditions and generation of transgenic lines

Plants were grown in a glasshouse under long day conditions (16 hr light, 8 hr dark, 23°C, 42% humidity). Three weeks after germination seedlings were transferred to 3 litre pots and grown to maturity. To generate transgenic lines the C-terminal domain of the human TET3 DNA dioxygenase (hTET3c) (**Figure 1 Supplement 1A-B)** was cloned in pGREEN 0179 (Hollwey, Watson et al. 2016) and transferred into Agrobacterium for the transformation of tomato. More than ten independent transgenic lines were generated after culture in kanamycin selection media and PCR genotyped using oligos TET3.FW: 5'-CAACGCGGTGGTGGAGAGCTACTC-3' and TET3.RV: 5'-GGGCTCAGGCTGTTGTGCAG-3' to select lines with low copy transgene insertions. The features of the inserted transgene in selected lines were determined using whole-genome bisulfite sequencing data (**Figure 1 Supplement 1C-D**).

Reference genome construction

The composite reference used for analysis comprised *Solanum lycopersicum* Heinz 1706 SL4.0 assembly (Sol Genomics build 4.00) (Hosmani, Flores-Gonzalez et al. 2019), the chloroplast genome

(GenBank NC_007898.3), the mitochondrial genome (GenBank MF034192.1) and the T-DNA sequence containing the hTET3cd (5591 bp). ITAG 4.1 gene models, the organellar annotations and the transgene features were concatenated into a unified GFF3 then converted to GTF for RNA-seq analysis.

Classification of chromatin

Euchromatic regions were defined as follows. The genome was divided into 100 kbp bins. The average H3K27ac ChIP-seq signal was calculated for each bin, followed by a moving average calculation over a 2 Mbp window (20 bins) centered on each bin. Bins with a moving average H3K27ac signal higher than 11 RPGC were considered euchromatic, and the rest were classified as heterochromatic. Heterochromatic regions smaller than 500 kbp were reclassified as euchromatin (**Figure 1 Supplement 5**).

Low input Whole Genome Bisulfite Sequencing

For the production of low-imput whole genome bisulfite sequencing, purified genomic DNA was subjected to bisulfite conversion following a random priming-based approach (Smallwood, Lee et al. 2014). Briefly, DNA was combined with 0.4 mM dNTPs and 0.4 µM oligo 1 ((Biotin)CTACACGACGCTCTTCCGATCTNNNNNNNNN), then incubated at 65 °C for 3 minutes and immediately placed on ice. Klenow exo– polymerase (50 U; Sigma) was added, and samples were incubated at 4 °C for 5 minutes, followed by a thermal ramp of +1 °C every 15 seconds to 37 °C, and held at 37 °C for 30 minutes. Following this, samples were denatured at 95 °C for 1 minute and quickly cooled on ice. A fresh aliquot of oligo 1 (10 pmol), 25 U Klenow exo–, and 1 nmol dNTPs in a total volume of 2.5 µl was added. The thermal priming and extension cycle (4 °C for 5 min, +1 °C/15 s to 37 °C, 37 °C for 30 min) was repeated for a total of five rounds. Samples were then treated with 40 U of Exonuclease I (NEB) at 37 °C for 1 hour to remove unincorporated primers. DNA was purified using 0.8× Agencourt AMPure XP beads (Beckman Coulter) according to the manufacturer's instructions and eluted in 10 mM Tris-Cl (pH 8.5). The eluted DNA was incubated with washed M-280 Streptavidin Dynabeads (Life Technologies) for 20 minutes at room temperature with rotation. Beads were washed twice with 0.1 N NaOH to remove non-biotinylated DNA, followed by two washes with 10 mM Tris-Cl (pH 8.5). Beads were resuspended in 48 µl of 0.4 mM dNTPs, 0.4 µM oligo 2 (TGCTGAACCGCTCTTCCGATCTNNNNNNNNN), and 1× Blue Buffer. Samples were denatured at 95 °C for 45 seconds and placed on ice. Subsequently, 100 U of Klenow exo– was added, followed by incubation at 4 °C for 5 minutes, ramping +1 °C every 15 seconds to 37 °C, and holding at 37 °C for 90 minutes. Beads were washed with 10 mM Tris-Cl (pH 8.5), then resuspended in 50 µl of PCR master mix containing 0.4 mM dNTPs, 0.4 µM PE1.0 forward primer (AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT), 0.4 µM indexed iPCRTag reverse primer, 1 U KAPA HiFi HotStart DNA Polymerase, and 1× HiFi Fidelity Buffer (KAPA Biosystems). Library amplification was performed under the following conditions: 95 °C for 2 minutes; 12–13 cycles of 94 °C for 80 seconds, 65 °C for 30 seconds, and 72 °C for 30 seconds; followed by a final extension at 72 °C for 3 minutes and hold at 4 °C. Amplified libraries were purified using 0.8× AMPure XP beads and assessed for quality using the Agilent Bioanalyzer with High-Sensitivity DNA chips and quantified using the KAPA Library Quantification Kit for Illumina. Final libraries were pooled and sequenced on an NextSeq 500 platform (Illumina) using 75 bp single-end reads

Raw illumina BCL files were demultiplexed with bcl2fastq. Single-end reads were quality checked using FastQC (v0.11.5) (Andrews, Krueger et al. 2010), and adapter- and quality-trimmed using fastp (v0.23.2) (Chen 2023), removing bases whose four-base sliding-window mean quality fell

below 20 and discarding reads shorter than 40 bp (--cut_right --cut_window_size 4 --cut_mean_quality 20 --length_required 40). Quality was assessed again with FastQC again and MultiQC (v1.17) (Ewels, Magnusson et al. 2016) was used to aggregate quality reports. Trimmed reads were aligned to the previously converted and indexed genome using Bismark (v0.22.3) (Krueger and Andrews 2011) with parameter --non_directional, as due to the library preparation, most reads come from the Complementary of the original Top Strand (CTOT) and the Complementary of the original Bottom Strand (CTOB). PCR duplicates were removed with deduplicate_bismark. Cytosine methylation was extracted with bismark_methylation_extractor, outputting context-resolved genome-wide reports (--cytosine_report --CX). Mapping, deduplication and methylation extraction quality was assessed using Bismark internal scripts. Coverage BigWig files were created from deduplicated and sorted BAM files using BAMscale (v1.0) (Pongor, Gross et al. 2020). BigWig files representing methylation levels were generated using a custom script, CX_report2bigwig.sh, which converts CX reports into BigWigs for visualisation. Methylation levels were averaged within genomic bins of 100 kilo base pairs (kbp) using a custom Perl script, CX_report2mC_bins.v1.0.pl which calculates methylation percentages for each context. Plants were classified as TET3 positive when ≥10 uniquely mapped reads per million (RPM) aligned to the TET3 T-DNA contig. Samples with <14 million covered cytosines were excluded.

Standard Whole Genome Bisulfite Sequencing

For the production of standard whole-genome bisulfite sequencing libraries, two fully expanded leaves from tomato plants were pooled for each sample. Genomic DNA was extracted with the DNeasy Plant Mini Kit (Qiagen, Germany). DNA libraries were generated using the Illumina TruSeq Nano kit (Illumina, CA, USA). The bisulfite treatment was carried out using the Epitect Plus DNA Bisulfite Conversion Kit (Qiagen, Germany) followed by the insertion of adaptors by ligation in a thermal cycler. After clean-up of the bisulfite conversion reaction, library enrichment was done using Kapa Hifi Uracil+ DNA polymerase (Kapa Biosystems, USA). Libraries were sequenced with 2 × 150 bp paired-end reads on an HiSeq 3000 platform (Illumina).

Reads were analysed in the same way as the low input WGBS except without the --non-directional parameter of Bismark and in paired-end mode.

To investigate genome-wide DNA methylation differences between the wild-type and TET3 plants we employed the DMRcaller package (v1.22.0) for DMR identification (Catoni, Tsang et al. 2018). Initial analyses were conducted to optimize the method and parameters suitable for our dataset. Methylation reports for all contexts (CX_reports) generated by Bismark were imported into R (v4.0.3). Biological replicates for each genotype were individually assessed for coverage per cytosine and spatial correlation. Coverage per cytosine was evaluated in each context (CG, CHG, CHH). The spatial correlation of methylation levels between neighbouring cytosines was assessed using distances ranging from 1 bp to 50,000 bp. High spatial correlation was observed in the CG and CHG contexts over short distances (up to 200 bp), suggesting regional methylation patterns. The CHH context exhibited low spatial correlation. Based on these assessments, biological replicates were pooled to increase statistical power and coverage depth. We tested noise filter, bins, and neighbourhood DMR identification methods provided by DMRcaller on chromosome 6 (47 Mbp, the smallest in Tomato genome) using different DMR sizes (10 to 5000 bp) and minimum proportion differences (0.1 to 0.9) to determine the most effective approach for our data. As spatial correlation for CG and CHG was high, then Noise filter method was chosen with a sliding window of 100 bp, and a minimum methylation difference of 0.5, while for CHH what has a lower correlation, we chose the bins method with a 50 bp bins and a threshold of 0.3. Contiguous DMRs separated by

<200 bp (CG/CHG) or <100 bp (CHH) were merged, and only DMRs p < 0.01 and ≥4 cytosines per region were kept.

Small RNA sequencing

Tomato plants (*Solanum lycopersicum* cv. M82) were grown from seeds in compost (Levington M3) in plant growth chambers (24/18°C, 16 hr/8 hr day/night regime). Each sample for sRNA analysis is a pool of expanded leaves from 3 plants that are 4-week-old. 4 samples were collected for sRNAseq. Samples were grinded with Ns(l) mortar and pestle. Total RNA isolation was performed with TRIzol reagent (Invitrogen) and the Direct-zol RNA Miniprep (Zymo Research) according to the manufacturer's instructions. sRNA library preparation and sequencing was outsourced to Novogene (Cambridge, UK). Small RNA reads (fastq format) were subjected to 3' Illumina universal adaptor removal (trimming) using cutadapt (v1.18). Sequences with <15nt and >40nt in length were discarded. Reads were processed with BBDuk from BBMap (v38.22) (Bushnell 2014) to remove ribosomal sequences. The remaining sequences were mapped to the reference genome using Bowtie (v1.1.2) with 3 different approaches: 1/keeping uniquely mapping reads with no mistmaches, 2/ multi mapping reads with no mistmaches, 3/ multi mapping reads with one mistmach. Mapping files were converted to bigwig using samtools (v1.3) for all sRNAs and also specifically for sRNAs sizes 21, 22 and 24 nt.

Transcriptome Analysis

For mRNA-seq analysis, total RNA (1.33 µg) from each sample was used to purify polyA+ mRNA, which was then used for cDNA synthesis and amplification. RNA-seq libraries were prepared using the TruSeq RNA Sample Preparation Kit from Illumina (San Diego, CA). Sequencing of RNA libraries was carried out on an Illumina Nextseq500 (75 bp single end). Initial quality assessment of raw RNA-seq data was performed using FastQC (v0.11.5) (Andrews, Krueger et al. 2010). Adapter trimming and quality filtering were conducted using fastp (v0.23.2) (Chen 2023) with the parameters --cut_right, --cut_window_size 4, --cut_mean_quality 20, --length_required 40, and --overrepresentation_analysis. Post-trimming quality was reassessed with FastQC to ensure the effectiveness of trimming. MultiQC (v1.17) (Ewels, Magnusson et al. 2016) was used to aggregate quality reports from FastQC and fastp.

Trimmed reads were aligned to the genome using STAR aligner (v2.7.9a) (Dobin, Davis et al. 2013). The reference genome was indexed with STAR using the gene annotation file and the parameter --sjdbOverhang 75. Alignment was performed with the parameters '--outSAMtype BAM SortedByCoordinate, --quantMode GeneCounts --outWigType bedGraph' which generated sorted BAM files, gene count files, and coverage bedGraph files that were converted to BigWig format using bedGraphToBigWig. The --outWigStrand Unstranded option was used due to the unstranded nature of the libraries. Differential expression analysis was performed using DESeq2 (v1.42.1) (Love, Huber et al. 2014) from the unstranded counts. Genes were classified as significantly differentially expressed at an FDR < 0.01 and |log2 fold-change| > 1.

Transposable-element (TE) expression was assessed with TEtranscripts (v2.2.3) (Jin, Tam et al. 2015). The trimed reads used for gene expression analysis were realigned to the genome with STAR (parameters --outFilterMultimapNmax 100 --winAnchorMultimapNmax 200) so that up to 100 best multihits per read were retained. The resulting BAM files were provided to TEcount (part of TEtranscripts) in multi-mode with the libraries treated as unstranded (--stranded no). Gene models (ITAG 4.1) and a curated GTF of RepeatModeler-predicted repeats were supplied via the --GTF and --TE options, respectively, and reads were presorted on the fly (--sortByPos). TEcount simultaneously generated gene and TE read-count tables for each sample; these matrices were

later merged with the STAR-derived gene counts and subjected to differential expression analysis alongside the mRNA-seq data.

Data Analysis and Visualization

For visualizing genomic regions we used JBrowse2 (Diesh, Stevens et al. 2023). For figures, we used R version 4.0.5 (www.r-project.org) with packages ggplot2 (v3.5.1) (Wickham 2016), eulerr (v7.0.2) (Larsson 2024), ComplexHeatmap package (v2.8.0) (Gu 2022), tidyHeatmap (v 1.10.1) (Mangiola and Papenfuss 2020) and EnrichedHeatmap (v1.32.0) (Gu, Eils et al. 2018).

Declarations:

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Acknowledgements: We thank Gary Grant for help with plant growth; Liliana M. Costa for discussions and comments on the manuscript.

Funding: Supported by BBSRC grants (BB/N00194X/1 and BB/P02601X/1) to J.G-M.

Author's contributions: J.G-M. and P.M. conceived the project. J.A-S., J.E., K.O., N.H. and JG-M designed and conducted experiments. J.A-S. and J.E. analysed the data. JG-M supervised the work. J.A-S. and JG-M wrote the manuscript with input from the rest of the authors.

Competing interest: All authors declare that they have no competing interests.

Availability of Data and Materials: Sequencing data (WGBS and RNA-seq) that support the findings of this study have been deposited at the NCBI GEO under the accession code GSE303457.

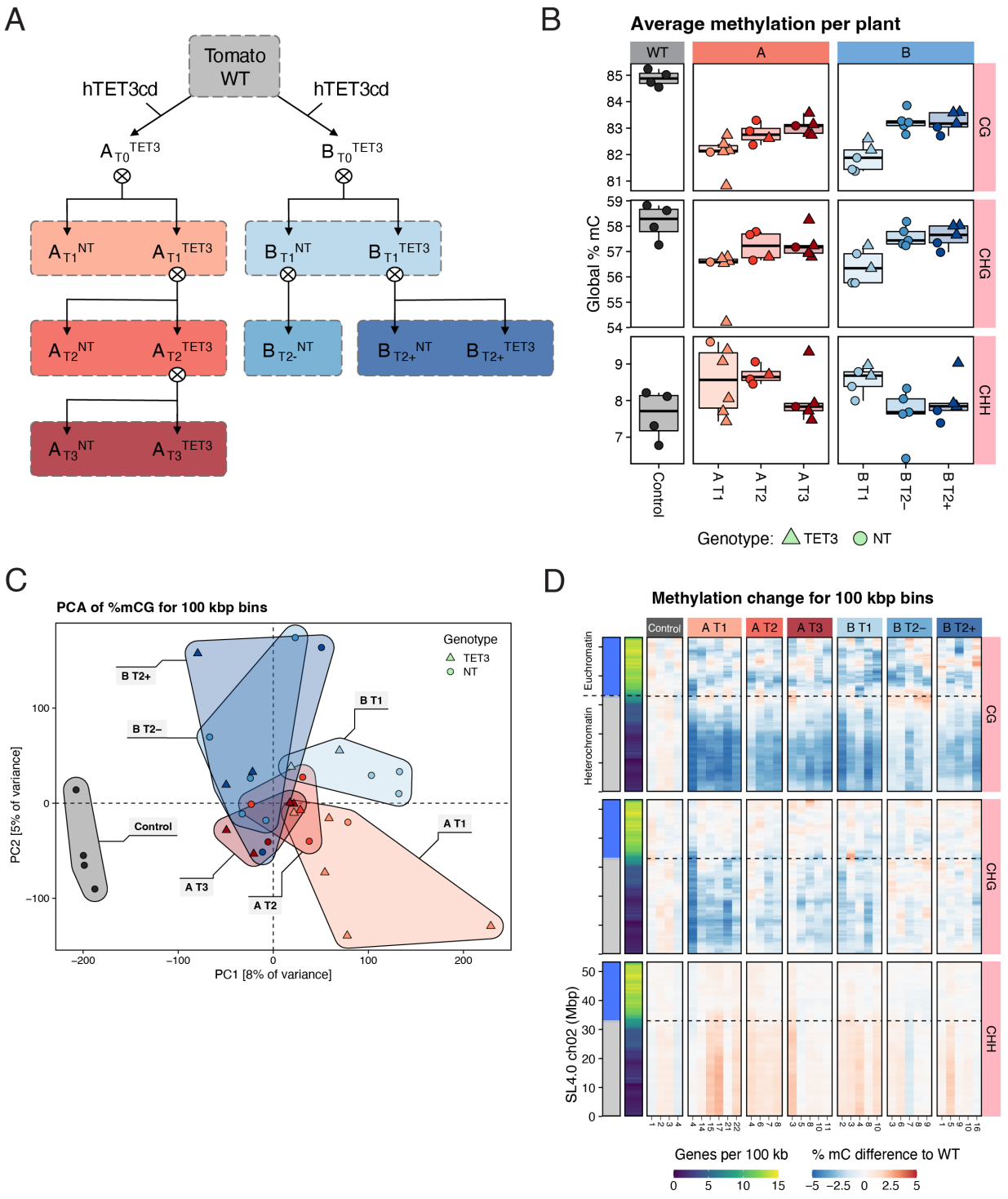


Figure 1: Characterization of Methylation Changes in Transgenic Tomato Lines Expressing hTET3cd:

- **(A)** Schematic representation of the pedigree of the two independent TET3 transgenic lines (Line A and Line B) and their propagation strategy across generations (T1, T2, T3).
- **(B)** Boxplots showing the methylation rate for each line and each generation (A T1, A T2, A T3, B T1, B T2-, B T2+) in different methylation contexts (CG, CHG, CHH). Each dot represents the

516 average methylation rate for the whole genome of an individual plant. Triangles represent
517 plants containing the transgene (TET3), and circles represent non-transgenic siblings (NT).

518 • **(C)** Principal Component Analysis (PCA) of CG methylation levels for 100 kbp bins. The
519 percentage of variance explained for each Principal Component (PC) is shown in square
520 brackets. Convex hulls surround each group (Control WT, A T1, A T2, A T3, B T1, B T2-, B T2+).
521 Triangles represent plants containing the transgene (TET3), and circles represent non-
522 transgenic siblings (NT).

523 • **(D)** Heatmap showing the percentage change in methylation difference between each plant and
524 the average of four control WT plants in different methylation contexts (CG, CHG, CHH) across
525 Chromosome 2 in 100 kbp bins. Chromosome 2 was chosen to illustrate the contrast between
526 repeat-rich heterochromatin and gene-rich euchromatin. The dotted line indicates the
527 boundary between heterochromatin and euchromatin. On the left, two heatmaps display the
528 compartmentalization of the genome: one showing heterochromatin and euchromatin as
529 defined by H3K27ac, and the other showing gene density per 100 kbp bins.

530

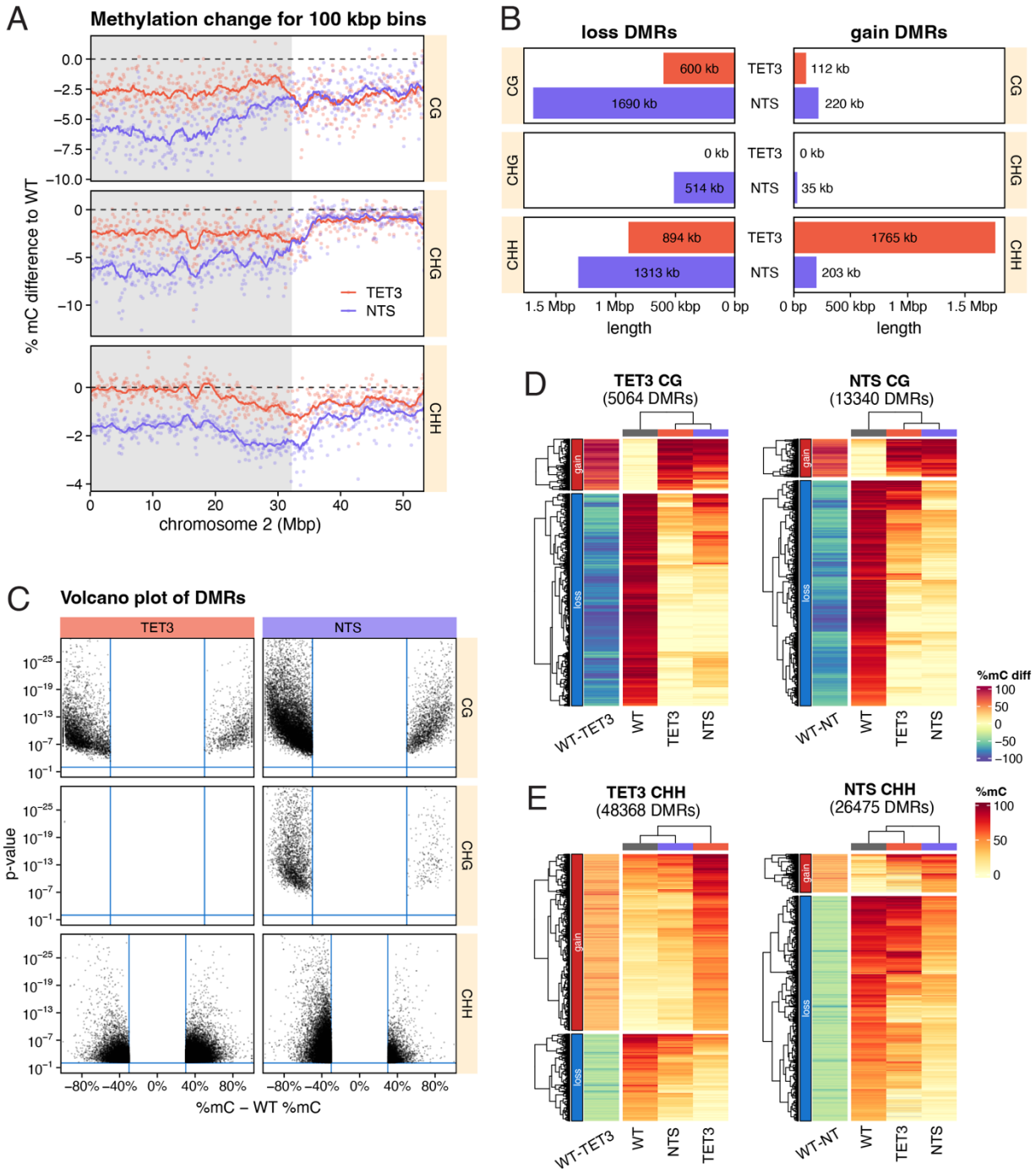


Figure 2: Detailed Methylation Analysis of A.T1 Plants:

- **(A)** Line plots showing the percentage methylation difference from the WT across Chromosome 2 for A.T1 plants (A.T1 TET3 and A.T1 NT) in three methylation contexts (CG, CHG, CHH). Dots represent average for 100 kbp bins, while lines represent moving averages across 2 Mbp, and the shaded region indicates heterochromatin.
- **(B)** Column plots displaying the number of Differentially Methylated Regions (DMRs) for each genotype in each methylation context (CG, CHG, CHH), categorized by gain and loss of methylation. The DMRs were identified using the Noise Filter method for CG and CHG contexts

and the Bins method for CHH context, with different minimum proportion difference (mpd) thresholds (**see Methods**).

- **(C)** Volcano plots of the DMRs for both samples (A.T1 TET3 and A.T1 NT) in the three methylation contexts (CG, CHG, CHH). The plots show the p-value against the percentage methylation change compared to WT, highlighting significant DMRs with blue lines indicating the thresholds for significance.
- **(D)** Heatmaps displaying the methylation levels of each DMR for CG contexts A.T1 plants compared to WT. Each heatmap shows the hierarchical clustering of DMRs, with colour gradients representing the level of methylation. The top and side dendrograms illustrate the clustering of samples and DMRs, respectively, indicating patterns of methylation gain and loss across different genomic regions. The column to the left of each heatmap represents the difference in methylation between WT and the A.T1 plant in each comparison.
- **(E)** Heatmaps displaying the methylation levels of each DMR for CHH context in A.T1 plants compared to WT. Each heatmap shows the hierarchical clustering of DMRs, with colour gradients representing the level of methylation. The top and side dendrograms illustrate the clustering of samples and DMRs, respectively, indicating patterns of methylation gain and loss across different genomic regions. The column to the left of each heatmap represents the difference in methylation between WT and the A.T1 plant in each comparison.

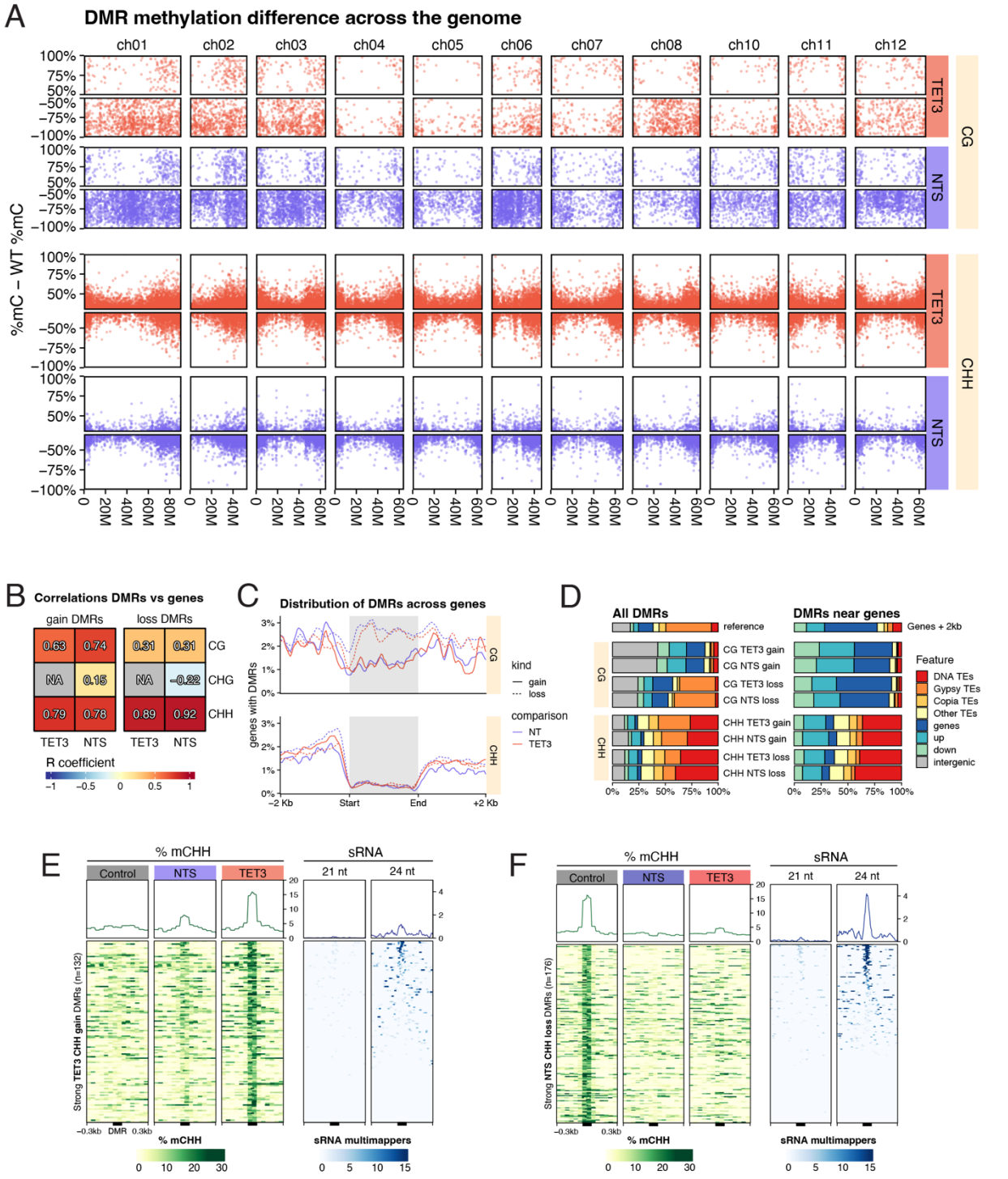


Figure 3: DMR Methylation Difference and Distribution Across the Genome

- **(A)** Plot showing the methylation differences for the previously computed DMRs across different chromosomes in both A.T1 TET3 (red) and A.T1 NT (blue) plants for CG and CHH contexts. Each dot represents a specific DMR.
- **(B)** Heatmap showing the correlation coefficients (R values) between the density of DMRs and genes for 3 Mbp windows, split by plant, direction and context.

- **(C)** Metagene plot showing the distribution of DMRs relative to gene positions showing gain (solid line) and loss (dotted line) for both A.T1 TET3 (red) and A.T1 NT (blue) plants for CG and CHH contexts. Grey shaded area represent gene body region, with upstream and downstream regions extending 2 kbp on either side.
- **(D)** DMR location for DNA TEs, Gypsy TEs, Copia, TE, other TEs, genic, intergenic, 2 kb upstream of genes and 3 kb downstream of genes across CG, and CHH contexts for each plant split by gain and loss of methylation. Left side shows all DMRs and right side only those withing 2 kb of genes.
- **(E and F)** Heatmaps showing the CHH methylation in WT, NTS and TET3, and the coverage of 21nt and 24nt small RNAs (sRNAs). Centred around TET3 CHH gain DMRs **(E)** and NTS CHH loss DMRs **(F)**. The x-axis spans from 300 bp upstream and downstream of DMRs.

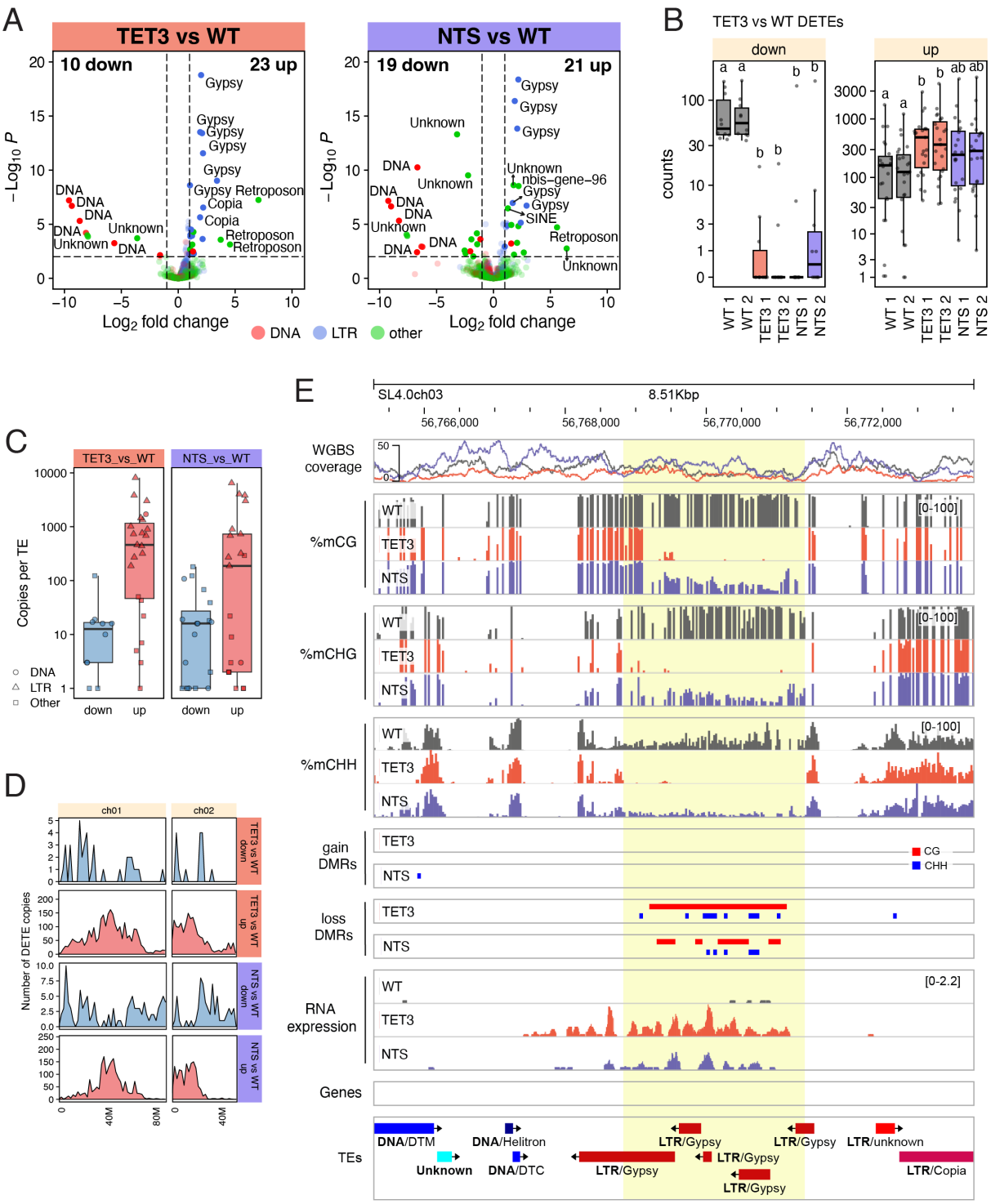


Figure 4: TEs are affected in TET3 progenies

- **(A)** Volcano plots showing the distribution of Differentially Expressed TEs (DETEs) in TET3 vs WT (left) and NTS vs WT (right). TEs with a log2 fold change > 1 and padj < 0.01 are highlighted, with colors and labels indicating the main TE classifications. Red: DNA TEs, Blue: LTRs TE, Green: other TEs
- **(B)** Boxplots of normalised expression of DETEs in TET3 vs WT in a semilogarithmic scale, split by downregulated and upregulated. Letters above each box show Tukey HSD groups (one-way

ANOVA, $P < 0.05$); samples that share a letter are not significantly different within the corresponding panel.

- **(C)** Boxplots of copy number of DETEs classified by comparison and direction. Shape represent group of TEs. Circle: DNA TEs, Triangle: LTRs TE, Square: other TEs.
- **(D)** Plot showing the distribution of DETEs copied across the chromosomes 1 and 2 in windows of 2 Mbp for both TET3 and NTS upregulated and downregulated.
- **(E)** Genome Browser view around a demethylated region in TET3 and NTS plants (yellow box) with Transcription activation of Transposons. Tracks for methylation in CG, CHG and CHH context and expression values for WT (Grey), TET3 (Red) and NTS (Blue) alongside DMRs, Gene and TE annotation are shown.

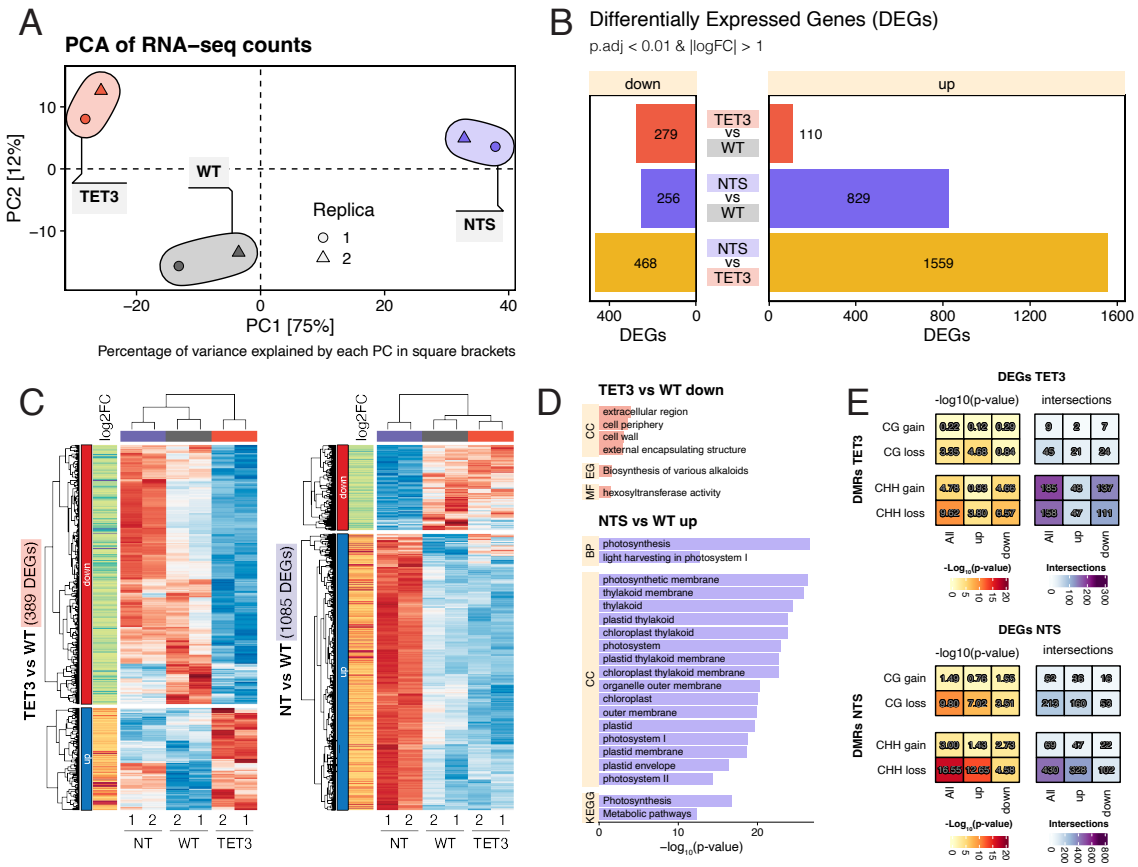
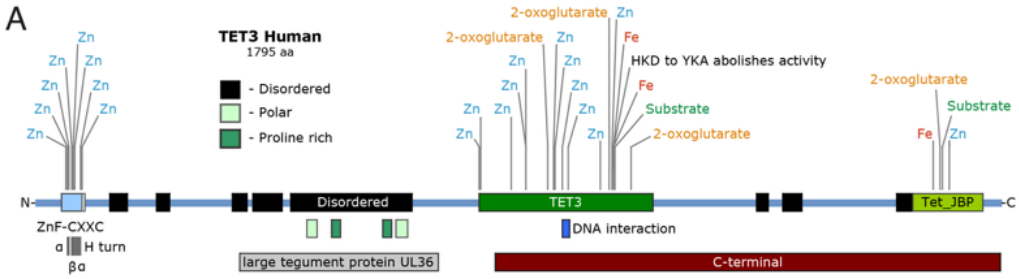
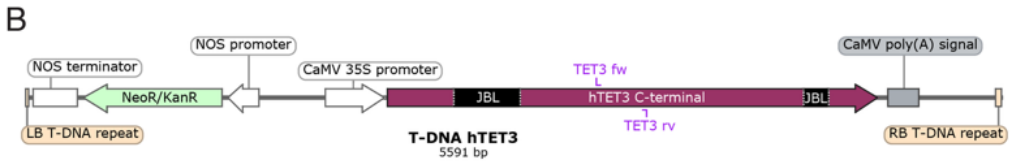


Figure 5: Transcriptome Analysis of TET3 and NT Plants and Their Connection to DMRs

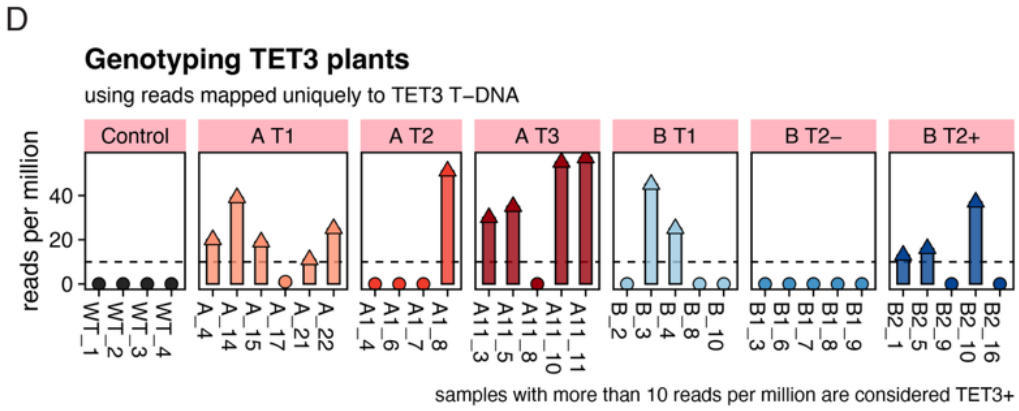
- **(A)** Principal Component Analysis (PCA) of RNA-seq counts showing variance between replicates and plant types (A.T1 TET3, A.T1 NT, and WT). Each point represents an individual sample, with replicates indicated by shape (circle: replicate 1; triangle: replicate 2).
- **(B)** Bar plot displaying the number of differentially expressed genes (DEGs) in pairwise comparisons: TET3 vs WT, NT vs WT, and NT vs TET3. Genes are separated into downregulated (left) and upregulated (right) categories. Criteria: padj < 0.01 and |log2 fold change| > 1.
- **(C)** Heatmaps of DEGs for TET3 vs WT (left) and NT vs WT (right), with rows representing genes and columns representing individual samples. Log2 fold changes (log2FC) are colour coded. Clustered groups highlight differences between plant types.
- **(D)** Gene Ontology (GO) and KEGG pathway enrichment analysis of DEGs. Top enriched categories for TET3 downregulated genes (left) and NT upregulated genes (right) are displayed with corresponding -log10(p-value).
- **(E)** Matrix plots showing the intersections between DEGs (downregulated and upregulated) and DMRs (CG gain/loss, CHG gain/loss, and CHH gain/loss) for both TET3 (top) and NT (bottom). Colour intensity represents significance (-log10(p-value)), and intersection counts are displayed within the matrix cells.





C

Line	Libraries sequenced	Libraries passed QC	TET3 positive
WT	6	4	0
A T1	8	6	5
A T2	5	4	1
A T3	5	5	4
B T1	8	5	2
B T2 -	6	5	0
B T2 +	9	5	3



<u>Figure 1 Supplement 1</u>: Generation and Initial Screening of Transgenic Tomato Plants Expressing hTET3cd:

- **(A)** Schematic representation of the human TET3 protein, highlighting the domains and the C-terminal catalytic domain (residues 859 to 1795) used for the demethylation of tomato plants.
- **(B)** Genetic map of the T-DNA insertion used in tomato plants, which includes the resistance gene and the C-terminal domain of the human TET3 (hTET3cd).
- **(C)** Table summarizing the number of plants per line used for the low-input whole-genome bisulfite sequencing (WGBS) experiment. The table shows the number of sequenced plants, the number of plants that passed quality control, and the number of TET3-positive plants.
- **(D)** Plot showing the Reads Per Million (RPM) mapped to the T-DNA in each plant for the low-input WGBS experiment. Samples with more than 10 reads per million were considered TET3 positive.

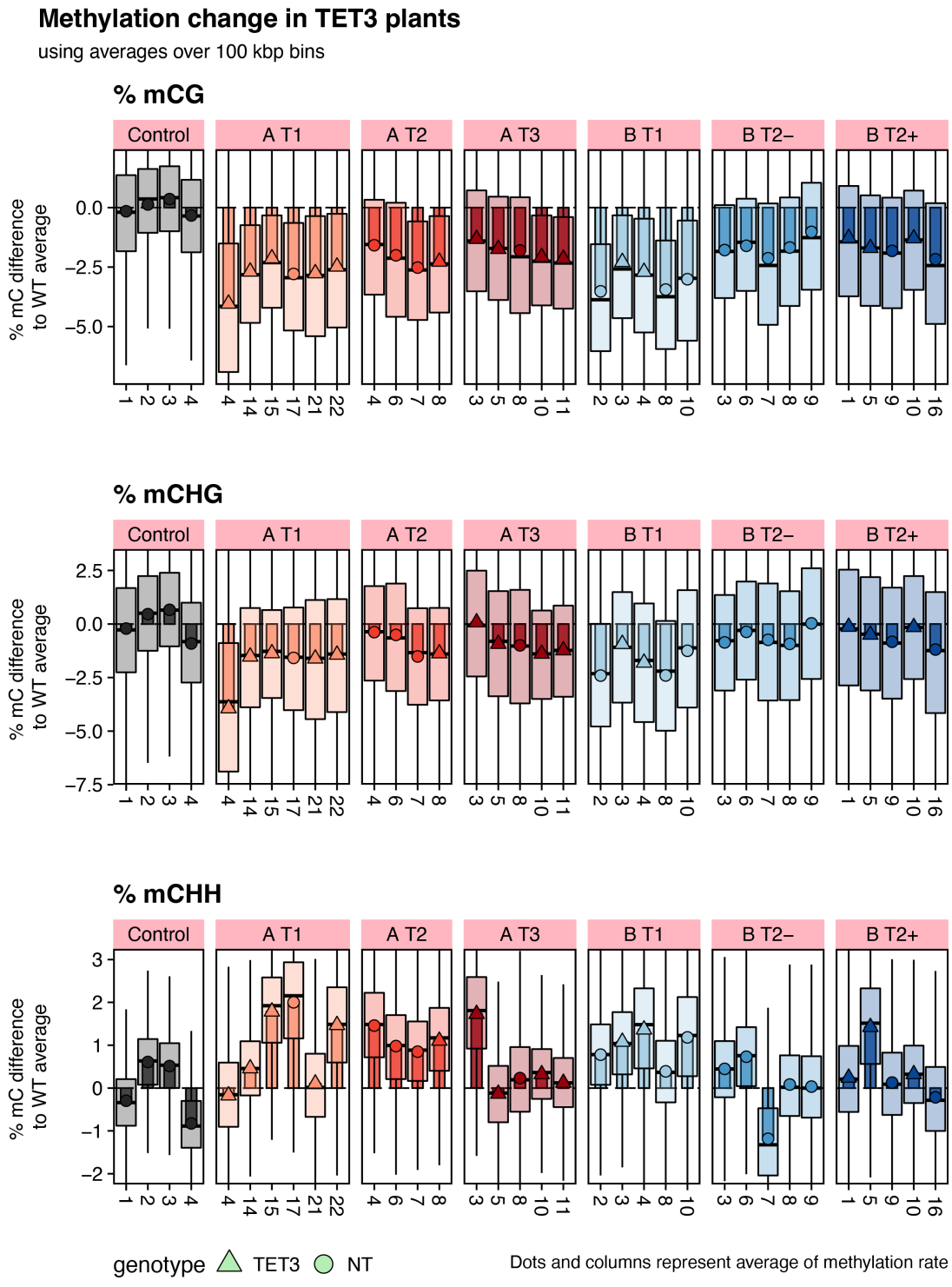
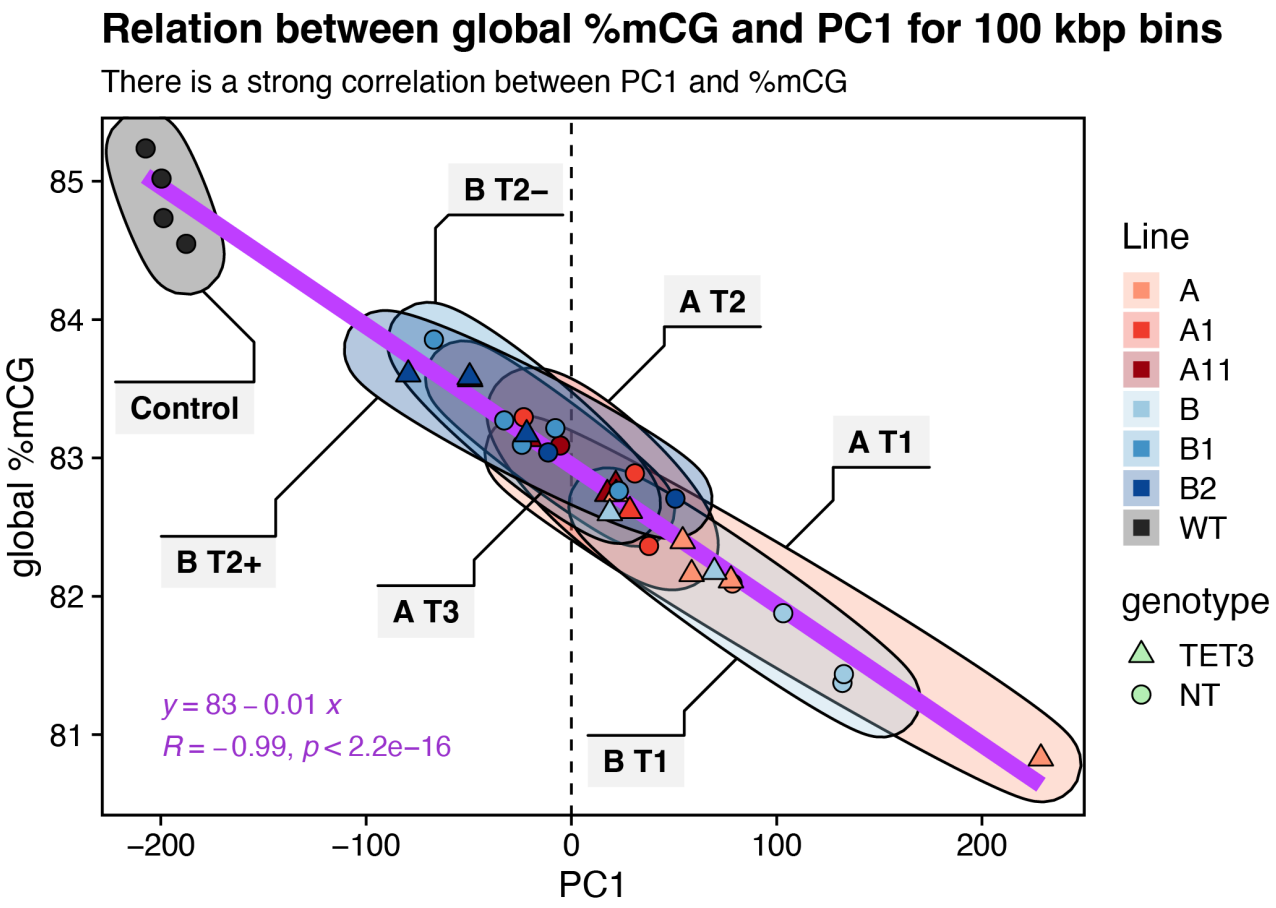


Figure 1 Supplement 2: Methylation Change in TET3 Plants

- Boxplots showing the percentage difference in methylation levels from the WT average for individual plants in different methylation contexts (CG, CHG, CHH) using averages over 100 kbp bins. Each dot and column represent the average methylation rate for an individual plant. Triangles represent plants containing the transgene (TET3), and circles represent non-transgenic siblings (NT). Groups for each generation of the transgenic lines (A T1, A T2, A T3, B T1, B T2-, B T2+) are displayed.



<u>Figure 1 Supplement 3:</u> Correlation Between Global %mCG and PC1 for 100 kbp Bins:

- Scatter plot showing the relationship between global CG methylation (%mCG) and the first principal component (PC1) for 100 kbp bins. Each point represents an individual plant, with the genotype indicated by triangles (TET3) and circles (NT). The data points are color-coded by group (control, A T1, A T2, A T3, B T1, B T2, B T2-). Convex hulls surround each group to indicate clustering. The regression line (purple) shows a strong Pearson correlation between PC1 and %mCG.

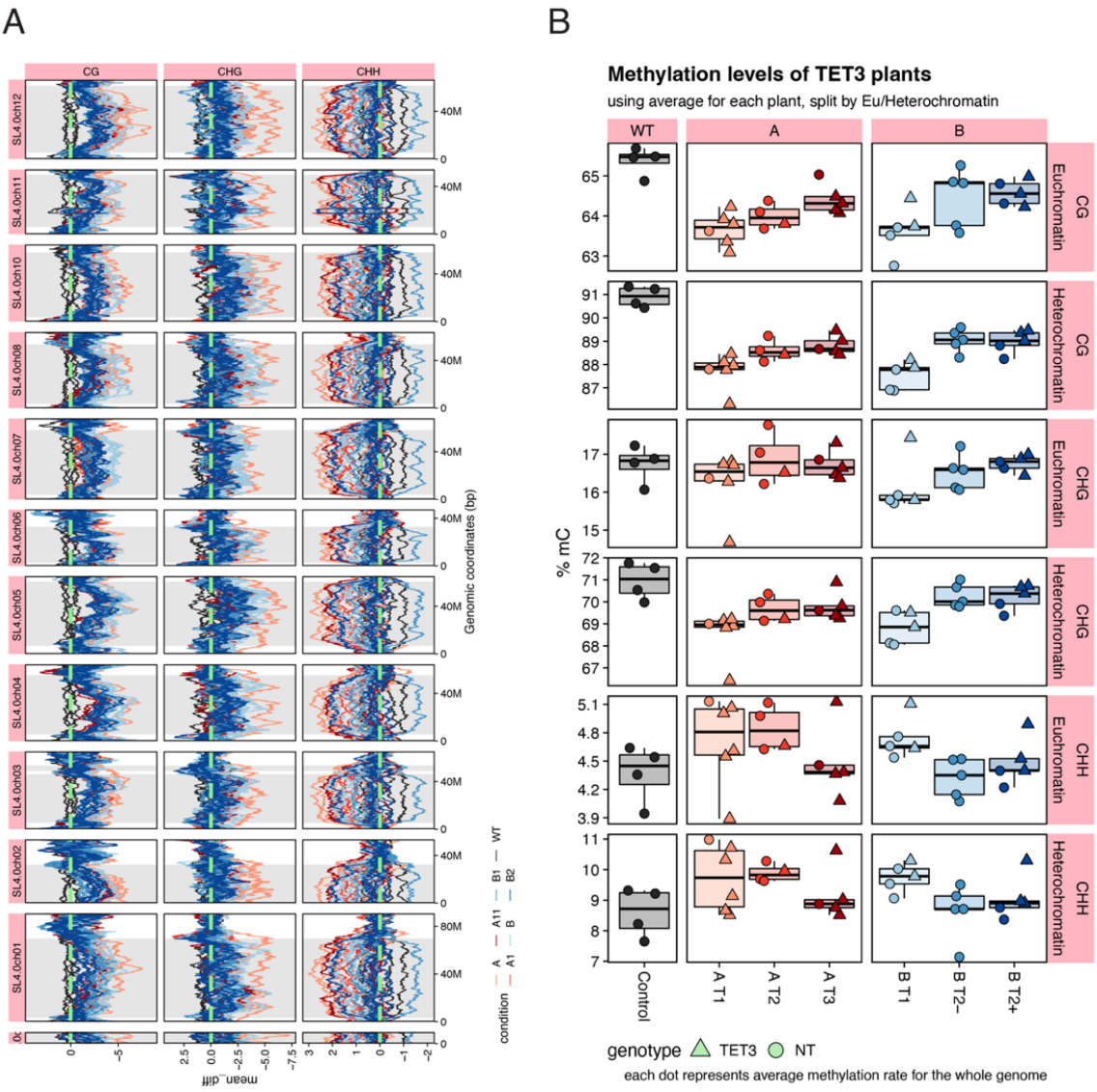


Figure 1 Supplement 4: Methylation Change in TET3 Plants Separated by Chromatin Context:

- **(A)** Line plots showing the percentage change in methylation levels compared to the WT average for individual plants in different methylation contexts (CG, CHG, CHH) across each chromosome for 100 kbp bins. Each line represents an individual plant, color-coded by transgenic line and generation. Shaded areas represent Heterochromatin.
- **(B)** Boxplots showing the percentage difference in methylation levels from the Control WT average for individual plants in different methylation contexts (CG, CHG, CHH), separated by euchromatin and heterochromatin regions. Each dot represents the average methylation rate for the whole genome of an individual plant. Triangles represent plants containing the transgene (TET3), and circles represent non-transgenic siblings (NT).

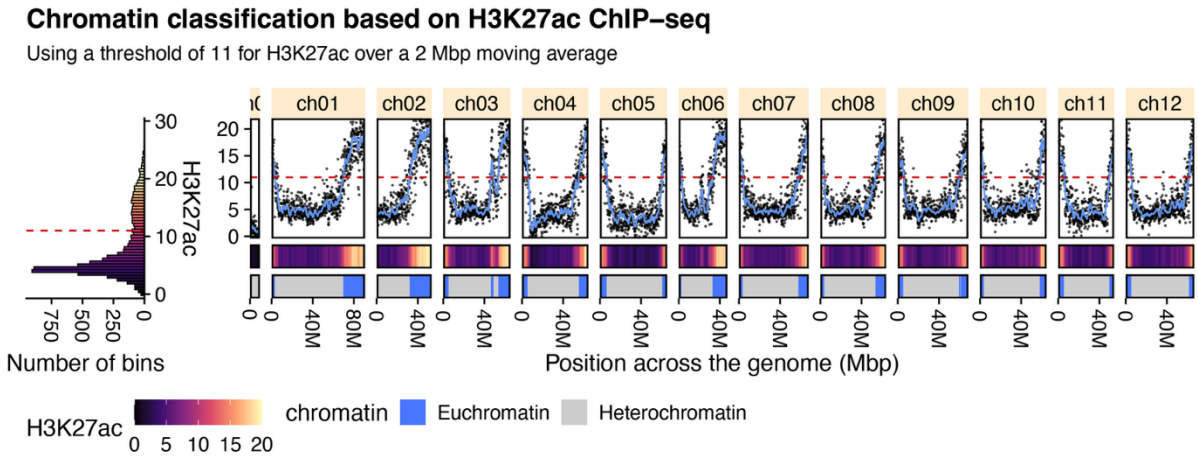
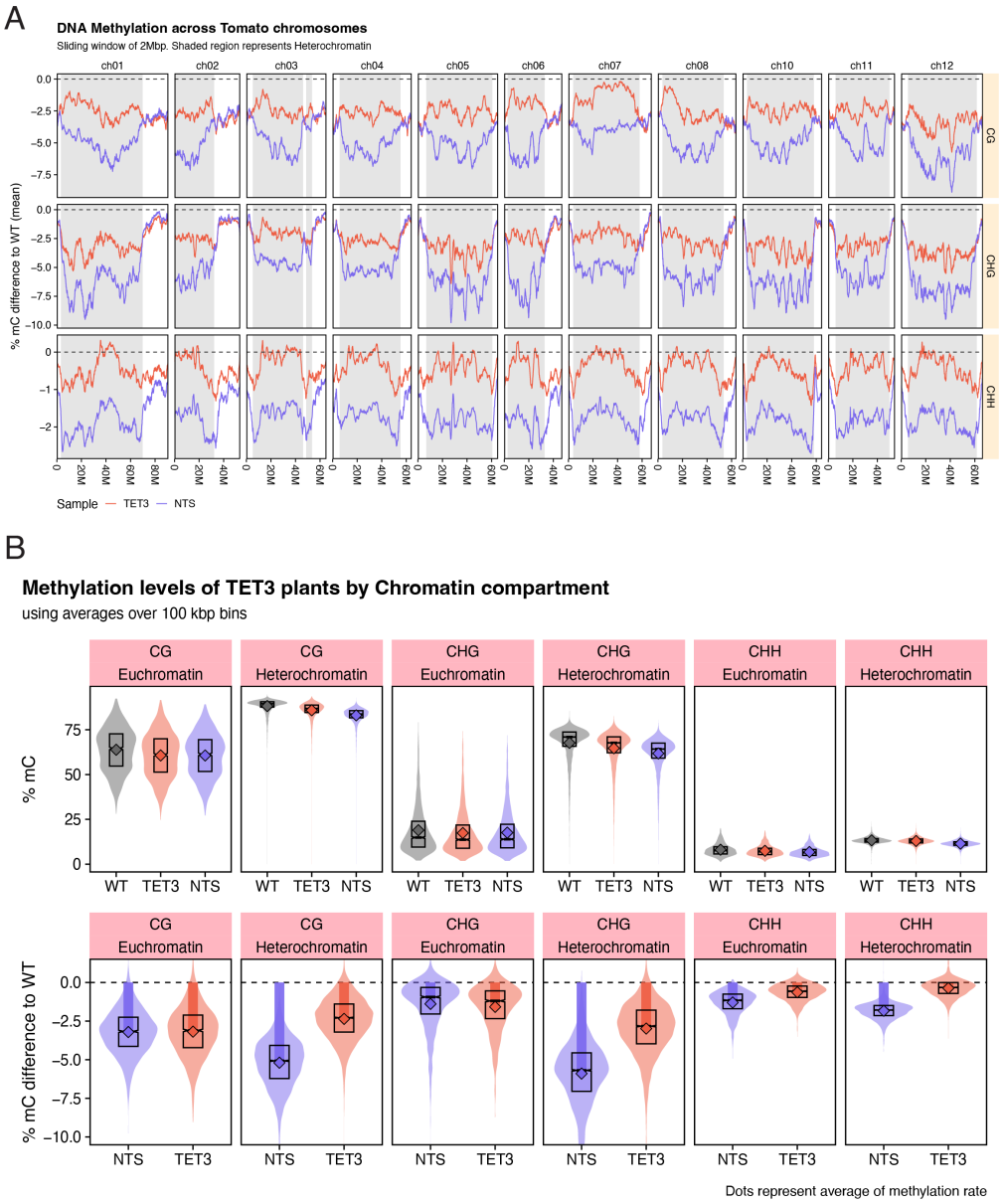


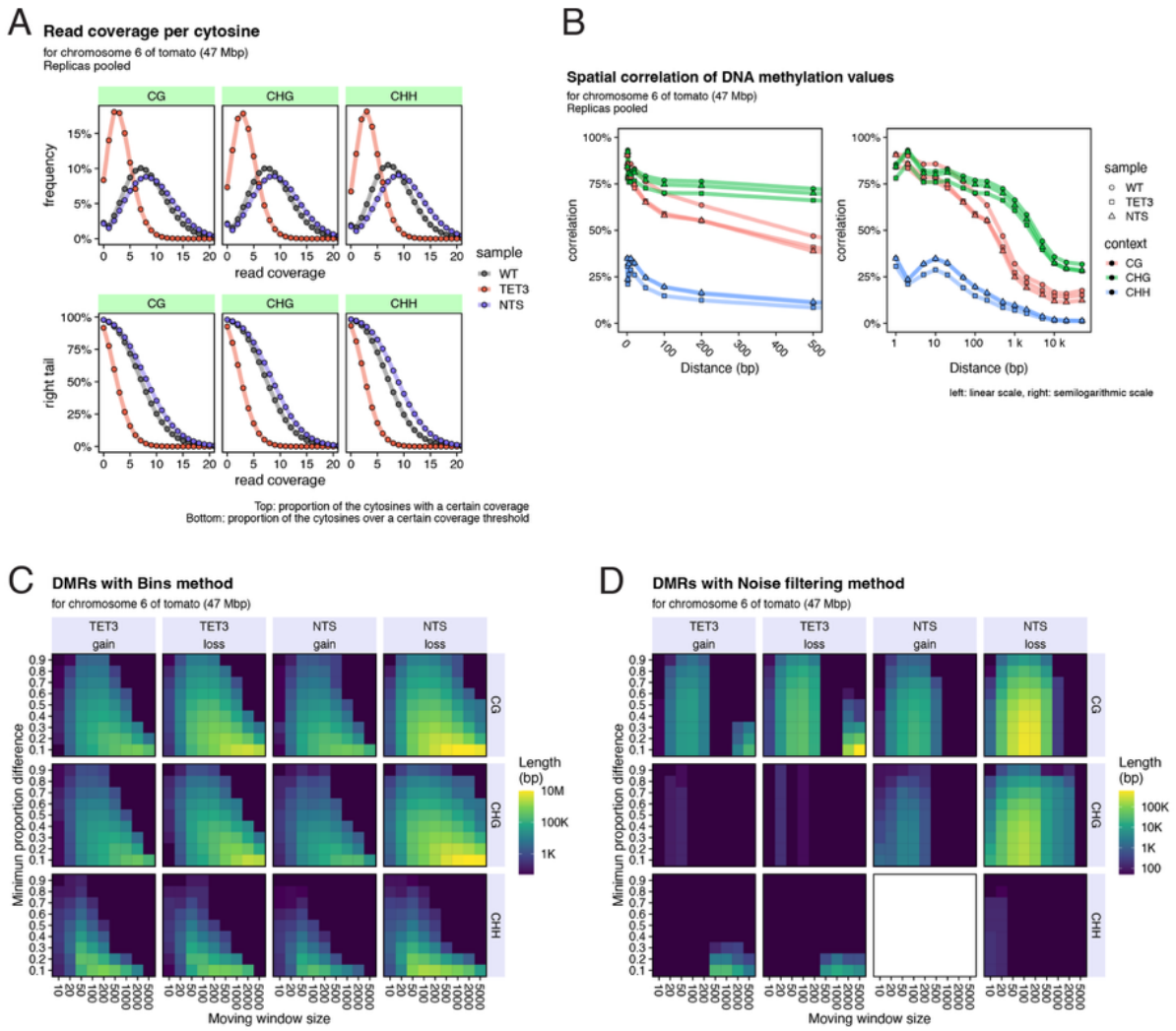
Figure 1 Supplement 5: Chromatin Classification Based on H3K27ac ChIP-seq

- Chromatin classification across the tomato genome based on H3K27ac ChIP-seq data, showing the H3K27ac signal (y-axis) across the genomic position (x-axis). A 2 Mbp moving average was used with a threshold of 11 for H3K27ac to differentiate between euchromatin and heterochromatin regions. The histogram on the left displays the distribution of H3K27ac signal across bins, with the red dashed line indicating the threshold. Below each chromosome plot, the classification into euchromatin and heterochromatin is shown, with blue representing euchromatin and grey representing heterochromatin.



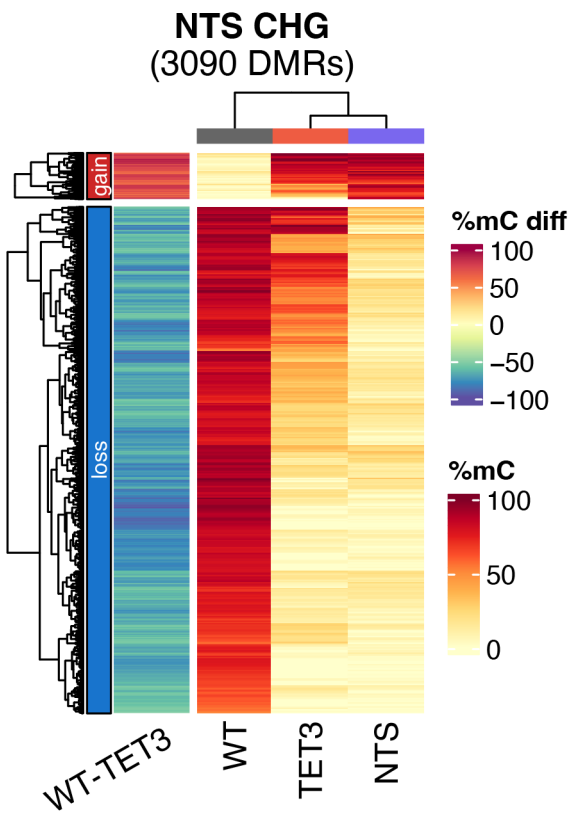
<u>Figure 2 Supplement 1:</u> Detailed DNA Methylation Analysis Across All Chromosomes and Chromatin Compartments:

- **(A)** Line plots showing the percentage methylation difference from the WT average in 100 kbp bins across all tomato chromosomes for the two A.T1 plants (A.T1 TET3 and A.T1 NT) in three methylation contexts (CG, CHG, CHH). The lines represent moving averages across 2 Mbp. Red lines indicate the TET3 positive plant (A.T1 TET3), and blue lines indicate the non-transgenic sibling (A.T1 NT). The shaded regions represent heterochromatin.
- **(B)** Violin plots showing methylation levels for each chromatin compartment (euchromatin and heterochromatin) in the three methylation contexts (CG, CHG, CHH). The top row shows the absolute methylation levels, and the bottom row shows the percentage difference in methylation compared to the WT control. Each dot represents the average methylation rate for 100 kbp bins. Red violins and dots represent the TET3 positive plant (A.T1 TET3), and blue violins and dots represent the non-transgenic sibling (A.T1 NT).



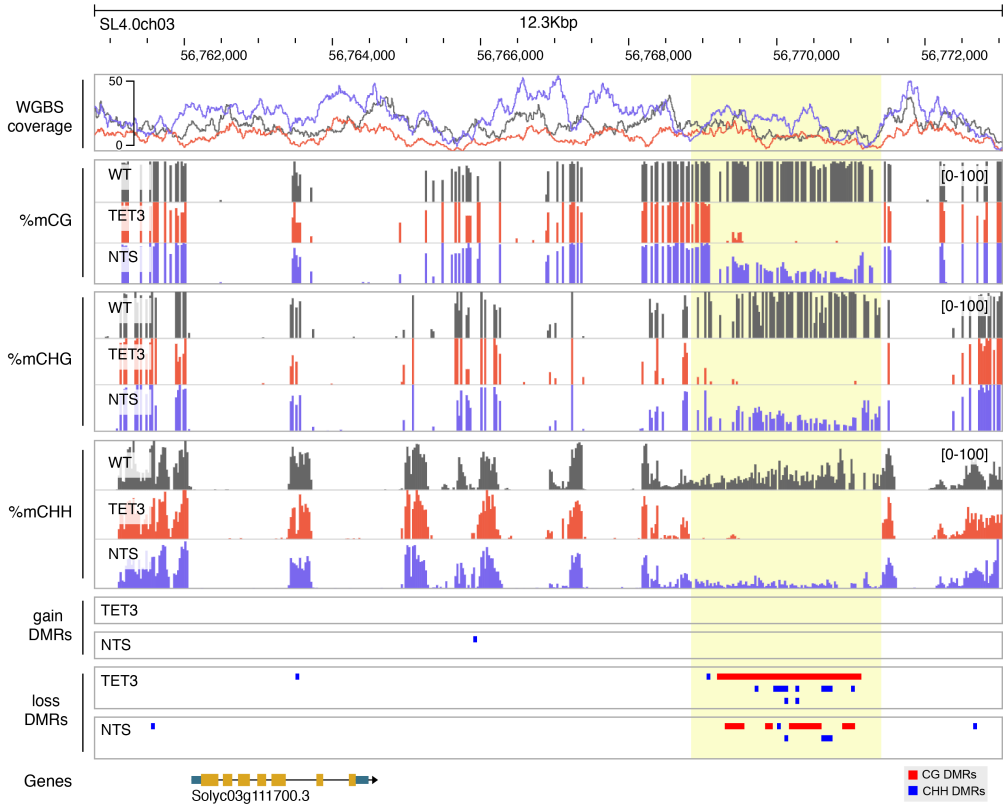
<u>Figure 2 Supplement 2</u>: Methylation Statistics and DMR Analysis:

- **(A)** Frequency distribution of read coverage per cytosine for chromosome 6 for the three methylation contexts (CG, CHG, CHH). The top row shows the proportion of cytosines with a certain coverage, and the bottom row shows the proportion of cytosines with coverage over a certain threshold. Red lines indicate A.T1 NT, blue lines indicate A.T1 TET3, and black lines indicate WT.
- **(B)** Spatial correlation of DNA methylation values: Spatial correlation of DNA methylation values for chromosome 6 for the three methylation contexts (CG, CHG, CHH). The left plot uses a linear scale, and the right plot uses a semi-logarithmic scale to show the correlation decay with distance. Red lines indicate CG, green lines indicate CHG, and blue lines indicate CHH.
- **(C)** Heatmaps showing the effect of varying the moving window size and the minimum proportion difference (mpd) on the length of DMRs detected using the Bins method for chromosome 6 (47 Mbp). DMRs are categorized by gain and loss of methylation in three contexts (CG, CHG, CHH) for both A.T1 NT and A.T1 TET3.
- **(D)** Heatmaps showing the effect of varying the moving window size and the minimum proportion difference (mpd) on the length of DMRs detected using the Noise filtering method for chromosome 6 (47 Mbp). DMRs are categorized by gain and loss of methylation in three contexts (CG, CHG, CHH) for both A.T1 NT and A.T1 TET3.



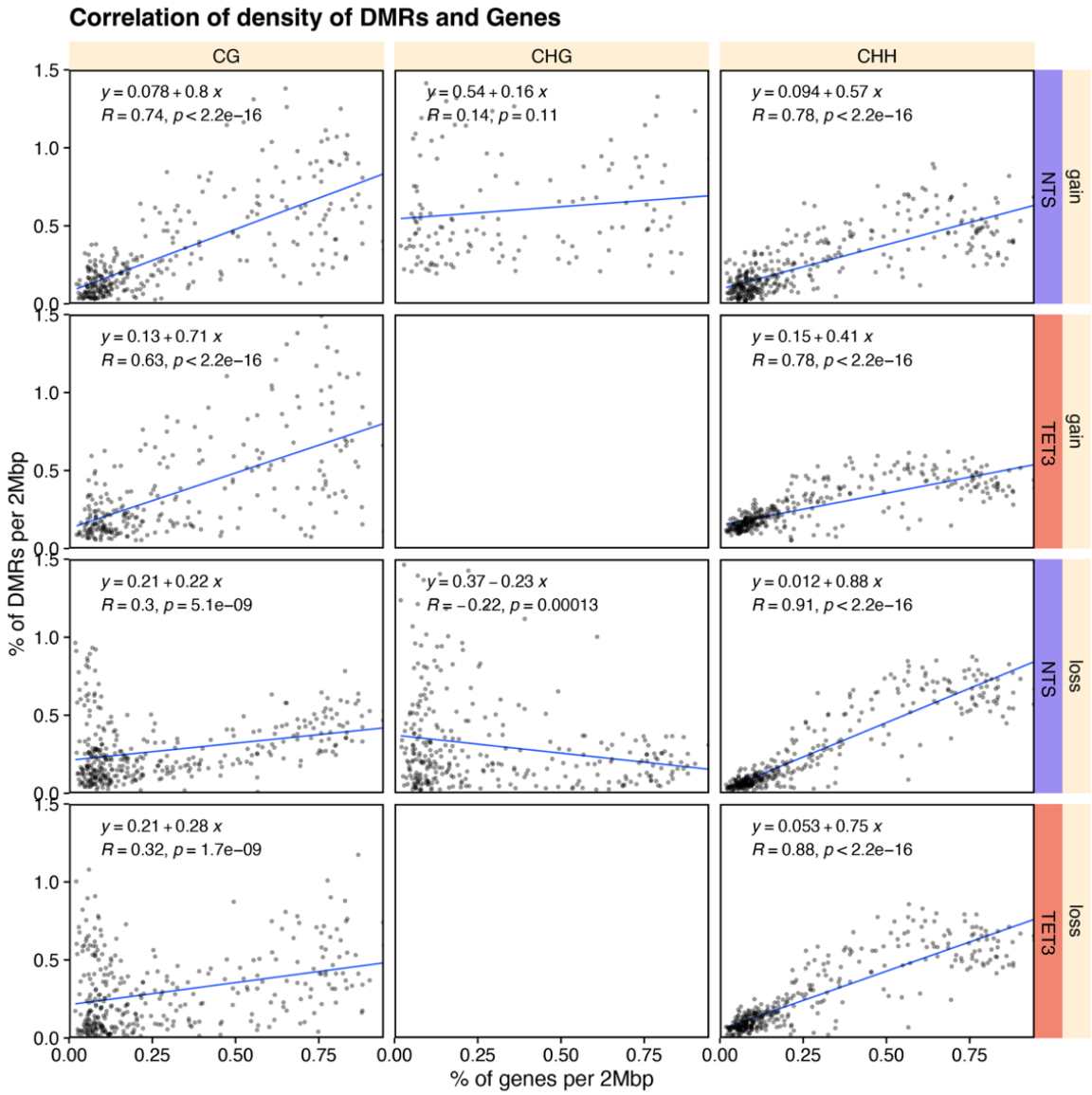
<u>Figure 2 Supplement 3:</u> Heatmap of Methylation Levels for CHG DMRs:

- Heatmap showing the methylation levels of differentially methylated regions (DMRs) in the CHG context for NTS compared to the WT control. The analysis identified 3090 DMRs. Each column represents an individual plant, and each row represents a DMR. The colour gradient indicates the level of methylation, with hierarchical clustering applied to both DMRs and samples to illustrate patterns of methylation gain and loss. The column to the left of the heatmap represents the difference in methylation between WT and NTS. No CHG DMRs were found in CHG context for the TET3 plant.



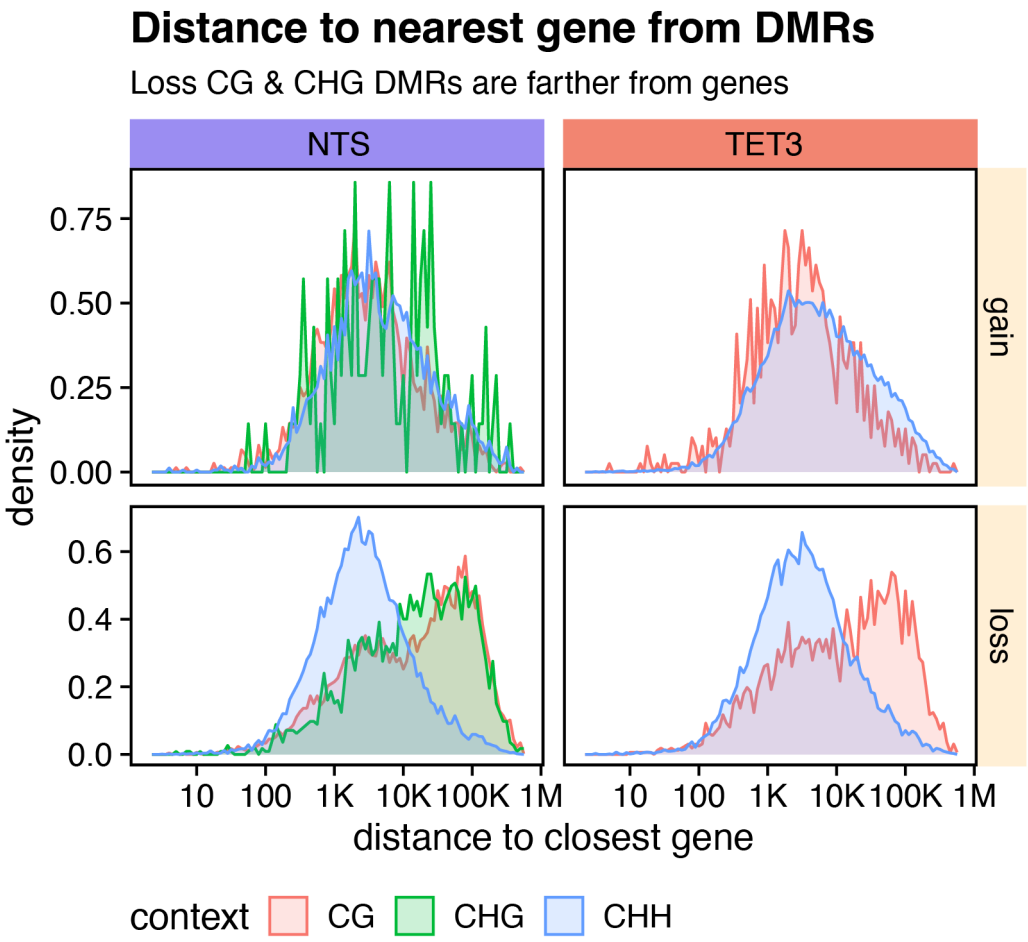
<u>Figure 2 Supplement 4</u>: Genome Browser View of a Differentially Methylated Region (DMR):

- Genome browser view of an example region on chromosome 3. The top panel shows WGBS coverage for three samples: WT (grey), TET3 (red), and NTS (blue). Tracks below display cytosine methylation levels in the CG, CHG, and CHH contexts (range 0–100% methylation) for each sample. The gain DMRs and loss DMRs panels are split my sample with red for CG DMRs and blue for CHH DMRs. Genes are annotated at the bottom. Highlighted region exhibits reduced DNA methylation across all cytosine contexts in TET3 compared to WT, with a smaller reduction in NTS plants.



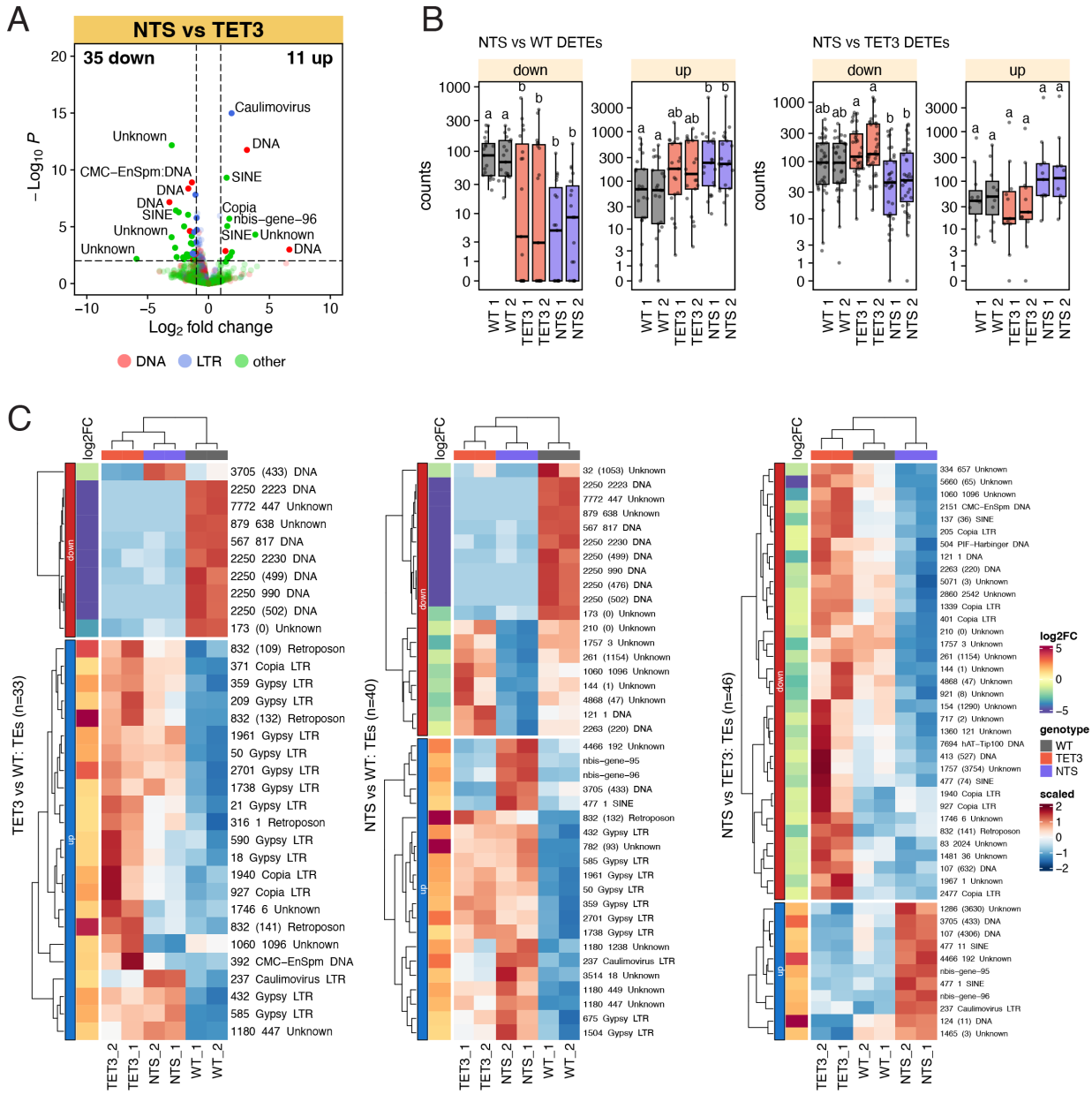
<u>Figure 3 Supplement 1</u>: DMR and Gene Density Across the Genome:

- Scatter plot showing correlation coefficients (R values) between the density of DMRs and genes for 3 Mbp windows, split by plant, direction and context. Each dot represents a 3 Mbp bin, Blue lines represent best-fit regression line, with their equations, R coefficient and p-value displayed.



<u>Figure 3 Supplement 2</u>: Distance form DMR to Genes:

- Density plots showing the distance of DMRs from the nearest gene in a semilogarithmic scale, split by plant (NTS and TET3), direction (gain and loss) and context (CG, CHG, CHH).



739

740 **Figure 4 Supplement 1: TE expression:**

741 • **(A)** Volcano plot showing the distribution of Differentially Expressed TEs (DETEs) in NTS vs TET3.
742 TEs with a log2 fold change > 1 and padj < 0.01 are highlighted, with colors and labels indicating
743 the main TE classifications. Red: DNA TEs, Blue: LTRs TE, Green: other TEs.

744 • **(B)** Boxplots of normalised expression of DETEs in NTS vs WT (left) and NTS vs TET3 (right) in a
745 semilogaritmic scale, split by downregulated and upregulated. Letters above each box show
746 Tukey HSD groups (one-way ANOVA, P < 0.05); samples that share a letter are not significantly
747 different within the corresponding panel.

748 • **(C)** Heatmaps of DETEs for TET3 vs WT (left), NT vs WT (centre) and NTS vs TET3 (right), with
749 rows representing TE families and columns representing individual samples. Log2 fold changes
750 (log2FC) are represented on the left of each heatmap. TE families are split between up and
751 down regulated and bot TE families and samples and clustered.

752

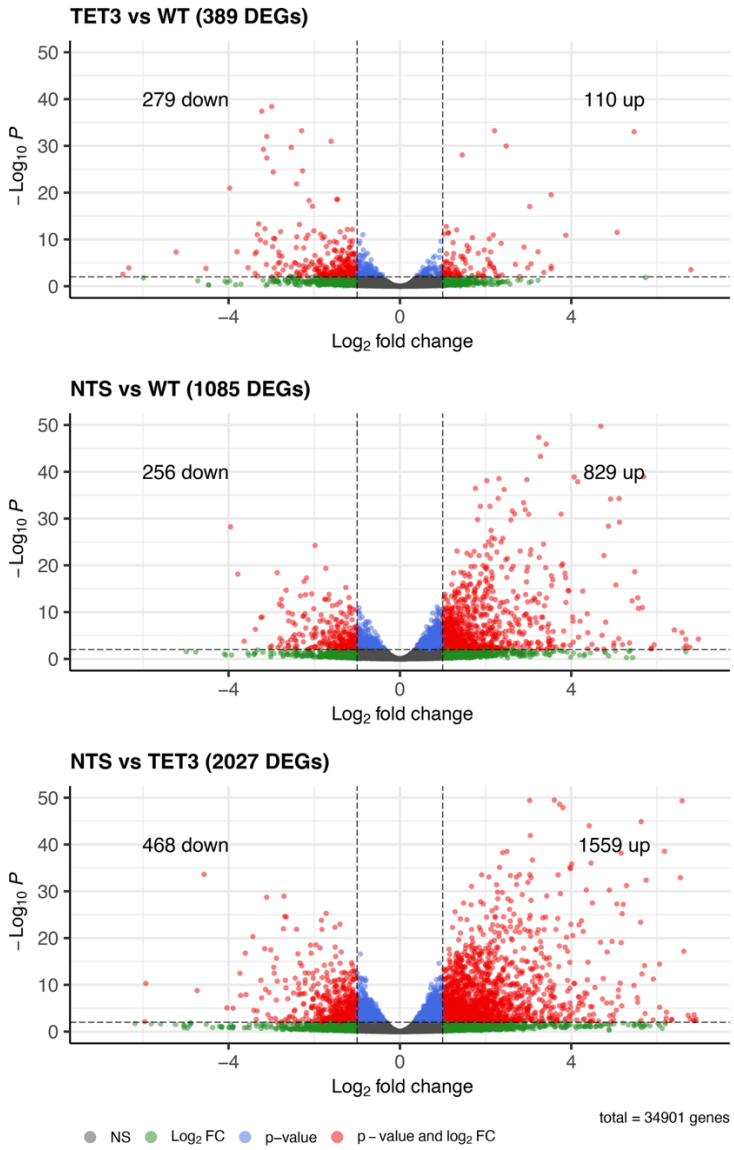


Figure 5 Supplement 1: Volcano plots of Genes

- Volcano plot showing the distribution of Differentially Expressed Genes (DEGs) in TET3 vs WT (top), NTS vs WT (middle) and NTS vs TET3 (bottom). Genes with a log2 fold change > 1 and padj < 0.01 are highlighted in red.

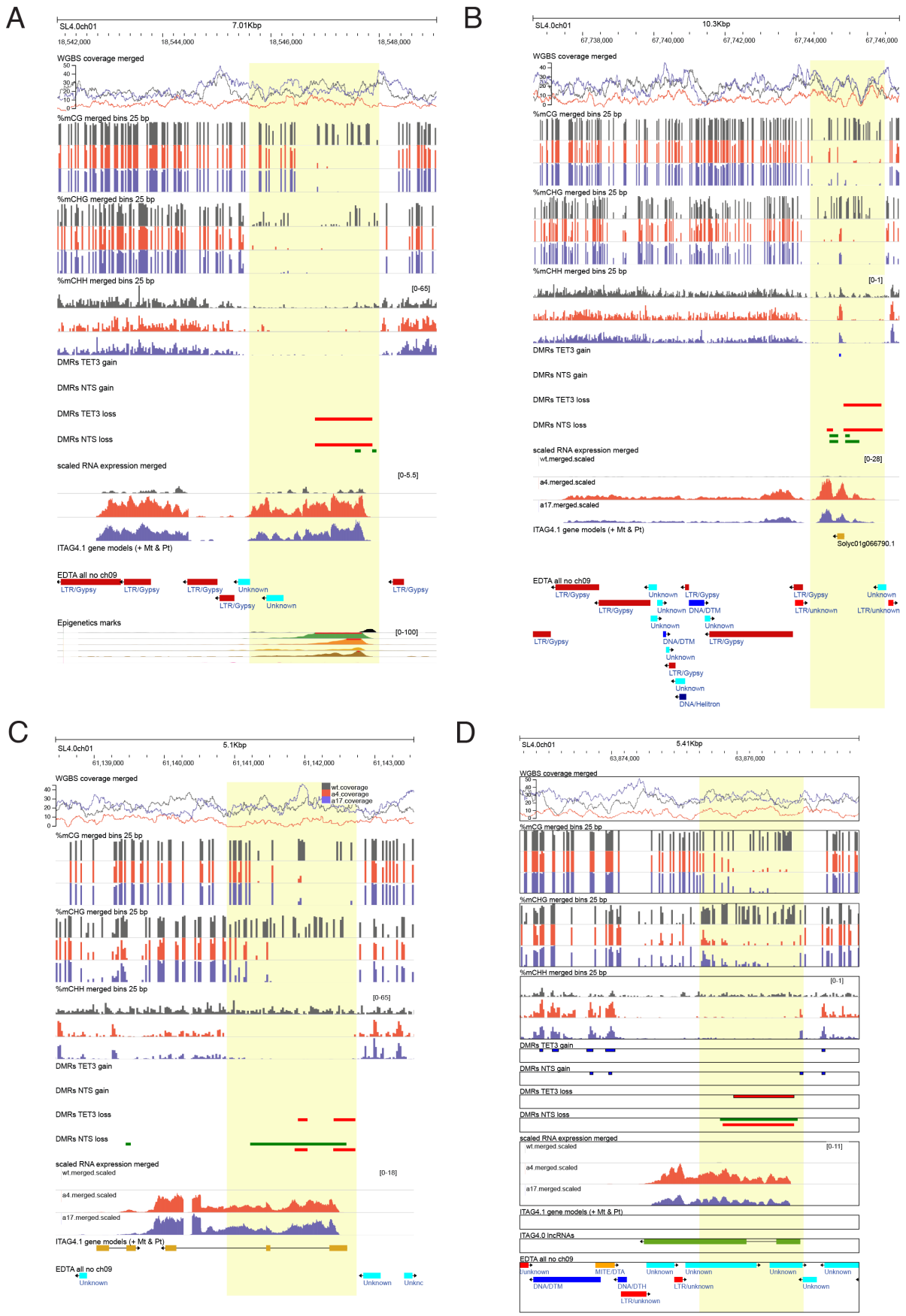

A
SL4.0ch01
7.01Kbp
18,542,000
18,544,000
18,546,000
18,548,000
WGBS coverage merged
%mCG merged bins 25 bp
%mCHG merged bins 25 bp
%mCHH merged bins 25 bp
[0-65]
DMRs TET3 gain
DMRs NTS gain
DMRs TET3 loss
DMRs NTS loss
scaled RNA expression merged
[0-5.5]
ITAG4.1 gene models (+ Mt & Pt)
EDTA all no ch09
LTR/Gypsy
Unknown
Epigenetics marks
[0-100]
B
SL4.0ch01
10.3Kbp
67,738,000
67,740,000
67,742,000
67,744,000
67,746,000
WGBS coverage merged
%mCG merged bins 25 bp
%mCHG merged bins 25 bp
%mCHH merged bins 25 bp
[0-1]
DMRs TET3 gain
DMRs NTS gain
DMRs TET3 loss
DMRs NTS loss
scaled RNA expression merged
wt.merged.scaled
[0-28]
a4.merged.scaled
a17.merged.scaled
ITAG4.1 gene models (+ Mt & Pt)
Solyc01g066790.1
EDTA all no ch09
LTR/Gypsy
Unknown
DNA/DTM
LTR/unknown
C
SL4.0ch01
5.1Kbp
61,139,000
61,140,000
61,141,000
61,142,000
61,143,000
WGBS coverage merged
wt.coverage
a4.coverage
a17.coverage
%mCG merged bins 25 bp
%mCHG merged bins 25 bp
%mCHH merged bins 25 bp
[0-65]
DMRs TET3 gain
DMRs NTS gain
DMRs TET3 loss
DMRs NTS loss
scaled RNA expression merged
wt.merged.scaled
[0-18]
a4.merged.scaled
a17.merged.scaled
ITAG4.1 gene models (+ Mt & Pt)
EDTA all no ch09
Unknown
D
SL4.0ch01
5.41Kbp
63,874,000
63,876,000
WGBS coverage merged
%mCG merged bins 25 bp
%mCHG merged bins 25 bp
%mCHH merged bins 25 bp
[0-1]
DMRs TET3 gain
DMRs NTS gain
DMRs TET3 loss
DMRs NTS loss
scaled RNA expression merged
wt.merged.scaled
[0-11]
a4.merged.scaled
a17.merged.scaled
ITAG4.1 gene models (+ Mt & Pt)
ITAG4.0 lncRNAs
EDTA all no ch09
Unknown
MITE/DTA
DNA/DTM
DNA/DTH
LTR/unknown

Figure 5 Supplement 2: Genome views of DMRs

- **(A-D)** Genome Browser views around demethylated regions in TET3 and NTS plants (yellow boxes) with Transcription activation of Genes and Transposons. Tracks for methylation in CG, CHG and CHH context and expression values for WT (Grey), TET3 (Red) and NTS (Blue) alongside DMRs, Gene, long non-coding RNAs and TE annotation are shown.

References

Akimoto, K., H. Katakami, H. J. Kim, E. Ogawa, C. M. Sano, Y. Wada and H. Sano (2007). "Epigenetic Inheritance in Rice Plants." Annals of Botany **100**(2): 205-205.

Amabile, A., A. Migliara, P. Capasso, M. Biffi, D. Cittaro, L. Naldini and A. Lombardo (2016). "Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing." Cell **167**(1): 219-219.

Andrews, S., F. Krueger, A. Segonds-Pichon, L. Biggins, C. Krueger and S. Wingett (2010). FastQC: a quality control tool for high throughput sequence data. Babraham, UK, Babraham Institute.

Andrews, S., F. Krueger, A. Segonds-Pichon, L. Biggins, C. Krueger and S. Wingett (2010). "FastQC: a quality control tool for high throughput sequence data." Babraham Institute.

Bartee, L., F. Malagnac and J. Bender (2001). "Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene." Genes & Development **15**(14): 1753-1758.

Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. United States.

Calarco, J. P., F. Borges, M. T. Donoghue, F. Van Ex, P. E. Jullien, T. Lopes, R. Gardner, F. Berger, J. A. Feijo, J. D. Becker and R. A. Martienssen (2012). "Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA." Cell **151**(1): 194-205.

Catoni, M., J. M. F. Tsang, A. P. Greco and N. R. Zabet (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**.

Chen, S. (2023). "Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp." iMeta **2**(2): e107.

Chen, S. (2023). "Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp." iMeta **2**(2): e107-e107.

Choi, Y., M. Gehring, L. Johnson, M. Hannon, J. J. Harada, R. B. Goldberg, S. E. Jacobsen and R. L. Fischer (2002). "DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis." Cell **110**(1): 33-42.

Christman, J. K. (2002). "5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy." Oncogene 2002 21:35 **21**(35): 5483-5495.

Diesh, C., G. J. Stevens, P. Xie, T. D. J. Martinez, E. A. Hershberg, A. Leung, E. Guo, S. Dider, J. Zhang, C. Bridge, G. Hogue, A. Duncan, M. Morgan, T. Flores, B. N. Bimber, R. Haw, S. Cain, R. M. Buels, L. D. Stein and I. H. Holmes (2023). "JBrowse 2: a modular genome browser with views of synteny and structural variation." Genome Biology **24**: 1-21.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson and T. R. Gingeras (2013). "STAR: ultrafast universal RNA-seq aligner." Bioinformatics **29**: 15-21.

Dowen, R. H., M. Pelizzola, R. J. Schmitz, R. Lister, J. M. Dowen, J. R. Nery, J. E. Dixon and J. R. Ecker (2012). "Widespread dynamic DNA methylation in response to biotic stress." Proceedings of the National Academy of Sciences **109**(32): E2183-E2191.

Erdmann, R. M. and C. L. Picard (2020). "RNA-directed DNA Methylation." PLOS Genetics **16**(10): e1009034-e1009034.

Ewels, P., M. Magnusson, S. Lundin and M. Käller (2016). "MultiQC: summarize analysis results for multiple tools and samples in a single report." Bioinformatics **32**(19): 3047-3048.

Finnegan, E. J., W. J. Peacock and E. S. Dennis (1996). "Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development." Proceedings of the National Academy of Sciences **93**(16): 8449-8454.

Gallego-Bartolomé, J., J. Gardiner, W. Liu, A. Papikian, B. Ghoshal, H. Y. Kuo, J. M. C. Zhao, D. J. Segal and S. E. Jacobsen (2018). "Targeted DNA demethylation of the arabidopsis genome using the human TET1 catalytic domain." Proceedings of the National Academy of Sciences of the United States of America **115**(9): E2125-E2134.

Ghosh, A., A. U. Igamberdiev and S. C. Debnath (2021). "Tissue culture-induced DNA methylation in crop plants: a review." Molecular biology reports **48**(1): 823-841.

Gong, Z., T. Morales-Ruiz, R. R. Ariza, T. Roldán-Arjona, L. David and J. K. Zhu (2002). "ROS1, a repressor of transcriptional gene silencing in Arabidopsis, encodes a DNA glycosylase/lyase." Cell **111**(6): 803-814.

Gouil, Q. and D. C. Baulcombe (2016). "DNA Methylation Signatures of the Plant Chromomethyltransferases." PLOS Genetics **12**(12): e1006526-e1006526.

Gu, Z. (2022). "Complex heatmap visualization." iMeta **1**: e43.

Gu, Z., R. Eils, M. Schlesner and N. Ishaque (2018). "EnrichedHeatmap: An R/Bioconductor package for comprehensive visualization of genomic signal associations." BMC Genomics **19**: 234–234.

Gutierrez-Marcos, J. F. and H. G. Dickinson (2012). "Epigenetic reprogramming in plant reproductive lineages." Plant Cell Physiol **53**(5): 817-823.

Hackett, J. A. and M. Azim Surani (2013). "DNA methylation dynamics during the mammalian life cycle." Philosophical Transactions of the Royal Society B: Biological Sciences **368**(1609).

Heard, E. and R. A. Martienssen (2014). "Transgenerational epigenetic inheritance: myths and mechanisms." Cell **157**(1): 95-109.

Henderson, I. R. and S. E. Jacobsen (2007). "Epigenetic inheritance in plants." Nature **442**(7106): 993-1000.

Hollwey, E., A. Briffa, M. Howard and D. Zilberman (2023). "Concepts, mechanisms and implications of long-term epigenetic inheritance." Current Opinion in Genetics & Development **81**: 102087-102087.

Hollwey, E., S. Out, M. R. Watson, I. Heidmann and P. Meyer (2017). "TET3-mediated demethylation in tomato activates expression of a CETS gene that stimulates vegetative growth." Plant Direct **1**(4): e00022.

Hollwey, E., M. Watson, P. Meyer, E. Hollwey, M. Watson and P. Meyer (2016). "Expression of the C-Terminal Domain of Mammalian TET3 DNA Dioxygenase in Arabidopsis thaliana Induces Heritable Methylation Changes at rDNA Loci." Advances in Bioscience and Biotechnology **7**(5): 243-250.

Hosmani, P. S., M. Flores-Gonzalez, H. Van De Geest, F. Maumus, L. V. Bakker, E. Schijlen, J. Van Haarst, J. Cordewener, G. Sanchez-Perez, S. Peters, Z. Fei, J. J. Giovannoni, L. A. Mueller and S. Saha (2019). An improved de novo assembly and annotation of the tomato reference genome using single-molecule sequencing, Hi-C proximity ligation and optical maps, Cold Spring Harbor Laboratory.

Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang (2011). "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." Science **333**(6047): 1300-1303.

Ji, L., W. T. Jordan, X. Shi, L. Hu, C. He and R. J. Schmitz (2018). "TET-mediated epimutagenesis of the Arabidopsis thaliana methylome." Nature Communications **9**(1): 1-9.

Jin, Y., O. H. Tam, E. Paniagua and M. Hammell (2015). "TEtranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets." Bioinformatics **31**(22): 3593-3599.

Johannes, F., E. Porcher, F. K. Teixeira, V. Saliba-Colombani, M. Simon, N. Agier, A. Bulski, J. Albuisson, F. Heredia, P. Audigier, D. Bouchez, C. Dillmann, P. Guerche, F. Hospital and V. Colot (2009). "Assessing the impact of transgenerational epigenetic variation on complex traits." PLoS Genet **5**(6): e1000530.

Kankel, M. W., D. E. Ramsey, T. L. Stokes, S. K. Flowers, J. R. Haag, J. A. Jeddeloh, N. C. Riddle, M. L. Verbsky and E. J. Richards (2003). "Arabidopsis MET1 cytosine methyltransferase mutants." Genetics **163**(3): 1109-1122.

Krueger, F. and S. R. Andrews (2011). "Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications." bioinformatics **27**(11): 1571-1572.

Larsson, J. (2024). "eulerr: Area-Proportional Euler and Venn Diagrams with Ellipses."

Law, J. A. and S. E. Jacobsen (2010). "Establishing, maintaining and modifying DNA methylation patterns in plants and animals." Nature Reviews Genetics **11**(3): 204-220.

Leeke, B. J., W. Varsally, S. Ogushi, J. Zohren, S. Menchero, A. Courtois, D. M. Snell, A. Teissandier, O. Ojarikre, S. K. Mahadevaiah, F. Decarpentrie, R. J. Oakey, J. L. VandeBerg and J. M. A. Turner (2025). "Divergent DNA methylation dynamics in marsupial and eutherian embryos." Nature.

Liu, X. S., H. Wu, X. Ji, Y. Stelzer, X. Wu, S. Czauderna, J. Shu, D. Dadon, R. A. Young and R. Jaenisch (2016). "Editing DNA Methylation in the Mammalian Genome." Cell **167**(1): 233-247.e217.

Love, M. I., W. Huber and S. Anders (2014). "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." Genome Biology **15**(12): 550-550.

Mangiola, S. and A. T. Papenfuss (2020). "tidyHeatmap: an R package for modular heatmap production based on tidy principles." Journal of Open Source Software **5**: 2472.

Matzke, M. A. and R. A. Mosher (2014). "RNA-directed DNA methylation: an epigenetic pathway of increasing complexity." Nature Reviews Genetics **15**(6): 394-408.

Penterman, J., D. Zilberman, J. H. Huh, T. Ballinger, S. Henikoff and R. L. Fischer (2007). "DNA demethylation in the Arabidopsis genome." Proceedings of the National Academy of Sciences **104**(16): 6752-6757.

Pongor, L. S., J. M. Gross, R. Vera Alvarez, J. Murai, S.-M. Jang, H. Zhang, C. Redon, H. Fu, S.-Y. Huang, B. Thakur and others (2020). "BAMscale: quantification of next-generation sequencing peaks and generation of scaled coverage tracks." Epigenetics & chromatin **13**: 1-13.

Reinders, J., B. B. Wulff, M. Mirouze, A. Mari-Ordonez, M. Dapp, W. Rozhon, E. Bucher, G. Theiler and J. Paszkowski (2009). "Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes." Genes Dev **23**(8): 939-950.

Saze, H., K. Tsugane, T. Kanno and T. Nishimura (2012). "DNA Methylation in Plants: Relationship to Small RNAs and Histone Modifications, and Functions in Transposon Inactivation." Plant and Cell Physiology **53**(5): 766-784.

Seisenberger, S., J. R. Peat and W. Reik (2013). "Conceptual links between DNA methylation reprogramming in the early embryo and primordial germ cells." Current Opinion in Cell Biology **25**(3): 281-288.

Smallwood, S. A., H. J. Lee, C. Angermueller, F. Krueger, H. Saadeh, J. Peat, S. R. Andrews, O. Stegle, W. Reik and G. Kelsey (2014). "Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity." Nat Methods **11**(8): 817-820.

Stroud, H., T. Do, J. Du, X. Zhong, S. Feng, L. Johnson, D. J. Patel and S. E. Jacobsen (2014). "Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis." Nature structural \& molecular biology **21**(1): 64-72.

Tonosaki, K., R. Fujimoto, E. S. Dennis, V. Raboy and K. Osabe (2022). "Will epigenetics be a key player in crop breeding?" Frontiers in Plant Science **13**: 958350-958350.

Wang, Y., X. Xue, J. K. Zhu and J. Dong (2016). "Demethylation of ERECTA receptor genes by IBM1 histone demethylase affects stomatal development." Development (Cambridge) **143**(23): 4452-4461.

Wang, Z. and D. C. Baulcombe (2020). "Transposon age and non-CG methylation." Nat Commun **11**(1): 1221.

Wang, Z., D. C. Baulcombe, Z. Wang and D. C. Baulcombe (2020). "Transposon age and non-CG methylation." Nature Communications 2020 11:1 **11**(1).

Wibowo, A., C. Becker, J. Durr, J. Price, S. Spaepen, S. Hilton, H. Putra, R. Papareddy, Q. Saintain, S. Harvey, G. D. Bending, P. Schulze-Lefert, D. Weigel and J. Gutierrez-Marcos (2018). "Partial maintenance of organ-specific epigenetic marks during plant asexual reproduction leads to heritable phenotypic variation." Proceedings of the National Academy of Sciences **115**(39): E9145-E9152.

Wibowo, A., C. Becker, G. Marconi, J. Durr, J. Price, J. Hagmann, R. Papareddy, H. Putra, J. Kageyama, J. Becker, D. Weigel and J. Gutierrez-Marcos (2016). "Hyperosmotic stress memory in arabidopsis is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by dna glycosylase activity." eLife **5**(MAY2016).

Wibowo, A. T., J. Antunez-Sanchez, A. Dawson, J. Price, C. Meehan, T. Wrightsman, M. Collenberg, I. Bezrukov, C. Becker, M. Benhamed, D. Weigel and J. Gutierrez-Marcos (2022). "Predictable and stable epimutations induced during clonal plant propagation with embryonic transcription factor." PLOS Genetics **18**(11): 1-19.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. New York, Springer-Verlag.

930 Williams, B. P., L. L. Bechen, D. A. Pohlmann and M. Gehring (2022). "Somatic DNA
931 demethylation generates tissue-specific methylation states and impacts flowering time." The
932 Plant Cell **34**(4): 1189-1206.

933 Wu, X. and Y. Zhang (2017). "TET-mediated active DNA demethylation: mechanism, function
934 and beyond." Nature Reviews Genetics 2017 18:9 **18**(9): 517-534.

935 Yamamuro, C., D. Miki, Z. Zheng, J. Ma, J. Wang, Z. Yang, J. Dong and J. K. Zhu (2014).
936 "Overproduction of stomatal lineage cells in Arabidopsis mutants defective in active DNA
937 demethylation." Nature Communications 2014 5:1 **5**(1): 1-7.

938 Yang, Y., K. Tang, T. U. Datsenka, W. Liu, S. Lv, Z. Lang, X. Wang, J. Gao, W. Wang, W. Nie, Z.
939 Chu, H. Zhang, A. K. Handa, J.-K. Zhu and H. Zhang (2019). "Critical function of DNA
940 methyltransferase 1 in tomato development and regulation of the DNA methylome and
941 transcriptome." Journal of Integrative Plant Biology **61**(12): 1224-1242.

942 Zemach, A., M. Y. Kim, P. H. Hsieh, D. Coleman-Derr, L. Eshed-Williams, K. Thao, S. L. Harmer
943 and D. Zilberman (2013). "The arabidopsis nucleosome remodeler DDM1 allows DNA
944 methyltransferases to access H1-containing heterochromatin." Cell **153**(1): 193-205.

945 Zhang, H., Z. Lang and J. K. Zhu (2018). "Dynamics and function of DNA methylation in plants."
946 Nature Reviews Molecular Cell Biology **19**(8).

947 Zhang, H. and J.-K. Zhu (2025). "Epigenetic gene regulation in plants and its potential
948 applications in crop improvement." Nature Reviews Molecular Cell Biology **26**(1): 51-67.

949 Zhang, X., Y. Zhang, C. Wang and X. Wang (2023). "TET (Ten-eleven translocation) family
950 proteins: structure, biological functions and applications." Signal Transduction and Targeted
951 Therapy 2023 8:1 **8**(1): 1-20.

952 Zhong, X., J. Du, C. J. Hale, J. Gallego-Bartolome, S. Feng, A. A. Vashisht, J. Chory, J. A.
953 Wohlschlegel, D. J. Patel and S. E. Jacobsen (2014). "Molecular mechanism of action of plant
954 DRM de novo DNA methyltransferases." Cell **157**(5): 1050-1060.

955 Zhu, J. K. (2009). "Active DNA Demethylation Mediated by DNA Glycosylases." Annual review of
956 genetics **43**: 143-143.

957