Investigating the effects of a mammalian demethylase in plant species

Elizabeth Helen Hollwey

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Biology

August 2017

Intellectual Property and Publication Statement

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Work from jointly authored publications forms part of Chapter Two within this thesis.

For Chapter Two, the relevant publication is “Hollwey, E., Watson, M., and Meyer, P. (2016). Expression of the C-terminal domain of mammalian TET3 DNA dioxygenase in Arabidopsis thaliana induces heritable methylation changes at rDNA loci, Adv. Biosci. Biotechnol. 07, 243–250”. Figures 3 and 4 are directly attributable to the candidate. Other authors for the publication were Michael Watson, who contributed Figures 1 and 2, and Peter Meyer, who wrote the manuscript.

Acknowledgements

Many people are owed thanks for the help and support they have given over the course of my PhD. Firstly, I want to thank my project supervisor Professor Peter Meyer for his constant advice and guidance which have greatly improved me as a scientist. In the Netherlands, I would like to thank my supervisor Iris Heidmann for her invaluable support and ideas, as well as the rest of the team at Enza, particularly Suzan Out and Marieke Ykema for all their assistance.

I am extremely grateful to other members of the group and the office, particularly Valya Vassileva, Michael Watson, Sam Brocklehurst, Sue Marcus, Najihah Mohamed Ghazani and Mohd Ahmad Mokhtar for their support both technical and emotional.

Finally I owe a huge debt to my family, including Jo, Mike and Alex Hollwey and to my husband, Jonny Royle, for their support in getting me to this point, for sitting through many practice presentations and reading many drafts, not to mention designing software to make data collection simpler, but also for always reducing my stress levels and cheering me up.

Abstract

DNA methylation is one of a variety of epigenetic modifications that occurs in plant and animal genomes. It consists of a methyl group added to a cytosine base within the DNA. The established main effect of DNA methylation is to affect the expression of genes and other genetic elements by influencing the likelihood of transcription, although it may also have other effects for example by affecting splicing in maize. The removal of methylation, DNA demethylation, is carried out differently in plants and animals. In this project, the catalytic domain of the mammalian demethylase enzyme TET3, which removes methylation, has been transformed into plants both to investigate the mechanism of TET3-mediated demethylation and identify genes in plants which are sensitive to methylation changes. It is shown that this can result in both hypo- and hypermethylation in plants and in the production of oxidative derivatives of TET3-mediated demethylation. Methylation changes caused by TET3 in plants could be inherited in the absence of the transgene. In tomato, TET3 produces a heritable phenotypic change in shoot architecture, resulting from the demethylation and activation of a tomato gene which has not previously been functionally characterised. Transformation of TET3 into tomato also resulted in defects in shoot apical meristem maintenance allowing investigation of the causes of this phenotype, which can cause production losses for the tomato industry. TET3 was silenced in tobacco and lettuce, suggesting that expression of this gene may have detrimental effects in some species.

Contents

List of Figures

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Chapter 5:

Chapter 6:

Appendix:

Abbreviations

A	adenine
AID	Activation-Induced Deaminase
AMD	Apical Meristem Decline
APOBEC1	Apolipoprotein B RNA Editing Catalytic Component 1
ASR2	ABSCISIC ACID STRESS RIPENING 2
ATC	ARABIDOPSIS THALIANA CENTRORADIALIS HOMOLOGUES
bp	base pairs
BFT	BROTHER OF FT AND TFL1
C	cytosine
caC	carboxylcytosine
cDNA	complimentary DNA
CEN	CENTRORADIALIS
CETS	CENTRORADIALIS/TERMINAL FLOWER 1/SELF-PRUNING
CLE9	CLAVATA3/ESR RELATED PROTEIN 9
CLV1	CLAVATA1
CLV2	CLAVATA2
CLV3	CLAVATA3
CMT2	CHROMOMETHYLTRANSFERASE 2
CMT3	CHROMOMETHYLTRANSFERASE 3
Cnr	Colourless non ripening
Col-0	Arabidopsis Columbia
CRN	CORYNE
CZ	central zone
DDM1	DECREASE IN DNA METHYLATION 1
DME	DEMETER
DML2	DEMETER-LIKE 2
DML3	DEMETER-LIKE 3
DNA	deoxyribonucleic acid

DNMT1	DNA Methyltransferase 1
DNMT3A	DNA Methyltransferase 3A
DNMT3B	DNA Methyltransferase 3B
DRM1	DOMAINS REARRANGED METHYLTRANSFERASE 1
DRM2	DOMAINS REARRANGED METHYLTRANSFERASE 2
EDTA	ethylenediaminetetraacetic acid
ESC	embryonic stem cell
fC	formylcytosine
FIS2	FERTILIZATION-INDEPENDENT SEED 2
FT	FLOWERING LOCUS T
FWA	FLOWERING WAGENINGEN
G	guanine
GUS	B-glucuronidase
HI	heat-inducible
hmC	hydroxymethylcytosine
IDM1	INCREASED DNA METHYLATION 1
IDM2	INCREASED DNA METHLYLATION 2
IAA	isoamyl alcohol
LB	lysogeny broth
MBD7	METHYL-CPG-DNA BINDING DOMAIN PROTEIN 7
MBD4	Methyl-CpG-Binding Domain Protein 4
mC	methylcytosine
MET1	METHYLTRANSFERASE 1
MS	Murashige and Skoog
ncRNA	non-coding ribonucleic acid
NPTII	NEOMYCIN PHOSPHOTRANSFERASE II
NRPE1	NUCLEAR RNA POLYMERASE E1
OC	organising centre
OD	optical density
PCR	polymerase chain reaction

PEBP	phosphatidylethanolamine-binding protein
PGCs	primordial germ cells
PTGS	post-transcriptional gene silencing
PZ	peripheral zone
RdDM	RNA-directed DNA methylation
rDNA	ribosomal DNA
RNA	ribonucleic acid
RNAi	RNA interference
ROS1	REPRESSOR OF SILENCING 1
RPK2	RECEPTOR-LIKE PROTEIN KINASE 2
RPS	repetitive sequence
RT-PCR	reverse transcription PCR
RZ	rib zone
SAM	shoot apical meristem
SDS	sodium dodecyl sulphate
SFT	SINGLE FLOWER TRUSS
SP	SELF PRUNING
SP5G	SELF PRUNING 5G
SP9D	SELF PRUNING 9D
T	thymine
TAE	Tris-acetate-EDTA
TDG	Thymine DNA Glycosylase
T-DNA	transfer DNA
TET	Ten-Eleven Translocation
TET3c	TET3 catalytic domain
TFL1	TERMINAL FLOWER 1
TSS	Transformation Storage Solution
WT	Wild type
WUS	WUSCHEL

1. Introduction

1.1. General Introduction to Epigenetics

The term "epigenetics" was first developed by Charles Waddington in 1942 to describe the study of the network of processes that occurs between the genotype and the phenotype (Waddington, 1942). Modern definitions of epigenetics have since evolved, and are contentious and often debated, but attempts at agreeing a consensus definition have been made including *"The study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence"* (Riggs et al., 1996). Epigenetic changes are known to underlie a variety of phenomena including transgene silencing (Meyer and Heidmann, 1994), imprinting (DeChiara et al., 1991), and position effect variegation (Spofford, 1967) and are likely to be involved in more.

Several different types of epigenetic modification exist, including changes at the RNA, DNA and chromatin level. Non-coding RNAs (ncRNAs) can cause post-transcriptional gene silencing (PTGS) (Meister and Tuschl, 2004) and RNA modifications (such as RNA methylation) affect gene expression by altering RNA stability (Motorin et al., 2010). RNA methylation and ncRNAs can act at the RNA level without interaction with DNA but RNA is also intimately linked to the establishment and targeting of epigenetic modifications at the DNA and chromatin level, for example through RNA-direction DNA methylation in plants (RdDM)(Wassenegger et al., 1994) or dosage compensation in *Drosophila* (Smith et al., 2000).

Epigenetic modifications to DNA and chromatin can alter the likelihood of transcription by influencing chromatin organisation and structure and thus the compaction and accessibility of DNA within chromatin (Bannister and Kouzarides, 2011; Klose and Bird, 2006), as well as affecting long-range interactions between different regions of DNA (Bartkuhn and Renkawitz, 2008). The packaging of DNA into chromatin allows compaction of DNA into the nucleus and contains a

fundamental repeating subunit, the nucleosome, constructed from eight histone proteins (Kornberg, 1977). Covalent modifications to histone proteins (such as acetylation)(Bannister and Kouzarides, 2011) or the incorporation of alternate histone proteins into the nucleosome (Bönisch and Hake, 2012) affect how chromatin is packaged by influencing the relationship of one nucleosome to another, recruiting remodelling enzymes and binding proteins, or altering the structure of the nucleosome itself. Epigenetic chromatin modifications have predominantly been identified on the N-terminal tails of histone proteins (which are exposed on the surface of the nucleosome) although modifications have also been found throughout histone proteins, both within the core and on C-terminal tails (Zhang et al., 2003). Epigenetic changes to the chromatin are interlinked with epigenetic modification of DNA, namely DNA methylation (Cedar and Bergman, 2009) on which the focus of this project will be.

1.2. DNA Methylation

DNA methylation is an evolutionarily ancient, widespread epigenetic modification which consists of the addition of a methyl ($-CH_3$) group to a DNA base. It is found in both prokaryotic and eukaryotic species but there are key differences in the role it plays in different organisms (Piccolo and Fisher, 2014). In bacteria, DNA methylation acts to protect DNA from the bacterium's own restriction enzymes, which act to digest invading viruses (Wilson and Murray, 1991).

In eukaryotes, DNA methylation is currently best understood in plants, animals and fungi, where it is most often associated with gene silencing. Many species of these kingdoms possess DNA methylation, although not all (Piccolo and Fisher, 2014) and DNA methylation can also be found elsewhere, for example the marine diatom, *Phaeodactylum tricornutum*, has 6% total methylation (Veluchamy et al., 2013). Similarities in the methods of establishing and maintaining DNA methylation suggest that there were ancestral processes which were subsequently lost in many species, including *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Caenorhabditis elegans*. DNA methylation can occur in adenine and/or cytosine

DNA bases in different species (Ratel et al., 2006) and differences in the patterns and roles of DNA methylation are found across eukaryotes (Feng et al., 2010; Zemach et al., 2010). In plants and animals, although adenine DNA methylation has been reported in mouse embryonic stem cells (Wu et al., 2016), cytosine DNA methylation is the most prevalent and best understood, and is the focus of this project.

1.2.1. Patterns of DNA methylation differ in different species

As mentioned above, the level and distribution of cytosine DNA methylation varies in different species. While methylation across the genome is largely continuous in mammals, methylation patterns in invertebrates and some plant species such as the model plant *Arabidopsis thaliana* are mosaic, with densely methylated regions interspersed with regions of very little methylation (Suzuki and Bird, 2008; Zhang et al., 2010)(Figure 1.1a). Other plant species such as maize have a higher and more continuous distribution of methylation (Gouil and Baulcombe, 2016) which may be the result of a high density of transposons (SanMiguel et al., 1996). Methylation of transposons and repetitive sequences is usually enriched in plant species compared to the rest of the genome (Feng et al., 2010) and is present in a large number of species including many fungal and vertebrate species as well as plants (Zemach et al., 2010)(Figure 1.1a).

DNA methylation at different types of loci can have different functions. Methylation of promoter regions is found in most plants, vertebrates and fungi where it usually results in gene silencing. Gene body methylation is evolutionarily ancient and has been observed in several species of plants and animals, including Arabidopsis, mouse and honeybee but is absent in some early diverging land plants such as *Selaginella moellendorffii* (Zemach et al., 2010) and appears to be absent in most fungi, although not all (Jeon et al., 2015; Mishra et al., 2011; Zemach et al., 2010)(Figure 1.1a). The function of gene body methylation has not been fully established. In maize, methylation in the gene body at intron-exon boundaries is correlated with a reduced splicing efficiency at these sites while gene body

methylation within exons may deter insertion of transposons (Regulski et al., 2013). In Arabidopsis, approximately 20% of genes contain gene body methylation. These genes are most likely to be moderately expressed and it has been suggested that the purpose of gene body methylation here is to prevent initiation of transcription from cryptic sites (Zilberman et al., 2007). Gene body methylation appears to be associated with evolutionarily conserved genes, which suggests that it is functional (Takuno and Gaut, 2012) although levels of gene body methylation vary greatly between species (Takuno et al., 2016).

On a sequence level, cytosine DNA methylation can occur in different contexts: CG, CHG or CHH, where H can be adenine (A), cytosine (C) or thymine (T) but not guanine (G). In mammals, methylation occurs mainly in the CG context where 70-80% of cytosines can be methylated (Feng et al., 2010; Lister et al., 2009). Recent studies have identified that non-CG methylation, while rare, does occur in mice (Ramsahoye et al., 2000) and is consistently present across a variety of human tissues (Ziller et al., 2011) where it can be functional in the regulation of reporter gene expression (Inoue and Oishi, 2005) and protein binding (Chen et al., 2015). In plants, the distribution of methylation across the three contexts is more balanced, although CG remains the most highly methylated cytosine context (Niederhuth et al., 2016). In Arabidopsis, 22-30% of cytosines in the CG context are methylated, in comparison to 6-9% of CHG and 1.5-4% of CHH sites, giving an overall methylation level of 5% (Cokus et al., 2008; Feng et al., 2010; Niederhuth et al., 2016)(Figure 1.1b). Other plant species contain higher levels of methylation, for example rice, where 14-18% of cytosines are methylated (Feng et al., 2010; Zemach et al., 2010), or tomato, where 22-24% of cytosines are methylated (Zhong et al., 2013). This increase occurs across all contexts, with 73-85% of CG sites, 52-56% of CHG sites and 8-14% of CHH sites being methylated in tomato (Zhong et al., 2013)(Figure 1.1b). Some species of plants do lack methylation in non-CG contexts such as the green algae *Volvox carteri*, but these appear to be the exception (Feng et al., 2010; Zemach et al., 2010)(Figure 1.1). Some common fungi, such as *S. cerevisiae*, contain no DNA methylation, but a diverse range of other fungi do, suggesting that these

species may have lost DNA methylation independently (Zemach et al., 2010). Methylation in fungi appears to be mostly in the CG and CHH contexts, although small levels of methylation in the CHG context have been detected (Jeon et al., 2015). Overall methylation levels are so far lower than those discovered in plants (Jeon et al., 2015; Mishra et al., 2011), with *Laccaria bicolor,* a mycorrhizal fungus, having the highest methylation levels observed in fungi at this point: 5% of all cytosines methylated and 20% of cytosines in the CG context methylated (Zemach et al., 2010).

a

	Methylation			Location			Distribution
	CG	CHG	CHH	Promoter methylation	Gene body methylation	Transposon methylation	
Plants							
Arabidopsis thaliana	✓	✓	✓	✓	✓	✓	Mosaic
Zea mays (maize)	✓	✓	✓	✓	✓	✓	Continuous
Volvox carteri (green algae)	✓	X	X	✓	✓	✓	Mosaic
Vertebrates							
Mus musculus (mouse)	✓	✓	✓	✓	✓	✓	Continuous
Danio rerio (zebrafish)	✓	✓	✓	✓	✓	✓	Continuous
Invertebrates							
Apis mellifera (honeybee)	✓	X	X	X	✓	X	Mosaic
Drosophila melanogaster	X	X	X	X	X	X	-
Fungi							
Saccharomyces cerevisiae	X	X	X	X	X	X	-
Laccaria bicolor	✓	✓	✓	✓	X	✓	Mosaic

b

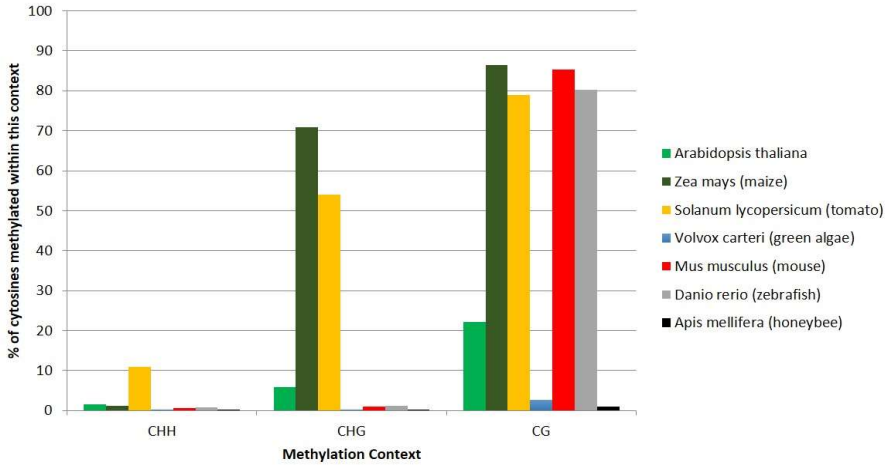


Figure 1.1. Location, level and distribution of methylation varies between species

(a) Chart indicating the presence or absence of methylation in different contexts and at different sites across different species. A tick for methylation indicates the presence of methylation at this context (CG, CHG or CHH). Location indicates whether methylation of the promoter region, methylation of gene bodies or methylation of transposons and other repetitive elements occurs. Distribution refers to whether methylation occurs globally as it does in mouse (continuous) or if densely methylated regions are interspersed with sparsely methylated ones, as they are in Arabidopsis (mosaic). A dash indicates that no methylation is present in this species.

(b) A graph showing the methylation levels at the three different contexts in a variety of species.

The context of DNA methylation can also be subdivided beyond the traditional three contexts. Methylation patterns vary in subcontexts of CG, CHG and CHH methylation sites, for example CGT methylation levels are lower than other CG subcontexts in Arabidopsis and tomato, but higher than other CG subcontexts in rice (Gouil and Baulcombe, 2016). Within the CHG context, CCG sites are less methylated compared to CAT/CTG sites in Arabidopsis, tomato, maize and rice while in the CHH context, levels were frequently higher in CAA, CAT or CTA contexts compared to CCA, CCC and CCT sites, although this varied across the genome (Gouil and Baulcombe, 2016). It has been hypothesised that these differences reflect differences in the enzymes responsible for methylation of these sites. Methylation levels at different subcontexts of non-CG methylation in humans also vary, being consistently higher at CA sites across the different tissues (Ziller et al., 2011).

1.2.2. Establishment and maintenance of DNA methylation

Most of the enzymes responsible for the establishment and maintenance of DNA methylation are conserved between plants, animals and fungi. *De novo* methylation is carried out by DNA methyltransferase 3A/3B/3L (DNMT3A/DNMT3B/DNMT3L) in animals and by a homolog of DNMT3, DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), in plants (Okano et al., 1999). Different families of

enzymes are thought to be responsible for the maintenance of DNA methylation in the three different contexts (Law and Jacobsen, 2010). CG methylation is maintained by DNA methyltransferase 1 (DNMT1) restoring methylation at hemimethylated sites in animals (Bestor, 1992). Homologues of DNMT1 exist in plants where they again maintain CG methylation (METHYLTRANSFERASE 1 genes)(Saze et al., 2003)(Figure 1.2) and in fungi, where the DNMT1 family has split into three subfamilies (Masc2, Dim-2, and Masc1 genes) (Huang et al., 2016). A family of enzymes unique to plants, the CHROMOMETHYLTRANSFERASEs (CMTs), maintains CHG methylation (Lindroth et al., 2001), supported by a feedback loop involving lysine methylation of histone 3 (Lindroth et al., 2004)(Figure 1.2). Non-CG methylation in mammals has recently been shown to be dependent on the DNMT3 enzymes (Shirane et al., 2013; Ziller et al., 2011), and CHH methylation in plants is similarly established by the DNMT3 homolog, DRM2, which acts through the RNA-directed DNA methylation (RdDM) pathway to establish *de novo* DNA methylation in all contexts (Chan et al., 2005)(Figure 1.2). This process is especially important to maintain silencing of transposable elements.

An increasing amount of evidence suggests that the roles of the DNA methyltransferases are not as strictly divided as previously thought. MET1 can be responsible for methylation of non-CG sites, both in conjunction with CMTs within the context of CCG sites (Yaari et al., 2015) and in the context of MET1-dependent dense methylation, high levels of methylation in all three contexts that is lost in *met1* mutants (Watson et al., 2014). CMT3 is known to methylate CHH sites at some loci as well as carrying out CHG methylation (Bartee et al., 2001), as does CMT2 (Stroud et al., 2014; Zemach et al., 2013) which may lead to the differences in CHH subcontext methylation discussed earlier (Gouil and Baulcombe, 2016). CMT3 also appears to be required for the establishment of CG gene body methylation in angiosperms (Bewick et al., 2016). Activity of all three methyltransferase families is required to maintain methylation in any context at the RPS transgene in Arabidopsis (Singh et al., 2008) and multiple sites in the Arabidopsis genome require interplay

between different methyltransferase pathways for maintenance of methylation (Stroud et al., 2013).

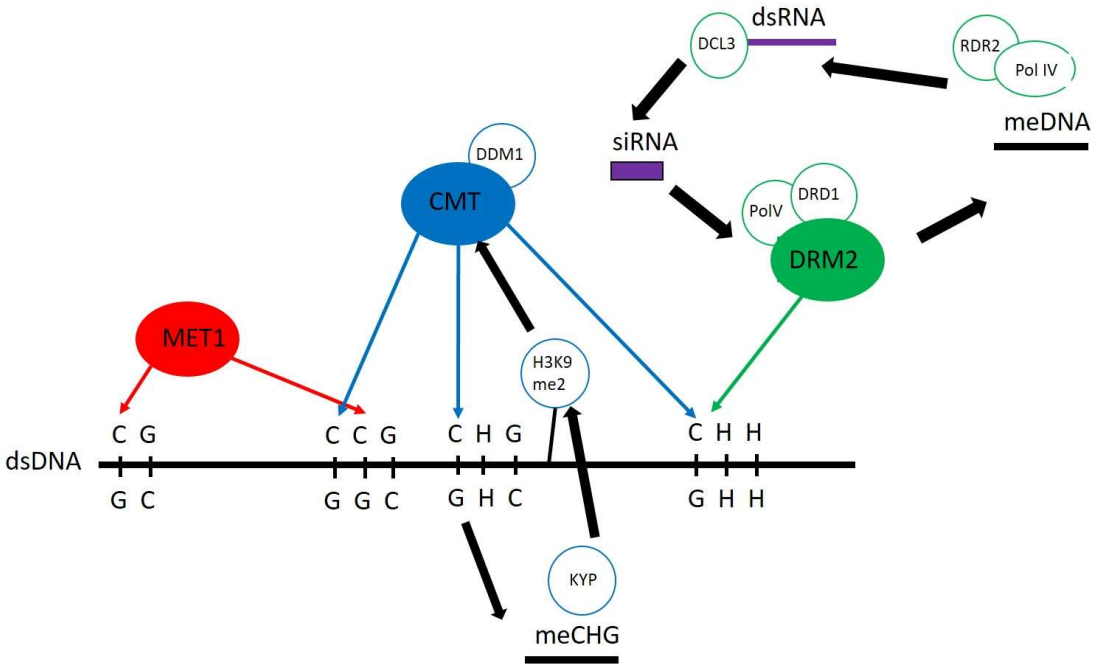


Figure 1.2. Multiple processes contribute to DNA methylation at different contexts in plants

MET1 maintains methylation at CG sites by binding to hemimethylated CG regions when the cell divides. CMT (chromomethyltransferase) enzymes can cause methylation at CHG and CHH sites, and can act in a positive feedback loop with histone modification H3K9me2 (demethylation of lysine 9 of histone 3), where the CMT binds this modification, causing methylation at nearby cytosines. CMT mediated methylation requires the DDM1 (DECREASE IN DNA METHYLATION 1) chromatin remodeler. Methylation at CHG sites recruits KRYPTONITE (KYP), the histone methyltransferase and thus this maintains the loop. Methylation at CHH sites can occur due to CMT enzymes independently of small RNAs or can be small RNA dependent through the RNA-directed DNA methylation pathway. In the RNA-directed DNA methylation pathway, double stranded RNA is produced by RNA POLYMERASE IV (Pol IV) and RNA DEPENDENT RNA POLYMERASE II (RDR2) and cut into small RNAs by DICER LIKE 3 (DCL3). These target the DRM2 methyltransferase, which requires the chromatin remodeler (DRD1) and RNA POLYMERASE V (Pol V)(Matzke et al., 2015). This process is especially important to maintain silencing of transposable elements. Arrows indicating methylation are coloured. Black arrows indicate a feedback loop which causes methylation. Filled circles are methyltransferases, open circles are other proteins or epigenetic modifications involved in the process.

1.3.DNA Demethylation

DNA methylation patterns are highly dynamic (Yamagata et al., 2012), with patterns of DNA methylation constantly changing as the cell divides and genes are transcribed. The process of the removal of DNA methylation (DNA demethylation) is as important as its addition. Cytosine DNA demethylation can occur passively or actively in both plants and animals (Gehring et al., 2009) and also plays an important role in a variety of processes such as plant development, acquisition of pluripotency and cancer. Downregulation of the DNA methyltransferases that maintain DNA methylation can lead to passive demethylation as cells divide and DNA methylation is not restored (Rougier et al., 1998), or passive DNA demethylation may be caused by shortages of enzymes and other components of the methylation process such as S-adenosyl methionine, preventing restoration of DNA methylation on hemimethylated DNA (the donor of the methyl group)(Chiang et al., 1996). During the plant life cycle, downregulation of the RNA-directed DNA methylation machinery causes passive loss of methylation during seed germination (Kawakatsu et al., 2017).

The active removal of DNA methylation occurs through the action of enzymes in both plants and animals. Active DNA demethylation is important for maintaining normal methylation patterns throughout the life cycle of an organism and is also a crucial element in reproduction for both plants and animals as well as in many other processes (Gehring et al., 2009). Active DNA demethylation is independent of DNA replication and occurs more rapidly than passive DNA demethylation (Mayer et al., 2000).

1.3.1. Active DNA demethylation in plants occurs through the base excision repair pathway

In plants, active DNA methylation acts through the base excision repair pathway and requires the ROS/DME family of DNA glycosylases which in Arabidopsis consists of REPRESSOR OF SILENCING 1 (ROS1), DEMETER (DME), DEMETER-LIKE 2 (DML2)

and DEMETER-LIKE 3 (DML3) (Gehring et al., 2009)(Figure 1.3). These remove 5-methylcytosine by excising it (Agius et al., 2006), leaving a cleaved ribose sugar which is removed by an endonuclease (Li et al., 2015b), and then subsequently replaced by an unmethylated cytosine base (Li et al., 2015c). They can also remove thymine in the case of a thymine-guanine mismatch, but show greater affinity for 5-methylcytosine when both bases are located in a CG-dense context (Morales-Ruiz et al., 2006). This process is essential for many processes in Arabidopsis; for example *DME*, which is expressed in the female gametes, is required for the imprinting of genes such as *MEDEA* (Choi et al., 2002)*, FWA* (Kinoshita et al., 2004) and *FIS2* (Jullien et al., 2006) in the endosperm (Figure 1.4). This process involves demethylation in the companion cells (the vegetative nucleus and the central cell) and results in methylation of the gametes (Ibarra et al., 2012)(Figure 1.4). *ROS1*, *DML2* and *DML3* are expressed in vegetative tissues where they demethylate the 5' and 3' ends of genes (Penterman et al., 2007a)(correlating with the characteristic reduction in methylation at these sites (Feng et al., 2010)), while ROS1 is also known to target transposons and intergenic regions (Tang et al., 2016) where it can prevent spreading of DNA methylation beyond the edge of these regions (Yamamuro et al., 2014)(Figure 1.4).

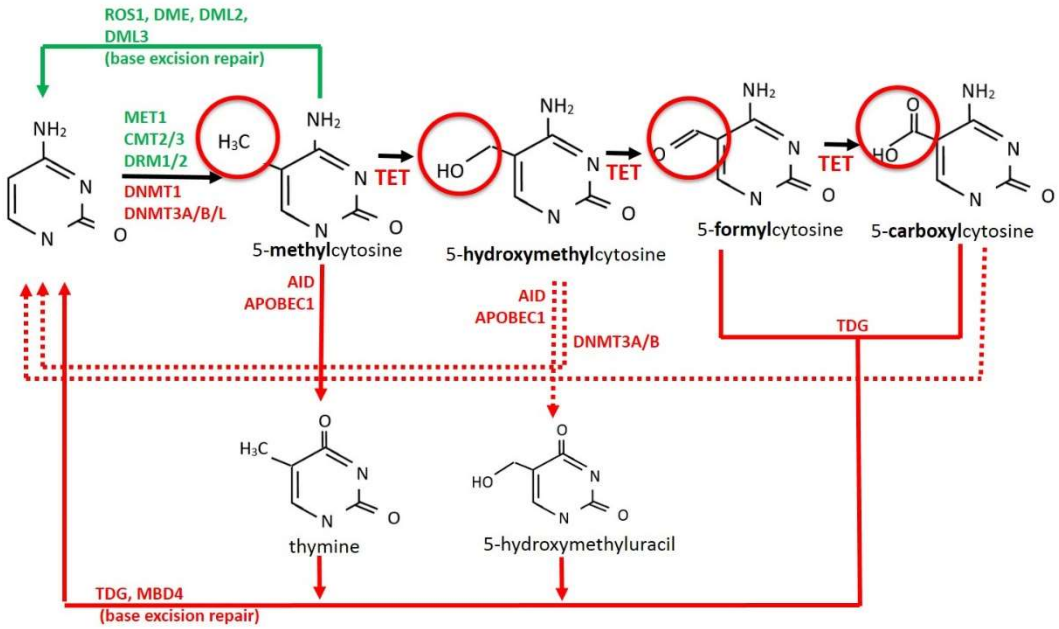


Figure 1.3. Enzymes carrying out DNA methylation and demethylation in plants and mammals

Homologous methyltransferase enzymes add DNA methylation in both plants and mammals. DNA demethylation in plants is carried out by DNA glycosylases through the base excision repair pathway. In mammals, TET enzymes can remove methylation by oxidation of 5-methylcytosine, which leads to demethylation in a variety of ways. Deaminases AID and APOBEC1 can convert 5-methylcytosine to thymine, allowing its removal by the base excision repair pathway.

Enzymes which carry out methylation and demethylation in plants are indicated in green, and in red for mammals. Dotted lines indicate that it is unknown if this mechanism contributes to demethylation in vivo.

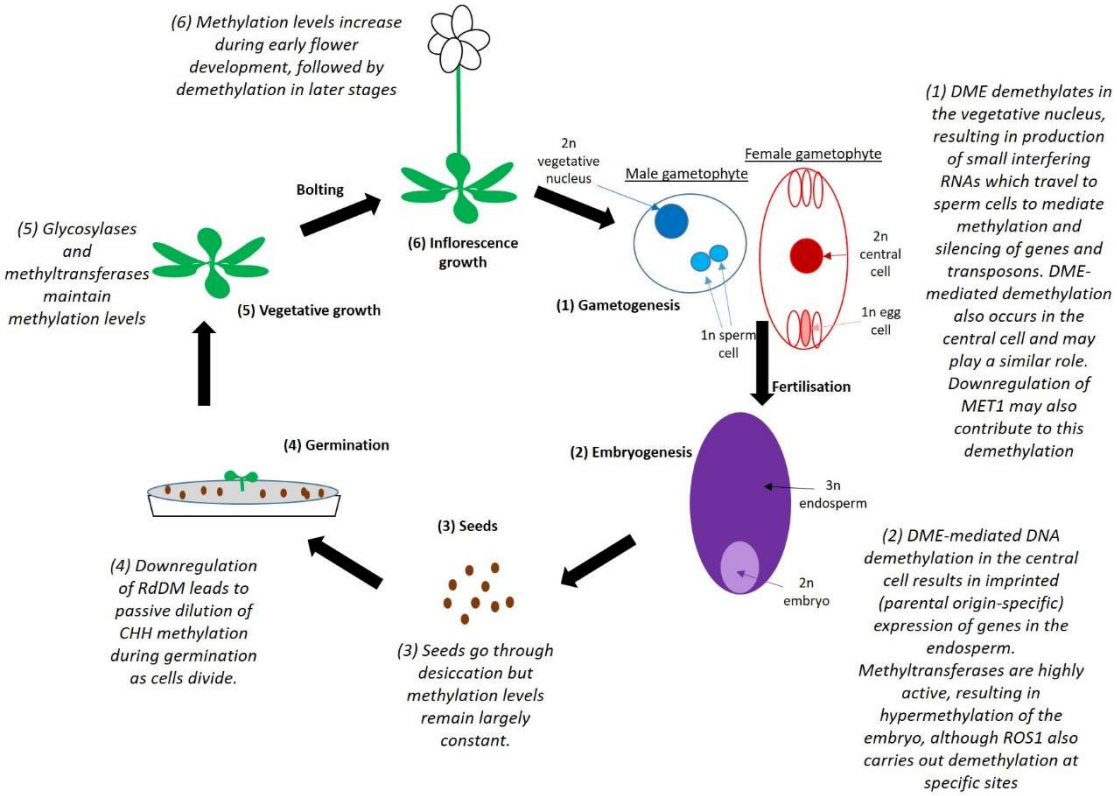


Figure 1.4. DNA methylation changes during the Arabidopsis thaliana lifecycle

During gametogenesis, demethylation of companion cells (the vegetative nucleus and the central cell) by the DNA glycosylase DME causes targeting of methylation in the gametes (sperm and egg cells). One sperm cell fuses with the diploid central cell to produce the triploid endosperm, while the other fuses with the egg cell to produce the embryo. Embryogenesis then occur, and genes such as MEDEA and FIS2 are expressed from maternal but not paternal alleles. Methylation levels increase during embryogenesis, rising to their highest point in mature embryos. Methylation levels remain constant during seed desiccation. Global DNA demethylation then occurs upon germination by passive loss during cell division and low activity of the CMT2 and RdDM pathways. DNA glycosylases and methyltransferases act during vegetative growth to ensure DNA methylation levels are maintained. Widespread gene expression changes occur during the transition to flowering, and the time at which this occurs can be affected by methylation changes e.g. hypomethylation of the FWA gene. Methylation levels increase during organogenesis and are followed by demethylation in late flower development. These changes largely correlate with gene expression levels (Yang et al., 2015).

It is known that some loci in Arabidopsis are targeted by only a single DNA glycosylase while others are targeted by several in combination, and that the 5' and 3' ends of genes are specifically demethylated by DML enzymes (Penterman et al., 2007a). How DNA glycosylases are targeted towards the regions where they act has not been fully established, but some pathways have been elucidated. Several pieces of evidence demonstrate that ROS1 can interact with the RdDM pathway, removing methylation originating from RdDM (Penterman et al., 2007b). These include the fact that expression of the *ROS1* gene is promoted by methylation caused by RdDM (Williams et al., 2015) and that an RNA binding protein, ROS3, functions in the ROS1 demethylation pathway (Zheng et al., 2008). Loci targeted by ROS1 can also be RdDM-independent (Penterman et al., 2007b; Tang et al., 2016). *ROS1* expression is reduced in a *met1* mutant through a different mechanism to in mutants of the RdDM pathway (Rigal et al., 2012; Williams et al., 2015) and ROS1 can be recruited to methylated DNA by a protein complex containing a methyl-binding protein METHYL-CPG-DNA BINDING DOMAIN PROTEIN 7 (MBD7)(Li et al., 2015a; Wang et al., 2015), the histone acetyltransferase INCREASED DNA METHYLATION 1 (IDM1) (Qian et al., 2012) and an alpha-crystallin domain protein INCREASED DNA METHYLATION 2 (IDM2) (Zhao et al., 2014).

1.3.2. Active DNA demethylation in animals occurs differently to active DNA demethylation in plants

Animals also possess DNA glycosylases capable of excising 5-methylcytosine *in vitro* such as methyl-CpG-binding domain protein 4 (MBD4)(Zhu et al., 2000a) and thymine-DNA glycosylase (TDG)(Zhu et al., 2000b). Similarly to the plant DNA glycosylases, these two DNA glycosylases can also excise thymine when it occurs in a thymine-guanine mismatch. Unlike the plant DNA glycosylases, TDG and MBD4 have only weak activity when excising 5-methylcytosine and do not seem to carry out this role *in vivo* (Gehring et al., 2009). Active DNA demethylation must therefore occur in a different manner in animals compared to in plants.

One mechanism for DNA demethylation in animals is deamination of the modified cytosine. Two enzymes known to have cytosine deaminase activity (which converts cytosine to uracil) were found to also act on 5-methylcytosine (5mC), converting this to thymine (Morgan et al., 2004)(Figure 1.3). The new thymine base is still paired with a guanine, resulting in a mismatch that can be excised by TDG or MBD4. The genes for these two deaminase enzymes, *AID* (activation-induced deaminase) and *APOBEC1* (apolipoprotein B RNA editing catalytic component 1) are expressed in oocytes and primordial germ cells (where large amounts of active DNA demethylation occur) along with a nearby cluster of pluripotency genes (Morgan et al., 2004). Demethylation in mouse primordial germ cells is greatly reduced in *aid* mutant cells but substantial demethylation does still occur (Popp et al., 2010).

The family of enzymes primarily responsible for DNA demethylation in mammals is known as the TET (Ten-Eleven-Translocation) family, and has three members (Schuermann et al., 2016). All three contain a catalytic domain capable of oxidising 5-methylcytosine, while only TET1 and TET3 contain the DNA-binding CXXC domain (Tan and Shi, 2012)(Figure 1.5). TET enzymes actively demethylate DNA by oxidising 5mC to 5-hydroxymethylcytosine (5hmC)(Ito et al., 2010; Tahiliani et al., 2009) and then further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) in a stepwise manner (Ito et al., 2011)(Figure 1.3). The TET enzymes are expressed differently in various tissues, where they are likely to have distinct roles (Figure 1.6). For example, *TET1* and *TET2* (but not *TET3*) are expressed in embryonic stem cells (ESCs) where they have a role in determining cell fate (Koh et al., 2011) and act at distinct regions, with TET1 targeted primarily to transcriptional start sites and TET2 to gene bodies (Huang et al., 2014). A large variety of proteins are thought to interact with the TET enzymes, affecting their genomic location (Delatte et al., 2014). These include OGT (a glucosamine transferase important in embryogenesis), HDAC2 (a histone deacetylase) and UHRF1 (a ubiquitin ligase which binds hemimethylated DNA)(Delatte et al., 2014). In addition to targeting TET activity, these interacting partners may alter the effect of TET enzymes, as discussed below (Zhang et al., 2015).

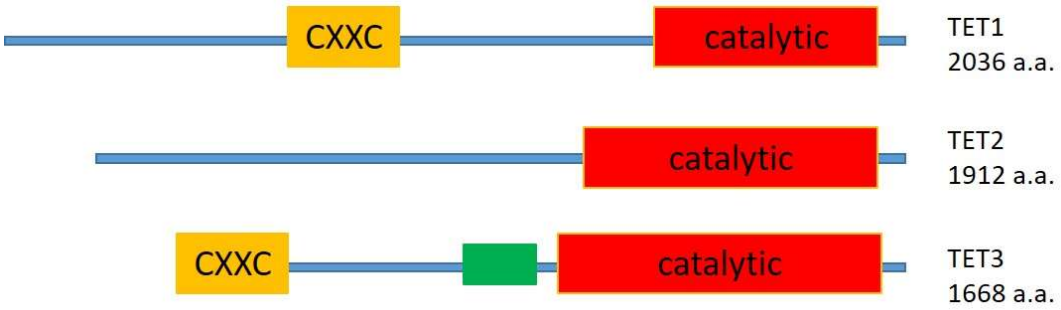


Figure 1.5. The TET gene family

There are three TET enzymes in mammals, TET1, TET2 and TET3. TET1 and TET3 contain CXXC domains (orange), DNA binding domains which bind to unmethylated cytosines. All three have a catalytic domain at the other end with methyldioxygenase activity (red). TET3 also has a domain of unknown function between these two domains (green).

Mutation of TET1 and TET2 is not lethal. Homozygous *tet1* mice are smaller, but are viable and fertile (Dawlaty et al., 2011), as are homozygous *tet2* mice, which also show an increased propensity to develop myeloid malignancies with characteristics resembling different leukaemia types (Li et al., 2011b). Double knockouts of *TET1* and *TET2* are more harmful than single mutants, with many mice dying shortly after birth with severe abnormalities and growth defects. Those mice which survive to adulthood are largely normal by 2 months of age, although they possess some alterations in methylation at imprinted regions and reduced fertility (Dawlaty et al., 2013). Homozygous mutation of *TET3*, unlike the other two, is lethal (Gu et al., 2011). *TET3* mRNA is not found in ESCs, but is present in the oocyte and zygote, and may be involved in the global demethylation that occurs in the male pronucleus (Gu et al., 2011)(Figure 1.6). Other reports suggest that TET3 is only required to maintain an active demethylation wave which has already begun (Amouroux et al., 2016), although no alternative explanation for this active demethylation wave has been suggested. The TET enzymes may be able to compensate for each other, given that triple *TET* knockout ESCs differentiated abnormally and were unable to support development (Dawlaty et al., 2014).

It is known that TET enzymes contribute to different processes, for example TET2 acts in the specification of hematopoietic cells, where it prevents differentiation into the myeloid lineage over the lymphoid lineage (Delhommeau et al., 2009). Hematopoietic stem cells with a mutated TET2 enzyme are therefore more likely to differentiate into cells of the myeloid lineage (Delhommeau et al., 2009)(Figure 1.6). TET enzymes also function in adult mammalian tissues, for example TET1 acts in the mouse brain to regulate proliferation of neural progenitor cells, which give rise to new neurons in adult brains (Zhang et al., 2013)(Figure 1.6). Mice in which TET1 is knocked out have impaired hippocampal neurogenesis, accompanied by poor learning and memory (Zhang et al., 2013). TET proteins also have functions independent of their catalytic activity, for example in the inflammation response, where TET2 causes repression of a key cytokine *IL-6* by recruiting HDAC2 (Histone Deacetylase 2). This results in the removal of histone acetylation and therefore increased condensation of the heterochromatin (Zhang et al., 2015)(Figure 1.6). Other established functions of TET proteins should be reinvestigated to determine whether catalytic function is required (Lian et al., 2016), for example in neural development, where a catalytically inactive version of TET3 is able to partially restore some function in Xenopus as long as the DNA-binding CXXC domain is present (Xu et al., 2012).

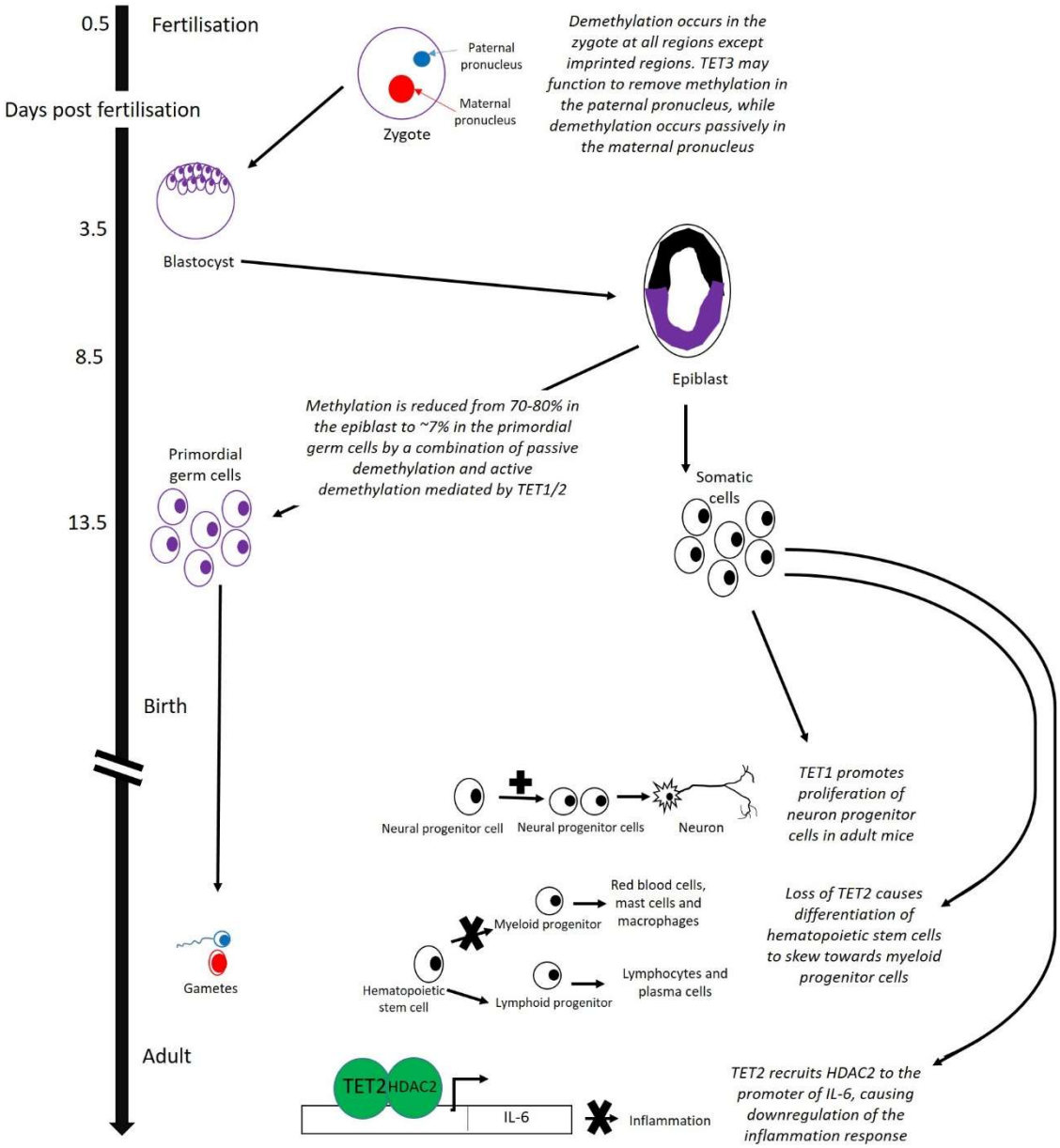

0.5
Fertilisation
Days post fertilisation
3.5
8.5
13.5
Birth
Adult
Paternal pronucleus
Maternal pronucleus
Zygote
Demethylation occurs in the zygote at all regions except imprinted regions. TET3 may function to remove methylation in the paternal pronucleus, while demethylation occurs passively in the maternal pronucleus
Blastocyst
Epiblast
Methylation is reduced from 70-80% in the epiblast to ~7% in the primordial germ cells by a combination of passive demethylation and active demethylation mediated by TET1/2
Primordial germ cells
Somatic cells
Gametes
Neural progenitor cell
Neural progenitor cells
Neuron
TET1 promotes proliferation of neuron progenitor cells in adult mice
Red blood cells, mast cells and macrophages
Myeloid progenitor
Hematopoietic stem cell
Lymphoid progenitor
Lymphocytes and plasma cells
Loss of TET2 causes differentiation of hematopoietic stem cells to skew towards myeloid progenitor cells
TET2
HDAC2
IL-6
Inflammation
TET2 recruits HDAC2 to the promoter of IL-6, causing downregulation of the inflammation response

Figure 1.6 TET enzymes have many different roles throughout the mammalian life cycle
Many methylation changes occur during embryogenesis. The involvement of TET begins immediately post-fertilization in the pronuclei where TET3 is required to either carry out or maintain demethylation in the paternal pronucleus. Methylation levels then begin to increase in the blastocyst until they reach 70-80% in the epiblast. The primordial germ cells, which will become the gametes, are then specified and demethylation occurs passively at first to remove methylation globally, while active TET1 and TET2-mediated demethylation is required to remove methylation specifically at imprinted regions. In adult tissues, TET enzymes still have roles to play in processes including the proliferation of neural progenitor cells, specification of cells in the hematopoietic cell lineage and, independently of its catalytic ability, downregulating the immune response.

5mC oxidation carried out by TET enzymes can result in DNA demethylation in multiple ways. Passive DNA demethylation can occur, as 5hmC is not a good substrate for the maintenance methyltransferase DNMT1 (Valinluck and Sowers, 2007) and therefore methylation levels are not maintained as the cell divides (Inoue and Zhang, 2011). 5hmC may be deaminated to 5-hydroxymethyluracil, which can be removed by the base excision repair pathway (Cortellino et al., 2011), but it is unclear if the mammalian deaminases have 5hmC deaminase activity *in vivo* (Guo et al., 2011; Nabel et al., 2012)(Figure 1.3). Removal of 5hmC could also occur through the *de novo* methyltransferases, DNMT3A/B, which have been shown to be capable of converting 5hmC to C *in vitro,* although it is unknown if this takes place *in vivo* (Chen et al., 2012)(Figure 1.3). Further oxidation of 5hmC to 5fC and 5caC allows active demethylation through the base excision repair pathway, as the DNA glycosylase TDG is able to remove these intermediates (He et al., 2011; Maiti and Drohat, 2011)(Figure 1.3). An involvement for TDG is supported by the facts that TET1 has been found to physically and functionally interact with TDG *in vivo* (Müller et al., 2014; Weber et al., 2016), and that targeting of TDG to DNA can result in demethylation (Gregory et al., 2012). In addition, mouse ESC extracts have been shown to be capable of decarboxylating 5caC to C, suggesting this could also be a mechanism for removal of 5caC (Schiesser et al., 2012)(Figure 1.3).

The discovery of the TET-mediated demethylation pathway led to increased research into the derivatives of TET-mediated demethylation. This has resulted in suggestions that these derivatives may be an epigenetic mark in their own right (Bachman et al., 2014, 2015; Iurlaro et al., 2016). Evidence for this includes the fact that 5hmC (Bachman et al., 2014) and 5fC can be stable (Bachman et al., 2015) and the identification of proteins which bind specifically to either 5hmC, 5fC or 5caC including DNA helicases, p53 and chromatin remodelling proteins (Spruijt et al., 2013).

1.4. Epigenetic Reprogramming

Removal of DNA methylation is particularly important in the process of epigenetic reprogramming, a global change in epigenetic marks. This occurs twice in the mammalian lifecycle (Messerschmidt et al., 2014), in primordial germ cells (PGCs) during embryogenesis (Hajkova et al., 2002) and also in the zygote immediately after fertilisation (Mayer et al., 2000)(Figure 1.6). Reprogramming in the zygote removes all methylation except at imprinted regions (Reik and Walter, 2001; Tremblay et al., 1997), while reprogramming in PGCs removes methylation both globally and at imprint control regions to reach an eventual level of ~7%, compared to 70-80% in the epiblast cells from which they are derived (Seisenberger et al., 2012)(Figure 1.6). Methylation in the PGCs is then restored in a sex-specific manner, including imprints (Davis et al., 2000; Messerschmidt et al., 2014). Imprinted regions control the expression of genes from a single allele in a parent of origin specific manner (Ferguson-Smith, 2011). In mammals, epigenetic marks at imprinted regions are usually established in the germline (Davis et al., 2000) and imprinted expression can persist in adult tissues (Ferguson-Smith, 2011; Rougeulle et al., 1997). In comparison, imprinted genes in plants are almost exclusively found in the endosperm (Gehring and Satyaki, 2017), although transient imprinted expression from a small number of genes has also been described in the earliest stages of the embryo (Raissig et al., 2013)(Figure 1.4). Given that epialleles have been observed to be inherited in plants, it was originally suggested that plants did not undergo

epigenetic reprogramming, but it is now known that considerable epigenetic reprogramming does occur during plant sexual reproduction, although this must be incomplete to allow occasional inheritance of parental epigenetic states (Kawashima and Berger, 2014). Methylation changes during embryogenesis are largely a result of the RNA-directed DNA methylation pathway, and cause global increase of DNA methylation particularly at CHH contexts (Bouyer et al., 2017)(Figure 1.4).

1.5. DNA Methylation Changes Can Affect Plant Phenotypes

In plants, natural variations in epigenetic patterns, or epialleles, occur between different strains and ecotypes of a species, where they can result in differences in phenotype. Epialleles often form as a result of nearby transposable elements (Slotkin and Martienssen, 2007). One example of an epiallele occurs in Arabidopsis, where hypomethylation upstream of the *FWA* gene results in activation of transcription and late flowering (Soppe et al., 2000). These epigenetic changes can also occur as the result of mutations in the genes involved in DNA methylation for example loss of DDM1 (DECREASE IN DNA METHYLATION 1), a chromatin remodelling protein, can cause the *FWA* hypomethylation described above (Soppe et al., 2000). A different epiallele exists in tomato, where hypermethylation at the *Cnr* (colourless non-ripening) locus results in tomato fruit which do not ripen (Zhong et al., 2013).

Epialleles can remain stable over multiple generations (Manning et al., 2006; Soppe et al., 2000), demonstrating that, as discussed above, epigenetic reprogramming in plants is not complete. However, many methylation changes produced by mutation of the DNA methylation machinery are not stable, and can return to wild type levels or vary stochastically over subsequent generations in different plants (Reinders et al., 2009). Epigenetic changes often have this erratic nature, meaning that not all plants altered in the same epigenetic pathway will show the same level of phenotype.

Changes in methylation as a result of mutation in the DNA methylation machinery have been well studied in Arabidopsis, where mutants of maintenance methylation (*met1)*(Kankel et al., 2003), triple mutants of the DNA glycosylases (*ros1-3 dml2-1 dml3-1)*(Penterman et al., 2007a) or triple mutants of three of the methyltransferases (*drm1-2, drm2-2, cmt3-11)*(Cao and Jacobsen, 2002a) are viable, although plants display pleiotropic phenotypes including delayed flowering, sterility and reduced plant size. These phenotypes can become more severe over subsequent generations (Kakutani et al., 1996). In other species, phenotypes from mutations in the DNA methylation machinery are often more severe. As discussed earlier, loss of the TET3 demethylase enzyme in mice is lethal (Gu et al., 2011), as is mutation of the *MET1* homologue *DNMT1* in mice embryos or human ESCs (Li et al., 1992; Liao et al., 2015). Some plant species can also react more severely to alterations in the DNA methylation machinery than Arabidopsis does. For example, *met1* mutants in rice display seedling lethality (Hu et al., 2014), *cmt3* mutants have developmental abnormalities such as reduced fertility and dwarf phenotypes (Cheng et al., 2015), and *drm2* mutants have growth defects, delayed heading and sterility (Moritoh et al., 2012) which contrasts with the lack of phenotypes in single *cmt3* (Lindroth et al., 2001) or *drm2* (Cao and Jacobsen, 2002b) mutants of Arabidopsis. In tomato, mutation of one of the components of the RdDM pathway, RNA Polymerase V, appears to be lethal (Gouil and Baulcombe, 2016) and tomato transformants containing a RNAi construct directed against the *MET1* gene could not be regenerated (Watson, 2013).

In addition to mutations in the enzymes of methylation, DNA methylation levels can vary as a result of other causes. In mammals, DNA methylation can show circadian variation (Lim et al., 2014), which could also be the case in plants. DNA methylation levels in plants can also be altered by biotic and abiotic stresses including bacterial infection (Dowen et al., 2012), osmotic stress (Kovarik et al., 1997) and cold stress (Steward et al., 2002). These changes can correlate with the activation of stress-inducible genes, for example osmotic stress in tomato roots correlates with hypomethylation and increased expression of *ABSCISIC ACID STRESS RIPENING 2*

(*ASR2*) (González et al., 2013), a member of the *ASR* gene family which is known to be important in osmotic stress responses (Golan et al., 2014). These changes may result in an increase in stress tolerance. Some stress-induced changes in DNA methylation have been demonstrated to be heritable, for example when rice plants are subjected to nitrogen-deficiency, many regions become differentially methylated, and some of these novel epigenetic states can be inherited (Kou et al., 2011). The progeny of plants which have inherited epigenetic modifications grow better in nitrogen-deficient conditions than progeny of plants which have not, suggesting that epigenetic changes in a stressed plant may function to increase the tolerance of their progeny to stress (Kou et al., 2011). However, characterisation of loci at which heritable epigenetic changes have resulted in increased stress tolerance has not yet occurred (Meyer, 2015).

1.6.Thesis Objective

As described above, plants demethylate DNA using a different mechanism to mammals. Demethylation mediated by plant DNA glycosylases has been shown to function in mammalian cells (Parrilla-Doblas et al., 2017) but the reverse, demethylation mediated by the TET mammalian demethylase enzymes in plants, has not been demonstrated. By transforming the mammalian demethylase TET3 into plants, two research questions can be investigated. Firstly, more can be learned about how TET enzymes act in isolation. By analysing levels of 5-methylcytosine, and the oxidised derivatives produced by TET3, it can be seen whether the catalytic domain of TET3 (TET3c) can carry out demethylation at sites within plant DNA and how far along the oxidative pathway it will proceed. Secondly, *TET3c* can be transformed into plants where altering the endogenous methylation systems has proved severe. As TET3 has no essential function in these species, plants with altered methylation patterns may be more likely to survive, allowing us to identify genes and processes where methylation plays an important role. In addition, TET3c will act outside of the usual systems for maintenance of normal methylation patterns. This may allow TET3c to generate novel epigenetic changes which are not

reversed and therefore remain as heritable epialleles. The aims of this project are to investigate whether the mammalian demethylase *TET3* can carry out demethylation in plants, using *Arabidopsis thaliana* as a proof of concept system, and then expand this work into crop species such as tomato in order to identify phenotypes linked to methylation changes. This will allow us to compare the epigenetic changes and phenotypes observed as a result of *TET3* expression in Arabidopsis and tomato.

2. Expressing a Mammalian Demethylase in *Arabidopsis thaliana* Induces Heritable Changes in DNA Methylation

2.1. Introduction

Active DNA demethylation occurs very differently in mammals compared to plants, as discussed in the introduction. In plants, active demethylation is carried out through the base excision repair pathway (Gehring et al., 2009). In the model plant *Arabidopsis thaliana,* this involves one of four different DNA glycosylases: REPRESSOR OF SILENCING 1(ROS1), DEMETER (DME) (Morales-Ruiz et al., 2006) DEMETER-LIKE 2 (DML2) and DEMETER-LIKE 3 (DML3) (Ortega-Galisteo et al., 2008), which make up the DME family. These remove 5-methylcytosine (5mC) by excising the methylated base and the gap is then filled by an unmethylated cytosine.

In mammals, active DNA demethylation can be carried out through the activity of the TET (Ten-Eleven-Translocation) family of enzymes (Ito et al., 2010, 2011; Tahiliani et al., 2009)(as discussed in the Introduction, pages 29-33). The family consists of three enzymes, TET1, TET2 and TET3. These enzymes can convert 5-methylcytosine to 5-hydroxymethylcytosine (5hmC) (Ito et al., 2010) and can also oxidise 5-hydroxymethylcytosine further to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Ito et al., 2011). Thymine-DNA glycosylase (TDG) can then excise 5fC and 5caC (He et al., 2011; Maiti and Drohat, 2011).

No homologues of the TET enzymes have been found in plants (Iyer et al., 2009), and there is no conclusive evidence for the presence of 5hmC in plants (Erdmann et al., 2014), suggesting that in plants this pathway for active DNA demethylation does not exist. This difference provides an opportunity to use plants for the study of mammalian demethylases, and to use the mammalian demethylases to induce novel changes in plant DNA methylation. In order to investigate if mammalian DNA demethylases could cause demethylation in plants, the catalytic domain of TET3

(*TET3c*) had previously been cloned behind the constitutive 35S promoter and transformed into Arabidopsis (Hollwey et al., 2016; Watson, 2013)(Figure 2.1).

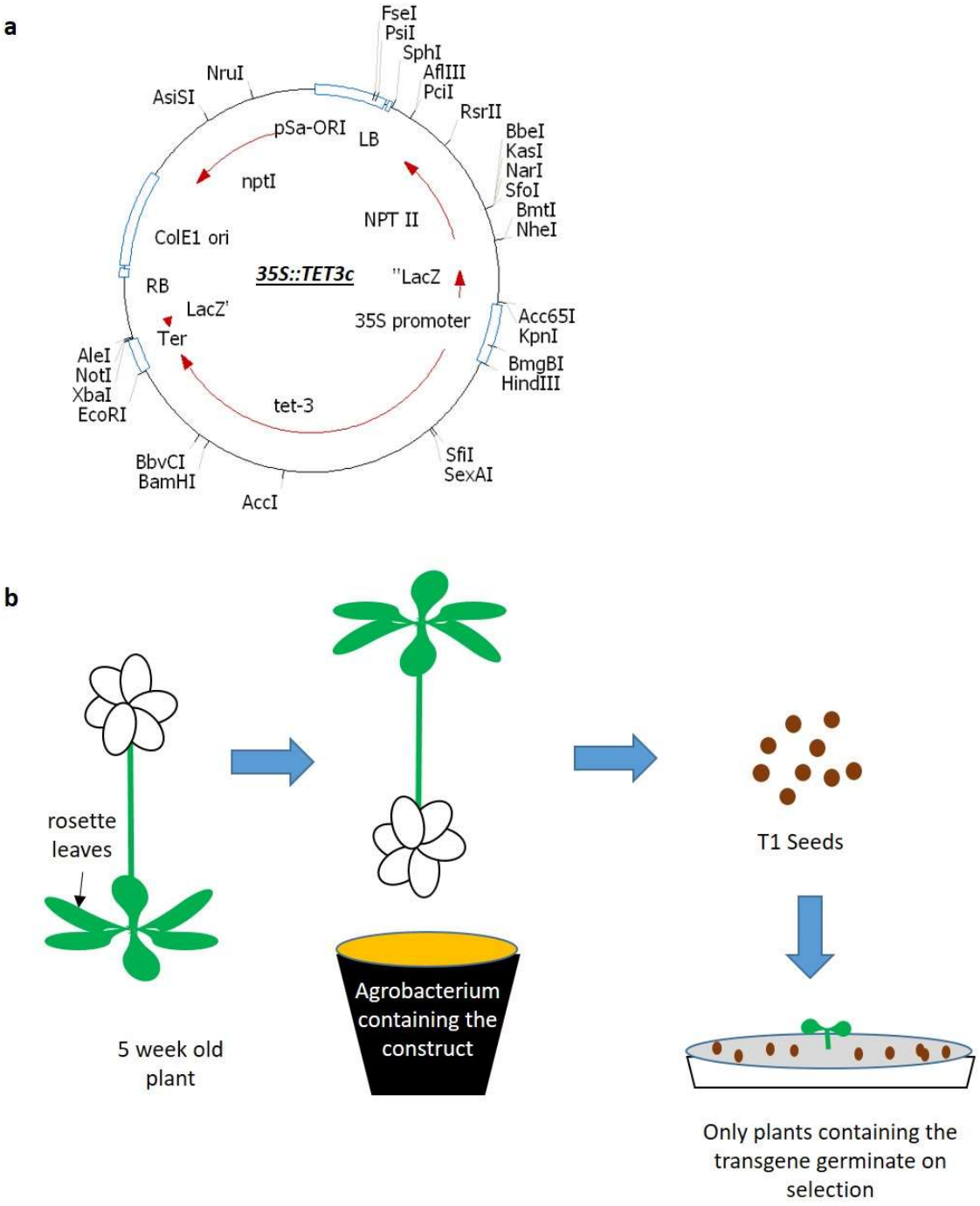


Figure 2.1. 35S::TET3c was transformed into Arabidopsis by floral dip

(a) Map of the plant expression vector pGreen 0029 35S containing the genomic DNA of the catalytic domain of the mouse TET3 gene.

(b) Arabidopsis were transformed by floral dip by inverting 4 week old plants with open flowers into an Agrobacterium solution. These plants constitute the T0, and the seeds which are collected are the T1. T1 seeds were grown on media containing antibiotic. Approximately 1 in 1000 seeds contained the transgene and were therefore resistant to the antibiotic and grew.

Using both bisulfite sequencing and methylation sensitive Southern blot analysis, *35S::TET3c* has been shown by Dr Michael Watson to cause changes in methylation in the gene body of the ribosomal DNA encoding the 18S subunit (Hollwey et al., 2016; Watson, 2013). Two lines (Line A and Line B) showed large, antagonistic changes with methylation levels appearing to decrease in Line A, and to increase in Line B.

In this chapter, *TET3c* plants of both Line A and Line B were analysed further in order to obtain a better understanding of the methylation changes occurring.

2.2. Results

2.2.1. TET3 causes epigenetic changes at the rDNA locus in *TET3c*+ Arabidopsis

Demethylation by TET enzymes involves the production of oxidative intermediates (5hmC, 5fC and 5caC)(Figure 2.2a). An analysis of methylation using bisulfite sequencing will result in 5hmC being classified as 5mC, while 5fC and 5caC will appear as unmethylated cytosines (Figure 2.2b). Oxidative bisulfite sequencing allows the separation of 5mC and 5hmC by oxidising 5hmC to 5fC (Booth et al., 2012). Bisulfite sequencing can therefore be used in combination with oxidative bisulfite sequencing to quantify levels of 5mC and 5hmC individually (Figure 2.2c). This was done for a 311bp section of the 18S rDNA (Figure 2.3, Figure 9.1). 10 clones were sequenced for each plant and three plants were analysed per line. Overall levels of 5mC have reduced in Line A plants containing the transgene (Line A

TET3c+), as had previously been seen by Dr Michael Watson, but remained unchanged in transgenic Line B plants (Line B *TET3c*+)(Figure 2.3a, Figure 9.1a). Significant levels of 5hmC were not present in wild type Arabidopsis, as expected, but were present in *TET3c+* plants of both Line A and Line B (Figure 2.3b, Figure 9.1b). Cumulative changes are shown in Figure 2.3; Figure 9.1 shows the same data in more detail.

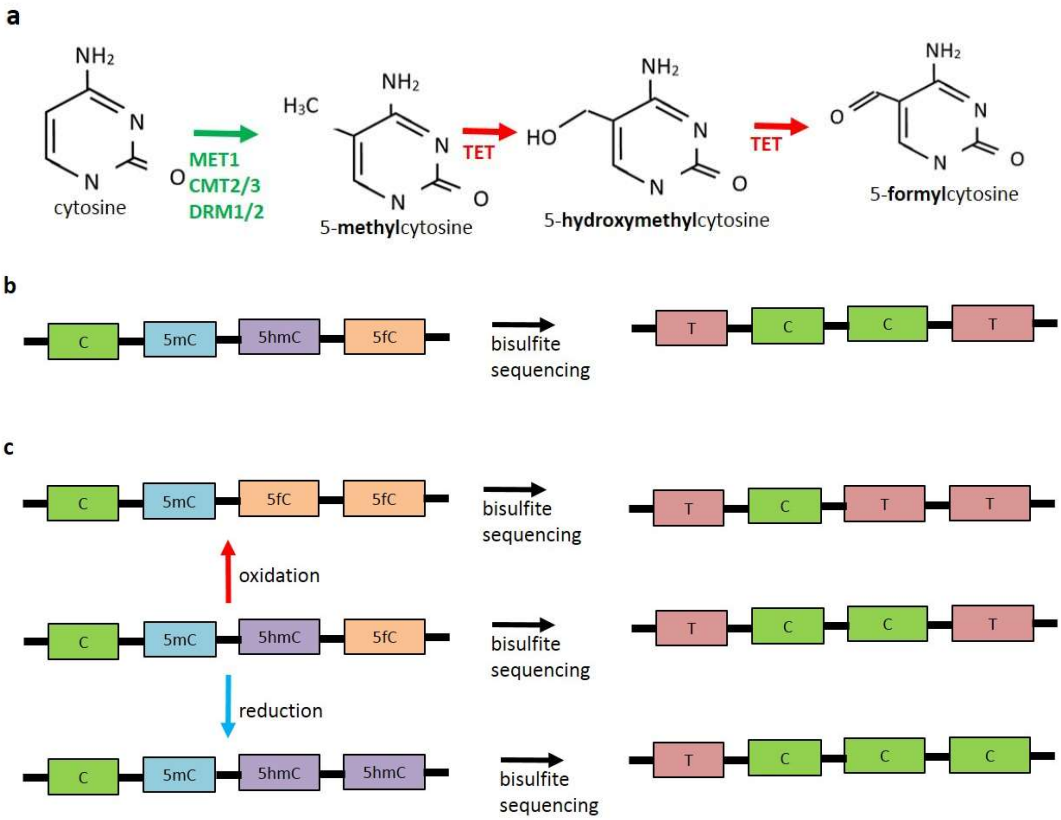


Figure 2.2. TET-mediated demethylation produces intermediates which can be distinguished by oxidative and reduced bisulfite sequencing

(a) TET demethylates DNA by progressively oxidising 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC) and further to 5-formylcytosine (5fC).

(b) Bisulfite sequencing is used to analyse levels of 5mC, which is read as cytosine after bisulfite treatment. Unmethylated cytosine is deaminated to uracil and read as thymine. 5hmC, like 5mC, is not converted to uracil and is therefore read as cytosine while 5fC and 5caC are deaminated and appear as thymine.

(c) Oxidation converts 5hmC to 5fC. Subsequent bisulfite sequencing results in this being read as thymine. Subtraction of the oxidative bisulfite sequencing methylation levels from standard bisulfite sequencing methylation levels calculates levels of 5hmC. Similarly, reduction converts 5fC to 5hmC, subsequent bisulfite sequencing of which results in it being read as cytosine. Subtraction of the standard bisulfite sequencing methylation levels from reduced bisulfite sequencing methylation levels calculates levels of 5fC.

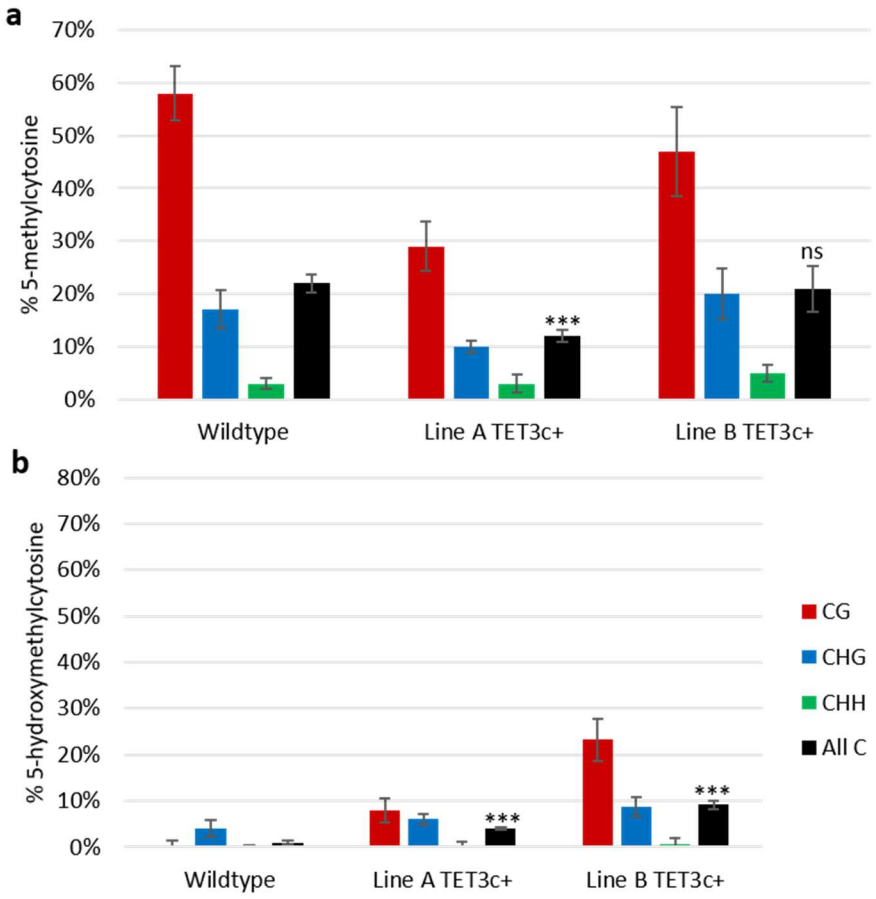


Figure 2.3. Methylation levels in a region of the 18S rDNA of Arabidopsis were altered in TET3c plants

(a) Levels of 5-methylcytosine were reduced in Arabidopsis plants containing TET3c in Line A (Line A TET3c+) but did not change in Line B (Line A TET3c+). The % of 5-methylcytosine was calculated as the number of methylated cytosine bases divided by the total number of cytosine bases in the region. Total levels of 5-methylcytosine across all contexts are indicated in black.

(b) 5-hydroxymethylcytosine could be detected in Line A and Line B plants containing TET3c. Total levels of 5-hydroxymethylcytosine across all contexts are indicated in black.

*Three biological replicates were analysed by bisulfite sequencing for each sample and 10 clones for each replicate. DNA was taken from the rosette leaves of 5 week old plants. Graphs show averages with error bars representing standard error. The significance of a change from WT is indicated by asterisks: ns= not significant, ***=P<0.005, calculated using Student's two-tailed t-test.*

2.2.2. Epigenetic changes caused by TET3 in Arabidopsis can be retained in the absence of the transgene

In order to analyse if the changes in DNA methylation caused by *TET3c* were heritable, heterozygous T1 plants were selfed and plants from which the *TET3c* transgene had been segregated away were selected using PCR. Levels of 5mC and 5hmC were analysed, again by using a combination of bisulfite sequencing and oxidative bisulfite sequencing. Demethylation caused by *TET3c* in Line A was shown to be heritable, while in Line B a significant increase in methylation was observed (Figure 2.4a, Figure 9.2a). 5hmC was no longer present in *TET3c*- plants of either line (Figure 2.4b, Figure 9.2b), indicating that the *TET3c* transgene was responsible for its production. Cumulative changes are shown in Figure 2.4; Figure 9.2 shows the same data in more detail.

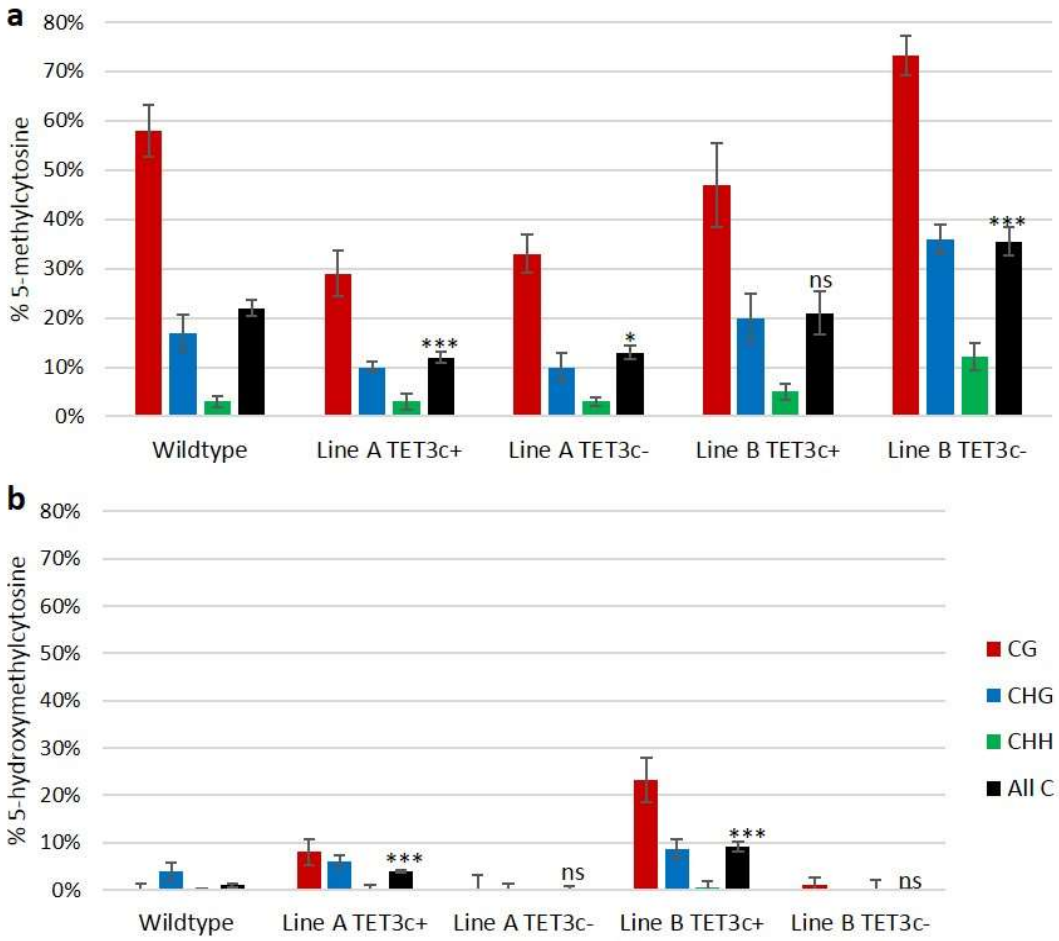


Figure 2.4. Methylation changes observed in Line A TET3c+ were heritable

(a) The reduction in methylation seen in Line A TET3c+ was still present after the transgene had been segregated away (Line A TET3c-). In Line B, an increase of methylation was seen after the TET3c transgene was no longer present (Line B TET3c-). Total levels of 5-methylcytosine across all contexts are indicated in black.

(b) 5-hydroxymethylcytosine was not present in Line A and Line B plants after the TET3c transgene had been segregated away (Line A TET3c- and Line B TET3c-). Total levels of 5-hydroxymethylcytosine across all contexts are indicated in black.

*Three biological replicates were analysed by bisulfite sequencing for each sample and 10 clones for each replicate. Graphs show averages with error bars representing standard error. The significance of a change from WT is indicated by asterisks: ns= not significant, * = P <0.05, ***=P<0.005, calculated using Student's two-tailed t-test.*

2.2.3. 5fC is produced at the rDNA locus in *TET3c*+ Arabidopsis *rdd* glycosylase mutants

Mammalian DNA glycosylases are known to have strong activity for the excision of 5fC and 5caC (He et al., 2011; Iurlaro et al., 2013; Maiti and Drohat, 2011), but not 5hmC (He et al., 2011; Maiti and Drohat, 2011). Three of the four plant DNA glycosylases (DME, DML3 and ROS1) have been shown *in vitro* to excise 5hmC*,* but not 5fC (Brooks et al., 2014; Jang et al., 2014). To investigate their behaviour *in vivo,* triple glycosylase mutants (*ros1-3; dml2-1; dml3-1*) had been transformed with *TET3c* by Dr Michael Watson. These were examined using oxidative bisulfite sequencing to investigate levels of 5mC and 5hmC. Methylation levels at the analysed region of the 18S rDNA were similar in the triple glycosylase mutants (*rdd*) in comparison to wild type *Col-0* Arabidopsis (Figure 2.5a, Figure 9.3a). In *rdd TET3c+* plants, methylation levels at this region were reduced, similarly to the hypomethylation observed in the Line A *TET3c+* plants (Figure 2.5a, Figure 9.3a), demonstrating that TET3c-mediated demethylation at this locus does not require these three plant DNA glycosylases. 5hmC was absent in *rdd* plants without the *TET3c* transgene, and present in *rdd TET3c*+ plants, again at similar levels to that seen in Line A plants (Figure 2.5b, Figure 9.3b), suggesting that these three DNA glycosylases do not have strong 5hmC excision activity *in vivo*.

In mammals, TET-mediated demethylation oxidises 5mC to 5hmC and then continues to oxidise 5hmC to 5fC and from 5fC to 5caC (Ito et al., 2011). 5fC and 5caC appear as unmethylated cytosines in both bisulfite sequencing and oxidative bisulfite sequencing. However, levels of 5fC can be analysed at a single-base resolution using reduced bisulfite sequencing (Booth et al., 2014). This reduces 5fC to 5hmC which appears as a methylated base in bisulfite sequencing. 5fC therefore also appears as a methylated base using reduced bisulfite sequencing, and a combination of reduced bisulfite sequencing and bisulfite sequencing can be used to calculate levels of 5fC (Figure 2.2c). This was done for the *TET3c*+ plants, as well as wild type and *rdd* plants (Figure 2.5c, Figure 9.3c). Significant levels of 5fC were

not found in any plants except *rdd TET3c+* plants, where 5fC was found only in the CG context (shown in red). Cumulative changes are shown in Figure 2.5; Figure 9.3 shows the same data in more detail.

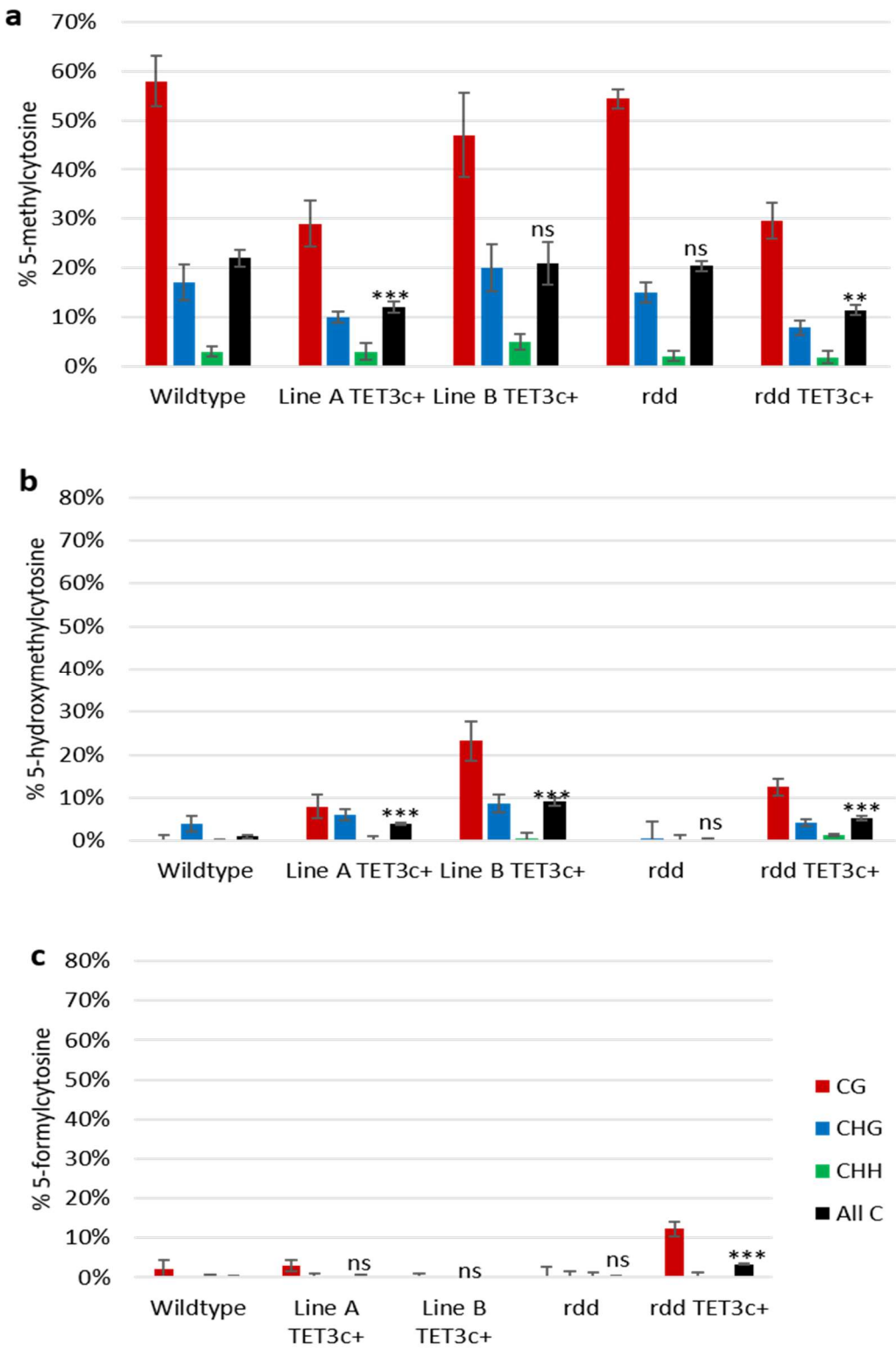


Figure 2.5. Triple glycosylase mutants (ros1-3;dml2-1;dml3-1) with TET3c (rdd TET3c) and without (rdd) could still be demethylated by TET3c and showed increased levels of 5-fC

(a) TET3c caused demethylation in rdd plants, to similar levels as that seen in Line A TET3c+. Total levels of 5-methylcytosine across all contexts are indicated in black.

(b) Levels of 5-hydroxymethylcytosine were not increased in rdd plants in comparison to TET3c+ plants in the wild type background, suggesting that these glycosylases do not excise 5-hydroxymethylcytosine. Total levels of 5-hydroxymethylcytosine across all contexts are indicated in black.

(c) 5-formylcytosine was absent in all plants except rdd TET3c+ plants, where it was present only in the CG context. Total levels of 5-formylcytosine across all contexts are indicated in black.

*Three biological replicates were analysed by bisulfite sequencing for each sample and 10 clones for each replicate. Graphs show averages with error bars representing standard error. The significance of a change from WT is indicated by asterisks: ns= not significant, ** = P <0.01, ***=P<0.005, calculated using Student's two-tailed t-test.*

2.2.4. Investigating the phenotypic behaviour of *TET3c* and *TET1* lines

TET3c+ plants were phenotypically normal except a slight delay in flowering time which was observed (Figure 2.6a). This is shown as the age and number of leaves at bolting (i.e. when the transition from vegetative growth to flowering occurs). As discussed in the introduction, this can be a result of hypomethylation and increased expression of *FWA* (Soppe et al., 2000). Expression levels of *FWA* were analysed but could not be detected in *TET3c+* lines at 30 cycles (Figure 2.6b). Another flowering gene, *TFL1 (*a homologue of the *CEN1.1* gene analysed in Chapter Three) causes delayed flowering when ectopically expressed (Ratcliffe et al., 1998) and is indicated to have dense methylation upstream of the promoter according to the epigenome browser (http://neomorph.salk.edu/epigenome/epigenome.html). *TFL1* expression levels were also analysed (Figure 2.6b) but again could not be detected. Delayed flowering would be expected to result from increased expression of these genes, but neither was expressed at levels high enough for detection. A more

sensitive analysis could be done using quantitative RT-PCR, but expression levels above 30 cycles are unlikely to be biologically significant.

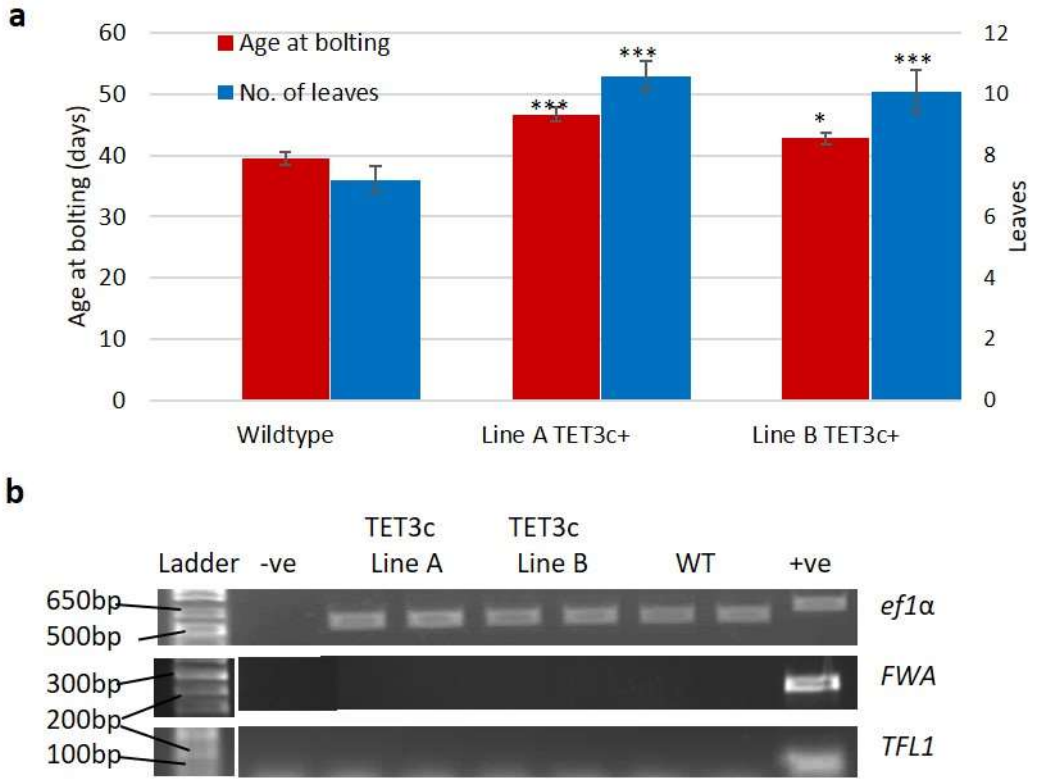


Figure 2.6. Bolting was delayed in TET3c Arabidopsis

(a) Age and number of leaves at bolting of Arabidopsis plants containing TET3c (n=20) was recorded, and delayed bolting in comparison to wild type Arabidopsis was observed.

(b) Plants which contain the TET3c transgene were analysed for FWA (267bp) and TFL1 (181bp) expression using semi-quantitative RT-PCR of cDNA pools (n=10). No expression was detected. 2 technical replicates of each pool were done. cDNA levels were normalised using the constitutively expressed elongation factor 1α (556bp). –ve indicates a negative H_2O control, +ve indicates a positive DNA control. WT indicates a Col-O wild type control. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

*Graphs show averages with error bars representing standard error. The significance of a change from wild type is indicated by asterisks: ns= not significant, * = P <0.05, ***=P<0.005, calculated by Student's two-tailed t-test.*

TET1 and TET2 have been found to have greater oxidative activity in comparison to TET3 *in vitro* (Ito et al., 2011) and therefore a construct containing the full length *TET1* cDNA (including both the catalytic and the CXXC domain) under the constitutive 35S promoter was made and transformed into Arabidopsis (Figure 2.7a) to investigate if more severe phenotypes could be produced. Expression of *TET1* was analysed using semi-quantitative RT-PCR and three lines which express *TET1* were chosen for further analysis (Figure 2.7b). Expression of TET3c had previously been confirmed to be strong in both Line A and Line B of *TET3c* Arabidopsis (Figure 2.7b). *35S::TET1* plants were grown under long day conditions alongside wild type Arabidopsis and analysed for any abnormal phenotypes. However, the only phenotype observed was again a slight delay in flowering time (Figure 2.7c), demonstrating that neither the TET3 nor the TET1 mammalian demethylase has a strong phenotypic effect on *Arabidopsis thaliana.*

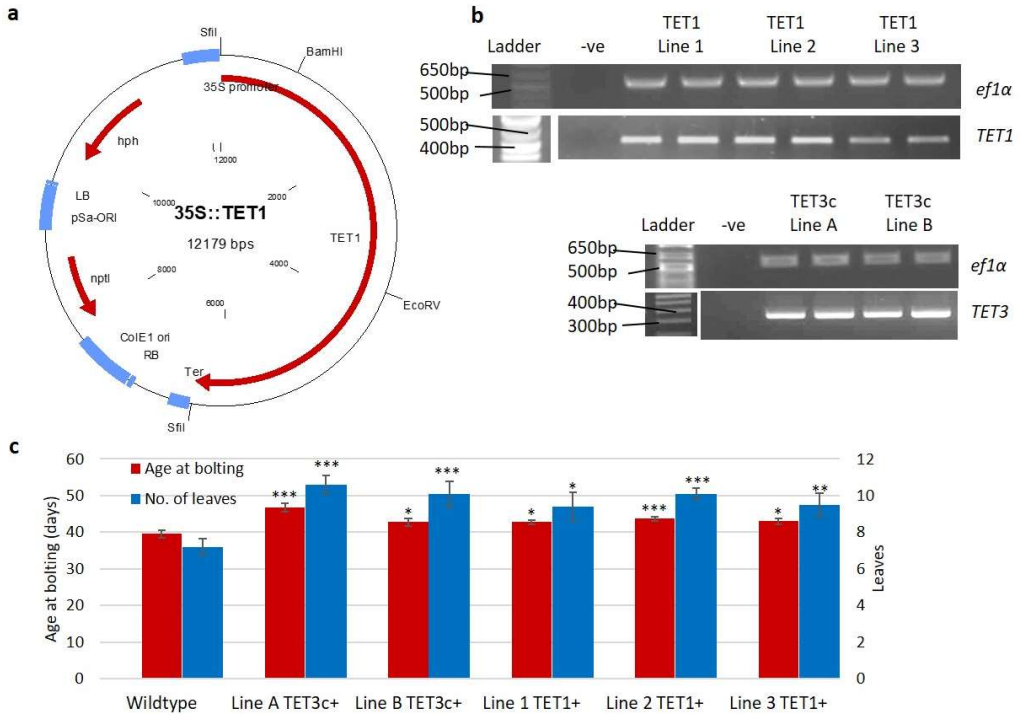


Figure 2.7. Transformants were generated which expressed TET1 but they did not show a more severe phenotype than TET3c Arabidopsis

(a) Map of the plant expression vector pGreen 0029 35S containing the cDNA encoding the full length TET1 protein.

(b) Plants which contain the TET1 cDNA were shown to express it using semi-quantitative RT-PCR of cDNA pools (n=10, 468bp) at 25 cycles. Expression levels were similar to those seen in TET3c lines (330bp). 2 technical replicates of each pool were done. cDNA levels were normalised using the constitutively expressed elongation factor 1α (556bp). –ve indicates a negative H_2O control. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

*(c) Both Arabidopsis plants containing TET1 and those containing TET3c showed delayed flowering in comparison to wild type Arabidopsis, but TET1 plants did not show a more extreme phenotype than TET3c. The age and number of leaves at bolting was recorded. Graphs show averages (n=11-26) with error bars representing standard error. The significance of a change from wild type is indicated by asterisks: ns= not significant, * = P <0.05, ** = P<0.01 ***=P<0.005, calculated by Student's two-tailed t-test.*

2.3. Discussion

Both hypo- and hypermethylation were caused by *TET3c* at the same locus in Arabidopsis. *TET3c* achieves this by oxidising 5-methylcytosine in Arabidopsis as it does in mammals, shown by the presence of 5hmC at the rDNA locus in *TET3c*+ Arabidopsis. The hypomethylation observed in Line A was heritable, appearing when the transgene is present and remaining after it has been lost. In Line B, the hypermethylation was not present in transgenic plants, and only appeared when the transgene had been lost. However, the total amount of methylation and demethylation intermediates (5mC+5hmC) was increased in Line B *TET3c*+ plants, as was previously observed by Dr Michael Watson using a methylation sensitive Southern blot which does not discriminate between 5mC and 5hmC (Hollwey et al., 2016; Watson, 2013). Methylation levels therefore have increased in Line B *TET3c*+ plants in comparison to wild type, but as the 5mC is constantly being oxidised to 5hmC by *TET3c*, overall 5mC levels are unchanged as long as the *TET3c* transgene is present. Once the transgene is lost, both 5mC and 5hmC appear to be retained as 5mC, leading to a hypermethylated state.

Both the new hypomethylated and the new hypermethylated state were maintained across all biological replicates in the two lines, and regardless of the presence of the *TET3c* transgene. This suggests that the presence of *TET3c* initially destabilised the equilibrium level of methylation at this locus, and a new, different equilibrium has been reached. In Line A, the new equilibrium point is hypomethylated; in Line B, the new equilibrium point is hypermethylated. Despite the oxidising action of *TET3c,* complete demethylation did not occur and these new levels of methylation were maintained, and remained after transgene loss. Further work would need to be done to investigate how stable these new equilibria are, and if specific conditions are required to change the equilibrium. Whether the equilibrium changes to a hypo- or hypermethylated state may depend on a variety of factors during the initial event which are likely to vary stochastically. These include the rate at which methylation is restored by MET1 when the cell divides, or

the rate at which TET3 or the glycosylases remove methylation. ROS1, DML2 and DML3 were not required to reach a hypomethylated state, as hypomethylation was observed in the *rdd TET3c*+ line. The fourth plant DNA glycosylase, DME, is thought to act mainly during gametogenesis, where it is required for endosperm gene imprinting (Choi et al., 2002), but may be capable of contributing to the demethylation that occurs here, as we do not know at which stage this occurs.

5hmC was present in all *TET3c*+ plants analysed. While much higher levels of 5hmC were observed in Line B *TET3c*+ plants in comparison to Line A or *rdd* plants, this reflected a higher overall level of methylation. The amount of 5hmC as a percentage of 5mC remained largely consistent: 33% in Line A *TET3c*+, 43% in Line B *TET3c*+, 46% in *rdd TET3c*+. The proportion of 5hmC to 5mC was slightly higher in *rdd TET3c*+ plants. This would be expected if the glycosylases were excising 5hmC in Line A and Line B *TET3c*+ plants *in vivo* as they have been shown to do *in vitro.* However, more lines would need to be investigated to determine if this is significant. If the glycosylases are excising 5hmC, it must be being produced by *TET3c* more quickly than it is excised, suggesting that plant DNA glycosylases do not have a high activity for excising 5hmC *in vivo*.

5fC was detected only in the *rdd TET3c*+ line. This suggests that *in vivo,* 5fC is excised by one of the three glycosylases mutated in this line (DML2, DML3 and ROS1). It has been shown that ROS1 and DML3 do not excise 5fC *in vitro* (Brooks et al., 2014; Jang et al., 2014) and it is currently unknown if DML2 can excise 5fC *in vitro.* DML2 may therefore be responsible for 5fC excision in *TET3c+* plants, or DML3 and ROS1 may behave differently *in vivo.* 5fC was observed in cytosines of the CG context alone. This may mean that *TET3c* only produces 5fC in the CG context or that 5fC is oxidised more quickly to 5caC in the CHG and CHH contexts. However, given the lower levels of methylation in the CHG and CHH contexts in comparison to the CG context, it is also possible that levels of 5fC in the CHG and CHH contexts were too low to be detected using reduced bisulfite sequencing.

TET3c expression and *TET1* expression was strong in all lines analysed. Despite the higher levels of oxidative activity observed for TET1 *in vitro* compared to TET3 (Ito et al., 2011), phenotypes in *TET1+* Arabidopsis were not more severe than those seen in the *TET3c+* Arabidopsis. There are several possible reasons for this. The full length *TET1* cDNA was used, in comparison to the *TET3c* transformants which contained only the catalytic domain of TET3. It is possible that expressing the full length version of *TET1* may reduce the processivity of TET1 in plants. The full length version of *TET1* will contain the CXXC domain, which is also present in DNMT1 and TET3. In the DNMT1 protein, the CXXC domain binds to unmethylated DNA and has an inhibitory effect (Song et al., 2011), while in the TET3 protein, the CXXC domain binds to 5caC and unmethylated cytosines, restricting its activity to particular regions (Jin et al., 2016). If the CXXC domain of TET1 also has a negative role on its activity, the TET1 cDNA might be expected to be less effective than the catalytic domain of TET3. Alternatively, although TET1 has a greater oxidative activity *in vitro,* this may not translate to a greater oxidative activity *in vivo* and in a plant model. Finally, it may be that *TET1* has a greater oxidative activity than *TET3c* in plants, but this does not translate to more severe phenotypic changes in Arabidopsis. Methylation levels at the ribosomal DNA in *TET1* Arabidopsis were not analysed, as it had already been shown that methylation changes could occur, but *TET1* may cause more severe changes than were observed in *TET3c* lines. In order to distinguish between these possibilities, it would be useful to express the catalytic domain of TET1 alone in Arabidopsis, or to investigate the expression of *TET1* and *TET3c* in plants that are more sensitive to changes in DNA methylation.

During the writing of this thesis, data on the transformation of the catalytic domain of TET1 was made available online (Ji et al., 2017). The work largely agreed with our data, showing significant demethylation in the CG context and to a lesser extent in the CHG and CHH context. *35S::TET1cd* Arabidopsis also showed a delay in flowering, but in these lines this could be attributed to a change in the expression of *FWA.* Ji et al did not find evidence for 5-hydroxymethylcytosine in their *35S::TET1cd* Arabidopsis. However, this difference may be due to a difference in

methods, as we used oxidative bisulfite sequencing, while Ji and colleagues used TAB (TET-assisted bisulfite) sequencing. If 5-hydroxymethylcytosine occurs at low levels or has a variable location between biological replicates, then this method may not detect it.

The data in this chapter demonstrates that the catalytic domain of the mammalian DNA demethylase TET3 can be active in plant species, and can cause heritable changes in DNA methylation. Changes in methylation were observed only at a region of the ribosomal DNA, not a protein-coding gene and only minor phenotypes were observed. This may be due to the resilience of Arabidopsis when it comes to changes in methylation, as discussed in the introduction. To investigate the utility of *TET3c* as a tool for generating epigenetic diversity in plants it is important to find *TET3c* mediated changes in methylation at protein-coding genes, which can be linked to changes in gene expression. In order to do this, transformation of *TET3c* into other plant species can be carried out, which is discussed in later chapters.

3. *TET3c* Expression in Tomato Identifies a New Gene Promoting Vegetative Growth Whose Expression Correlates with Methylation Changes

3.1. Introduction

The role of DNA methylation is likely to vary in different species, even within the plant kingdom. Methylation levels are low in Arabidopsis thaliana (~5% total methylation)(Cokus et al., 2008; Feng et al., 2010) in comparison to rice (~15%)(Feng et al., 2010) or tomato (~25%) (Messeguer et al., 1991; Zhong et al., 2013). Homozygous mutation of *MET1* in rice results in seedling lethality (Hu et al., 2014) while Arabidopsis plants containing an RNAi construct against *MET1* (Finnegan et al., 1996) or a mutated *MET1* gene (Kankel et al., 2003) are viable, although later generations show more severe phenotypes (Finnegan et al., 1996; Mathieu et al., 2007). Tomato transformants containing an antisense construct against tomato *MET1*, meanwhile, cannot be regenerated (Watson, 2013).

The catalytic domain of TET3, a mammalian DNA demethylase, can heritably alter DNA methylation in Arabidopsis (Hollwey et al., 2016). The only phenotypic effect of this in Arabidopsis is a mild delay in flowering, as discussed in Chapter Two of this thesis. However, Arabidopsis may not be the best model plant in which to analyse the effects of the *TET3c* construct. As discussed above, mutation of the main methyltransferase is less harmful in Arabidopsis than in other plant species and levels of methylation are lower. Dense methylation (methylation in all three cytosine contexts (Watson et al., 2014)) may be more common in tomato in comparison to in Arabidopsis, as 13.2% of tomato genes containing introns are densely methylated specifically in the intron (i.e. not in the exon), a huge increase on the 0.7% of Arabidopsis introns with intron-specific dense methylation (M. Wilson, personal communication, 2016).

Tomato is also an extremely important crop plant in which inhibition of DNA methylation is known to induce premature ripening (Zhong et al., 2013). Naturally

occurring epialleles such as *Cnr* and *sulfurea* result in defects in ripening (Manning et al., 2006) and pigment production respectively (Gouil et al., 2016). Mutants of the main methyltransferases do not yet exist in tomato, but a RNAi knockdown of *SlDML2*, a tomato orthologue of the Arabidopsis DNA glycosylase *ROS1* which carries out active DNA demethylation, has been analysed (Liu et al., 2015). Four DNA glycosylases exist in both Arabidopsis and tomato that can carry out active DNA demethylation. In Arabidopsis, triple mutants of three of the four of these show little change in phenotype or gene expression (Penterman et al., 2007a). In tomato, transgenic tomato lines containing an RNAi construct against *SlDML2* exhibit ripening inhibition, showing that active DNA demethylation controls fruit ripening in tomato (Liu et al., 2015). Null mutants of *SlNRPEI*, a component of the RdDM pathway, are lethal in tomato (Gouil and Baulcombe, 2016) while in Arabidopsis these mutants are fully viable and lack obvious morphological defects (Kanno et al., 2005). This, and the fact that *MET1* RNAi tomato transformants cannot be produced, suggests that altering the methylation systems of tomato produces a stronger phenotype than in Arabidopsis. These facts make tomato a good candidate for transformation with the *TET3c* construct, which produces changes in methylation at certain loci but milder phenotypes in comparison with mutation of a methyltransferase. *TET3c* was therefore transformed into tomato and the effects were analysed on a phenotypic and molecular level.

3.2. Results

3.2.1. The catalytic domain of TET3 was expressed at a high level in tomato

The *35S::TET3c* construct was transformed into tomato by Dr Michael Watson with the assistance of Dr Iris Heidmann and Suzan Out at Enza Zaden. 1220 explants were generated, resulting in the production of six independent transformant lines. Expression levels of the *TET3c* cDNA were investigated using semi-quantitative RT-PCR (Figure 3.1). *TET3c* was expressed at a high level in all six *TET3c* tomato lines in the T1 generation.

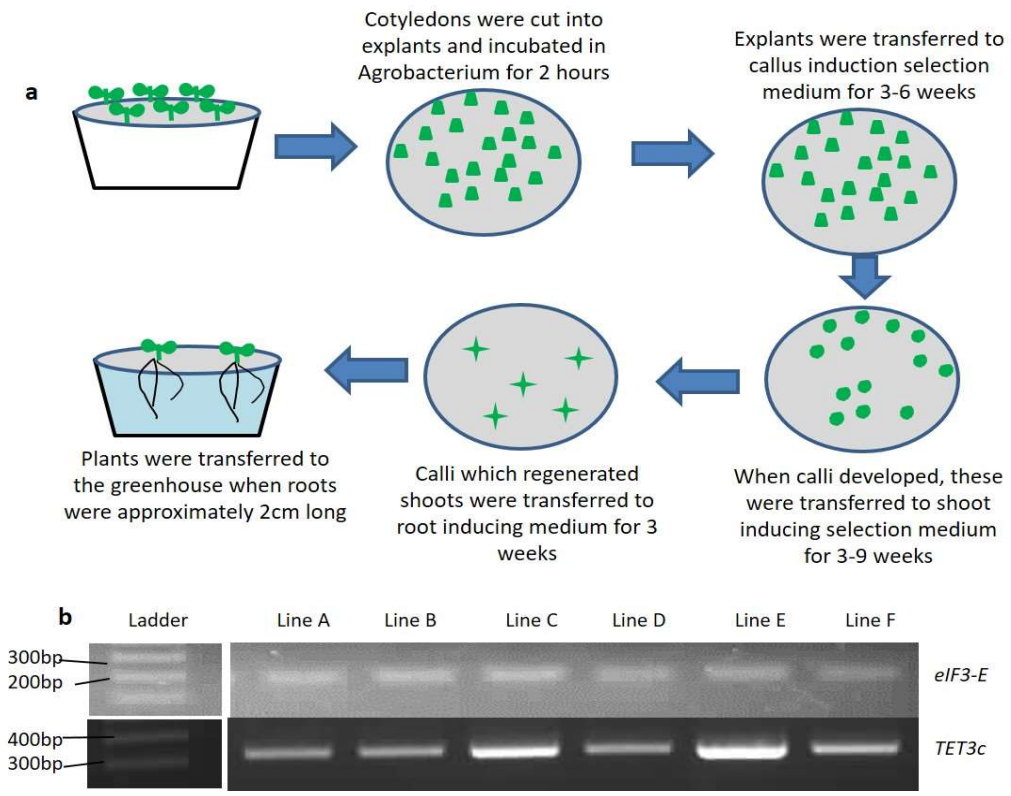


Figure 3.1. TET3c was transformed into tomato where TET3 expression was high

(a) Tomato were transformed with the 35S::TET3c construct. Cotyledons of 10 day old seedlings were dissected into small squares and cultured on Agrobacterium plates containing the construct. These were then placed on medium first to induce callus formation, then to induce shoot formation and finally to induce root growth. This process took a minimum of 3 months.

(b) Tomato containing the TET3c transgene were shown to express it using semi-quantitative RT-PCR at 25 cycles of cDNA pools (n>8). TET3c (330bp) was expressed more strongly than the constitutively expressed eukaryotic initiation factor 3E (150bp), used for normalisation of cDNA levels. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

3.2.2. Expressing TET3c in tomato caused multiple phenotypic effects

TET3c tomato displayed multiple phenotypes (Figure 3.2). The abnormalities with the shoot apical meristem (Figure 3.2b) are discussed further in Chapter Four. Changes in shoot architecture were observed in adult *TET3c* tomato plants including determinate growth (plants terminating with a flower) and high variability in the distance between leaves (Figure 3.2c). In order to investigate these phenotypes in more detail, 18 week old tomato plants of three *TET3c* lines (Line A, Line D and Line F) were analysed.

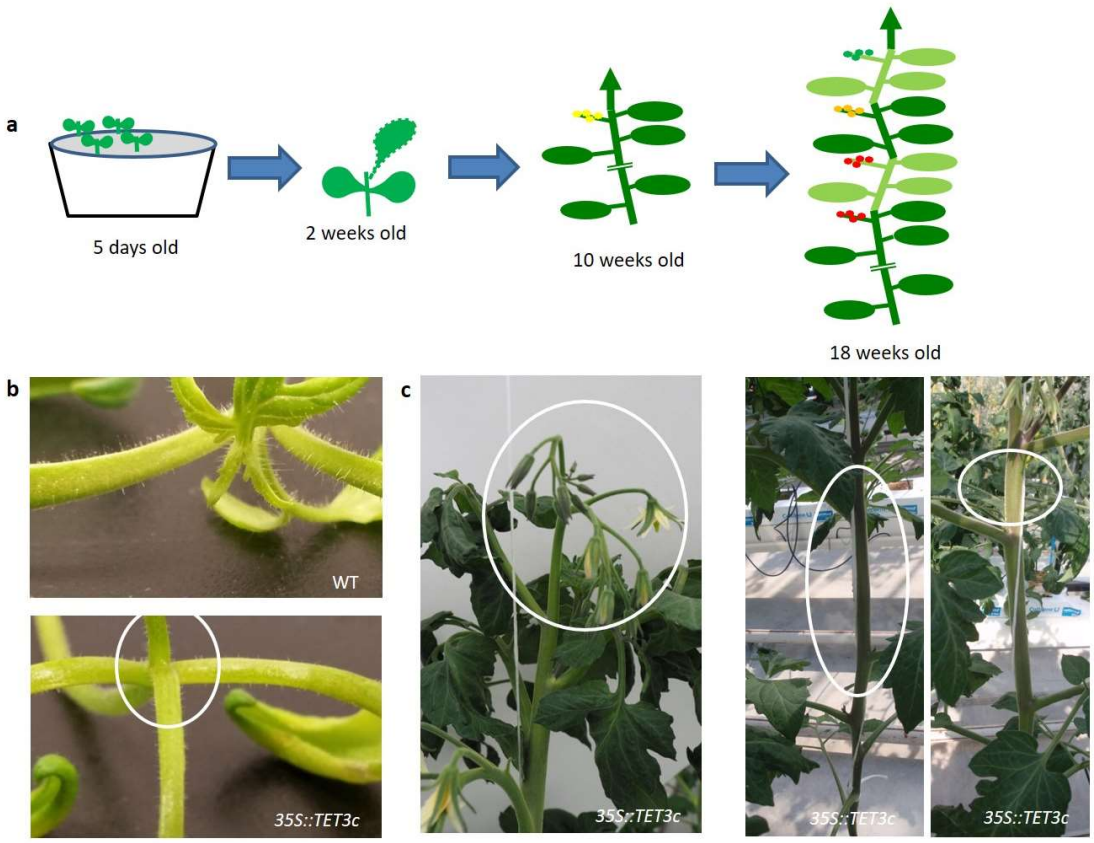


Figure 3.2. TET3c tomato had multiple phenotypes

(a) The tomato life cycle takes 4-6 months to complete, depending on season. Seedlings germinate and cotyledons expand, followed by the development of true leaves. The first inflorescence starts to develop between 8-10 weeks and the first fruit will be ripe at approximately 16 weeks.

(b) Five week old TET3c tomato seedlings showed abnormalities with shoot meristem maintenance. The white circle indicates the abnormal meristem.

(c) Adult TET3c tomato showed altered shoot architecture including terminal flowers and variation in the distance between leaves. The white circle indicates abnormalities, either the flower in the terminal position or variable lengths of the stem between leaves.

Tomato shows a sympodial growth pattern (Figure 3.3a) meaning that it is composed of a series of determinate units. Each unit comprises a period of vegetative growth, terminating in an inflorescence (McGarry and Ayre, 2012). Vegetative growth of the tomato then continues from the uppermost axillary meristem (Lifschitz et al., 2006). The initial growth of the tomato, known as the primary shoot, grows for 8-12 leaves before terminating (Lifschitz et al., 2006). Subsequent growth occurs in sympodial units, each of which contains three leaves on average and again terminates in an inflorescence (Lifschitz et al., 2006). Mutations in different tomato genes can result in changes in the number of leaves in both the primary stem and in sympodial units, for example mutations in *SELF PRUNING (SP)* decrease the number of leaves before the inflorescence (Shalit et al., 2009), whilst mutations in *SINGLE FLOWER TRUSS (SFT)* increase the number of leaves between each inflorescence (Lifschitz et al., 2006).

In the *TET3c* tomato plants, the number of leaves before the inflorescence in the primary stem and each sympodial unit was counted, as was the length of each unit. The distance between each individual leaf was also counted (internode length). A significant increase in both the length of and the number of leaves in the primary shoot was observed in *TET3c* tomato in comparison to control tomato plants containing only the kanamycin resistance gene (Figure 3.3b,d). Internode length also increased significantly in *TET3c* lines (Figure 3.3). *TET3c* tomato were obtained from the selfing of a heterozygous T1 *TET3c* tomato plant and therefore represent a population in which the transgene is segregating at a 3:1 ratio, with three-quarters of plants containing the transgene and one-quarter having lost it. *TET3c* were

genotyped for the presence of the transgene. Tomato which retained the transgene were denoted as *TET3c+* and tomato from which the transgene had been segregated were denoted as *TET3c-.* The number of leaves and length of the primary stem and subsequent sympodial units in *TET3c+* plants was compared to *TET3c-* tomato to investigate the heritability of the phenotype (Figure 3.3e). No significant change was observed, indicating that this phenotype is heritable.

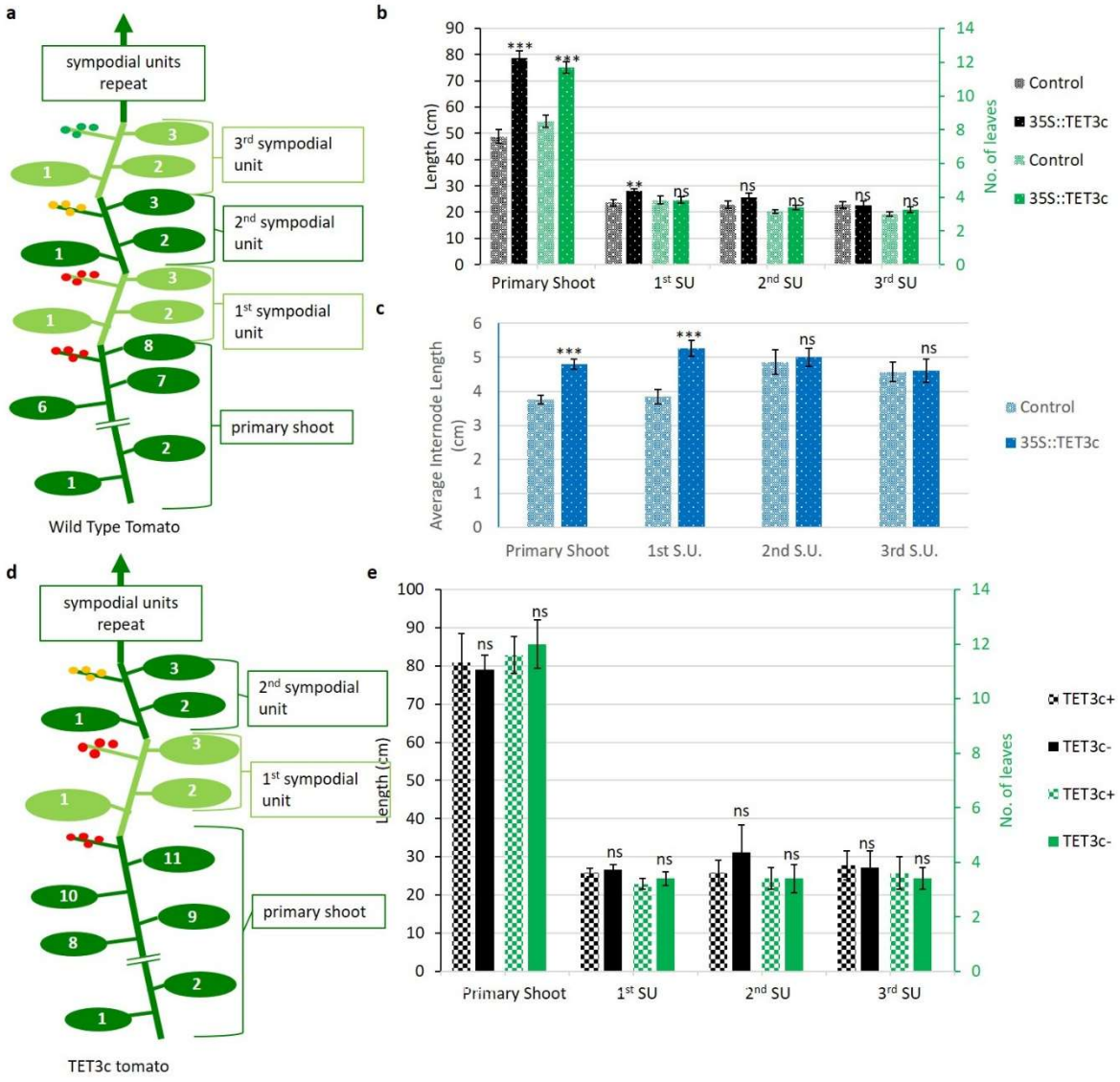


Figure 3.3. TET3c tomato had heritably increased length of and number of leaves in the primary shoot

(a) Tomato plants show a sympodial structure, made up of multiple units consisting of a period of vegetative growth terminating with a flower. Growth of the next sympodial unit begins from the axillary meristem of the uppermost leaf of the previous sympodial unit. Different colours of green have been used to indicate different sympodial units. Colour of fruits indicates ripeness.

(b) The length of shoot (black) and number of leaves (green) in the primary shoot increased in TET3c tomato (n=33) compared to the control (n=24).

(c) The distance between each leaf (internode length) was significantly increased in TET3c tomato in early sympodial units.

(d) The pattern of growth in a TET3 transformant was altered in comparison to the wild type tomato, as shown above.

(e) 35S::TET3c tomato were genotyped and the growth of plants with (TET3c+, n=5) and without (TET3c-, n=5) the transgene was compared. No significant difference in the length of shoot (black) and number of leaves (green) between inflorescences was observed between TET3c+ and TET3c- tomato and therefore these phenotypes in the TET3c tomato are heritable.

*Plants were measured at 18 weeks old. Graphs show averages with error bars representing standard error. The significance of a change from the control is indicated by asterisks: ns= not significant, **=P<0.01, ***=P<0.005, calculated using Student's two-tailed t-test.*

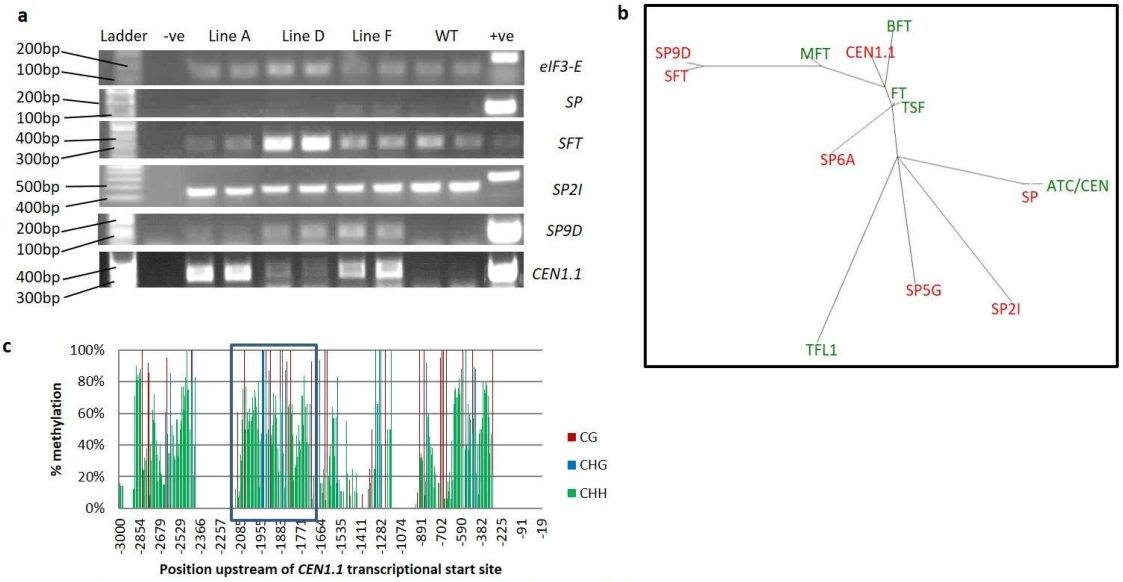
3.2.3. A new homologue of the *CETS* gene family was ectopically expressed in *TET3c* tomato

As discussed earlier, *sft* tomato mutants also show an increase in the size of the primary shoot. This phenotype is also observed when *SP* is constitutively expressed behind the 35S promoter (Pnueli et al., 1998). Based on gene sequence, *SP* and *SFT* both belong to the same family of flowering genes known as the *CETS (CENTRORADIALIS/TERMINAL FLOWER 1/SELF-PRUNING)* family which also contains the *Arabidopsis thaliana* flowering genes *FT* and *TFL1* (Carmel-Goren et al., 2003; Lifschitz et al., 2006; Pnueli et al., 1998)*(*Figure 3.4b). This family contains thirteen genes in tomato (Cao et al., 2016; Carmel-Goren et al., 2003), the expression patterns of five of which were investigated in *TET3c* tomato using semi-quantitative RT-PCR (Figure 3.4a). Expression of *SP* was not detected in *TET3c* or wild type tomato at 30 cycles, and therefore alterations in its expression pattern could not be analysed. Expression of two other genes was observed in these plants and was increased in *TET3* tomato lines. *SP9D (Solyc09g009560.1.1)* and *CEN1.1* (*Solyc03g026050.2.1)* have 69.1% and 69.3% identity (based on amino acid sequence) to *SP* respectively and are both *TFL1*-like members of the *CETS* family based on amino acid prediction (Cao et al., 2016)(Figure 3.4b). Expression of *SP9D* and *CEN1.1* was activated in leaves of *TET3c* tomato of all lines analysed in comparison to wild type leaves, where they were not expressed.

3.2.4. Hypomethylation and 5-hydroxymethylcytosine at the promoter of *TET3c* plants correlated with ectopic *CEN1.1* expression

According to the tomato methylation database (Zhong et al., 2013), high levels of methylation can be observed in the 3kb upstream of the *CEN1.1* transcriptional start site (Figure 3.4c). Across all contexts, 37% of cytosines are methylated, and 31% of all cytosines in the *CHH* context are methylated. This is an unexpectedly high number, as levels of methylation in the *CHH* context are the lowest of the three contexts in plants. In comparison, *SP9D* contained only 13% total methylation, and only 6% CHH methylation. Further work was therefore focused on *CEN1.1*, which in

addition to overall high levels of methylation contained a region of densely methylated DNA whose DNA sequence had high homology to the tomato RK01 TRIM retrotransposon (Figure 3.4d)(Witte et al., 2001).



d

Solanum lycopersicum retrotransposon RK01 TRIM, complete sequence

Sequence ID: EU177768.1 Length: 351 Number of Matches: 3

Range 1: 1 to 349 GenBank Graphics Next Match

Score	Expect	Identities	Gaps	Strand
483 bits(261)	2e-132	324/354(92%)	6/354(1%)	Plus/Plus

```
Query  1    TGTTAAAGAATGATAAAAGTCCTATATCGGTGGTTAATGAGATGGGTGGACTCCTTATAA  60
            ||||||||||||| |||||||| | ||||||| ||||||||||||||||||||||||||
Sbjct  1    TGTTAAAGAATGACAAAAGTCC-ACATCGGT-GTTAATGAGATGGGTGGACTCCTTATAA  58

Query  61   GGCTTGGACAATCCTCCTTCCTTTGAATTAACTTTTGGGGTGTGAGTTAGACCTAATACC  120
            ||||||| |||||||||| |||||||  || ||||||||||||||||||| ||||| |||
Sbjct  59   GGCTTGGGCAATCCTCCTCCCTTTGAGCTAGCTTTTGGGGTGTGAGTTAGGCCTAAGACC  118

Query  121  TCATTTTACATAGTATCAGAGCAGGGTCCGTCTCACCCGATGTTGGGACCCCCCAAAATC  180
            | ||||  |||| |||||| |||||||  || ||||||||||||||||| ||||||||||
Sbjct  119  TAATTTCACATGGTATCAAAGCAGGGCCCATCTCACCCGATGTTGGGA-CCCCCAAAATC  177

Query  181  AAAATTGACCACGCACCAGATGCTAAGCACTGGACGTGAGGTGGGGTGTTAAAGAATGAC  240
            |||||||  |||| ||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  178  AAAATTGCCCACACACCAGATGCTAAGCACTGGGCGTGAGGTGGGGTGTTAAAGAATGAC  237

Query  241  AAAAGTCATACATCGGTGGTTAATGAAATGGGTGGACTCCTTATAAGGCTT-GGCAATTC  299
            |||||||  |||||||| |||||||| |||||||||||||||||||||| |||||| |
Sbjct  238  AAAAGTC-CACATCGGT-GTTAATGAGATGGGTGGACTCCTTATAAGGCTTGGGCAATCC  295

Query  300  TCCTCCCTTTGAGCTAGCTTTTGGGGTGTGAGTTAGACCTAATACCTCATTTCA  353
            ||||||||||||||||||||||||||||||||||||||||||| |||| ||||||
Sbjct  296  TCCTCCCTTTGAGCTAGCTTTTGGGGTGTGAGTTAGGCCTAAGACCTAATTTCA  349
```

Figure 3.4. CEN1.1 showed ectopic expression in TET3c tomato leaves and was highly methylated in a repetitive region of the promoter

(a) TET3c tomato were analysed using semi-quantitative RT-PCR at 30 cycles of cDNA pools (n>8). CEN1.1 (301bp) and SP9D (194bp) were ectopically expressed in TET3c leaves. cDNA levels were normalised using the constitutively expressed eukaryotic initiation factor 3E (150bp). –ve indicates a negative H_2O control and +ve indicates a positive DNA control. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

(b) A phylogeny of a selection of tomato and Arabidopsis CETS family members, based on their protein sequence. Tomato genes are in red and Arabidopsis genes are in green.

(c) Methylation levels upstream of the CEN1.1 transcriptional start site are high according to the tomato epigenome database. The region in the box indicates the sequence analysed in (c).

(d) A sequence with high homology to the RK01 TRIM retrotransposon can be observed upstream of the CEN1.1 transcriptional start site. The homologous sequence is highlighted in a box in Figure 3.4c (-2089 to -1737) and its alignment to the RK01 TRIM retrotransposon is shown below. Base number 1 in the

Using primers for this region, quantitative PCR was carried out on DNA immunoprecipitated with an antibody directed against hydroxymethylcytosine. A 5hmC signal can also be observed in the wild type at an increase over the negative IgG control, indicating that either 5hmC is present in tomato, unlike in Arabidopsis, or that specificity of the 5hmC antibody is reduced in plant tissue in comparison to mammalian tissue where it has been tested. However, significant enrichment over the wild type was seen in the *TET3c+* samples in comparison to the wild type (Figure 3.5a) indicating an increase in 5-hydroxymethylcytosine levels and therefore potential *TET3c* activity. Bisulfite sequencing was used to analyse whether levels of methylation had changed in this region in leaves of *TET3c* lines where *CEN1.1* was being ectopically expressed in comparison to leaf tissue where *CEN1.1* was silenced (Figure 3.5b, Figure 9.4). Rather than using oxidative bisulfite sequencing, tissue

from *TET3c*- plants which ectopically expressed *CEN1.1* was used to ensure the absence of 5-hydroxymethylcytosine, which could otherwise conceal a reduction in methylation levels. Hypomethylation in this region was found, correlating with the activation of *CEN1.1* expression in *TET3c* plants (Figure 3.5b, Figure 9.4a). Cumulative changes are shown in Figure 3.5b; Figure 9.4a shows the same data in more detail.

3.2.5. *CEN1.1* showed a restricted expression pattern across tomato tissues, which correlated with upstream methylation

Expression analysis of *CEN1.1* was done in five week old wild type tomato in order to see where *CEN1.1* was naturally expressed (Figure 3.5c). *CEN1.1* expression was analysed in leaves, roots and shoot apices using semi-quantitative RT-PCR. *CEN1.1* was found to be expressed at high levels in the shoot apex, and also expressed at lower levels in root tissue. In order to see whether expression correlated with hypomethylation in wild type tissues as well as in *TET3c* plants, bisulfite sequencing was carried out on wild type leaf tissue and root tissue. DNA was extracted from shoot apical meristem (SAM) cells within the shoot apex using a dissection microscope and bisulfite sequencing was also carried out on these samples. DNA was shown to be hypomethylated at the *CEN1.1* promoter in DNA from roots and the SAM in comparison to DNA from leaves, where *CEN1.1* is silenced (Figure 3.5d, Figure 9.4b). CHH methylation in the SAM was also significantly reduced compared to the roots ($p=0.038$) in addition to the more significant reduction compared to leaves ($p=0.00023$). Demethylation was specifically found in both tissues for cytosines in the CHH context, a phenomenon which has also been observed in cucumber in response to low temperature conditions (Lai et al., 2017). Cumulative changes are shown in Figure 3.5d; Figure 9.4b shows the same data in more detail.

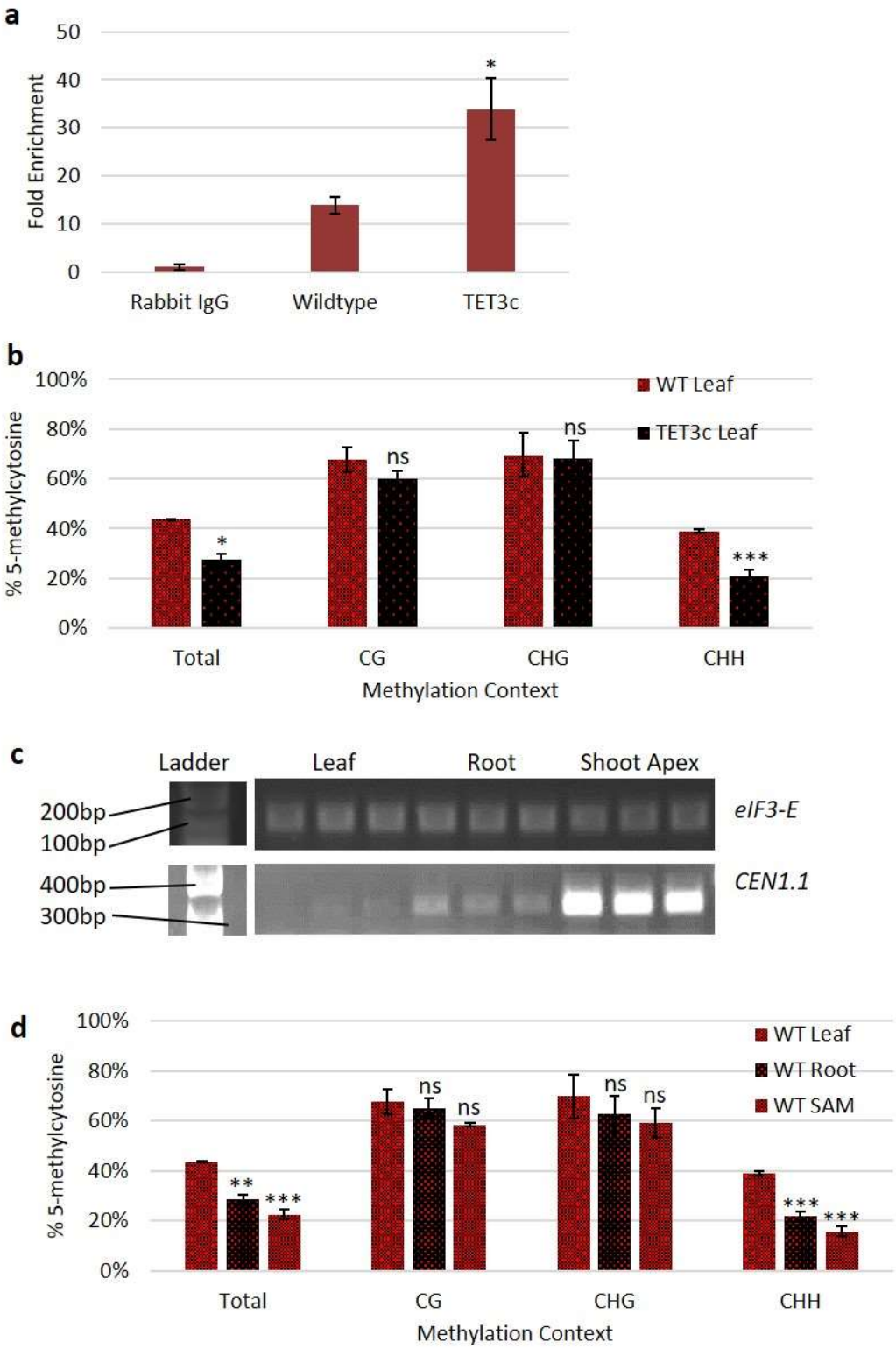

a
Fold Enrichment
50
40
30
20
10
0
*
Rabbit IgG
Wildtype
TET3c
b
% 5-methylcytosine
100%
80%
60%
40%
20%
0%
WT Leaf
TET3c Leaf
*
ns
ns
***
Total
CG
CHG
CHH
Methylation Context
c
Ladder
Leaf
Root
Shoot Apex
200bp
100bp
400bp
300bp
eIF3-E
CEN1.1
d
% 5-methylcytosine
100%
80%
60%
40%
20%
0%
WT Leaf
WT Root
WT SAM
**
***
ns
ns
ns
ns
***
***
Total
CG
CHG
CHH
Methylation Context

Figure 3.5. Expression of CEN1.1 correlated with hypomethylation in the promoter in wild type and TET3c tissues

(a) *Quantitative PCR was carried out on DNA immunoprecipitated with an antibody against 5-hydroxymethylcytosine. Significantly greater amplification of a 206bp region within the repetitive sequence upstream of the CEN1.1 transcriptional start site was seen, indicating the presence of 5hmC in this region. Three biological replicates were carried out for each sample. Fold enrichment was calculated by dividing the quantity of DNA in the immunoprecipitated DNA sample (calculated using a standard curve of input DNA) by the quantity of DNA in the negative rabbit IgG control.*

(b) *Levels of 5-methylcytosine were reduced in the same 206bp region upstream of the CEN1.1 transcriptional start site in TET3c tomato leaves compared to wild type tomato leaves. Three biological replicates were analysed for each tissue and 10 clones for each replicate. The % of 5-methylcytosine is calculated as the number of methylated cytosine bases divided by the total number of cytosine bases in the region.*

(c) *Expression of CEN1.1 (301bp) in leaf, root and shoot apex tissues from wild type tomato was analysed using semi-quantitative RT-PCR of cDNA pools (n>3) at 30 cycles. CEN1.1 was silenced in the leaf, weakly expressed in the root and strongly expressed in the shoot apex. cDNA levels were normalised using the constitutively expressed eukaryotic initiation factor 3E (150bp). Three technical replicates were done for each tissue. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).*

(d) *Levels of 5-methylcytosine were reduced in the same 206bp region upstream of the CEN1.1 transcriptional start site in both of the wild type tomato tissues which express the CEN1.1 gene. Three biological replicates were analysed for each tissue and 10 clones for each replicate. The % of 5-methylcytosine is calculated as the number of methylated cytosine bases divided by the total number of cytosine bases in the region.*

*Graphs show averages with error bars representing standard error. The significance of a change from the control is indicated by asterisks: ns= not significant, *=P<0.05, **=P<0.01, ***=P<0.005, calculated using Student's two-tailed t-test.*

3.2.6. Expressing *CEN1.1* behind a constitutive promoter in tomato resulted in increased vegetative growth and delayed flowering

The *CEN1.1* gene is 1074 base pairs long and contains 4 exons. It forms a protein of 175 amino acids containing one predicted domain, the PBP domain (Figure 3.6a). This domain is a phosphatidylethanolamine binding protein domain, present in all members of the *CETS* gene family where it has been theorised to play a role in kinase-signalling inhibition (Banfield and Brady, 2000). CEN1.1 is predicted to be a floral repressor according to the predicted structure of the binding pocket (Cao et al., 2016) To investigate the effects of ectopically expressing *CEN1.1,* a construct containing the *CEN1.1* genomic DNA behind the constitutive 35S promoter was made (Figure 3.6b). The *35S::CEN1.1* construct was transformed into tomato with the assistance of Dr Iris Heidmann and Suzan Out at Enza Zaden. Successful transformants were selected by growth on antibiotic selection and genotyping by PCR, using primers which would amplify only the transgenic copy of *CEN1.1*, not the endogenous copy. These were designed so that the reverse primer would anneal to the nos terminator (Ter on the map, Figure 3.6b), while the forward primer was within the *CEN1.1* gene. The nos terminator is not found downstream of the endogenous *CEN1.1* gene and therefore amplification would only occur from DNA containing the transgenic *CEN1.1* copy. Expression levels were confirmed using semi-quantitative RT-PCR on true leaves. As the wild type leaf does not express *CEN1.1*, transgene-specific primers were not required to discriminate between the expression of the endogenous or transgenic copy. Independent lines which expressed the *CEN1.1* gene at high level were generated (Figure 3.6c).

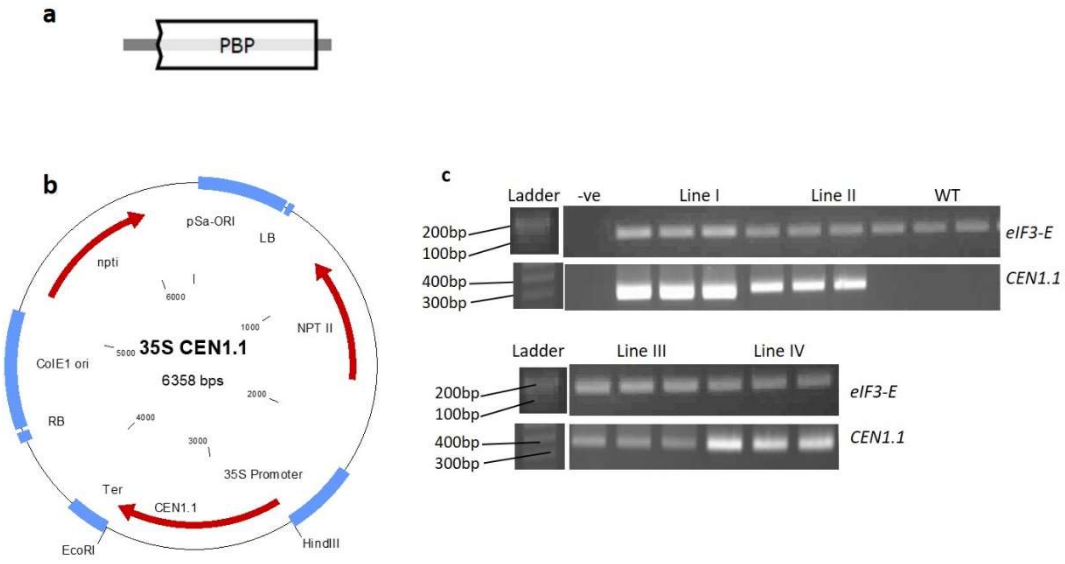


Figure 3.6. Transformants were generated which express CEN1.1

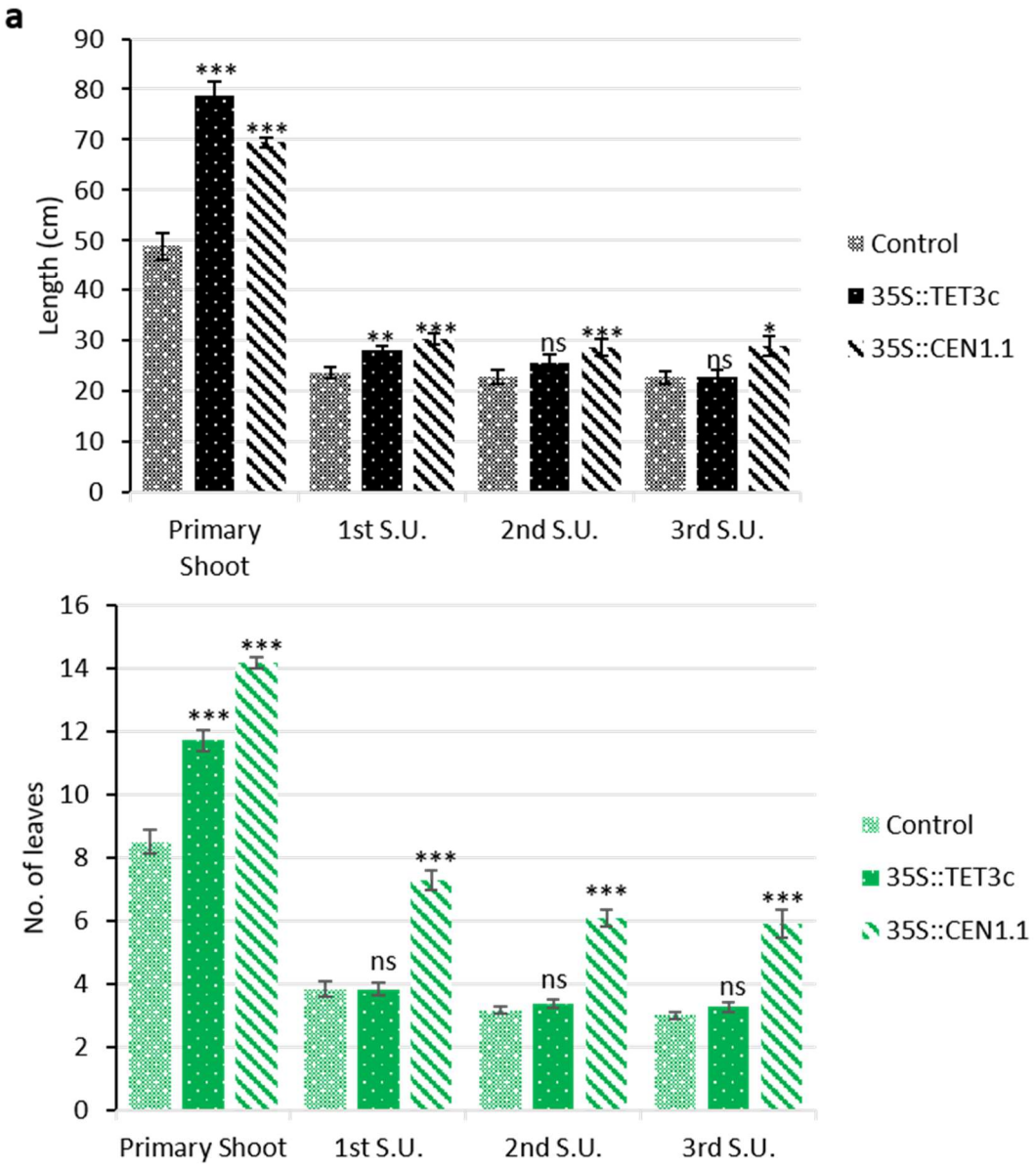
(a) The predicted structure of the CEN1.1 protein, which is 175 amino acids in length and contains one domain, the PBP/PEBP domain.

(b) Map of the plant expression vector pGreen 0029 35S containing the genomic DNA of the full length CEN1.1 gene.

(c) Plants which contain the CEN1.1 transgene were shown to express it (301bp) using semi-quantitative RT-PCR of cDNA from the T0 tomato at 25 cycles. 3 technical replicates of each pool were done. cDNA levels were normalised using the constitutively expressed eukaryotic initiation factor 3E (150bp). –ve indicates a negative H_2O control. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

18 week old plants from two of these lines (Lines I and II) were analysed. *CEN1.1* tomato again showed an increase in the length of and number of leaves in the primary shoot (Figure 3.7a), as the *TET3c* plants had. The increase in the number of leaves persisted across later sympodial units in *CEN1.1* tomato than *TET3c* tomato. This was to be expected given the on average lower expression of *CEN1.1* in *TET3c* plants in comparison to *CEN1.1* transformants (Figure 3.7b). No significant difference was observed between the two transformant lines (Figure 3.7c).

a


Length (cm)
90
80
70
60
50
40
30
20
10
0
***
***
**
***
ns
***
ns
*
Primary Shoot
1st S.U.
2nd S.U.
3rd S.U.
Control
35S::TET3c
35S::CEN1.1
No. of leaves
16
14
12
10
8
6
4
2
0
***
***
***
ns
ns
***
ns
***
Primary Shoot
1st S.U.
2nd S.U.
3rd S.U.
Control
35S::TET3c
35S::CEN1.1

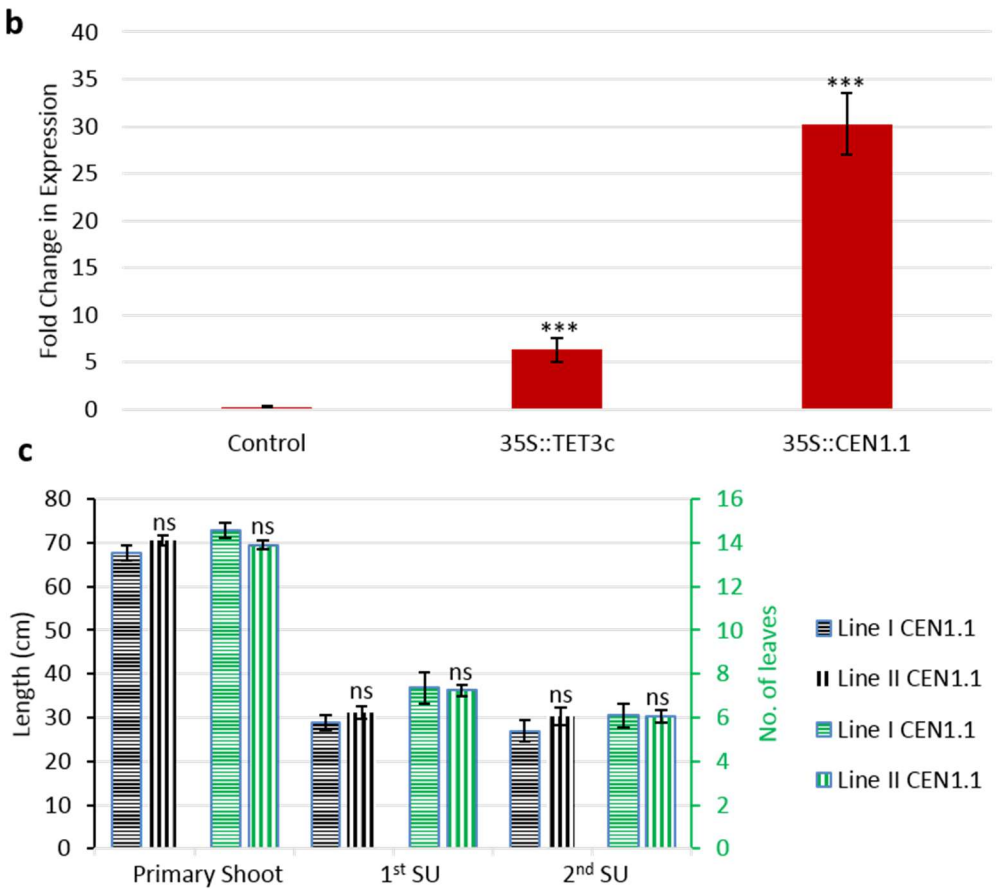


Figure 3.7. CEN1.1 tomato had an increased length of and number of leaves in the primary shoot

(a) The length of shoot (black, above) and number of leaves (green, below) between inflorescences increased in CEN1.1 tomato (n=44) compared to the control (n=24). The increase persisted into later inflorescences than in TET3c tomato (n=33).

(b) Quantitative RT-PCR was used to analyse expression of CEN1.1 in TET3c tomato and CEN1.1 tomato, and expression of CEN1.1 was on average higher in CEN1.1 tomato, correlating with the increase in phenotypes.

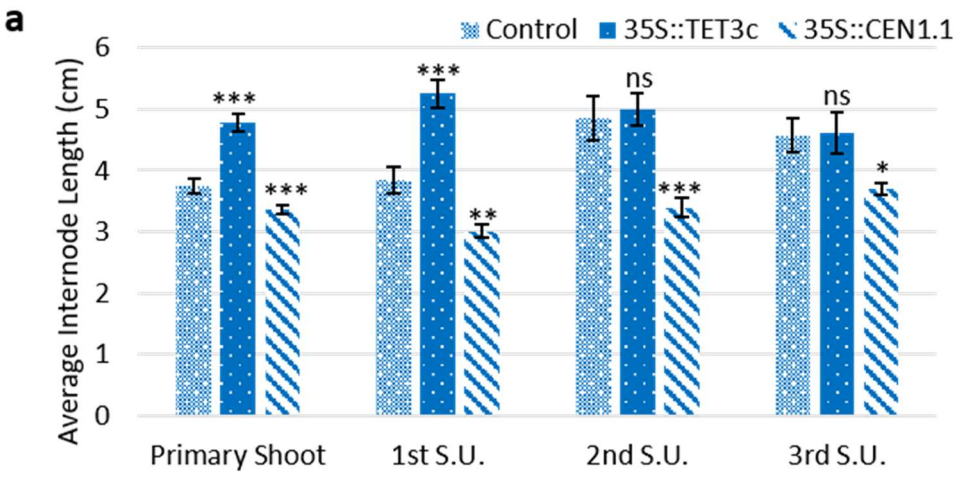
(c) Length of shoot (black) and number of leaves (green) between inflorescences did not change significantly between CEN1.1 Line I (n=17) and Line II (n=27).

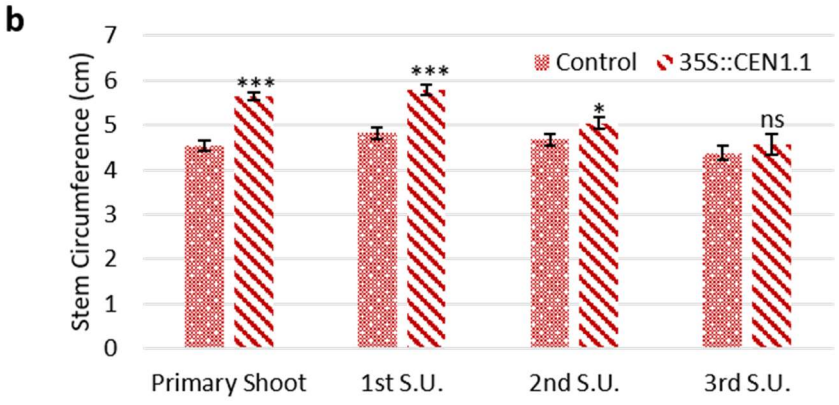
*Plants were measured at 18 weeks old. Graphs show averages with error bars representing standard error. The significance of a change from the control is indicated by asterisks: ns= not significant, *=P<0.05, **=P<0.01, ***=P<0.005, calculated using Student's two-tailed t-test.*

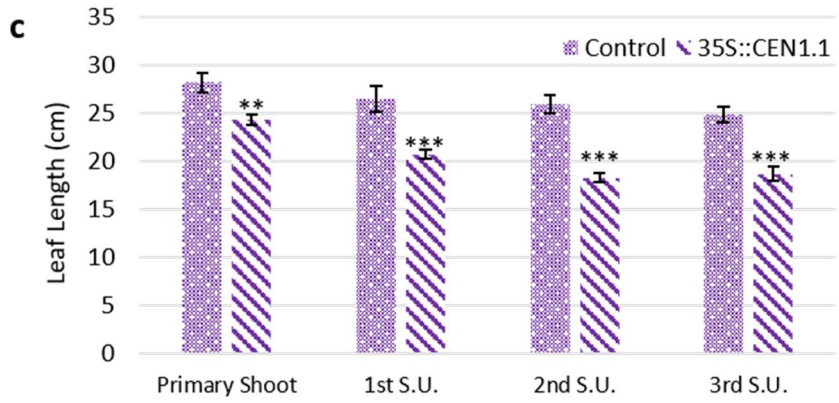
As well as length and number of leaves per unit, plants were also analysed for several other characteristics including:

a) Internode length (distance between leaves)
b) Stem width
c) Average leaf length
d) Number of fruit per inflorescence
e) Ripeness of fruit
f) Weight of fruit

This revealed a number of other phenotypes in the *CEN1.1* plants that had not been observed in the *TET3c* tomato. Internode length was significantly decreased in *CEN1.1* tomato, while it had significantly increased in *TET3c* tomato, indicating the involvement of additional genes in *TET3c* phenotypes (Figure 3.8a). Stem circumference was increased due to the large amount of fasciation seen in stems of *CEN1.1* tomato (Figure 3.8b). Leaf length was reduced in *CEN1.1* plants (Figure 3.8c) and ectopic meristems could be observed on the leaf rachis on 42% of *CEN1.1* leaves (n=196) compared to 0% of control leaves (n=33)(Figure 3.8d). Ectopic meristems also developed elsewhere on the plant, most notably on a small number of *CEN1.1* fruit (4.6%, n=608) where they could continue to grow and produced flowers of their own (Figure 3.8e). Ectopic meristems were not present on any control fruit (n=576).




b
Stem Circumference (cm)
0
1
2
3
4
5
6
7
Control
35S::CEN1.1
***
***
*
ns
Primary Shoot
1st S.U.
2nd S.U.
3rd S.U.


c
Leaf Length (cm)
0
5
10
15
20
25
30
35
Control
35S::CEN1.1
**
***
***
***
Primary Shoot
1st S.U.
2nd S.U.
3rd S.U.


Control
35S::CEN1.1


d
Control
35S::CEN1.1
Control
35S::CEN1.1

Figure 3.8. CEN1.1 tomato had multiple additional phenotypes

(a) The distance between leaves (internode length) decreased in CEN1.1 tomato (n=44) compared to the control (n=24). This contrasts to the increase observed in TET3c tomato (n=33).

(b) Stem circumference was increased in CEN1.1 tomato (n=44). 3 measurements were taken for each sympodial unit in each plant.

(c) Average leaf size was reduced in CEN1.1 plants (n=44) in comparison to the control (n=24), although leaf shape remains normal. Three leaves were measured for each sympodial unit in each plant.

(d) Vegetative meristems could grow from the rachis of leaves in CEN1.1 tomato.

(e) Vegetative meristems could grow from between the sepal and the fruit in CEN1.1 tomato.

*Plants were measured at 18 weeks old. Graphs show averages with error bars representing standard error. The significance of a change from the control is indicated by asterisks: ns= not significant, *=P<0.05, **=P<0.01, ***=P<0.005, calculated using Student's two-tailed t-test.*

3.2.7. Inflorescences of *35S::CEN1.1* tomato were large and leafy

85% of the inflorescences of *CEN1.1* tomato were abnormal in structure, with large quantities of vegetative material (Figure 3.9a). These leafy inflorescences did not commit to a floral identity but instead repeatedly switched between vegetative and inflorescent growth. The inflorescence resembled those which occur when *SP* is overexpressed (Pnueli et al., 1998), in *sft* or *jointless* mutants (Quinet et al., 2006), or *macrocalyx* mutants (Vrebalov et al., 2002). Quantitative RT-PCR was used to analyse the expression of these and other flowering or *CETS* genes in *35S::CEN1.1* plants (Figure 3.9b,c) but no change was observed, suggesting that if *CEN1.1* acts in the same pathway, it must either interact with these genes on a post-transcriptional level or act downstream of these genes. This reflects what is seen for other tomato *CETS* genes such as *SP* and *SFT,* where despite the role of SP in counteracting the effects of SFT, altered expression levels of *SFT* are not seen in *sp* mutants nor vice versa (Shalit et al., 2009).

a

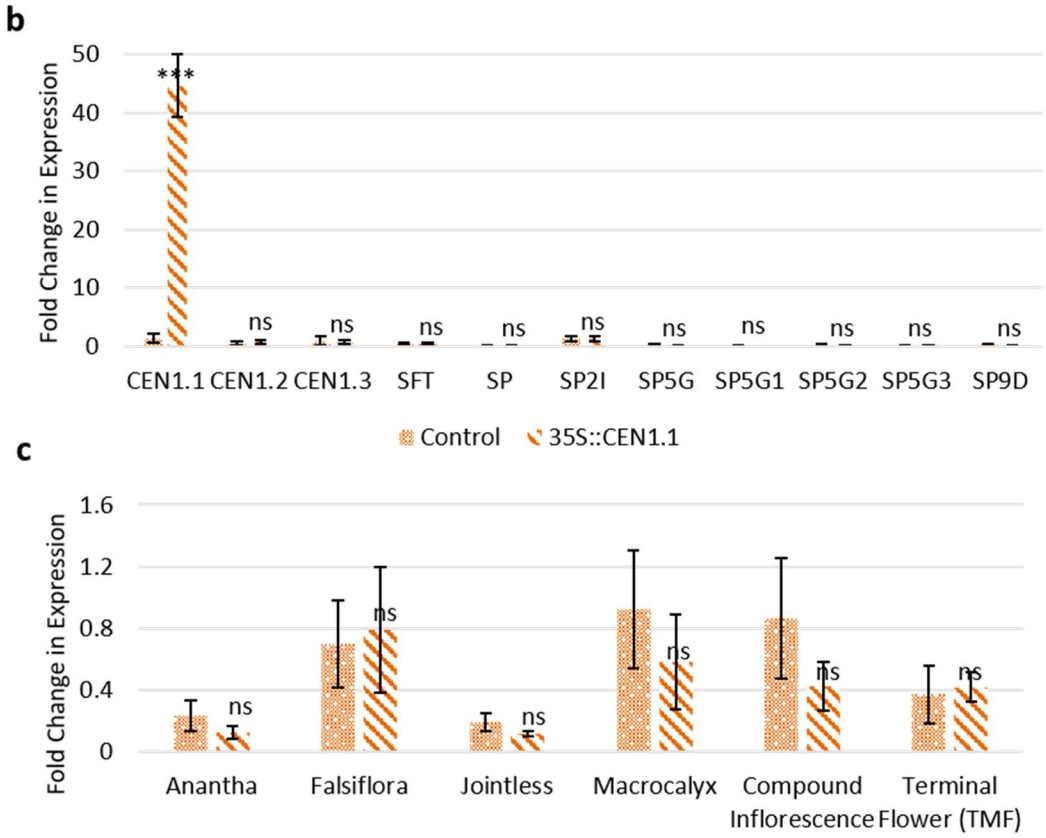


Figure 3.9. CEN1.1 tomato had abnormal branched inflorescences but did not show altered expression of other CETS or flowering genes

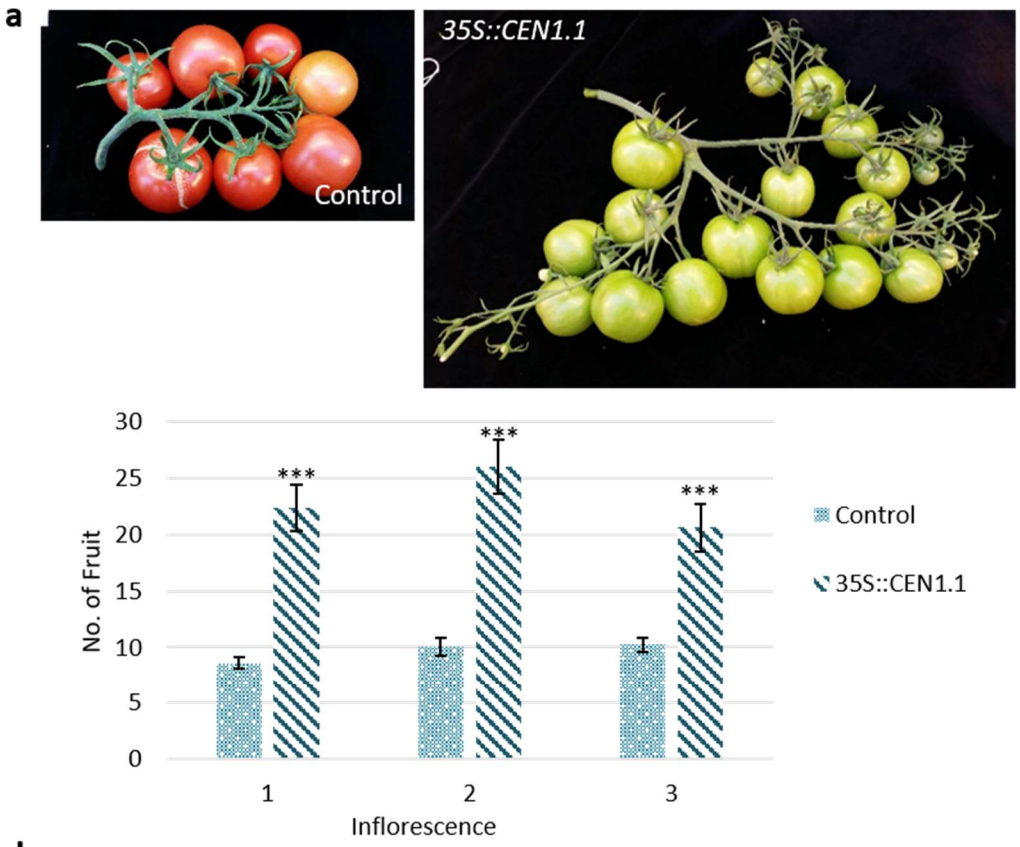
(a) Leafy abnormal inflorescences could develop on 35S::CEN1.1 tomato. Control inflorescences (above) did not contain leaf material while abnormal leafy inflorescences (below) produced large quantities of leaf material.

(b) Quantitative RT-PCR was used to analyse expression of CETS genes in control (n=10) and CEN1.1 tomato (n=28). No significant change was observed in any gene except CEN1.1.

(c) Quantitative RT-PCR was used to analyse expression of flowering genes in control (n=10) and CEN1.1 tomato (n=28). No significant change was observed in any gene.

*cDNA was extracted from leaves at 18 weeks old. Graphs show averages with error bars representing standard error. The significance of a change from the control is indicated by asterisks: ns= not significant, ***=P<0.005, calculated using Student's two-tailed t-test.*

One result of the leafy phenotype was that leafy inflorescences contain more flowers in comparison to a normal inflorescence (Figure 3.10a). *35S::CEN1.1* plants also developed more slowly than the control. Control plants contained more inflorescences and were taller than *CEN1.1* plants of an equivalent age (Figure 3.10b) Despite this, due to the high percentage of leafy inflorescences, *CEN1.1* plants on average produced more flowers in comparison to the control tomato plants (64 on the average *CEN1.1* plant, n=44, compared to 49 on the average control plant, n=23). However, the tomato fruit on the *CEN1.1* plant were at an earlier stage of ripening, presumably due to the slower development of *CEN1.1* plants (Figure 3.10c). On the first inflorescence of a control plant, the median fruit was fully ripe and ready to be removed from the plant (45 days post anthesis (DPA)), while the median fruit on the first inflorescence of a *CEN1.1* plant was still at the early stage of fruit growth, approximately 12 days post anthesis (DPA). Days post anthesis indicates the likely length of time since the first opening of the flower (anthesis). *CEN1.1* fruit ripened at the same rate as control tomato fruit, and these fruit were fully ripe within four weeks. Other than this, *CEN1.1* fruit were phenotypically normal in appearance and weight (Figure 3.10d).


a
Control
35S::CEN1.1
No. of Fruit
30
25
20
15
10
5
0
***
***
***
1
2
3
Inflorescence
Control
35S::CEN1.1
b
No. of inflorescences
6
5
4
3
2
1
0
***
Control
35S::CEN1.1
Height (m)
2.5
2
1.5
1
0.5
0
***
Control
35S::CEN1.1
c
% of Fruits at Each Stage
80%
70%
60%
50%
40%
30%
20%
10%
0%
Control
35S::CEN1.1
45 DPA
42 DPA
35 DPA
20 DPA
12 DPA
8 DPA
Flower
Bud

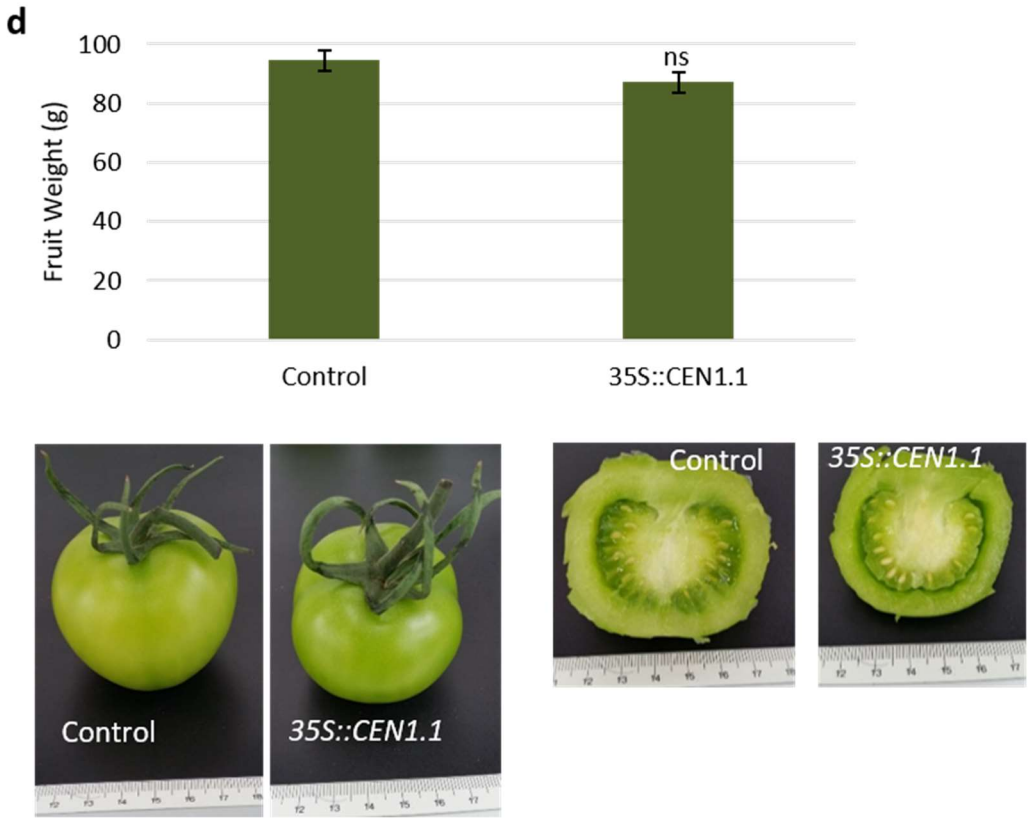


Figure 3.10. CEN1.1 plants produced an increased quantity of fruit but were developmentally delayed

(a) *Inflorescences on CEN1.1 plants (n=44) produced an increased number of fruit compared to the control (n=24).*

(b) *CEN1.1 tomato (n=44) contained fewer inflorescences and were shorter at 18 weeks of age than control tomato (n=24).*

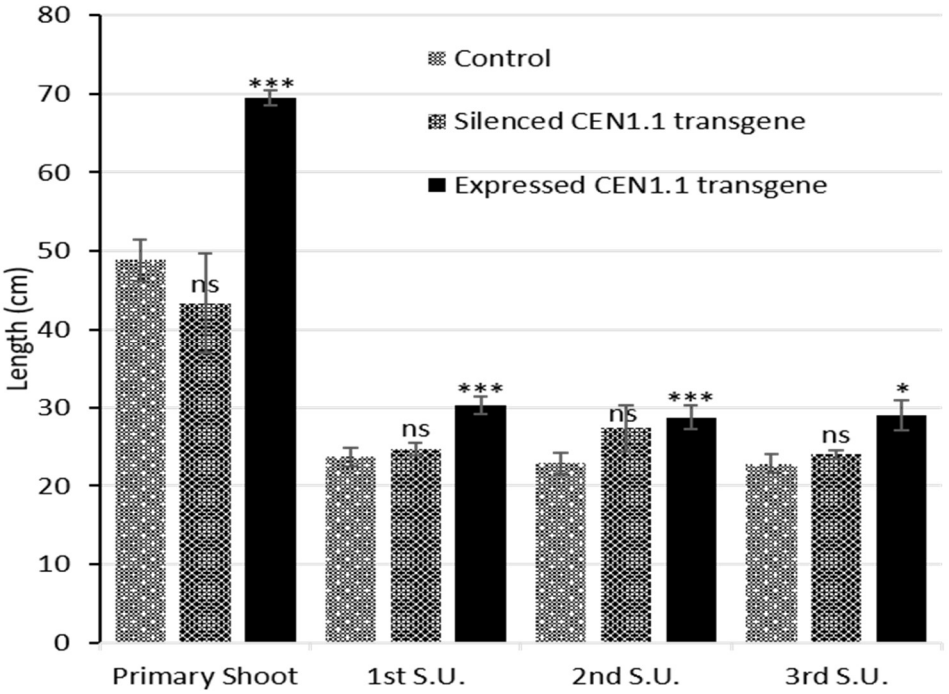
(c) *Fruits on the first inflorescence of CEN1.1 (n=831) and control tomato (n=187) were grouped according to ripeness. Fruits of the control tomato were on average more ripe than CEN1.1 fruits. Fruits were classified according to the DPA stage they resembled (days post anthesis, the first opening of the flower).*

(d) *Fruits of CEN1.1 tomato (n=20) were not significantly different in weight and did not appear obviously different to control tomato fruits (n=10).*

*Plants were measured at 18 weeks old. Graphs show averages with error bars representing standard error. The significance of a change from the control is indicated by asterisks: ns= not significant, ***=P<0.005, calculated using Student's two-tailed t-test.*

3.2.8. *CEN1.1* expression correlated with the observed phenotypes of increased vegetative growth

Transgenes which have successfully transformed into plants can subsequently become silenced (Meyer and Heidmann, 1994). Three of the *35S::CEN1.1* tomato plants had a silenced copy of the *CEN1.1* transgene, and did not show ectopic expression of *CEN1.1*. These plants were also phenotypically normal, showing no significant change in the length or number of leaves between inflorescences (Figure 3.11). This demonstrates that the increased vegetative growth observed in *35S::TET3c* and *35S::CEN1.1* tomato was a result of increased *CEN1.1* expression.



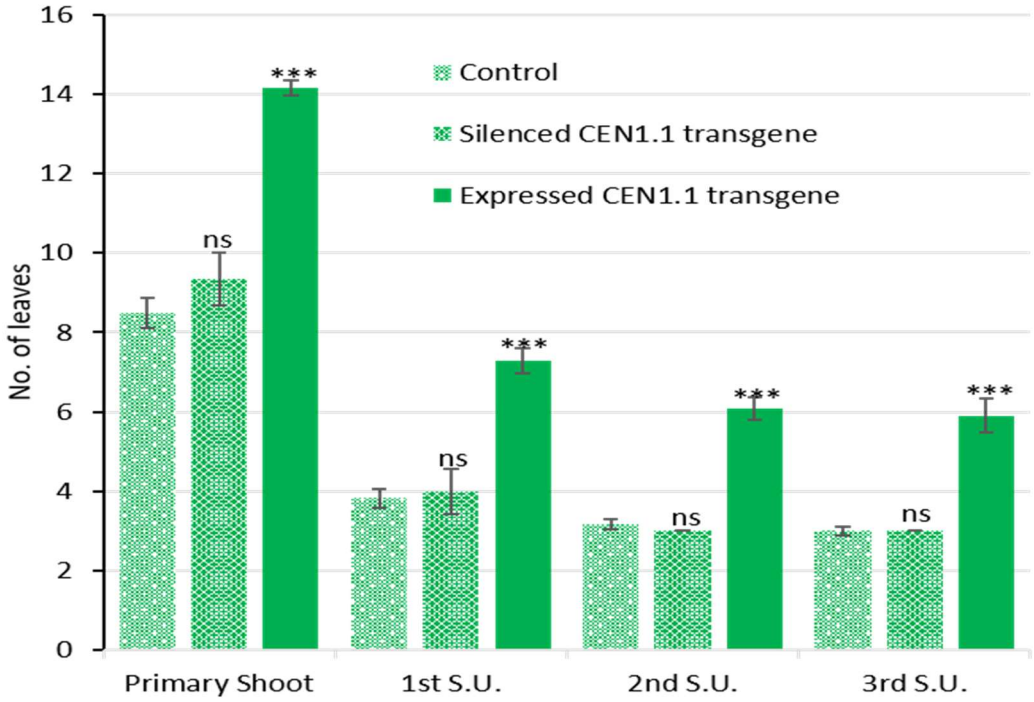


Figure 3.11. 35S::CEN1.1 tomato did not show abnormal phenotypes when CEN1.1 was not expressed

*The length of shoot (black, above) and number of leaves (green, below) between inflorescences was not increased compared to the control (n=24) in tomato with a silenced copy of the CEN1.1 transgene (n=3), demonstrating that phenotypes are caused by ectopic expression of CEN1.1. Plants were measured at 18 weeks old. Graphs show averages with error bars representing standard error. The significance of a change from the control is indicated by asterisks: ns= not significant, *=P<0.05, **=P<0.01, ***=P<0.005, calculated using Student's two-tailed t-test.*

3.2.9. Constitutive expression of *CEN1.1* in Arabidopsis also resulted in increased vegetative growth and delayed flowering

In order to investigate if *CEN1.1* would also have the same effect in a different species, the *35S::CEN1.1* construct was transformed into Arabidopsis. Successful transformants were identified and grown under long day conditions. When able to flower, *35S::CEN1.1* Arabidopsis were delayed in the floral transition in comparison

to wild type, bolting after 63 days and 58 leaves (Figure 3.12a,b). 40% of *35S::CEN1.1* Arabidopsis (n=15) did not bolt at all, reaching an age of 12 weeks before dying without having bolted. *CEN1.1* Arabidopsis plants which did bolt had fewer, thicker shoots and produced abnormal flowering structures without petals, stigma or stamen (Figure 3.12c). These were unable to produce seed and resembled structures which can be produced when the Arabidopsis *CETS* genes *TFL1* or *BFT* are overexpressed (Yoo et al., 2010).

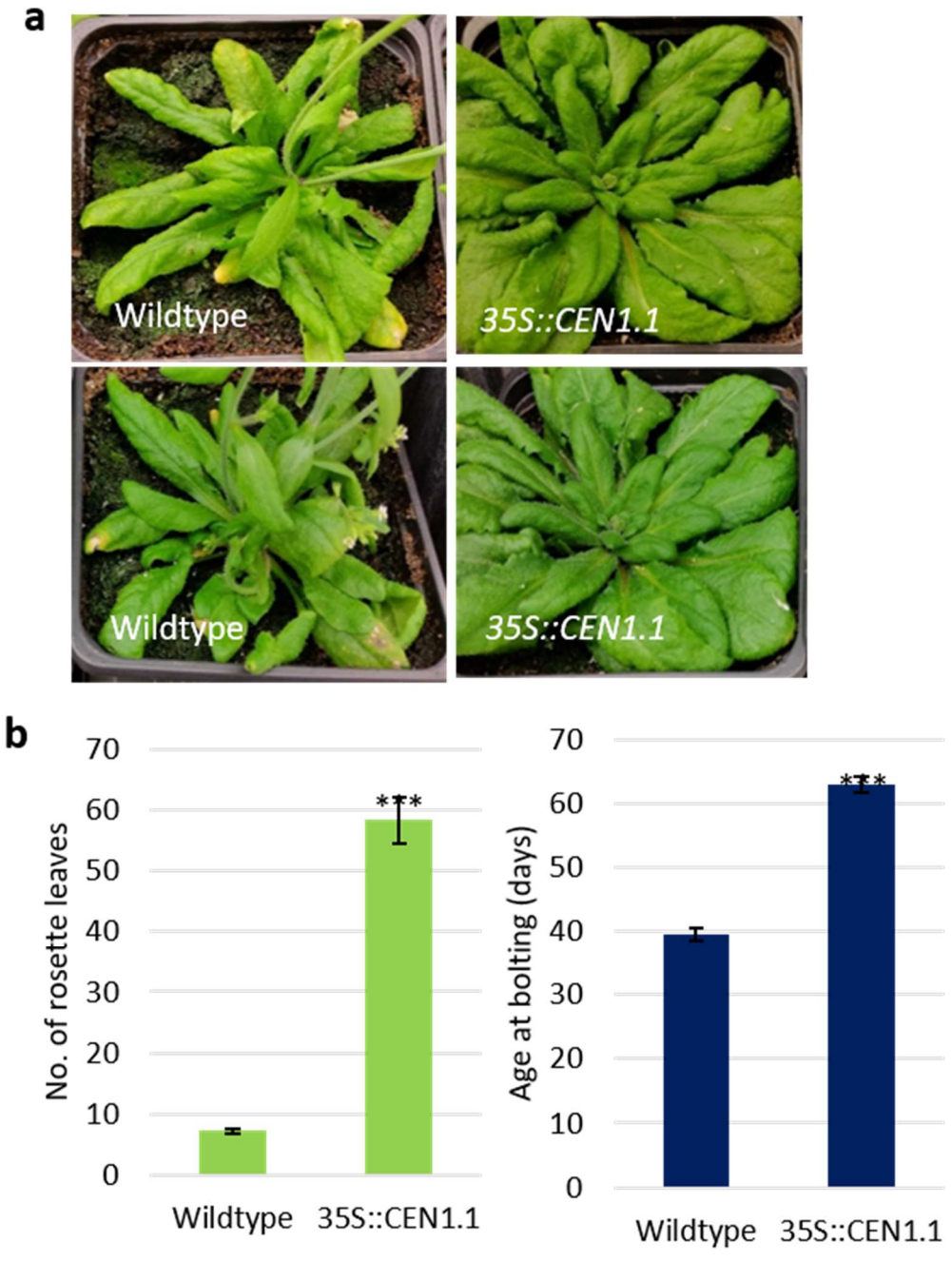


Figure 3.12. 35S::CEN1.1 Arabidopsis also had an increased amount of vegetative growth and delayed floral transition

(a) Nine week old 35S::CEN1.1 Arabidopsis had not bolted and produced an increased number of rosette leaves compared to wild type.

(b) Age and number of rosette leaves at bolting was counted for wild type (Col-0)(n=12) and 35S::CEN1.1 Arabidopsis (n=9). 35S::CEN1.1 Arabidopsis were delayed in the floral transition compared to the wild type.

(c) Flowering structures on 35S::CEN1.1 Arabidopsis were abnormal.

*Graphs show averages with error bars representing standard error. The significance of a change from the wild type is indicated by asterisks: ***=P<0.005, calculated using Student's two-tailed t-test.*

3.2.10. Knockdown of the CEN1.1 tomato gene using an inverted repeat cassette

The function of *CEN1.1* in tomato is unknown. To analyse this, a 313bp region of exon 4 of the *CEN1.1* gene was amplified and cloned into a vector containing two

cloning sites separated by an intron to generate an inverted repeat cassette. This was cloned behind the constitutive 35S promoter to create a *35S::CEN1.1ir* construct (Figure 3.13) which would cause RNAi (RNA interference) based silencing of the *CEN1.1* gene (Watanabe, 2011). Tomato transformations of this construct were carried out. 1129 explants were cultured with the *35S::CEN1.1ir* construct but no successful transformants were regenerated. In comparison, six transformant lines were produced using 1065 explants when cultured with the *35S::CEN1.1* construct. This indicates that *CEN1.1ir* may interfere with the regeneration of tomato transformants. An inducible version of *CEN1.1ir* would be required to confirm this.

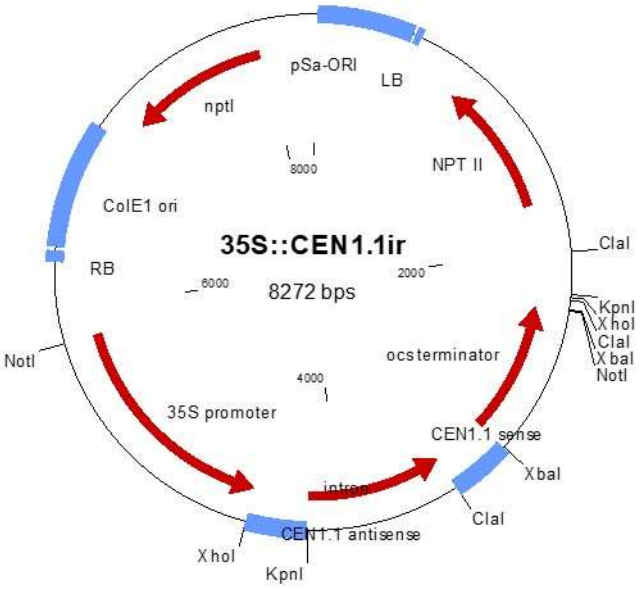


Figure 3.13. A construct containing an inverted repeat of part of the CEN1.1 gene was made

A map of the plant expression vector pGreen 0029 35S with an inverted repeat cassette containing 313bp region of the CEN1.1 gene. When expressed, the inverted repeat will anneal to form a stem loop which will then be processed by the plant's endogenous silencing systems into small RNAs (sRNAs). These result in the destruction of complementary mRNA sequences and therefore knockdown of CEN1.1 expression would occur.

3.2.11.*TET3c* tomato also possessed phenotypes which were not a result of ectopic *CEN1.1* expression

While constitutive expression of *CEN1.1* reconstituted several of the phenotypes observed in *TET3c* tomato, some phenotypes were not observed. These included the defects in shoot meristem maintenance (Figure 3.2b), the determinate growth (terminal flower phenotype) of tomato plants (Figure 3.2c) and an increase in internode length (the distance between leaves)(Figure 3.3c). The defects in shoot meristem maintenance will be discussed in more detail in Chapter Four. Determinate growth (Shalit et al., 2009) and internode length changes (Lifschitz et al., 2006) can also be a result of expression changes of the *CETS* genes and therefore the ectopic expression of *SP9D* observed in *TET3c* lines (Figure 3.4a) is a possible candidate to explain these phenotypes. As discussed above, methylation levels upstream of the transcriptional start site are not as high as they are in the *CEN1.1* gene (13% compared with 37%) but there are still regions with levels of CHH methylation that exceed 50% (Zhong et al., 2013). In addition, the second intron of *SP9D* contains dense methylation and may be a target of *TET3c.* A *35S::SP9D* construct was therefore made containing the cDNA of *SP9D* behind the 35S promoter (Figure 9.5a) and transformed into tomato. An additional construct containing the genomic DNA of *SP9D* behind the 35S promoter (Figure 9.5b) was made by Abirah Zulkifli and also transformed into tomato. Two lines containing *35S::SP9Dgenomic* and one line containing *35S::SP9DcDNA* were generated, all of which showed strong expression of the *SP9D* gene (Figure 9.5c). Phenotypic analysis of these lines was not possible within the time frame of the project.

3.3.Discussion

The use of *TET3c* in tomato has resulted in the characterisation of the *CEN1.1* gene as a member of the tomato *CETS* gene family whose expression correlates with DNA methylation levels. Ectopic expression of *CEN1.1* affected the floral transition in

tomato, a process controlled by many genes, including *JOINTLESS*, *SFT* and *SP*. *jointless* and *sft* mutants are similar to each other in several respects, including a delay in flowering and the presence of leafy inflorescences (Quinet et al., 2006). While *JOINTLESS* and *SFT* act to maintain the floral meristem identity and prevent a reversion to vegetative meristem activity (Quinet et al., 2006), *CEN1.1* had the opposite role, acting to stimulate vegetative meristem identity. As a result, inflorescences overexpressing *CEN1.1* were phenotypically similar to *jointless* and *sft* mutants, including the increase in the number of flowers. A similar, although less severe phenotype is also seen when *SP* is overexpressed (Pnueli et al., 1998) suggesting that *SP* has a role more similar to *CEN1.1* and antagonistic to *SFT* and *JOINTLESS*. Interactions between these genes are likely to be complex. *sp is* epistatic to *jointless* for the leafy inflorescence phenotype (Quinet et al., 2011; Szymkowiak and Irish, 2006) while *jointless* is epistatic to *sp* for the phenotype of delayed flowering (Quinet et al., 2011), as is *sft* (Molinero-Rosales et al., 2003). Expression levels of *SP*, *SFT* and *JOINTLESS* did not change in *35S::CEN1.1* plants (Figure 3.9), suggesting that they may have the same downstream result rather than affecting each other directly, or that if they do interact it is on a post-transcriptional level. In order to better understand the interaction of *CEN1.1* with *SFT*, *SP* and *JOINTLESS*, tomato plants containing the *35S::CEN1.1* construct could be crossed with *sft, sp* and *jointless* mutants.

Unlike *JOINTLESS,* which is a member of the MADS-box gene family (Mao et al., 2000), *CEN1.1* is not a transcription factor but instead contains a phosphatidylethanolamine binding domain, like *SFT, SP* and *CEN* (another founding member of the *CETS* gene family, which controls inflorescence architecture in *Antirrhinum* (Bradley et al., 1996)). Phosphatidylethanolamine-binding proteins (PEBPs) exist in many different organisms including bacteria, yeast, nematodes, insects, mammals, and plants, where they have been shown to play a role in diverse functions including serine protease activity (Hengst et al., 2001) and several signalling pathways such as the MAP kinase pathway in mammals (Corbit et al., 2003). This PEBP gene family in plants controls flowering in a large variety of species,

and can be divided into different branches based on their structure which largely also correlate with their activity either as floral repressors (*TFL1*-like, such as *SP* and *CEN1.1*) or floral activators (*FT*-like, such as *SFT*)(Cao et al., 2016; Chardon and Damerval, 2005). It has been theorised that PEBP proteins such as CEN and SP may also act through inhibition of kinase signalling pathway like their mammalian relatives (Banfield and Brady, 2000) but the target of their action is currently unknown.

Expression of *CEN1.1* correlated with upstream hypomethylation in wild type tissues, which has not been observed for any other tomato *CETS* genes. *SFT* and *SP* both have lower levels of methylation upstream in comparison to *CEN1.1* (12% and 28% respectively in comparison to 37%)(Zhong et al., 2013). *SP* expression did not change in *TET3c* tomato leaves, but increased expression of *SFT* was seen in one tomato line (Figure 3.4). This does not exclude the possibility of *SP* expression being controlled by methylation, as the *TET3c* construct does not carry out genome-wide demethylation. Screening the tomato epigenome database (Zhong et al., 2013) for methylation levels of *SP, SFT* and *CEN1.1* in ripe fruit reveals that methylation upstream of *SP* and *SFT* is similar in leaves and fruit, as is the overall genome methylation level (Zhong et al., 2013), despite the fact that *SP* and *SFT* are expected to be expressed in leaves and silenced in fruit (Carmel-Goren et al., 2003). Conversely, methylation levels upstream of *CEN1.1* increase in fruit tissues, with overall levels of methylation increasing from 37% to 50%, largely driven by an increase in CHH methylation from 31% to 47% (Zhong et al., 2013). This corresponds with the role proposed here for *CEN1.1* in stimulating vegetative meristematic growth, repression of which in the fruit is vital, and suggests that *SP* and *SFT* may not be regulated by DNA methylation. Further work could be done to investigate this, for example by carrying out targeted demethylation of *SP* or *SFT* upstream regions using a *TET3c* domain fused with a targeting domain.

CEN1.1 had the same effect when ectopically expressed in a heterologous system, as do other members of the *CETS* gene family, (Amaya et al., 1999; Mimida et al.,

2001). Ectopic expression of *CEN1.1* in Arabidopsis resulted in increased vegetative growth and delayed flowering. *35S::CEN1.1* flowered after 58 leaves on average. This phenotype was more severe than has previously been described when Arabidopsis *CETS* genes are constitutively expressed behind the 35S promoter in Arabidopsis (12.9 leaves for *35S::TFL1*, 12.6 leaves for *35S::ATC*, 20.9 leaves for *35S::BFT*)(Mimida et al., 2001; Yoo et al., 2010) or when the Antirrhinum gene *CEN* is overexpressed (13.6 leaves)(Mimida et al., 2001). Phenotypes observed in *35S::CEN1.1* Arabidopsis otherwise resembled phenotypes seen when TFL1-like genes are constitutively expressed (Yoo et al., 2010).

In addition to delaying the floral transition, the increase in vegetative growth seen in *35S::CEN1.1* tomato resulted in ectopic vegetative meristems developing from differentiated tissues such as leaves and fruit. These could continue to grow and produce flowers of their own. *CETS* genes are known to have a role in meristem maintenance during the floral transition. Overexpression of *SP* results in larger, more domed meristems (Tal et al., 2017) while overexpression of *SFT* in tomato leaf primordia results in arrest of the primary shoot apical meristem (Lifschitz et al., 2006). Part of the function of *SP* is to prevent termination of the meristem as a result of florigen signals which would terminate it (Tal et al., 2017) and *CEN1.1* may play a similar role, promoting meristematic growth and resulting in the presence of ectopic meristems in *35S::CEN1.1* tomato. Tomato containing the *35S::CEN1.1ir* construct could not be regenerated, but an inducible *CEN1.1ir* construct or a CRISPR-Cas9 construct targeted to the *CEN1.1* gene could be used to analyse the function of *CEN1.1* in wild type tomato.

An increased fruit yield of 30% was seen in *35S::CEN1.1* tomato, but there was also an 21% increase in the time needed to achieve ripe fruit and therefore whether this would be of use is debatable. Despite this, my results emphasise the importance of fully understanding the *CETS* gene family in tomato. Tomato *sp* mutants make up the majority of varieties used for mechanical harvesting (Jiang et al., 2013) and double *sp sp5g* mutants have been generated which also can improve harvest index

(total yield per plant weight) by decreasing the time needed to achieve ripe fruit (Soyk et al., 2017). Multiple patents exist for introgression lines altering the expression levels of *SFT*. In other species such as potato, *CETS* genes are involved in equally important transitions such as the transition to tuber formation (Navarro et al., 2011), and the role of DNA methylation in *CETS* genes of these species has yet to be analysed.

I have shown in this chapter that transformation of *TET3c* into tomato resulted in phenotypic changes including a delay in the floral transition. *TET3c* transformation also resulted in the hypomethylation and ectopic expression of a previously uncharacterised *CETS* gene, allowing us to identify and characterise this gene as a promoter of vegetative growth whose expression in different tissues correlates with DNA methylation levels upstream of the transcriptional start site. Additional phenotypes were observed in the *TET3c* tomato including a significant increase in internode length. This may be a result of ectopic expression of *SP9D* in *TET3c* lines or expression of another, unknown gene. Analysis of tomato lines constitutively expressing *SP9D* will allow investigation of this. *TET3c* tomato also had abnormalities with the shoot apical meristem which will be discussed in Chapter Four.

4. Transformation Can Increase the Incidence of the Blind Phenotype in Tomato

4.1. Introduction

The development of plants is highly plastic, allowing them to respond and acclimate to their surroundings. Growth of the plant continues throughout its lifecycle, which in the case of trees can be hundreds or thousands of years, and the majority of post-embryonic tissues in a plant are generated from either shoot or root meristems (Clark et al., 1995). The shoot apical meristem (SAM) is located at the centre of the growing tip and contains a pool of stem cells which are used for development of new tissues (Clark et al., 1995). Maintenance of the SAM is therefore extremely important to sustain continued growth and development, and a robust regulatory network is required to balance the differentiation of cells with the proliferation of stem cells and preserve the SAM at a constant size.

Many genes have been implicated in maintenance of the SAM, including a family of homeobox transcription factors called the *KNOX* genes (Byrne et al., 2002), the cytokinin signalling pathway (Werner et al., 2003) and the auxin response pathway (Zhao et al., 2010) but the central regulatory pathway that maintains stem cell size is the CLAVATA3 (CLV3)-WUSCHEL (WUS) feedback loop (Somssich et al., 2016). The CLV3-WUS pathway is best understood in Arabidopsis, but is also conserved in other plant species including rice (Chu et al., 2006) and tomato (Xu et al., 2015). Expression of *WUS*, a homeobox transcription factor, occurs in the organising centre (OC) of the SAM. This small group of cells is situated just below the central zone (CZ) of the SAM, which contains the pool of slowly dividing stem cells which maintain the SAM. Surrounding the CZ are the peripheral zones (PZ), in which cells divide rapidly and give rise to new organ primordia and below it is the rib zone (RZ) from which the stem tissue originates (Mayer et al., 1998).

The OC is defined by its expression of *WUS* (Mayer et al., 1998) which promotes stem cell identity and results in transcription of the *CLV3* gene (Schoof et al., 2000).

The CLV3 peptide is processed and secreted from the OC into the CZ (Kondo et al., 2006; Rojo et al., 2002), where it is detected by receptor-like kinases including homodimers of CLAVATA1 (CLV1)(Ogawa et al., 2008) or RECEPTOR-LIKE PROTEIN KINASE 2 (RPK2)(Kinoshita et al., 2010), or heterodimers of CLAVATA2 (CLV2), a receptor-like protein, and CORYNE (CRN), a receptor-like kinase (Müller et al., 2008). Perception of the CLV3 ligand activates a signalling pathway involving kinases and phosphatases (Betsuyaku et al., 2011; Song et al., 2006; Williams et al., 1997), resulting in repression of *WUS* transcription (Brand et al., 2000; Schoof et al., 2000), restricting it to the OC. This negative feedback loop ensures a balance between expression levels of *WUS*, which promotes cell proliferation and stem cell identity, and *CLV* genes, which reduce *WUS* transcription, and therefore a constant SAM size results (Brand et al., 2000; Schoof et al., 2000). *wus* mutants are unable to maintain shoot meristems (Mayer et al., 1998) while mutants of *CLV3*, *CLV1* and *CLV2* have enlarged meristems (Clark et al., 1993, 1995; Kayes and Clark, 1998) and an expanded zone of *WUS* activity (Schoof et al., 2000).

The CLV3-WUS system is robust, and has been shown to be able to buffer variations in *CLV3* levels over a 10-fold range and adapt to increased *CLV3* levels over time (Müller et al., 2006). Despite this, defects in meristem development have been observed in a wide variety of species (Hicklenton et al., 1993; de Jonge et al., 2016; Salter, 1957; Wetzstein and Vavrina, 2002). Spontaneous arrest of meristem growth has been called by many different names including blindness, apical meristem decline (AMD), toplessness and SAM arrest, among others (de Jonge, 2013; Wetzstein and Vavrina, 2002). Tomato seedlings with arrested meristems occur naturally at varying levels in different varieties and batches of seeds (Wetzstein and Vavrina, 2002). This is commonly called blindness by plant breeders, and can result in production losses as, even if the SAM arrest is temporary, growth of blind tomatoes will be delayed. Some varieties are particularly prone to blindness and the frequency of the blind phenotype can also be affected by environmental conditions such as nutrient supply (Wetzstein and Vavrina, 2002), high light intensity, and

growth temperature (de Jonge, 2013), suggesting that this may be an epigenetic phenomenon occurring as a result of stress.

As discussed in Chapter Three, the *35S::TET3c* construct was transformed into tomato plants to investigate its effects on genes sensitive to DNA methylation changes. *TET3c* tomato were abnormal in multiple respects including the shoot architecture alterations discussed in Chapter Three. *TET3c* tomato also showed an increased percentage of plants with the blind phenotype. Using blind *TET3c* tomato and wild type tomato, the cause of the blind phenotype in tomato was investigated.

4.2. Results

4.2.1. *TET3c* tomatoes had a high frequency of the blind phenotype

Four independent *TET3c* transformant lines were generated (Lines A-D) and the progeny were grown on kanamycin to select for the presence of the *TET3c* transgene. T1 tomato plants containing the *TET3c* construct were examined at five weeks old, and abnormalities with the shoot apical meristem were observed (Figure 4.1a). The shoot apical meristem of many *TET3c* tomato seedlings was not maintained but instead differentiated into leaf tissue or became callus (Figure 4.1b) and this phenotype was termed the “blind” phenotype. New meristems could develop to continue the growth of the plant, otherwise the growth of the plant was arrested. Meristematic growth could occur from the axil of leaves which have developed, centrally between the cotyledons or even out of otherwise differentiated rachis tissue (Figure 4.1c). Some plants were extremely bushy, possessing multiple leaves with very short internodes. These were classified as "compressed" (Figure 4.1d). The compressed phenotype observed at five weeks old only occurred in plants that had previously been classified as blind at three weeks old, indicating that this was a later manifestation of the blind phenotype. Compressed plants were therefore classified as blind in future experiments.

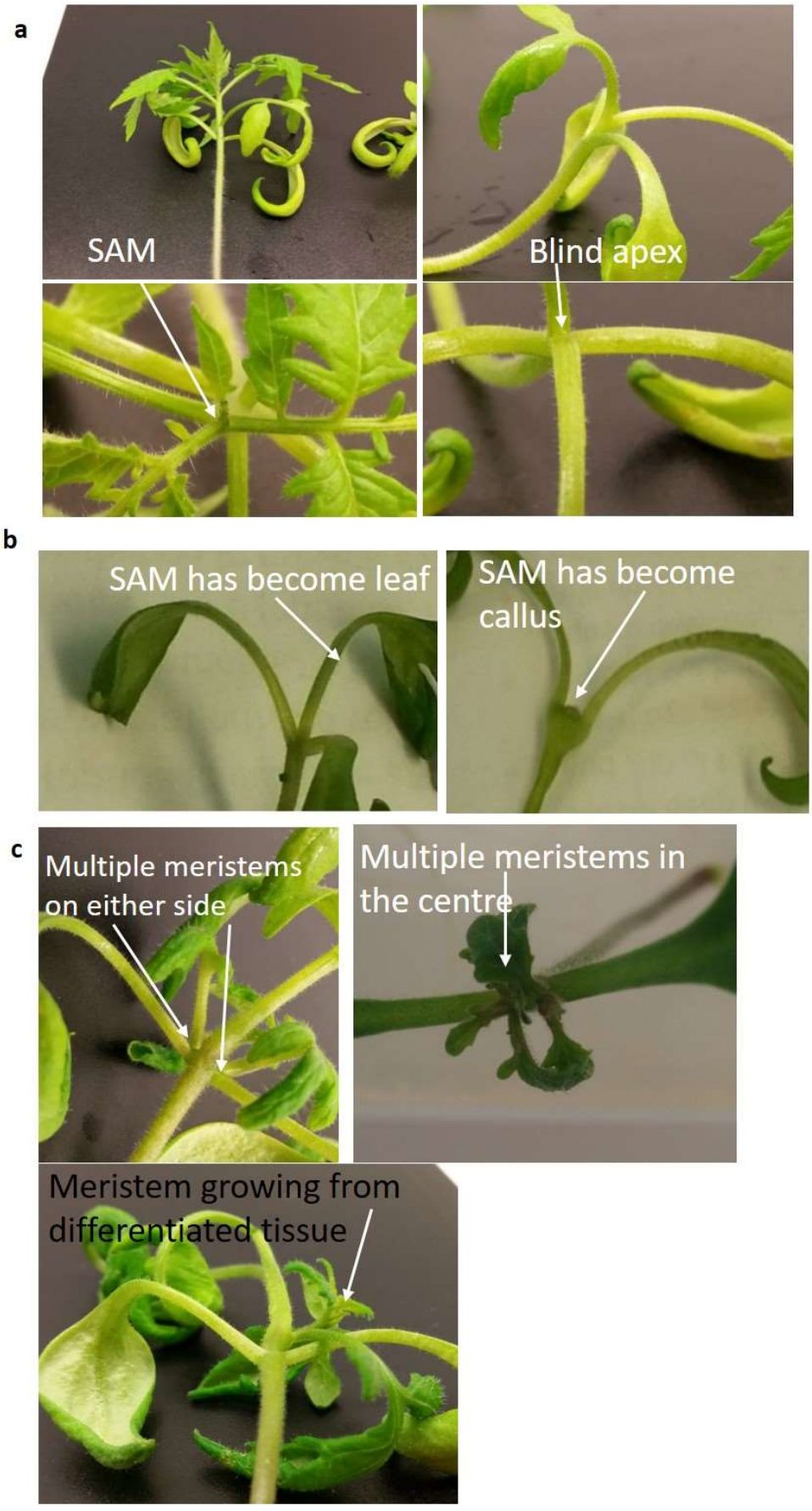

a
SAM
Blind apex
b
SAM has become leaf
SAM has become callus
c
Multiple meristems on either side
Multiple meristems in the centre
Meristem growing from differentiated tissue

Figure 4.1. TET3c tomato often displayed blindness, or SAM arrest

(a) Blindness is a loss of the shoot apical meristem (SAM). The shoot apical meristem was present in control seedlings (left) and absent in blind TET3c seedlings (right).

(b) In blind TET3c tomato, the SAM could differentiate into leaf or become callus.

(c) Meristems could subsequently develop on either side of the leaf, in the centre of the cotyledons or out of differentiated leaf tissue.

(d) "Compressed" TET3c plants produced many leaves, growing close together. The compressed phenotype developed from blind plants.

The cellular structure of blind meristems was investigated using tissue sections from tomato seedlings. Sections were mounted, stained with toluidine blue and examined using a light microscope (Figure 4.2). It could be seen from these sections that, in blind *TET3c* plants, the organisation of cells within the meristem had been disrupted (Figure 4.2c). The shoot apex was enlarged in comparison to the wild type but the precise cell layers and zones which normally exist within the shoot apical meristem were absent, including both the dense clusters of cells in the central and peripheral zones and the cells of the rib zone (RZ), usually identifiable by the parallel columns of cells (de Jonge, 2013). Blind *TET3c* plants were also delayed in comparison to the control. Despite being five weeks old, the apex of the *TET3c* seedling was still between the cotyledons, while control seedlings had produced 2-3 true leaves. As discussed above, blind *TET3c* plants could recover and develop new meristems which behaved and appeared normal, other than their unusual placement (Figure 4.2d).

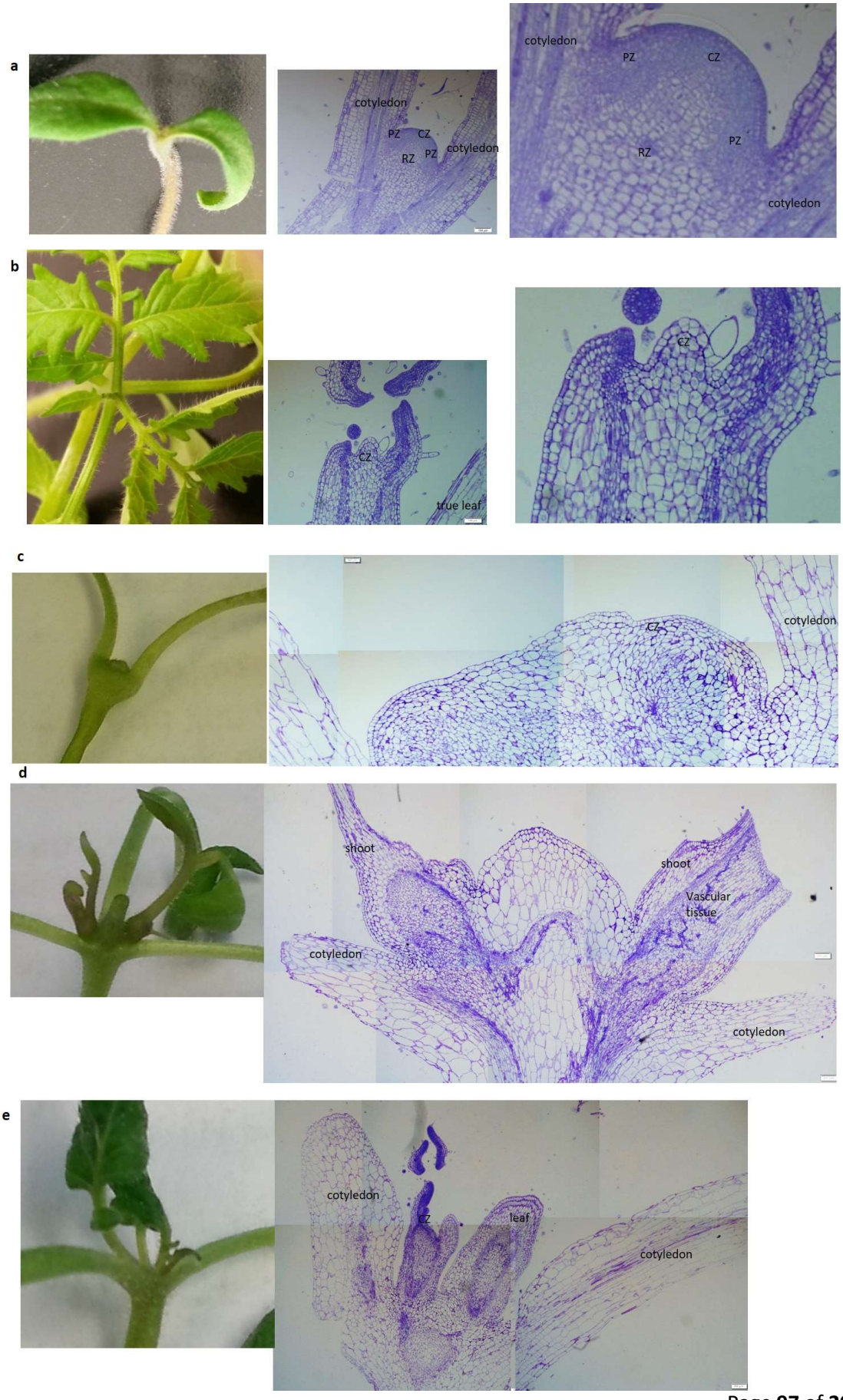

a
cotyledon
PZ
CZ
RZ
PZ
cotyledon
cotyledon
PZ
CZ
RZ
PZ
cotyledon
b
CZ
true leaf
CZ
c
CZ
cotyledon
d
shoot
shoot
Vascular tissue
cotyledon
cotyledon
e
cotyledon
CZ
leaf
cotyledon

Figure 4.2. SAMs of blind TET3c seedlings were disorganised and enlarged.

(a) SAM of a 2 week old control seedling. Labels indicate the different zones of the meristem. PZ = peripheral zones, CZ = central zone, both identifiable by the dense clusters of cells and separable by their location within the meristem, RZ = rib zone, identifiable by the parallel columns of cells. The image to the left is to scale with the others (scale bar in bottom right), while the image to the right is enlarged for clarity.

(b) SAM of a 5 week old control seedling. Distinct zones are more difficult to identify but the central zone can still be recognized by the reduction in cell size at the top of the meristem. The image to the left is to scale with the others (scale bar in bottom right), while the image to the right is enlarged for clarity.

(c) SAM of a 5 week old blind TET3c seedling. The meristem was disorganized and its separate zones were not identifiable. Cell size did not decrease towards the surface of the meristem. The image is to scale with the others (scale bar in the top left), but the meristem was much larger.

(d) SAM of a 5 week old blind TET3c seedling which had simultaneously developed 2 axillary meristems. These have grown out into shoots and therefore no longer have meristematic structure but instead the vascular structure can begin to be observed.

(e) SAM of a 5 week old blind TET3c seedling which had developed 2 leaves and a meristem in the centre of the cotyledons. Within the meristem that can be observed, the central zone can be identified by the dense clusters of cells, as in (b).

All plants are to scale (scale bar reads 100µm).

In order to investigate the frequency and heritability of the blind phenotype, T2 seeds of the transgenic T1 tomato plants were grown off selection and genotyped for the presence of the transgene. These plants were phenotyped at five weeks old (Figure 4.3b). The phenotype was followed through a further generation to investigate if the high levels of the blind phenotype remained stable in the T3 generation (Figure 4.3c). High levels were stable when the transgene has been lost, and increased when the transgene was still present. The increase in the levels of the

blind phenotype remained even when only offspring of plants which had a normal phenotype in the previous generation were considered (Figure 4.3d), indicating that *TET3c* continued to increase the frequency of the blind phenotype in successive generations.

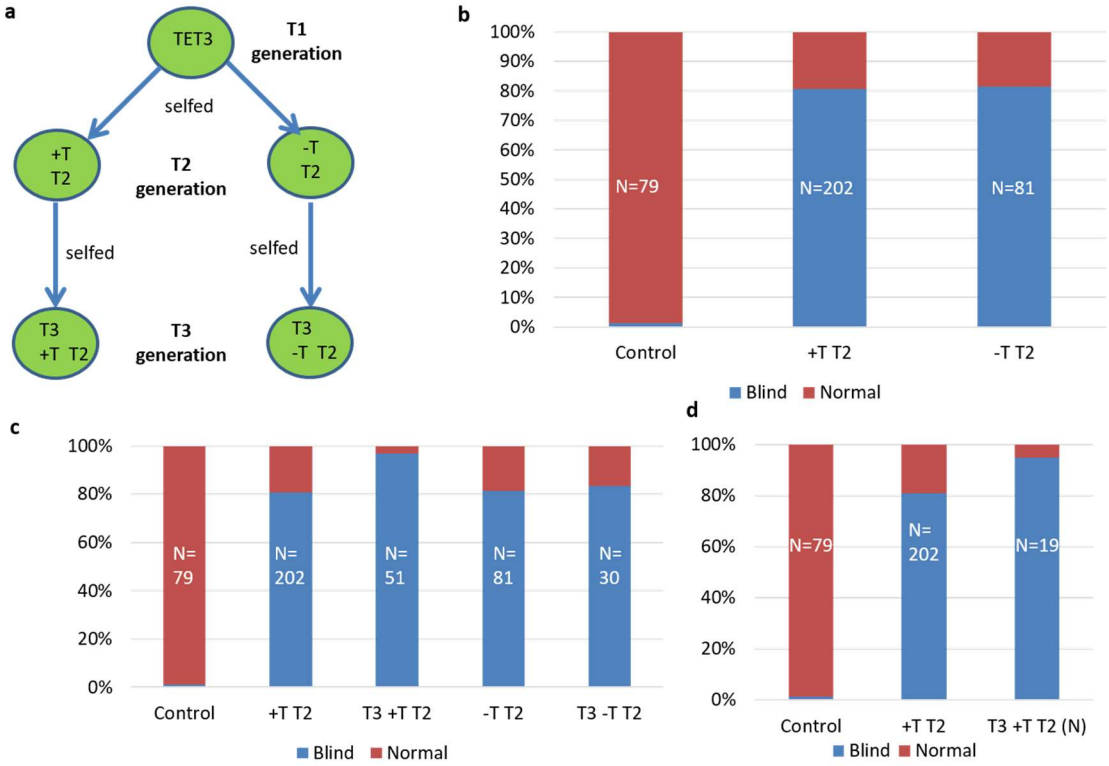


Figure 4.3. Likelihood of the blind phenotype was increased in TET3c tomato and this was heritable over two generations

(a) Different generations of the TET3c tomato were analysed. T1 plants containing the transgene were selfed. Seeds were grown off selection and genotyped to allow selection of T2 plants with the transgene (+T T2) and T2 plants which no longer contained the transgene (-T T2). Both +T T2 and –T T2 plants were selfed to obtain two sets of T3 plants, including those whose parents had contained the transgene (T3 +T T2) and those whose parents had not been transgenic (T3 –T T2).

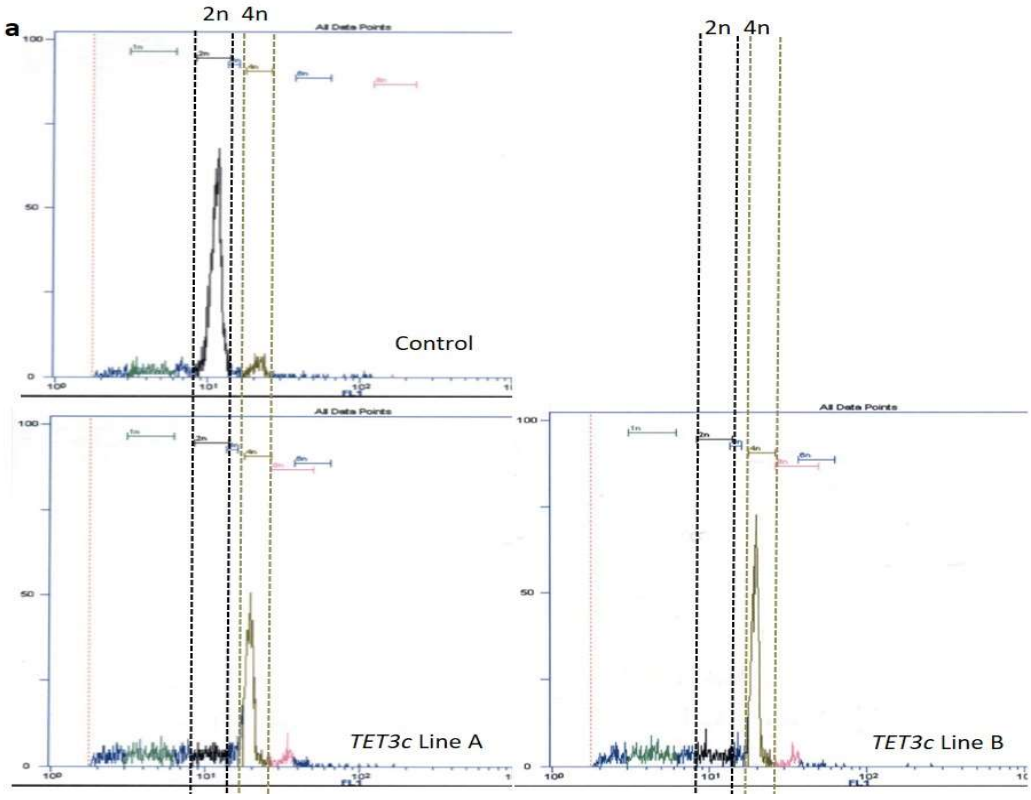
(b) High levels of the blind phenotype occurred in the TET3c tomato, and this was not affected by the presence or absence of the transgene in the T2 generation.

(c) High levels of the blind phenotype were maintained into the T3 generation when the transgene was absent, and increased when the transgene was present.

(d) The increase in the blind phenotype in the T3 generation when the transgene was present still occurred when only the offspring of plants which were normal in the T2 generation (T3 +T T2 (N)) were considered.

4.2.2. The frequency of the blind phenotype was higher in tetraploid *TET3c* tomato than diploid *TET3c* tomato

As well as the blind phenotype, *TET3c* tomatoes displayed other phenotypes. The plants possessed unusually thick leaves and could not be crossed with the wild type tomato. Dr Iris Heidmann carried out an analysis of the pollen, and found it to be diploid. This suggested that the *TET3c* tomato plants were tetraploid. Plants were analysed by cell flow cytometry, and this was confirmed (Figure 4.4a). It was possible that this was a result of the *TET3c* transgene, but tetraploidy is known to occur by chance during the tomato transformation process (Ellul et al., 2003). To investigate this, and the possibility that phenotypes observed in the *TET3c* tomato may be a result of the tetraploidy, two new *TET3c* transformant lines were generated (Lines E and F). The ploidy of these lines was confirmed to be diploid (Figure 4.4b).



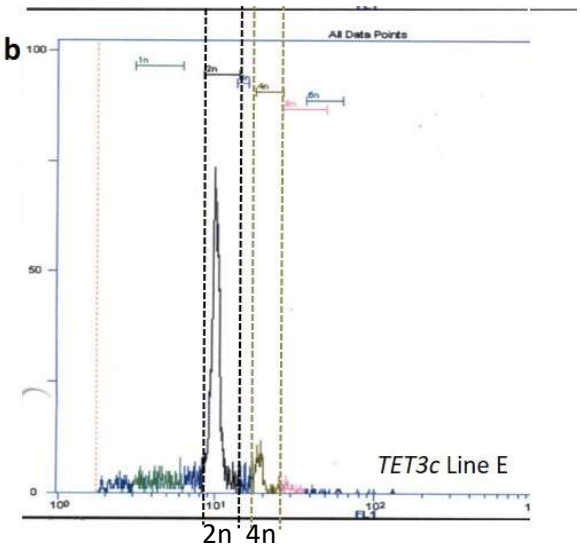


Figure 4.4. Ploidy was analysed using flow cytometry and the original TET3c lines were shown to be tetraploid

(a) Leaf samples from control and TET3c lines were analysed using flow cytometry. The majority of nuclei in the control (above) were diploid (a black peak) while most nuclei in the samples from TET3c lines A-D were tetraploid (a gold peak). Example graphs from lines A and B are shown (below). The boundaries of size for diploid (2n) lines is shown by dotted black lines, and the boundaries of size for tetraploid (4n) nuclei is shown by dotted gold lines. The majority of nuclei being within one of these boundaries classifies the tissue as diploid or tetraploid.

(b) Ploidy of plants from the new TET3c lines (E and F) were analysed by flow cytometry and were shown to be diploid. An example graph from line E is shown. The boundaries of size for diploid (2n) lines is shown by dotted black lines, and the boundaries of size for tetraploid (4n) nuclei is shown by dotted gold lines. The majority of nuclei being within one of these boundaries classifies the tissue as diploid or tetraploid.

Diploid *TET3c* plants were phenotyped for the presence of the blind phenotype. These lines also showed the blind phenotype at a higher frequency than the control, but not to the same extent as the tetraploid *TET3c* tomato (Figure 4.5).

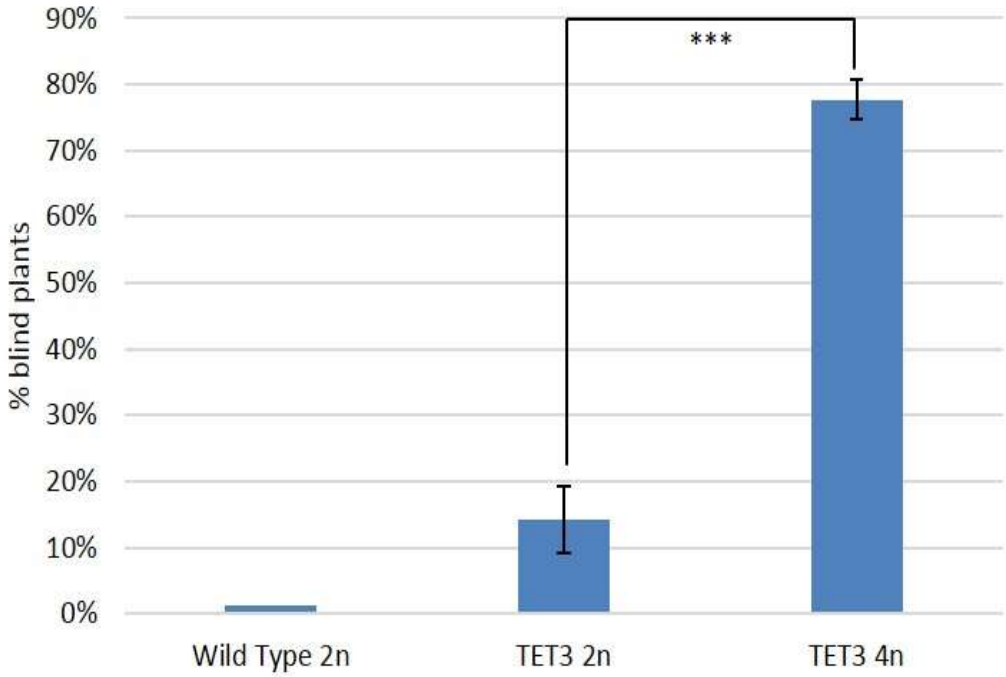


Figure 4.5. Diploid TET3c tomato were more likely to be blind than wild type tomato
*The likelihood of blind plants was increased in diploid TET3c tomato lines (n=112) but to a lesser extent compared to tetraploid TET3c tomato lines (n=299). Graphs show averages of different independent transformant lines with error bars indicating standard error. The significance of a change in the number of blind plants is indicated by asterisks: ns= not significant, ***=P<0.005, calculated by Student's two-tailed t-test.*

4.2.3. Tomato transformant lines containing *35S::CEN1.1* or *HI::GUS* also showed an increase in incidence of the blind phenotype

TET3c tomato were more likely to have the blind phenotype than the control, and this was increased if the *TET3c* tomato were tetraploid. There were therefore two separate effects to consider: the effect of *TET3c* on the frequency of the blind phenotype, and the effect of ploidy. Blindness is not reported to be a natural result

of tetraploidy in tomato (Nilsson, 1950), suggesting that tetraploidy alone was not the explanation for the increase in the blind phenotype.

TET3c may be causing the increase in blind phenotype incidence, and tetraploidy boosts this effect. This could be a result of more initial copies of *TET3c* during the transformation process, or epigenetic changes between the diploid and tetraploid tomato, which can occur between diploids and tetraploids (Róis et al., 2013). Alternatively, the two effects could be unrelated. In order to investigate this, the incidence of the blind phenotype in a variety of other transformants was analysed.

The frequency of the blind phenotype in the diploid *TET3c* tomato was compared to its frequency in three independent diploid transformants of an *HI::GUS* construct and three diploid *CEN1.1* transformant lines (discussed in Chapter Three) as well as the control (Figure 4.6a). The *HI::GUS* construct allowed us to determine if the increase in blind phenotype frequency which occurred in *TET3c* diploid tomato was a result of *TET3c* and the genes it activated (which includes *CEN1.1*) or a result of stress caused by transformation. Diploid *HI::GUS* tomato showed the blind phenotype at the same frequency as diploid *TET3c* and *CEN1.1* tomato. An increase in the frequency of the blind phenotype in this tomato variety can therefore be caused by the transformation of a transgene.

Whilst generating the *35S::CEN1.1* plants analysed previously in Chapter Three of this thesis, three tetraploid *35S::CEN1.1* lines were also generated. The frequency of the blind phenotype was analysed in these lines (Figure 4.6b). It was again increased in these tetraploid lines compared to the diploid *CEN1.1* tomato, but was significantly less than the frequency of the blind phenotype in tetraploid *TET3c* tomato. This indicates that although *TET3c* is not necessary for an increase in the likelihood of the blind phenotype in tetraploid lines, it is able to contribute to increasing the severity of this phenomenon.

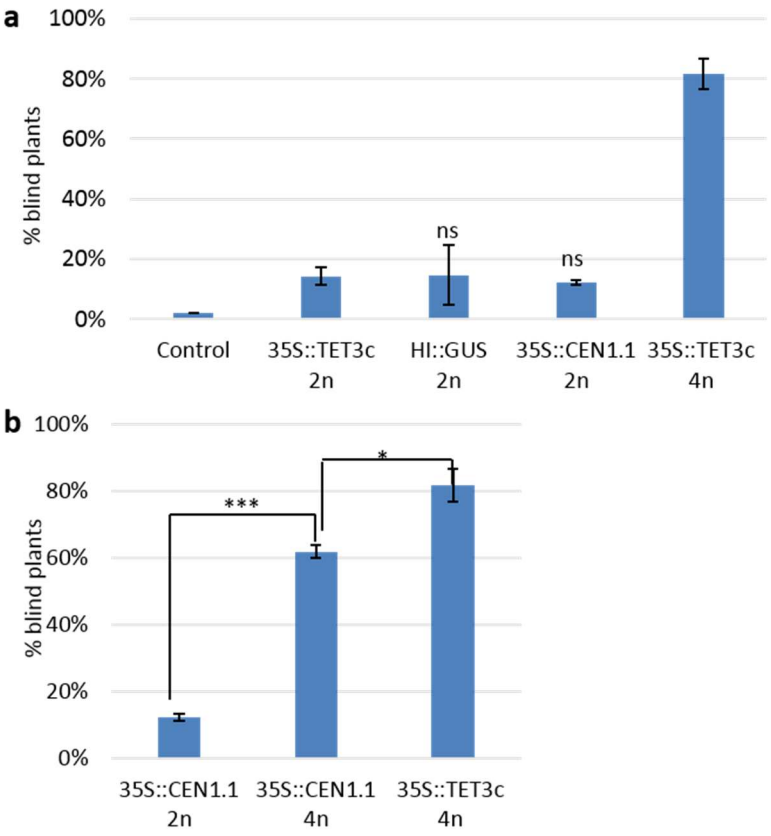


Figure 4.6. The frequency of the blind phenotype was increased in transformant lines and increased further in tetraploid lines

(a) The frequency of the blind phenotype was analysed in diploid transformants containing either the 35S::TET3c (n=112), 35S::CEN1.1 (n=41) or HI::GUS (n=34) construct. No significant difference between TET3c tomato and the other diploid transformants was observed, calculated using Student's two-tailed t-test.

*(b) The frequency of the blind phenotype was analysed in diploid and tetraploid 35S::CEN1.1 (n=34) transformants. A significant increase in the blind phenotype was seen in tetraploid CEN1.1 tomato compared to diploid, but this was significantly less than the frequency of the blind phenotype in tetraploid TET3c tomato (n=299). The significance of a change in the number of blind plants is indicated by asterisks: ns= not significant, * = P <0.05, ***=P<0.005, calculated by Student's two-tailed t-test.*

Graphs show averages of different independent transformant lines with error bars indicating standard error.

4.2.4. Ectopic expression of *WUS* correlated with the blind phenotype in diploids and tetraploids

Methylation changes caused by *TET3c* were not necessary to cause the blind phenotype in these tomato transformants. Despite this, the *TET3c* tomato lines with increased incidence of the blind phenotype are an excellent tool to investigate the cause of this phenotype, which can be highly problematic for tomato breeders.

RNA was collected from tomato containing either the *35S::TET3c*, *35S::CEN1.1* or *HI::GUS* construct in order to investigate which genes could be causing the blind phenotype. An initial analysis was done using semi-quantitative RT-PCR to check levels of *CEN1.1* expression. As the presence of the blind phenotype did not correlate with the presence of the transgene, it was not expected that expression of these genes would correlate with the blind phenotype. However, it was possible that endogenous *CEN1.1* expression had been activated in the *CEN1.1* tomato lines and was therefore causing the blind phenotype. Expression of *CEN1.1* did not correlate with the presence of the blind phenotype in *35S::TET3c*, *35S::CEN1.1* or *HI::GUS* tomato (Figure 4.7), and is therefore unlikely to be the cause of the blind phenotype. Expression of *CEN1.1* was also not restricted to transgenic *CEN1.1* lines, suggesting that expression of this gene may be activated by transformation stress.

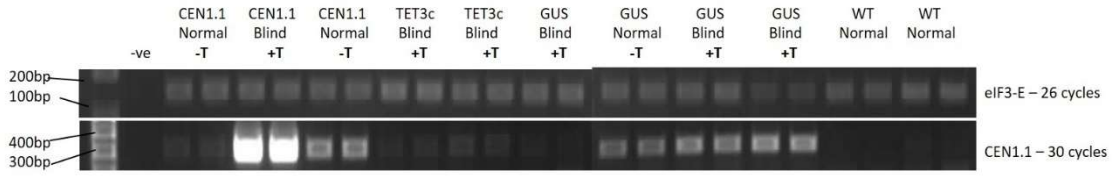


Figure 4.7. CEN1.1 expression did not correlate with the presence of the blind phenotype in tomato

Diploid tomato seedlings from the 35S::CEN1.1, 35S::TET3c and HI::GUS transformant lines were analysed for their expression of CEN1.1 (301bp) using semi-quantitative RT-PCR of cDNA. CEN1.1 could be ectopically expressed in tomato with the transgene (+T) and without (-T), although expression levels were highest when the 35S::CEN1.1 construct was present. The blind phenotype did not correlate with the presence or absence of CEN1.1 expression. 2 technical replicates of each sample were done. cDNA levels were normalised using the constitutively expressed eukaryotic initiation factor 3E (150bp). –ve indicates a negative H_2O control. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

An important regulatory pathway for control of SAM maintenance is the CLAVATA-WUSCHEL feedback loop, as discussed in the introduction (Somssich et al., 2016). The tomato homolog of *WUSCHEL, Solyc02g083950*, has been characterised by expression in Arabidopsis (Wang et al., 2012) and by laser ablation of the CZ (Reinhardt et al., 2003). Several potential tomato CLV family members have been identified (Zhang et al., 2014), but *Solyc11g071380* is most likely to be the functional homolog of *CLV3* and requires arabinosylation to function completely (Xu et al., 2015). Expression of three tomato *CLV* family members (*SlCLV3*, *SLCLE9* (which is able to rescue enlarged meristems lacking functional *SlCLV3* (Xu et al., 2015)) and *SlCLE13*, which is specifically expressed in the tomato shoot apex (Zhang et al., 2014)) and of the tomato homolog of *WUSCHEL* was therefore analysed using quantitative RT-PCR (Figure 4.8). Expression of *SlCLE3*, which is broadly expressed across tomato tissues (Xu et al., 2015) was also analysed for comparison (Figure 4.8). *35S::TET3c* tomato lines were used to analyse gene expression as they showed

the blind phenotype at the highest frequency and may therefore be expected to have the greatest changes in gene expression. Leaf material was used for the expression analysis to see if ectopic expression of these genes had occurred, as expression of *CLV3*, *CLE9*, *CLE13* and *WUS* is not found in the leaf in normal tomato (Wang et al., 2012; Xu et al., 2015; Zhang et al., 2014). One wild type blind plant was also observed at this time, and was used for this analysis (Figure 4.8).

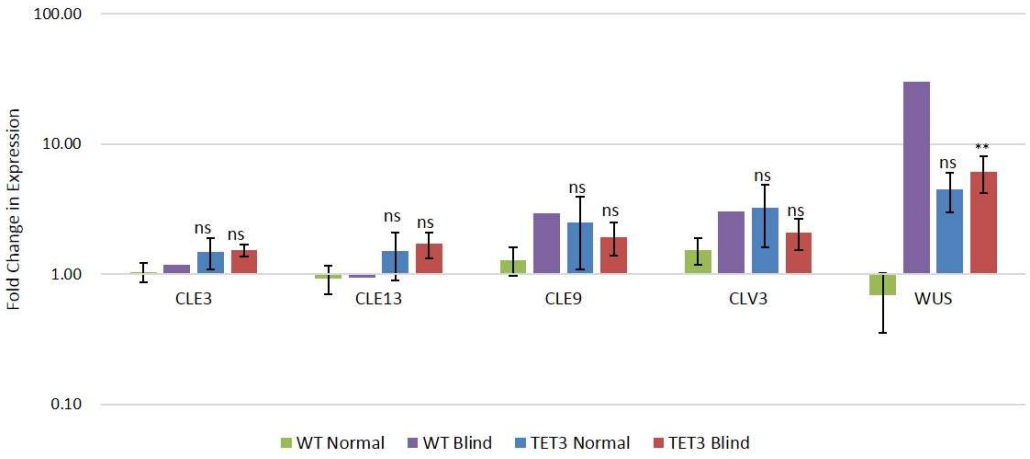


Figure 4.8. Expression levels of CLE9, CLV3 and WUS were raised in wild type blind tomato

*Expression levels of 4 CLAVATA tomato homologs, and a WUSCHEL tomato homolog were analysed in the leaves of normal and blind 5 week old tomato plants using quantitative RT-PCR. Samples were normalized to a single wild type normal seedling. Except WT blind, which was a single plant, graphs show averages of biological replicates (n=6-35) with error bars indicating standard error for all samples. The significance of a change in expression from WT normal is indicated by asterisks: ns= not significant, ** = P<0.01, calculated by Student's two-tailed t-test. Fold changes in expression were calculated using ΔCt method as described in the methods.*

Expression levels of *CLE9*, *CLV3* and *WUS* were all raised in the wild type blind plant (Figure 4.8). The significance of this could not be calculated due to the existence of only a single sample. Where significance could be calculated, only *WUS* expression had changed significantly, increasing in *TET3c* blind plants compared to wild type tomato (Figure 4.8). This is surprising given that the established effect of WUS is

maintenance of the shoot meristem, and therefore increased *WUS* expression would be expected to enlarge the shoot apical meristem, not result in its loss as is observed in the blind phenotype. Ectopic *WUS* expression in the leaves of tomato plants may indicate wider abnormalities in the expression of shoot meristem genes elsewhere in the shoot meristem maintenance pathway.

4.2.5. Expression of *WUS*, *CLV3* and *CLE9* was elevated prior to development of the blind phenotype in *TET3c* plants

Gene expression was analysed at five weeks old, at the same time as plants were phenotyped for the blind phenotype. Given that blind plants were frequently able to recover and continue to grow, the expression changes causing the blind phenotype are likely to be transitory. Therefore, expression analysis must be done in plants prior to the development of the blind phenotype. The earliest age at which plants could be classified as blind was 18 days after sowing and therefore RNA was collected from plants at 14 days old and quantitative RT-PCR was carried out (Figure 4.9). Expression of *WUS*, *CLV3* and *CLE9* was analysed, as these had increased in the wild type blind plant. An increase in the expression of *WUS*, *CLV3* and *CLE9* was observed in eight of the nine tomato seedlings. Due to the young age at which they were sampled, only three plants survived sampling (T1, T2 and T6). All three had shown ectopic expression of *WUS*, *CLV3* and *CLE9*, and did develop the blind phenotype on maturity. However, these sample sizes are too small to say conclusively that the ectopic expression of these genes causes the blind phenotype in these tomato lines.

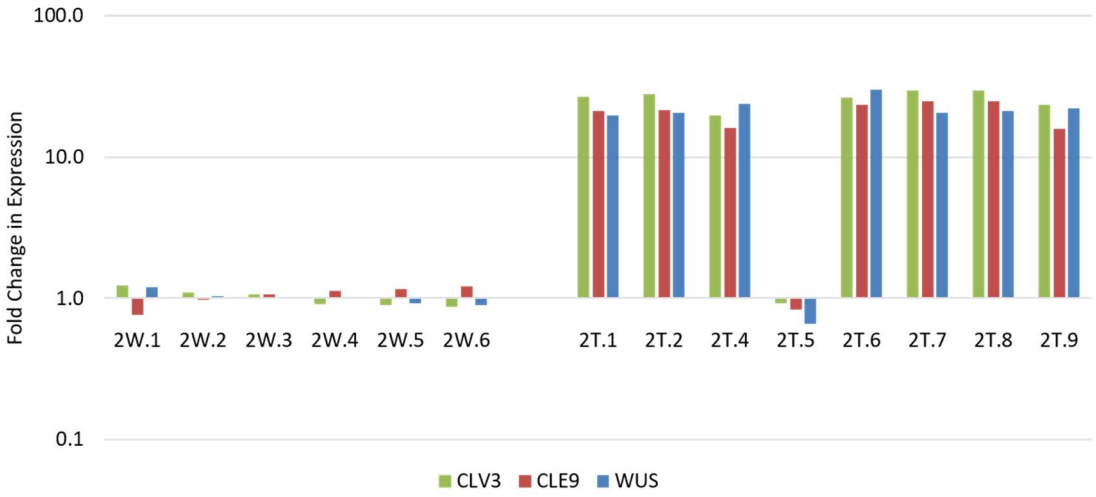


Figure 4.9. Expression levels of CLV3, CLE9 and WUS were raised in the majority of TET3c tomato seedlings prior to the development of the blind phenotype

Expression levels of CLV3, CLE9 and WUS were analysed in the leaves of 2 week old tomato plants using quantitative RT-PCR. Samples were normalized to a single wild type seedling and the constitutively expressed eukaryotic initiation factor 3E. Samples 2T.1, 2T.2 and 2T.6 were the only 2T seedlings to survive sampling and went on to develop the blind phenotype. Samples 2W.1, 2W.3, 2W. 4, 2W.5 and 2W.6 also survived sampling and did not develop the blind phenotype.

Three constructs were generated which contained either the *CLV3*, *CLE9* or *WUS* gene behind the constitutive 35S promoter (Figure 9.6a). An inducible construct was also made containing WUS fused to a glucocorticoid tag (Figure 9.6b), resulting in the production of WUS-GR fusion proteins. This allows induction of WUS effects, as the WUS-GR protein is sequestered in the cytoplasm, and only upon treatment of plants with dexamethasone is WUS transported to the nucleus and able to activate gene expression (Gallois et al., 2002). CLV3 and CLE9 constructs were not generated using the same induction system, as they are not transcription factors. Alternative induction systems such as the heat-inducible or ethanol-inducible system were considered, but have previously been shown to be problematic in tomato. High temperature conditions required for sustained induction of the HI promoter are lethal to tomato (Watson, 2013) and treatment of tomato plants with ethanol to

activate expression of an ethanol-inducible gene resulted in high levels of the blind phenotype in the offspring of ethanol-exposed control tomato which did not contain a transgene (Figure 9.6c). These constructs are currently being transformed into tomato.

4.2.6. The blind phenotype in natural tomato varieties also correlated with ectopic expression of *WUS, CLV3* and *CLE9*

All transformations were carried out in the EZCBT1 tomato variety, in which the blind phenotype occurs at a frequency of 1%. Other varieties of tomato where the blind phenotype is more common also exist, such as Naraam, where growth at high temperatures (25-26°C) results in an increased likelihood of the blind phenotype, with incidents of up to 20% of plants being blind reported (I. Heidmann, personal correspondence). In order to investigate if the molecular causes of the blind phenotype in these varieties could be similar to the causes in the transformant tomato lines, Naraam plants were grown under high temperature induction conditions in growth chambers and phenotyped (Figure 4.10). An increase in the percentage of blind plants was seen under high temperature conditions.

a

b

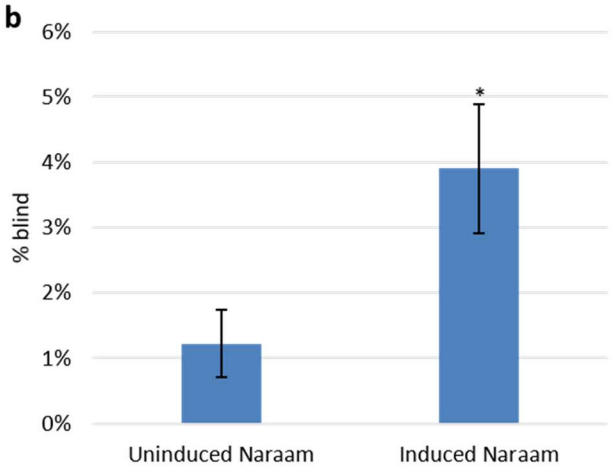


Figure 4.10. Tomatoes of the Naraam variety showed an increased likelihood of the blind phenotype under high temperature inductive conditions

(a) The blind phenotype was observed in Naraam tomato (right). It again appeared as a lack of shoot apical meristem, which is present in control plants (left). Images show the apex of 10 day old Naraam tomato seedlings.

*(b) An increased percentage of Naraam plants were blind when grown under inductive conditions of high temperature (25-26°C). Graphs show averages of 6 different seed batches, each containing 100 seeds with error bars indicating standard error. The significance of an increase from uninduced Naraam is indicated by asterisks: * = P <0.05, calculated using Student's one-tailed t-test.*

RNA was collected from both blind and normal plants grown at high temperatures (25-26°C). RNA was also collected from uninduced Naraam which was grown at 20-23°C, the same temperature used for growth of the other tomato plants. Using quantitative RT-PCR, expression levels of *CLV3*, *CLE3*, *CLE9*, *CLE15* and *WUS* were investigated (Figure 4.11). Significant upregulation of *WUS*, *CLE9* and *CLV3* was observed in blind Naraam plants. *WUS* expression was also significantly increased in plants which were grown at high temperatures compared to tomato plants grown at lower temperatures. This indicates that overexpression of *WUS* alone is not sufficient to cause the blind phenotype, as would be expected from its known biological role, but that ectopic *WUS* expression may be correlated to an increase in the likelihood of the blind phenotype.

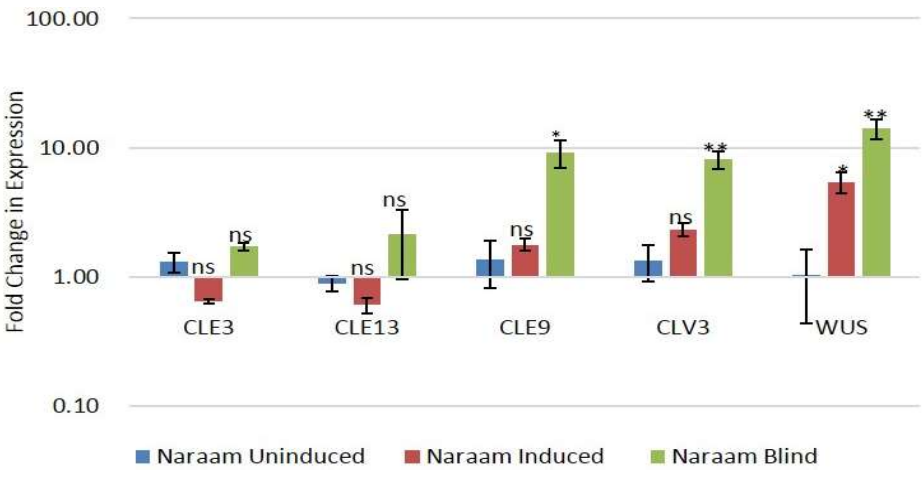


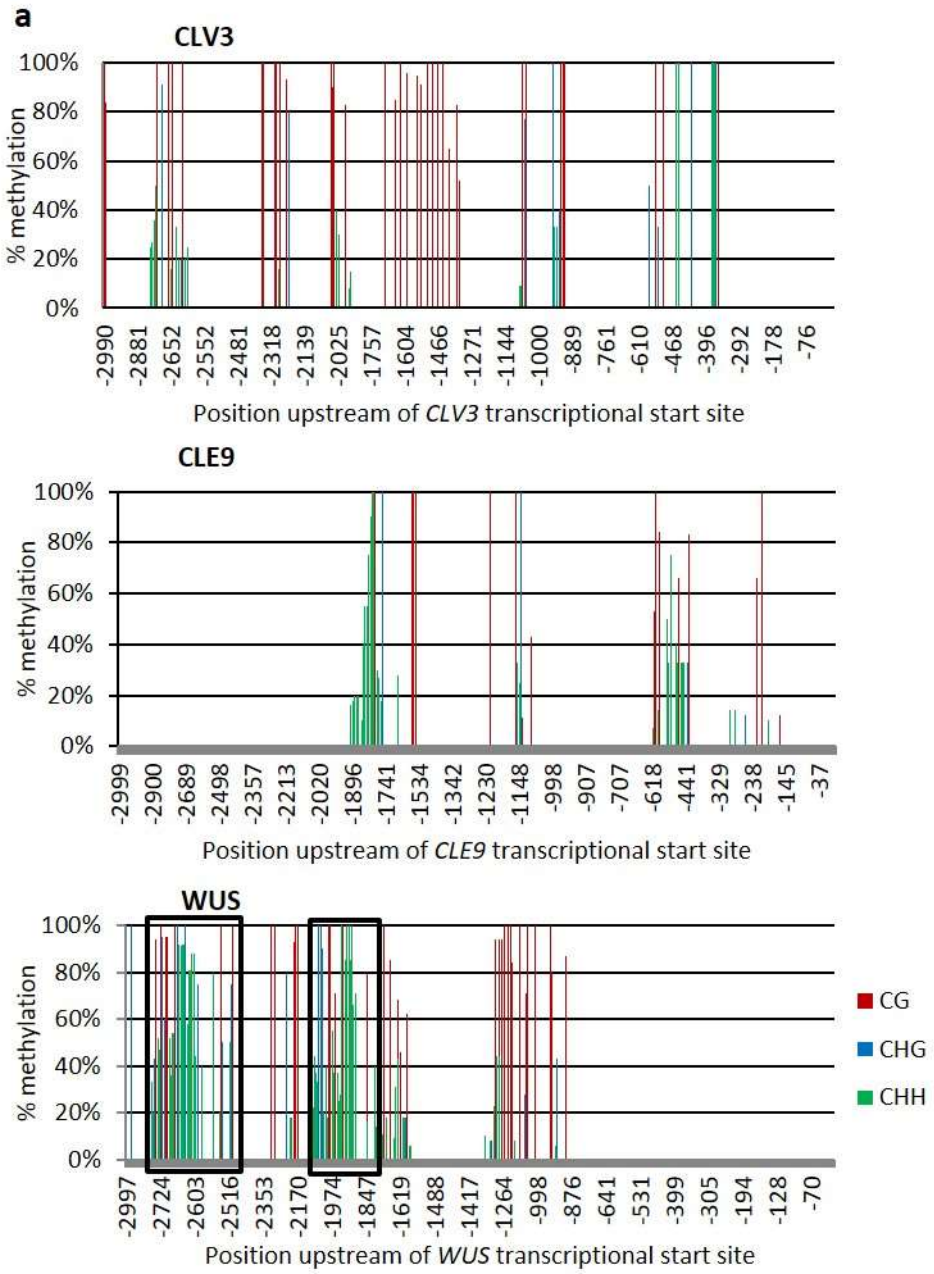
Figure 4.11. Expression levels of WUS, CLV3 and CLE9 were significantly raised in blind Naraam tomato

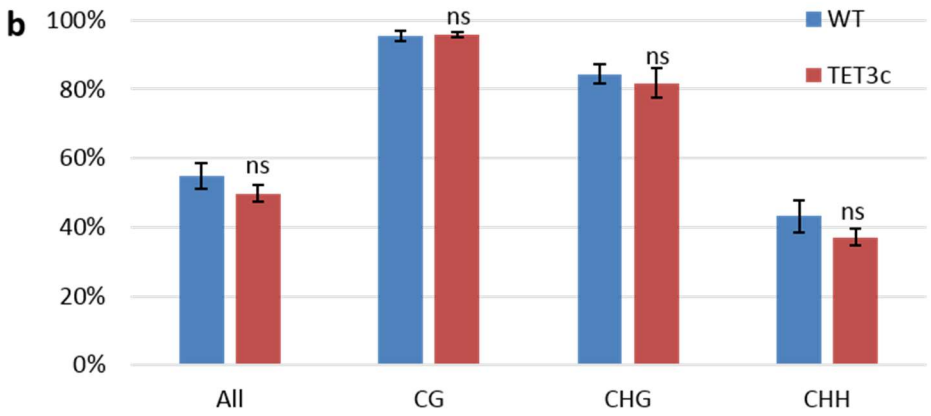
*Expression levels of 4 CLAVATA tomato homologs, and a WUSCHEL tomato homolog were analysed in the leaves of non-blind Naraam tomato grown at normal temperatures (Naraam Uninduced, n=3) and high temperatures (Naraam Induced, n=29), as well as the leaves of blind Naraam (Naraam Blind, n=3) using quantitative RT-PCR. Samples were normalized to a single Naraam seedling grown at normal temperatures. Graphs show averages of biological replicates with error bars indicating standard error for all samples. The significance of a change in expression from Naraam Uninduced is indicated by asterisks: ns= not significant, * = P <0.05, ** = P<0.01, calculated by Student's two-tailed t-test.*

4.2.7. Ectopic *WUS* expression did not correlate with significant methylation changes upstream of the promoter

The blind phenotype increased in frequency in tomato transformant lines, and this increase was independent of the presence of the transgene (Figure 4.3). Epigenetic changes are therefore likely to be the cause of the increase in the frequency of blind tomato. If *CLV3*, *CLE9* or *WUS* are causing the blind phenotype, their activation is likely to be linked to alterations in the epigenetic modifications surrounding these genes. Levels of DNA methylation upstream of *CLV3*, *CLE9* and *WUS* were investigated using the tomato epigenome database (Zhong et al., 2013)(Figure 4.12a). Upstream methylation levels were high for all three genes, supporting the possibility that they may be activated by epigenetic changes. *WUS* was the only gene significantly upregulated in five week old blind *TET3c* tomato (Figure 4.8) and *WUS* expression in Arabidopsis has been shown to be linked to methylation during shoot regeneration (Li et al., 2011a; Shemer et al., 2015). In addition, *WUS* expression was also significantly upregulated in Naraam tomato grown at temperatures known to induce the blind phenotype. This evidence suggests that *WUS* expression may be an earlier step in the induction of the blind phenotype, which can subsequently have the effect of activating *CLV3* and *CLE9* expression. Methylation levels upstream of the *WUS* gene were therefore analysed in wild type

and *WUS*-expressing *TET3c* tomato. As *TET3c* expression has been shown to cause changes specifically in CHH methylation in tomato (Chapter 3, Figure 3.5), methylation was analysed at two regions which contained 76% of the CHH methylation within the first 3kb upstream of the *WUS* transcriptional start sites. Although CHH methylation was slightly reduced, no significant changes in methylation were observed in any context (Figure 4.12b,c).




b
100%
80%
60%
40%
20%
0%
ns
ns
ns
ns
WT
TET3c
All
CG
CHG
CHH

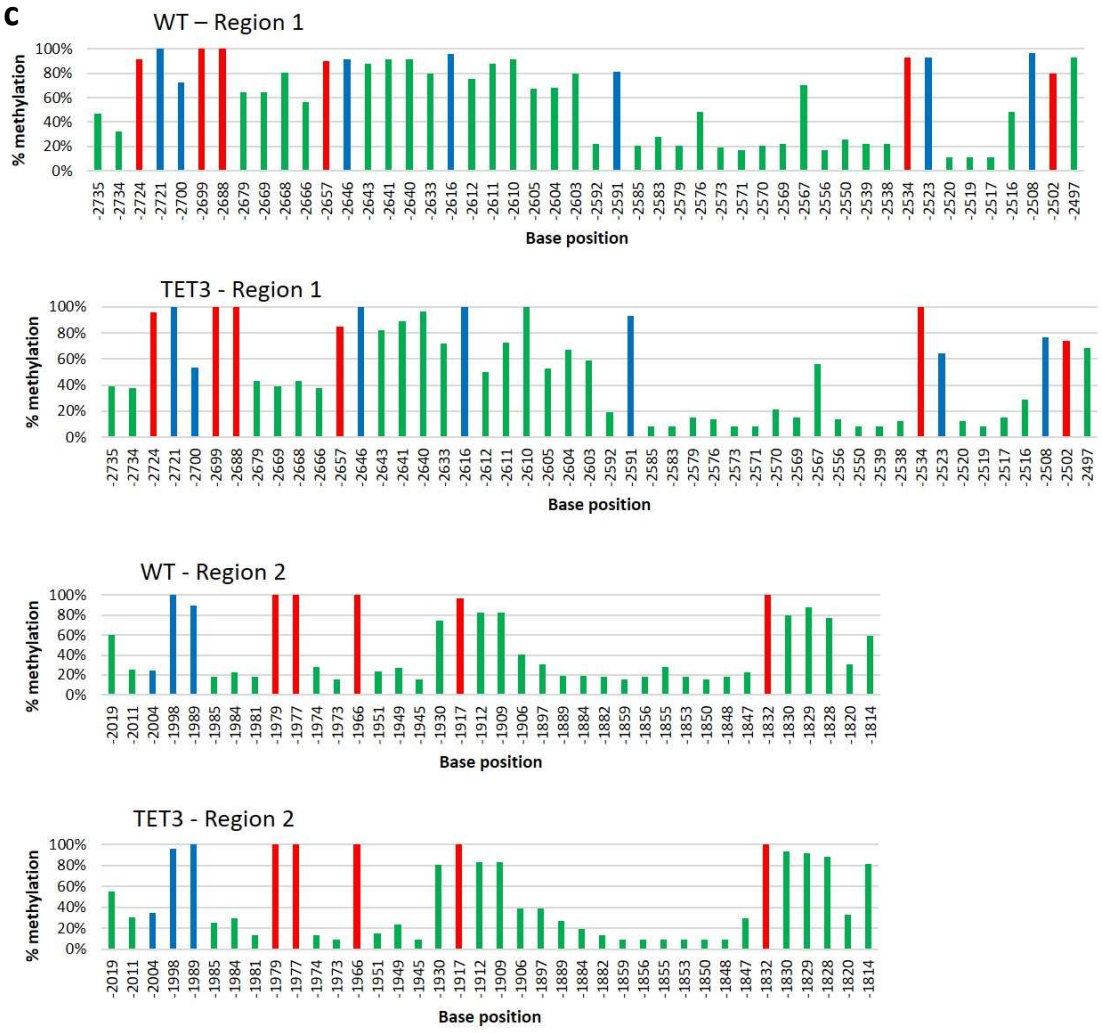

c
WT – Region 1
TET3 - Region 1
WT - Region 2
TET3 - Region 2
% methylation
Base position

Figure 4.12 Methylation levels upstream of WUS were high but did not change in TET3c plants which express WUS

(a) *Database methylation upstream of CLV3, CLE9, and WUS according to Zhong et al., 2013. Boxes indicate the region analysed in (b) and (c) (Region 1: -2735 to -2497 and Region 2: -2019 to -1814).*

(b) *Levels of 5-methylcytosine upstream of the WUS transcriptional start site were not significantly changed in TET3c tomato compared to the control. The % of 5-methylcytosine is calculated as the number of methylated cytosine bases divided by the total number of cytosine bases in the two regions. Three biological replicates were analysed by bisulfite sequencing for each sample and 10 clones for each replicate. Graphs show averages with error bars representing standard error. The significance of a change from WT is indicated by asterisks: ns= not significant, calculated using Student's two-tailed t-test.*

(c) *Levels of 5-methylcytosine upstream of the WUS transcriptional start site were not significantly changed in TET3c tomato compared to the control, shown on a site by site basis as it is in (a). The % of 5-methylcytosine was calculated as the number of methylated clones divided by the total number of clones at each site.*

4.3. Discussion

Diploid tomato plants of the EZCBT1 variety showed an increase in the incidence of the blind phenotype after transformation to 12.2% of plants on average. Tetraploid transformant lines, which can occasionally occur during leaf disc transformation, had a greater increase in the frequency of blind plants, with 71.8% of plants displaying the blind phenotype. Which genes cause the increase in the frequency of blind plants remains uncertain. Ectopic expression of *WUS* occurred in *TET3c* lines and in the Naraam tomato variety which is prone to the blind phenotype. Expression of *CLV3* and *CLE9* was also raised in blind Naraam tomato, as well as the single blind tomato of the EZCBT1 variety and in young blind *TET3c* tomato. This suggests that these genes, which are essential to maintain the shoot apical meristem at its normal

size, have defects in their regulation in plants with the blind phenotype. Ectopic expression of *WUS* was seen both in tomato with and without the blind phenotype in these lines, indicating that it is not sufficient to cause the blind phenotype. In the normal SAM, WUS induces expression of *CLV* genes. The same process may be occurring in the blind plants, with an initial activation of *WUS* expression resulting in the ectopic expression of *CLV3* and *CLE9*. Expression levels of *CLV3* and *CLE9* were very similar, showing a Pearson correlation of 0.96 and may be controlled by the same factors. Control of SAM maintenance is complex, involving many genes and the combined expression of *WUS* plus *CLV3* and/or *CLE9* could be required to cause the blind phenotype. Meristem size can remain stable even with large increases in *CLV3* expression (Müller et al., 2006), suggesting that even if *WUS, CLV3* and *CLE9* have a role in causing the blind phenotype, further feedback mechanisms are likely to be involved. Similarly, tomato plants containing an RNAi construct against the *WUSCHEL* gene are still able to form a functional meristem (Li et al., 2017). Changes in the expression of more than one gene are likely to be necessary to cause the blind phenotype.

Further investigation of this can be done using the four constructs described in Figure 4.10. *35S::WUS* transformant lines would allow analysis of whether constitutive expression of *WUS* behind the 35S promoter alone is sufficient to give an increase in the proportion of blind plants by activating expression of *CLV3* and *CLE9.* Alternatively, crosses of lines containing the *WUS* and *CLV* constructs may be required to generate an increase in the incidence of the blind phenotype (above the levels usually seen in transformant lines). The inducible *WUS::GR* construct may be especially useful if expression of *WUS* is preventing the formation of early meristems by activating *CLV3* and *CLE9* as this would be detrimental to transformation efficiency. However, increased *WUS* expression in Arabidopsis increases the rate of shoot regeneration (Li et al., 2011a) and therefore increased *WUS* expression in tomato may not be detrimental to transformation efficiency. The correct time point for induction of *WUS* movement in *WUS::GR* lines would also

have to be established, as it is unknown at which point the blind phenotype is caused in early seedling development.

Analysis of transformant lines containing alternative constructs demonstrated that transformation stress, not *TET3c,* was the cause of the increased incidence of the blind phenotype in diploid transformants. Tetraploidy could also increase the incidence of the blind phenotype in transformant lines of *TET3c* and *CEN1.1*. However, a significantly greater percentage of tetraploid *TET3c* plants were blind in comparison to tetraploid *CEN1.1* plants. In addition, in Figure 4.3c it was observed that the incidence of the blind phenotype increases in the T3 generation in comparison to the T2 generation, but only when the *TET3c* transgene is present. This may indicate a possible role for *TET3c* in amplifying the epigenetic changes caused by transformation stress and tetraploidy which have led to the blind phenotype. Subsequent generations of tetraploid *CEN1.1* plants should be analysed to investigate if the blind phenotype also increases in frequency in these lines.

Significant hypomethylation was not observed upstream of the *WUS* transcriptional start site in *TET3c* tomato, and therefore the ectopic expression of *WUS* was not a result of demethylation in this region. An increase in *WUS* expression may be caused by alternative epigenetic changes, on the RNA or chromatin level, or methylation changes outside of the analysed region. Alternatively, a different gene may cause ectopic expression of *WUS.* Hormones including auxin and cytokinin are known to affect the expression of *WUS* (Zhao et al., 2010)*,* and therefore genes involved in the synthesis, response to, or degradation of these hormones could be candidates for the cause of the blind phenotype. Changes in auxin or cytokinin levels would be expected to produce more global changes in addition to the blind phenotype, such as alterations in root length or structure. These were not observed, indicating that if a change in hormone levels or response causes the blind phenotype, it must be extremely localised. Simultaneous upregulation of *WUS, CLV3* and *CLE9* has previously been observed in tomato when enzymes responsible for processing of the CLV3 gene are mutated (Xu et al., 2015). These may therefore be important

candidates for investigation although mutants of these enzymes have enlarged meristems (Xu et al., 2015), suggesting that alternative mechanisms are involved.

In addition to being localised to the aerial tissue, the causes of the blind phenotype must be temporally localised. Blind *35S::TET3c* tomatoes were usually able to recover and continue growth, with meristems developing in a variety of locations where they would not usually be seen (Figure 4.1c,d). Other blind tomatoes, including Naraam tomatoes, were also able to recover, indicating that this is not unique to the *TET3c* blind tomato. Despite this, the increased level of vegetative meristematic behaviour seen in the *35S::CEN1.1* tomato could suggest that the ectopic activation of *CEN1.1* that is known to occur in *TET3c* tomato was enabling improved recovery of the blind plants. The recovery of plants ectopically expressing *CEN1.1* (100%, n=4) was compared to the recovery of plants which did not express *CEN1.1* (88%, n=8) but was not greatly increased. Blind tomato may alternatively be able to recover due to the downregulation of ectopically expressed genes or due to developmental stage specificity of the causes of the phenotype. In crop species such as rice, different CLV3 homologs regulate different meristem types (Suzaki et al., 2008) and there is no evidence that methylation control of WUS affects its role in floral meristem regulation in Arabidopsis (Cao et al., 2015).

Much is still unknown about the maintenance of SAM function in plant species, including how environmental stresses are integrated with the CLV3-WUS pathway (Somssich et al., 2016). Our work has shown that the blind phenotype in tomato can be caused not only by abiotic stresses such as high temperature, but also by the process of transformation and by the formation of tetraploids. Ectopic expression of tomato homologues of *WUSCHEL* and *CLAVATA* occurs in blind plants, indicating an involvement of this pathway in the blind phenotype. Further work is still required to investigate whether transformation induces the blind phenotype due to T-DNA integration or the process of leaf disc transformation and shoot regeneration, both processes which could be expected to disturb the epigenome of the tomato. Tetraploids formed by alternative processes such as colchicine treatment (Praça et

al., 2009) rather than leaf disc transformation must also be investigated to see if this can also result in an increased likelihood of the blind phenotype. As discussed above, ectopic expression of *WUS* may be part of the cause of the blind phenotype but more investigation of this would have to be done using constructs which overexpress *WUS*.

5. Expressing a Mammalian Demethylase in Different Plant Species

5.1. Introduction

Work in *Arabidopsis thaliana* has shown that the mammalian enzyme TET3 can cause changes in DNA methylation in plants (Hollwey et al., 2016), as discussed in Chapter Two. In Chapters Three and Four, the effects of expressing the catalytic domain of TET3 (TET3c) in tomato were analysed and appeared to be more widespread than in Arabidopsis, suggesting a greater sensitivity to methylation changes which correlates with the increased methylation levels in tomato. The possible applications of *TET3c* are not limited to tomato. As discussed previously, different plant species contain different levels of DNA methylation (Feng et al., 2010; Zemach et al., 2010) and may rely on DNA methylation to different extents. The result of disturbing the main methyltransferases have been analysed in other crop species, including in rice, where homozygous *met1* mutants have a severe delay in growth that is ultimately lethal (Hu et al., 2014), and in tobacco, where plants containing an antisense construct against tobacco *MET1* are reduced in methylation and display phenotypes including reduced leaf size and smaller internodes (Nakano et al., 2000).

Tobacco (*Nicotiana tabacum)* contains approximately 30% methylation (Kovarík et al., 2000; Messeguer et al., 1991), a higher methylation percentage than in tomato (~25%) (Messeguer et al., 1991; Zhong et al., 2013) or rice (~15%) (Feng et al., 2010). Tomato plants containing a *MET1* RNAi construct cannot be regenerated (Watson, 2013). This is a more severe phenotype than observed in tobacco *MET1* RNAi lines despite the lower levels of methylation in tomato, suggesting that the impact of disturbing the methylation systems of a plant is not always correlated to the levels of methylation in that species. This is also demonstrated by the severity of phenotypes for methylation mutants in mammals despite the comparatively low levels of total methylation in mammals (less than 5% of all cytosines are methylated in mouse, as methylation exists mainly in the CG context (Feng et al., 2010)). Rice,

too, has a lower level of total methylation than tobacco and, as described above, reacts very severely to mutation of *MET1* (Hu et al., 2014). This could be due to a *met1* mutation being more severe than an RNAi knockout of *met1*, which may reduce levels of the MET1 protein less effectively. The behaviour of tobacco *MET1* RNAi lines suggests that transformation of *TET3c* into tobacco may have only mild effects.

Lettuce (*Lactuca sativa*) is known to express transgenes poorly and at variable levels (McCabe et al., 1999a). Transgenes are also unstable over the generations, with unusually high levels of transgene inactivation occurring from one generation to the next (McCabe et al., 1999b). A combination of factors can affect this in plants, including position of integration into the genome (Pröls and Meyer, 1992), environmental conditions (Meyer et al., 1992) and methylation of transgenes (Meyer and Heidmann, 1994). Treatment with the methylation inhibitor 5-azacytidine improves transgene expression for some lettuce lines but not others, suggesting that transgene methylation is responsible for silencing in some lettuce lines (McCabe et al., 1999a). Given the demethylation activity of *TET3c*, transformation with the *TET3c* construct may result in more stable expression of the transgene.

The effects of *TET3c* were analysed in plants of both lettuce and tobacco. Tobacco was chosen due to the ease of transformation and culture of this species, as well as the large quantity of literature available on methylation in tobacco. Lettuce was chosen due to the usually poor expression of transgenes in this species, in order to investigate whether transformation with *TET3c* would result in stronger maintenance of transgene expression.

5.2. Results

5.2.1. The catalytic domain of TET3 was not expressed highly in lettuce

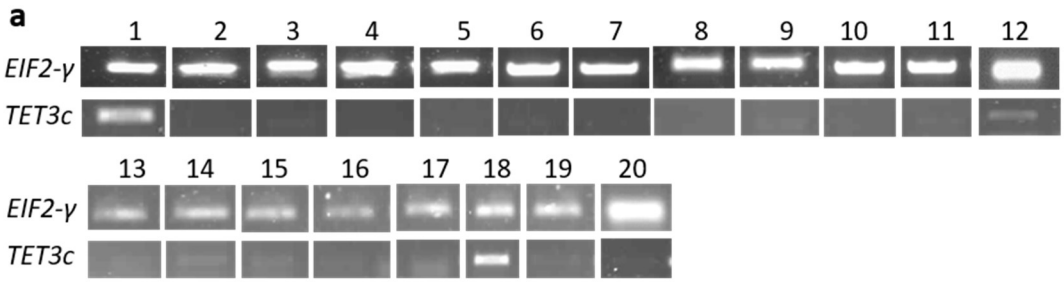
Lettuce (*Lactuca sativa* var. Pinokkio) was transformed with the *35S::TET3c* construct by Dr Iris Heidmann and Suzan Out at Enza Zaden. Twenty independent

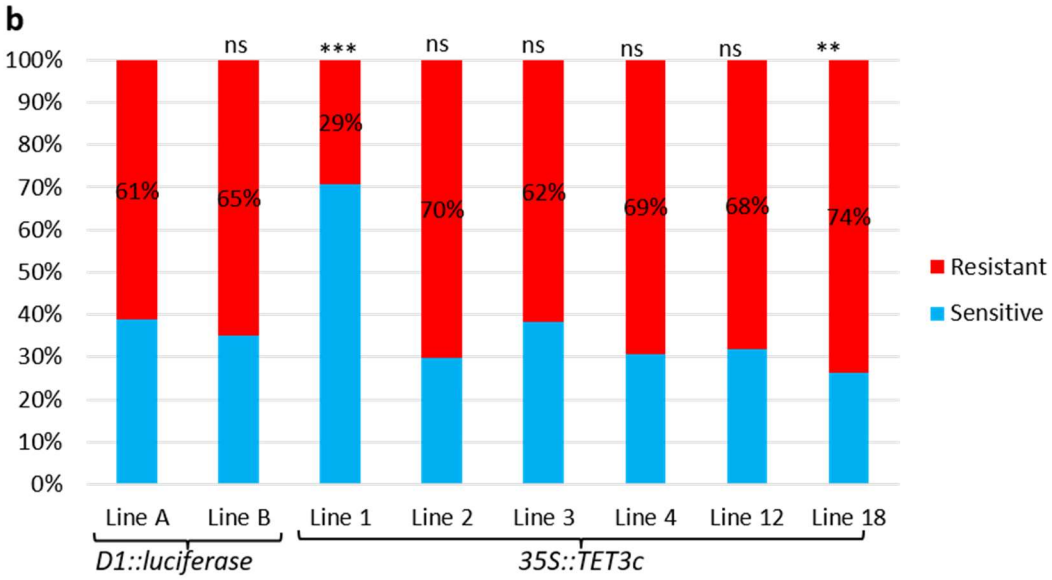
transformant lines were generated. T1 seeds of these lines were grown on kanamycin selection. DNA was extracted and plants were genotyped using PCR to confirm the presence of the transgene. RNA was extracted from plants which contained the transgene and converted to cDNA. Semi-quantitative RT-PCR on pools of cDNA from three plants was carried out to analyse the expression of *TET3c* in the lettuce plants (Figure 5.1a). *TET3c* expression was present in only three of the twenty lines: Lines 1, 12 and 18.

A germination assay on kanamycin selection was done to see if expression of *TET3c* correlated with reduced silencing of the *NPTII* gene in the T-DNA (for kanamycin resistance) and therefore possibly reduced methylation of the transgene. If no silencing had occurred, then 75% of plants should be resistant to kanamycin in the T1 generation. However, due to the high level of transgene silencing in lettuce, a lower percentage of kanamycin-resistant plants is expected. Two control lines containing a D1::luciferase construct (Line A and Line B) were used to calculate whether the percentage of kanamycin-resistant plants had increased in the *35S::TET3c* lines (Figure 5.1b). Six *TET3c* lettuce lines were analysed, including the three lines which expressed *TET3c*. The significance of any difference, either an increase or a decrease, from the control Line A was calculated, and only Line 1 and Line 18 showed a significant change. Line 1 showed an increase in the percentage of wild type plants, suggesting increased silencing of the *NPTII* transgene. Line 18 showed an increase in the percentage of kanamycin resistant plants, reaching the expected transgenic:wild type ratio of 3:1. This suggested reduced silencing of the *NPTII* transgene.

Five week old lettuce plants of Line 18 were examined but the plants were not phenotypically different from control lettuce (Figure 5.1c). RNA was collected from more plants of Line 18 to investigate whether *TET3c* expression was stable. *TET3c* expression was not seen in these plants (Figure 5.1d), suggesting that expression of *TET3c* in lettuce is no more stable than the expression seen for other transgenes.

These experiments demonstrated that *TET3c* did not improve the instability in transgene expression normally observed in lettuce.





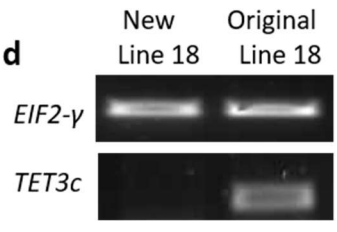


Figure 5.1. TET3c expression in lettuce was rare and unstable

(a) Expression levels of TET3c were analysed using semi-quantitative RT-PCR on cDNA pools from each line of lettuce leaves. TET3c was not expressed in all but three lines of TET3c transgenic lettuce. cDNA levels were confirmed with the housekeeping gene EIF2-γ at 25 cycles, while TET3c expression was checked at 30 cycles.

*(b) Lettuce seeds (n>19) of control (D1::luciferase) and TET3c lines were sown on kanamycin-containing media. The number of plants which germinated were recorded as kanamycin resistant, while seeds which did not germinate were assumed to be kanamycin sensitive. The percentage of plants which were kanamycin resistant is displayed in red. Line 18 showed a significant increase in the percentage of kanamycin resistant plants, indicating reduced silencing of the NPTII transgene. The significance of a change from control Line A is indicated by asterisks: ns= not significant, **=P<0.01, *** =P<0.005 calculated using the chi-squared test.*

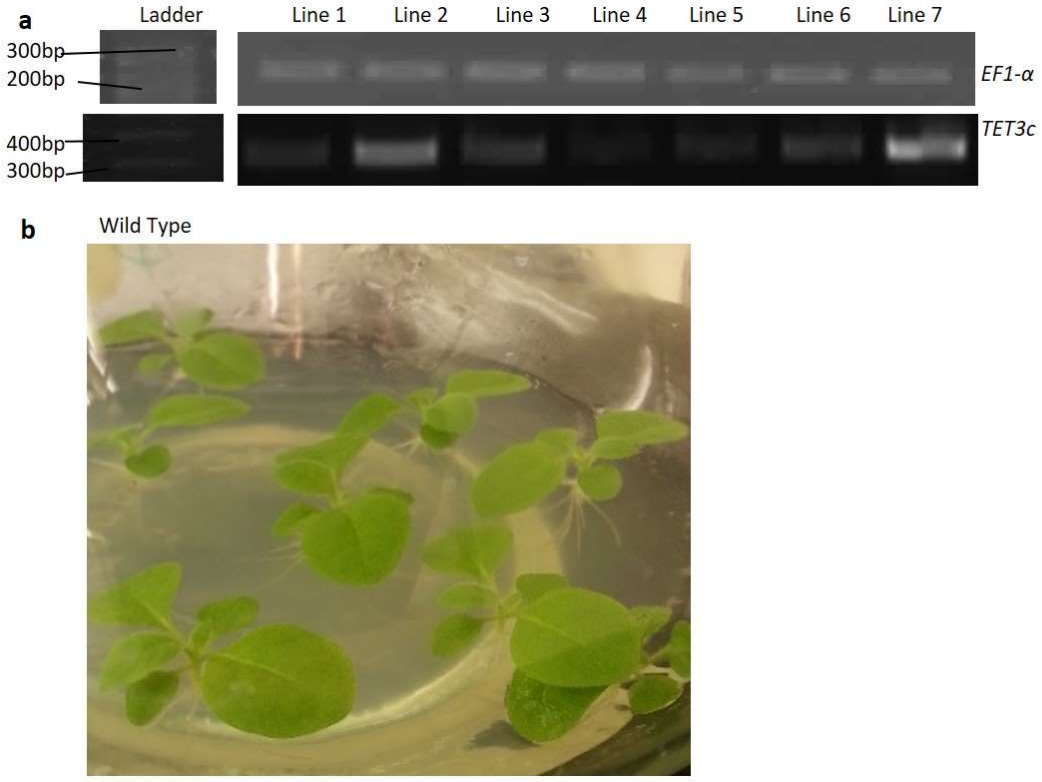
(c) Lettuce from Line 18 where TET3c was expressed showed no obvious morphological difference at the age of five weeks.

(d) Expression of TET3c in Line 18 was analysed using semi-quantitative RT-PCR in new plants of Line 18 but was no longer expressed. cDNA levels were confirmed with the housekeeping gene EIF2-γ at 25 cycles, while TET3c expression was checked at 30 cycles.

5.2.2. Expression of *TET3c* did not have a phenotypic effect on tobacco

The *35S::TET3c* construct was transformed into tobacco by Dr Iris Heidmann and Suzan Out at Enza Zaden. Seven independent transformant lines were produced. T1 seeds were grown on kanamycin selection. DNA and RNA was extracted from plants of four weeks old. Plants were genotyped and RNA from plants which contained the transgene was combined to generate cDNA. Semi-quantitative RT-PCR was done to analyse expression levels of *TET3c* in the tobacco lines (Figure 5.2a). Expression levels were extremely low in all but two lines, which did not appear phenotypically different (Figure 5.2b).

As discussed in Chapter Three, expression levels of *TET3c* in tomato were much higher than those seen in lettuce or tobacco as they were expressed more strongly than the housekeeping gene used for normalisation (Figure 5.2c) and *TET3c* tomato also displayed phenotypes.



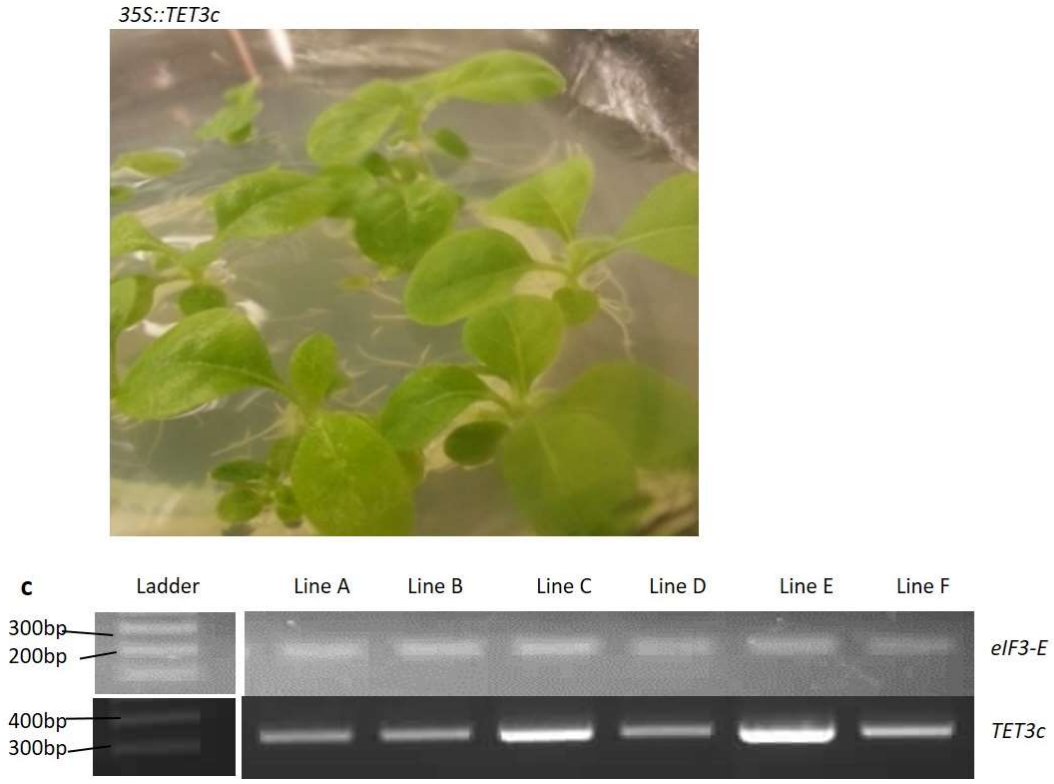


Figure 5.2. 35S::TET3c tobacco expressed TET3c at low levels

(a) Expression of TET3c (330bp) in the seven tobacco lines was checked using semi-quantitative RT-PCR on cDNA pools (n=3). TET3c expression was low in all samples. cDNA levels were normalised using EF1-α (200bp) at 25 cycles and TET3c expression was checked at 30 cycles. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

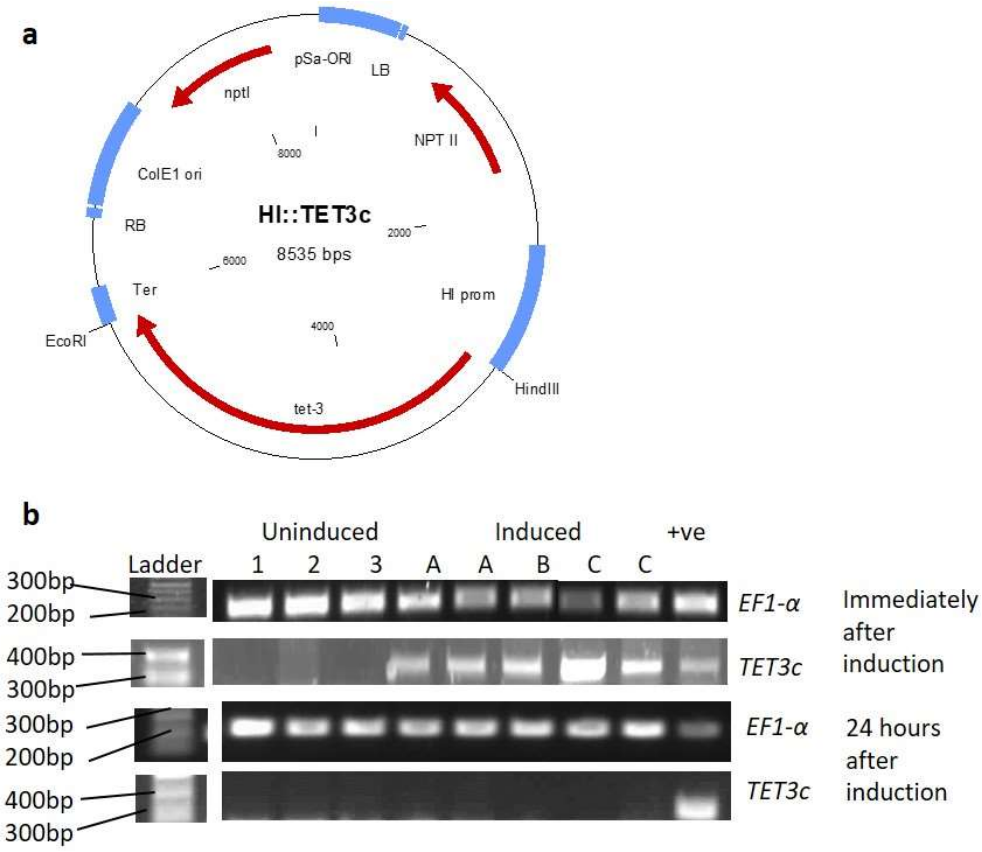
(b) Tobacco expressing TET3c showed no obvious morphological differences.

(c) Tomato expressed TET3c (330bp) more strongly than tobacco, as TET3c was expressed more strongly than the constitutively expressed eukaryotic initiation factor 3E (150bp), used for normalisation of cDNA levels. Tomato containing the TET3c transgene were shown to express it using semi-quantitative RT-PCR at 25 cycles of cDNA pools (n>8). Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

It is possible that *TET3c* was silenced in tobacco due to negative phenotypic effects and therefore a construct containing *TET3c* behind a heat inducible promoter (*HI::TET3c*) was made by Dr Michael Watson and transformed into tobacco by Dr

Iris Heidmann and Suzan Out at Enza Zaden, the Netherlands. This promoter is activated in *Arabidopsis thaliana* at 38°C (Gallois et al., 2002) and in tomato at 40°C (Watson, 2013). Three independent transformant lines containing *HI::TET3c* (Figure 5.3a) were generated.

Plants were grown for six weeks and then induced by incubation at 40°C for ten hours. RNA was collected immediately after incubation to analyse initial *TET3c* expression as well as 24 hours after the induction had ended to investigate whether *TET3c* expression was maintained (Figure 5.3b). *TET3c* expression was activated by the ten-hour induction, but was no longer being expressed after 24 hours (Figure 5.3b). The induction was carried out for 24 hours in order to investigate whether this resulted in *TET3c* being expressed for longer (Figure 5.3c).



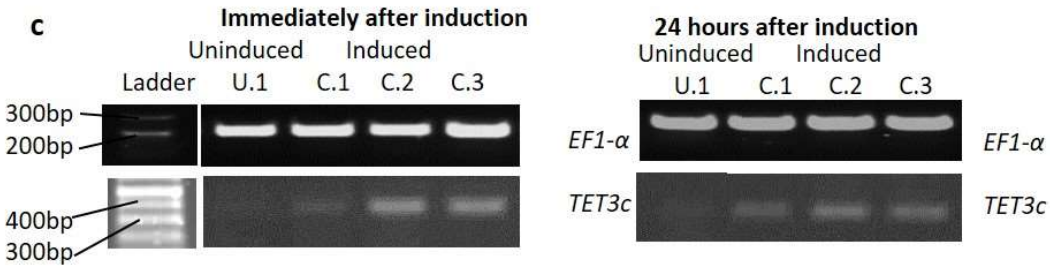


Figure 5.3. TET3c expression could be induced using a heat-sensitive promoter and incubation at 40°C

(a) TET3c was cloned under a heat inducible promoter and transformed into tobacco.

(b) Expression of TET3c (330bp) was checked in the three HI::TET3c tobacco lines using semi-quantitative RT-PCR both immediately after a ten hour induction and 24 hours later. TET3c expression was activated by ten-hour induction but was no longer present 24 hours after induction. cDNA levels were normalised using EF1-α (200bp) at 25 cycles and TET3c expression was checked at 30 cycles. +ve indicates a positive DNA control. Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

(c) Expression of TET3c (330bp) was present after a longer incubation of 24 hours using semi-quantitative RT-PCR and was maintained for at least 24 hours after induction. cDNA from induced tobacco from Line C (C1.1-C.3) was compared to cDNA of uninduced tobacco from Line C (U.1). Lines on the left indicate the size of bands in base pairs (bp) from the 1kb+ ladder (Invitrogen).

Although faint, *TET3c* cDNA was still detectable 24 hours after a longer 24 hour induction had ended. Over the next month, tobacco plants were analysed for any phenotypes such as the reduced leaf size seen in antisense *MET1* tobacco plants, but no change was observed (Figure 5.4a).

Repeated inductions were carried out (24 hour inductions, once a week), in order to maintain the *TET3c* expression. However, the stress of this treatment resulted in bleaching of leaves in both wild type and transgenic tobacco after three weeks

(Figure 5.4b) and the eventual death of the plants. No phenotypic difference between wild type and transgenic tobacco was observed.

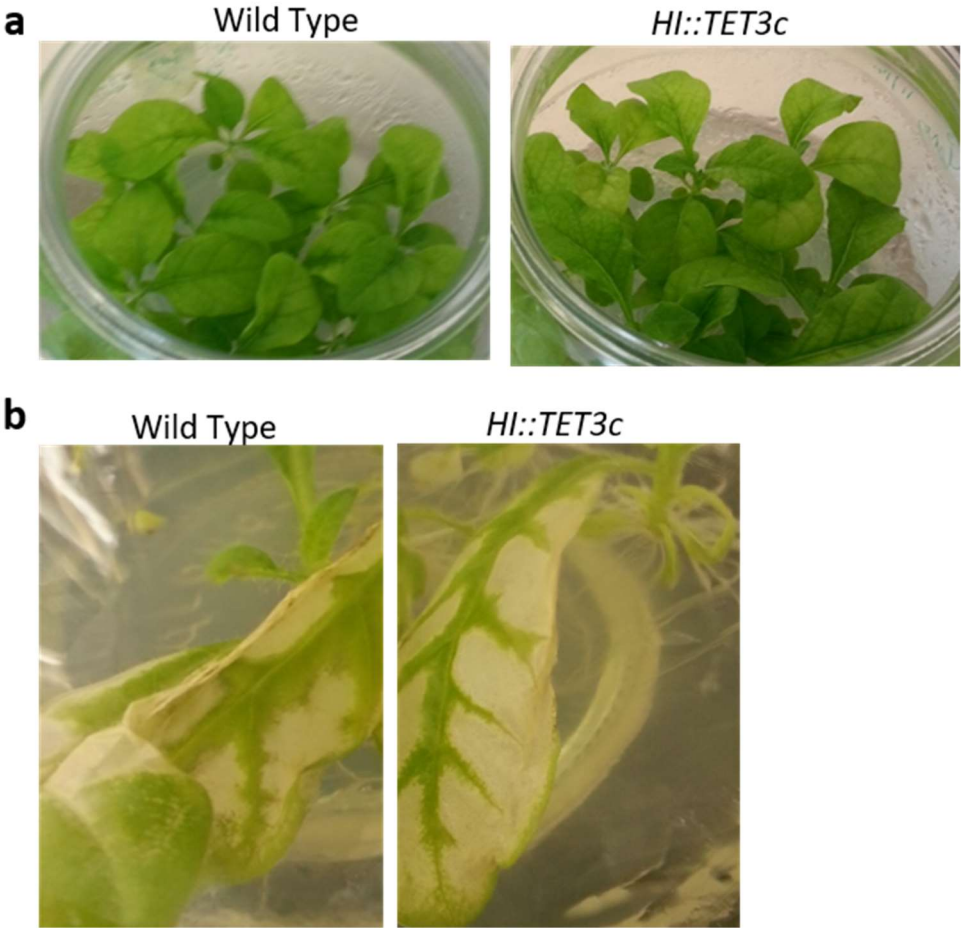


Figure 5.4. TET3c expression did not result in tobacco phenotypes but repeated heat induction was lethal to plants

(a) Induced HI::TET3c tobacco plants showed no phenotypes one month later.

(b) Repeated heat induction caused bleaching of tobacco leaves and eventual plant death.

5.3. Discussion

Expression of *TET3c* in *Arabidopsis thaliana* caused a slight delay in flowering, but no other change in phenotype while the effects of *TET3c* in tomato appear to be more diverse. The results of expressing *TET3c* in lettuce and tobacco were investigated, but no strong phenotypic changes were observed and expression of the *35S::TET3c* construct was low in both species. Low levels of expression in lettuce may be expected due to the fact that transgenes are commonly silenced in this species (McCabe et al., 1999a). In Line 18, the silencing of the transgene appeared to have been reduced while in Line 1 it has increased. The high level of kanamycin-sensitive plants in Line 1 could be a result of one of several possibilities. These include a truncated T-DNA insertion (Gheysen et al., 1990), silencing of the T-DNA due to insertion into a highly repetitive region (Pröls and Meyer, 1992) or a higher level of wild type plants due to the T-DNA inserting into a crucial region and therefore homozygous *TET3c* plants not surviving (Chen et al., 2001).

In Line 18 of the *TET3c* lettuce, 74% of plants germinated on kanamycin-containing media and therefore expressed the *NPTII* transgene, a 20% increase on the control. This suggests that *TET3c* may be able to relieve the silencing of the transgene, although more control lines would need to be investigated to be certain. Despite this, expression of *TET3c* was still not stable within this line, as other T1 plants of the same line did not show any expression of the *TET3c* transgene. Therefore even if *TET3c* can improve the level of transgene silencing seen in lettuce, this effect is both rare and inconsistent, even when a line which shows it has been found. This may be because the *35S::TET3c* construct contains only the catalytic domain of TET3 without any targeting domain, and therefore hypomethylation at the T-DNA may be unlikely to occur. To investigate this, a *TET3c* construct with a binding domain fused to the catalytic domain could be used to target TET3c-mediated demethylation activity to specific regions, including the T-DNA itself.

TET3c expression in *35S::TET3c* tobacco was silenced across multiple lines. This could suggest that the phenotypic effect of *TET3c* expression is strongly deleterious

to tobacco and therefore only transformants where the transgene is silenced survive. Despite this, expression of the *TET3c* construct behind an inducible promoter in tobacco did not result in phenotypic changes. It is possible that sustained expression of *TET3c* would have resulted in phenotypes, while the variable expression produced by the heat induction did not. It is also possible that the deleterious effect of *TET3c* expression only occurs at certain developmental stages. An induction of the *TET3c* construct at different stages of tobacco growth would allow investigation of this hypothesis.

6. General Discussion

The objective of this project was to investigate the effects of transforming the catalytic domain of TET3, a mammalian DNA demethylase, into plant species. Doing so has allowed us to demonstrate that TET3c can cause heritable methylation changes in plants which can lead to changes in gene expression and phenotype. From our data, we can learn more about the characteristics of TET3c-mediated demethylation in plants and about the processes controlled by DNA methylation in tomato.

6.1. TET3c-mediated Methylation Changes in Plants

The ease of transformation, rapid life cycle and wealth of epigenetic information available on *Arabidopsis thaliana* make it an ideal plant for the initial analysis of TET3c and its ability to demethylate DNA in plant species. Methylation changes had previously been observed in the ribosomal DNA using Southern blot (Hollwey et al., 2016; Watson, 2013) and this was confirmed using bisulfite sequencing. When *TET3c* was transformed into tomato, methylation changes were again observed. Unlike in Arabidopsis, in tomato hypomethylation was observed upstream of a protein-coding gene and correlated with the ectopic expression of that gene.

6.1.1. Oxidised derivatives of 5-methylcytosine were produced in *TET3c+* plants

Using oxidative bisulfite sequencing and hydroxymethylcytosine DNA immunoprecipitation (hMeDIP), the presence of 5-hydroxymethylcytosine was demonstrated at regions where methylation changes occur in both Arabidopsis and tomato. TET3c can therefore oxidise 5-methylcytosine in plants as it does in mammals (Figure 6.1a). Previous reports have suggested that the plant DNA glycosylases DME, ROS1 and DML3 can excise 5-hydroxymethylcytosine *in vitro* (Brooks et al., 2014; Jang et al., 2014). The ratio of 5hmC/5mC was slightly raised in triple glycosylase mutants (*ros1-3; dml2-1; dml3-1)* containing *TET3c* (*rdd TET3c+)*, but not in comparison to the variation seen between the two *TET3c* lines. This

suggests either that ROS1 and DML3 do not remove significant amounts of 5-hydroxymethylcytosine or that the production of 5hmC is regulated in response to the amount being removed. This would be different to what is observed in mammals, where mutation of TDG (which is capable of excising 5-formylcytosine) results in an increase in 5-formylcytosine levels (Iurlaro et al., 2016).

5-formylcytosine was observed only in the *rdd TET3c+* Arabidopsis, further demonstrating that TET3c can oxidise 5-methylcytosine in plants, and suggesting that at least one of these three enzymes is capable of excising 5-formylcytosine *in vivo* (Figure 6.1a). Despite contributing to TET3c-mediated demethylation by removing 5fC, none of these three DNA glycosylases were essential for TET3c-mediated demethylation in *TET3c* Arabidopsis at the analysed region of ribosomal DNA.

6.1.2. Methylation changes in *TET3c* plants were heritable

Hypomethylation observed in *TET3c* Arabidopsis remained when the transgene was no longer present. Similarly, expression of *TET3c* was not required for the ectopic expression of *CEN1.1* in *TET3c* tomato or for the resulting phenotypes. The hypomethylation upstream of the *CEN1.1* promoter was observed in plants which no longer contained the *TET3c* transgene. Methylation changes caused by TET3c were therefore heritable.

In contrast to the heritable decrease in methylation seen in Line A *TET3c* Arabidopsis, Line B *TET3c* Arabidopsis showed an increase in methylation. The changes in methylation caused by TET3c at the ribosomal DNA of Arabidopsis were stable after the initial change despite the continued presence of *TET3c*. This indicates that TET3c-mediated demethylation in plants occurs at a particular stage during transformation or development, at which point either an increase or a decrease may occur.

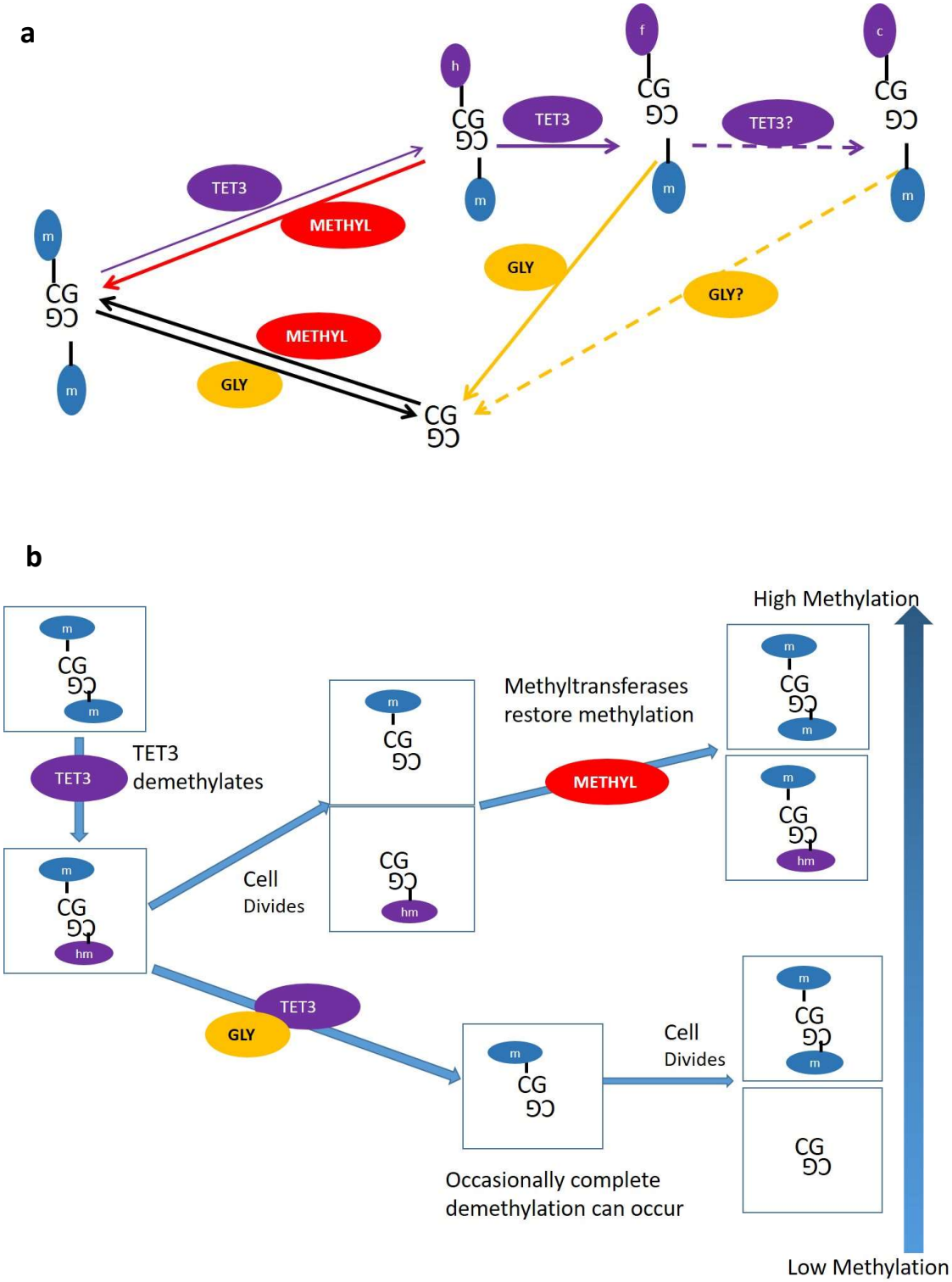

a
h
f
c
TET3
TET3?
m
TET3
METHYL
GLY
GLY?
METHYL
GLY
b
High Methylation
TET3
demethylates
Methyltransferases
restore methylation
METHYL
Cell
Divides
hm
TET3
GLY
Cell
Divides
Occasionally complete
demethylation can occur
Low Methylation

Figure 6.1. TET3 is able to oxidise 5-methylcytosine in plants to 5-hydroxymethylcytosine, which can result in either an increase or a decrease in DNA methylation

(a) A diagram explaining what we have learned about the functionality of TET3 in plants, and the involvement of other plant enzymes. Black lines indicate what was already known, solid coloured lines indicated what we have learned from this project and dotted lines represent unanswered questions. We have learned that TET3 can oxidise 5mC to 5hmC and 5fC, and that according to the rdd TET3c+ mutants, a plant glycosylase is capable of excising 5fC. The lack of complete demethylation indicates that a methyltransferase is still able to replace methylation at these sites despite the action of TET3. METHYL indicates a methyltransferase and GLY indicates a DNA glycosylase

(b) This schematic suggests how increases and decreases may occur stochastically. As methylation changes occurred only at a certain focus point, after which they were fixed, differences in methylation at these points would therefore be preserved. If plant methyltransferases are able to read 5hmC and replace it with 5mC, this would explain the increases in DNA methylation observed in Line B TET3c-. On the other hand, the action of either TET3 or DNA glycosylases may remove 5hmC before this occurs, resulting in a decrease in DNA methylation.

6.1.3. Methylation changes could occur in all contexts

In Arabidopsis, methylation changes at the rDNA locus occur specifically in CG and CHG contexts. CHH methylation levels, which were already low in this region (3% methylation) remain low in Line A *TET3c+* (3%) and Line B *TET3c+* (5%). Interestingly, Line B *TET3c-* Arabidopsis show an increase in CHH methylation (12%), suggesting that TET3c does have an effect at CHH sites, and that perhaps CHH methylation at this locus in Arabidopsis was too low to be able to evaluate this.

It might have been expected that TET3c would preferentially demethylate DNA in the CG context, as this makes up the majority of the methylation in mammals. The data from tomato suggests that this is not the case and demonstrates that TET3c can demethylate CHH sites which are highly methylated. Demethylation in *TET3c*

tomato upstream of the *CEN1.1* gene occurred specifically in the CHH context. In comparison to Arabidopsis, where CHH sites are either unmethylated or methylated at ~10% (Cokus et al., 2008), tomato CHH methylation levels are much higher, reaching 100% at some sites (Zhong et al., 2013). Therefore TET3c can demethylate plant DNA in all cytosine contexts, provided that methylation levels are sufficiently high.

6.2. TET3c-mediated Phenotypic Changes in Plants

6.2.1. *TET3c* Arabidopsis and tomato were delayed in flowering, in addition to other phenotypes observed in *TET3c* tomato

In tomato, multiple phenotypic changes including a delayed floral transition, vegetative growth on inflorescences, increased distance between leaves and a terminal flower phenotype were observed. Some of these phenotypes such as the delayed floral transition were a result of the ectopic expression of *CEN1.1,* but others are likely due to changes in the expression of alternative genes. A delay in flowering was also observed in Arabidopsis. The expression of *FWA,* a known cause of delayed flowering in Arabidopsis, was not altered in *TET3c* Arabidopsis. *TFL1,* the Arabidopsis homolog of *CEN1.1* is, interestingly, also densely methylated upstream of the transcriptional start site. However, no expression changes in *TFL1* were seen in *TET3c* Arabidopsis. The delay in flowering in *TET3c* transformants of the two species may, therefore, occur through a different mechanism.

One particularly obvious phenotype in the *TET3c* tomato was the high frequency of plants with the blind phenotype. This was demonstrated not to require the *TET3c* transgene but instead to result from a combination of transformation and tetraploidy in this tomato variety, although the presence of *TET3c* did increase the incidence of the blind phenotype in tetraploids.

6.2.2. No phenotypic changes were observed in tobacco and lettuce

Neither tobacco nor lettuce showed any obvious or consistent phenotypic changes in response to *TET3c* transformation, although developmental stages such as flowering were not examined in detail. Phenotypic changes at the cellular level, which occur in methylation mutants of Arabidopsis (Vassileva et al., 2016), could have occurred and gone unnoticed. Both also lacked strong and persistent expression of the *TET3c* transgene, making it likely that there were no phenotypic changes.

6.3. Outlook and Open Questions

6.3.1. Exploring the interactions of TET3c with DNA glycosylases

TET3c can oxidise 5-methylcytosine in plants and demethylate DNA. Although we demonstrated that 5-hydroxymethylcytosine was produced, the stability of this epigenetic mark in plants is unclear. 5hmC has been described as a stable epigenetic mark in mammals (Bachman et al., 2014), with proteins specifically capable of binding this modification (Spruijt et al., 2013). Similar genome-wide experiments in plants to identify plant proteins capable of binding 5hmC and 5fC have not been carried out, but would be of interest to analyse the full potential effects of TET3c-mediated demethylation in plants. Our results indicated that ROS1, DML2 and DML3 did not excise significant quantities of 5hmC *in vivo* but a comparison of the ratio of 5hmC/5mC in *TET3c+* and *rdd TET3c+* lines at more regions and in more lines is necessary to confirm this.

Our results demonstrated that while at least one of the three plant DNA glycosylases mutated in *rdd* lines could excise 5-formylcytosine *in vivo*, these enzymes were not essential for demethylation at the ribosomal DNA locus. Other loci may require this 5-formylcytosine excision activity. This could be investigated using a whole genome analysis of 5-methylcytosine levels in *TET3c+* and *rdd TET3c+* lines. Methylation levels in *TET3c+* lines would have to be analysed as a function of the methylation levels in wild type and *rdd* lines without the *TET3c* transgene, as

methylation levels are increased at multiple loci in *rdd* mutants (Penterman et al., 2007a). If 5-formylcytosine excision is essential for demethylation at a locus, then methylation levels would be expected to decrease at the locus between wild type and *TET3c+* lines but not between *rdd* and *rdd TET3c+* lines.

The contribution of the final plant DNA glycosylase, DME, to TET3c-mediated methylation changes is unknown. DME is expressed only in the female gametes prior to fertilisation (Choi et al., 2002) and therefore is not contributing to somatic demethylation but the developmental stage at which the initial TET3c-mediated methylation changes occur is also unknown. The 35S promoter is expressed during almost all plant developmental stages, other than early embryogenesis (Sunilkumar et al., 2002). Therefore TET3c-mediated demethylation may be occurring in the ovule and require DME. Tissue-specific versions of the *TET3c* construct could be utilised to investigate at which stage *TET3c* expression is required to result in the heritable methylation changes observed in Line A *TET3c+* and *rdd TET3c+* lines. If expression during gametogenesis is required, then DME may contribute to TET3c-mediated demethylation.

6.3.2. Understanding the stability and inheritance of TET3c-mediated changes

Arabidopsis plants descended from the same transformation event showed the same methylation phenotype, whether that was a decrease as seen in Line A or an increase as seen in Line B. This phenotype was present in all plants of that line and was heritable when the transgene was lost. This highlights the possibility of a focal point at which heritable methylation changes could occur. This is likely to be during the initial transformation rather than later in development; different methylation patterns would be observed between siblings if the focal point occurred later, for example during gametogenesis. The cause of this focal point could be stress, as environmental stress, both biotic and abiotic, is known to cause epigenetic changes (Meyer, 2015). Candidates for this stress include both the stress of T-DNA transfer and the stress of the Agrobacterium infection which facilitates it. These cause

significant changes in the expression of many genes (Jiang et al., 2003) and are likely to be associated with many epigenetic changes. To investigate if either allows heritable TET3c-mediated methylation changes to occur, Line B *TET3c+* Arabidopsis could be used. A comparison of rDNA methylation in Line B *TET3c+* Arabidopsis transformed with a new reporter construct and Line B *TET3c+* Arabidopsis dipped in Agrobacterium which did not contain a plant transformation vector could be carried out. A change in the steady state of methylation from the hypermethylated state to a hypomethylated state in some plants would be seen if this treatment gave an opportunity for heritable methylation changes to occur.

TET3c continues to oxidise 5-methylcytosine to 5-hydroxymethylcytosine and 5-formylcytosine during vegetative growth in Arabidopsis, and therefore the lack of complete demethylation must reflect replacement of methylation by DNA methyltransferases (Figure 6.1a). In addition, in Line B *TET3c+* Arabidopsis, both 5mC and 5hmC were inherited as 5mC when the transgene was no longer present (Figure 6.1b). This could either reflect a methyltransferase reading 5hmC (and replacing it as 5mC), or increased targeting of a methyltransferase to regions with large amounts of 5hmC (Figure 6.1b). The methyltransferases responsible for methylation at this region may determine whether this occurs. Some components of the methylation pathway such as VIM1, one of a family of proteins which can recognise hemimethylated DNA and act in MET1-mediated methylation (Kim et al., 2014), do not efficiently bind sites with 5hmC (Yao et al., 2012). Other components such as SUVH5 (a histone methyltransferase which can facilitate recruitment of CMT3 (Ebbs and Bender, 2006)) are capable of binding 5hmC at a similar affinity to 5mC (Rajakumara et al., 2016). A more complete analysis of the binding of different plant proteins to 5hmC would be useful to investigate this. Crossing the *TET3c* lines with mutants of methyltransferases or proteins which facilitate their recruitment would also allow us to investigate the interaction of TET3c with the endogenous methylation systems of plants. If the enzyme responsible for the replacement of methylation is mutated, an enhanced loss of methylation at any site where TET3c acts would be expected in plants which retain the transgene. If a protein is

responsible for targeting of methyltransferases to 5hmC, then crosses of Line B *TET3c+* plants with mutants of this enzyme would not be expected to show an increase in methylation when the transgene is lost.

6.3.3. Investigating at which loci TET3c-mediated demethylation occurs in plants

TET3c-mediated demethylation in tomato was seen specifically at CHH sites. This does not mean that TET3c can only achieve heritable demethylation in this context in tomato. Methylation changes were examined only at two loci, and were first discovered by identifying phenotypes and gene expression changes. At other loci, methylation changes may occur in all contexts. A whole genome bisulphite analysis of TET3c+ tomato would allow us to investigate whether, in tomato, TET3c does specifically demethylate regions with high CHH methylation. If this is the case, it could reflect that methylation at other contexts is more likely to be restored. Combining this with an expression analysis would allow us to examine whether demethylation of CHH is more likely to achieve gene expression changes than demethylation at other sequence contexts in tomato.

As discussed above, the loci which are affected by TET3c may be altered in a glycosylase mutant. The higher methylation levels seen at some loci in glycosylase mutants could also change which loci are affected, as increased methylation at a region may affect the action of TET3c regardless of the presence of the DNA glycosylases. The analysed rDNA region of Arabidopsis shows either an increase or a decrease in methylation but the factors affecting which occurs are unknown. A global analysis of TET3c+ and TET3c- methylation would determine if regions with higher levels of methylation are more likely to increase in methylation in comparison to similar regions with lower levels of methylation, or vice versa. A global analysis of methylation in TET3c- tomato could also identify if there are regions in tomato which increase in methylation, similarly to in Arabidopsis. Alternatively, whether a region increases or decreases in methylation may be a chance event depending on

the precise levels of TET3, DNA glycosylases and methyltransferases at that region during the focus point at which DNA methylation levels are destabilised (Figure 6.1b).

The methylation changes seen in both Arabidopsis and tomato are present at the same loci between biological replicates and in different transformant lines. This suggests that some loci are more susceptible to TET3c-mediated methylation changes. These loci may have distinct DNA methylation or genomic markers. A whole genome bisulfite analysis could be used to answer these questions. Alternatively, it may suggest that targeting of the TET3c protein is occurring despite the lack of a catalytic domain. This would suggest that proteins which target TET3 in mammals are conserved in plants and can carry out the same function here. Plant homologues of proteins known to interact with TET in mammals such as PRC2 (Polycomb Repressor Complex 2), HDAC2 (histone deacetylase 2) and LIN28A (Lin28 Homolog A)(Neri et al., 2013; Zeng et al., 2016, Zhang et al., 2015) could be identified. Using protein-protein interaction studies such as yeast two-hybrid analysis would allow us to discover whether TET3 can bind plant homologues of these proteins and therefore if they may be interacting with TET3c in plants.

6.3.4. The role of methylation in tomato

While the molecular cause of some of the phenotypes seen in the *TET3c* tomato was established in Chapter Three, others were not, including an increase in the distance between leaves and an increase in the number of plants with determinate growth (the terminal flower phenotype). In addition to analysis of constitutive *SP9D* expression in tomato, further work investigating the changes in gene expression in *TET3c* tomato may be required to investigate this.

The blind phenotype does not require *TET3c* but is epigenetic in nature. A comprehensive analysis of gene expression changes in the blind tomato produced using different methods could be used to identify which genes cause the blind

phenotype. Analysis of gene expression changes in the shoot apical meristem, rather than in leaf, may be required to identify these genes. The blind phenotype occurred only in the initial SAM (shoot apical meristem) of the tomato located between the cotyledons, and therefore an analysis of the differences in gene expression between different meristem stages could enable identification of the potential causes of this phenotype.

In addition, thorough analysis of all tissues of *TET3c* tomato at different developmental stages and under different conditions is likely to reveal additional phenotypes. This would give a greater understanding of the processes in tomato in which DNA methylation is involved.

6.3.5. The use of TET3c as a tool for demethylation

In this thesis, the catalytic domain of the mammalian demethylase TET3 has been used to cause methylation changes in Arabidopsis and tomato. This has demonstrated its utility as a tool for investigating DNA methylation. Transformation of *TET3c* into other crop species in which methylation is known to be important such as rice and maize would allow us to learn more about methylation in these species, in addition to being able to draw wider conclusions about how TET3c acts in plants. Less well studied plant species have also displayed phenotypes which depend on DNA methylation changes (Santo et al., 2017), illustrating the potential broad use of this system. In human cells, TET1 has been used in association with both transcription activator-like effector (TALE) repeat arrays (Maeder et al., 2013) and bacterial clustered regularly interspaced short palindromic repeats (CRISPRs)(Xu et al., 2016) to cause targeted changes in DNA methylation which result in changes in gene expression in these cells. CRISPR dCas9-TET1 fusions have also been demonstrated to function *in vivo* in mice embryos to increase gene expression (Morita et al., 2016). The fact that, in plants, TET3c can cause heritable demethylation opens up plants to targeted epigenome editing using these same tools.

7. Materials and Methods

7.1. Materials

7.1.1. Plant material

The *Arabidopsis thaliana ros1-3; dml2-1; dml3-1* triple mutant was provided by Robert Fischer, University of California, Berkeley.

All tomato, lettuce and tobacco material was provided by Enza Zaden, the Netherlands.

7.1.2. Bacterial strains

Plasmid cloning was carried out using *Escherichia coli* DH5α (New England Biolabs). Plant transformation was carried out using *Agrobacterium tumefaciens* GV3101:PMP90 (Hellens et al., 2000a).

7.1.3. Donated plasmids and DNA sequences

The pGreen collection (Hellens et al., 2000a, 2000b) was provided by Mark Smedley (John Innes Centre, Norwich Research Park, Colney, Norwich, NR4 7UH). pGreen 0029 containing the *35S-nos* cassette and the *35S::MET1ir* construct, containing the 35S promoter, *MET1* sense and antisense regions separated by an intron and the nos terminator, were provided by Dr Michael Watson (P. Meyer lab, Centre for Plant Sciences, University of Leeds, LS2 9JT). The *TET3* cDNA sequence was provided by Professor Richard Meehan (MRC Human Genetics Unit, MRC IGMM, University of Edinburgh Western General Hospital, Crewe Road, Edinburgh, EH4 2XU). The *TET1* cDNA was purchased from Cambridge Bioscience Limited. The glucocorticoid tag was amplified from a plasmid provided by Dr Barry Causier (B. Davies lab, Centre for Plant Sciences, University of Leeds, LS2 9JT).

7.1.4. Primer sequences

7.1.4.A. Primer sequences used for plasmid construction

Construct and Primer Name	Forward Primer	Reverse Primer
7.2.1.G.1: *35S::TET1*		
TET1ATGinsert	AAGGGCCGTCAAGGCCCACCATGGCGTCTCGATCCC	GTTCACTTTTAATGGGATCCAGGT
TET1SfiI	AGCTTTCTAAAGGCCGTCAAGGCCATGCGTGCTAGGCCTCATGGGCCAAACTCG	AATTCGAGTTTGGCCCATGAGGCCTAGCACGCATGGCCTTGACGGCCTTTAGAA
7.2.1.G.2: *35S::CEN1.1*		
BF78genclon	GGGAAGCTTGGCACGTTGATTGGTTTTTCG	GGGAATTCACAAGCAAATGAGTAGGACAAACA
7.2.1.G.3: *35S::SP9DcDNA*		
SP9DcDclon	GGGAAGCTTGGATGGCAAGAAGTTTAGAGCC	GGGAATTCCGCAGCGGTTTCACGTTG
7.2.1.G.4: *35S::CLV3*		
SlCLV3clon	GGGAAGCTTACTTGCTTTTTCATTTGCTTCTTGA	GGGAATTCTTGCATGATTATGGAGTTATGGAGT
7.2.1.G.5: *35S::CLE9*		
SlCLE9clon	GGGAAGCTTATCCCTGTTTGATGTTAATGGCG	GGGAATTCGCTATGGCATTATATTCCTGGGG
7.2.1.G.6: *35S::WUS*		
SlWUSclon	GGGAAGCTTATGGAACATCAACACAACATAG	GGGAATTCTTAGGGGAAAGAGTTGAGAGTA
7.2.1.G.7: *35S::CEN1.1ir*		
BF78RNAiex4cl	GG**ATCGAT**GACGCCAAAGCCGGTAATTG	GG**TCTAGA**GCTTGAGAATGACACTTAGAATCTC
BF78RNAiex4cl	GG**GGTACC**GACGCCAAAGCCGGTAATTG	GG**CTCGAG**GCTTGAGAATGACACTTAGAATCTC

7.2.1.G.9: *35S::WUS:GR*		
GRtagclon	GGGGATCCGAAGCTCGAAAAACAAAGAAAAAAATC	GGGGATCCTCATTTTTGATGAAACAGAAGC
SlWUSclonDEX	GGGAAGCTTATGGAACATCAACACAACATAG	GGGGATCCTATACTATATGTATAGAATAAAGGAAAATTTG

7.1.4.B. Primer sequences used for genotyping

Target	Forward Primer	Reverse Primer
35S::CEN1.1	GAGACGCCAAAGCCGGTAAT	CAACACATGAGCGAAACCCTAT
35S::SP9DcDNA	CTATGAGACCCCAAGGCCAAA	CAACACATGAGCGAAACCCTAT

7.1.4.C. Primer sequences used for expression analysis

Target	Forward Primer	Reverse Primer
Arabidopsis thaliana EF1a	GCGTGTCATTGAGAGGTTCG	GTCAAGAGCCTCAAGGAGAG
TET3c transgene	CTCTACGGGAAGAAGCGCA	CTCAACACATGAGCGAAACCC
TET1 transgene	CTGCCAACCTTAGGGAGTA	GGCAAGGTCTTGAGGAGAGG
Solanum lycopersicum eIF3-E	GAGCGATGGATGGTGAATCT	TTGTACGTGCGTCCAGAAAG
S. lycopersicum WUS	CCAGCAACTTACCCTTTTCTTG	TAAAGCAGAGTTACCCCTTTGG
S. lycopersicum CLV3	AAAGGAAGTTGCTCCTGTGAA	CCTCTTAGCTCCCAATCAGC
S. lycopersicum CLE9	CAATGCAAGCACAATCCTCT	CCTGCATCCTGGCTTATTCT

S. lycopersicum CLE3	CTGCTGAGATTTTAGTAAAGCCTG	GAATGCCTTTCTGTTTCTATTATCC
S. lycopersicum CEN1.1	GACCCTGATGCTCCAAGTCC	TGGCTGCAGTTTCTCTCTGG
S. lycopersicum SFT	TTTGTTTATCGTCAACCATC	CTTGGCTGGTTAATAACTTGG
S. lycopersicum SP	CGACAAATTAAAAGCA	GATGATATTACATTACAT
S. lycopersicum SP9D	GGTGAGCTATGAGACCCCAAG	CAGCGGTTTCACGTTGTGC
S. lycopersicum SP5G	CTTGAGGCCTTCACAAGTTGTC	GAGTTCGAGAACGGATCGC
Nicotiana tabacum EF1-a	CTCTCAGGCTCCCACTTCAG	AAGAGCTTCGTGGTGCATCT
Lactuca sativa EIF2-γ	GGTGCTTCCAAGACCTACCC	TGGTAGTCAGTGCGGTTGAC

7.1.4.D. Primer sequences used for bisulfite sequencing

Target	Forward Primer	Reverse Primer
Arabidopsis thaliana rDNA	GGGGAGGTAGTGAYAATAAATAA	CACTCTAATTTCTTCAAARTAACA
-2800 to -2528 from *S. lycopersicum WUS*	GTTAAATGAAGGGTATATGTGAGYYATTTTG	AAAATTTTRTCCCRCCACA
-1837 to -1651 from *S. lycopersicum CEN1.1*	GTGAGGTGGGGTGTTAAAGAATGA	CACCRATRTAACACTCCACCT

7.2. Methods

7.2.1. DNA analysis and cloning techniques

7.2.1.A. Isolation of genomic DNA from plants

DNA extraction was done according to the method from (Vejlupkova and Fowler, 2003) with some modifications. 560µL of DNA Extraction Buffer (200mM NaCl; 200mM Tris HCl pH 8; 0.07M EDTA pH 8; 18mM NaHSO3) and 180µL of 5% sarcosyl were added to ground tissue. Two phenol:chloroform:IAA (24:24:1) extractions were performed. Nucleic acids were precipitated with 325µL isopropanol and 325µL high salt buffer (0.8M sodium citrate; 1.2M NaCl). The pellet was washed with 70% ethanol and resuspended in dH2O + RNase (20mg/L).

7.2.1.B. Isolation of plasmid DNA from *Escherichia coli*

Mini-prep isolation of plasmid DNA was carried out using the alkaline lysis method described in (Sambrook et al., 1989) with some modifications. Individual colonies were grown overnight at 37°C in 3mL of LB broth (10g/L bactotryptone; 5g/L bacto-yeast extract; 10g/L NaCl) with the appropriate antibiotics. 1mL of this culture was pelleted, and resuspended in 100µL of Solution I (50mM glucose; 25mM Tris HCl pH 8; 10mM EDTA pH 8). 200µL of fresh Solution II (0.2M NaOH; 1% SDS) and 150µL of Solution III (3M CH_3COOK adjusted to pH 8 with glacial CH_3COOH) were added prior to incubation on ice for 10-30 minutes. Cell debris was removed by centrifugation and plasmid DNA was precipitated with 450µL of isopropanol. The pellet was washed with 70% ethanol and resuspended in 50µL of dH2O + RNase (20mg/L).

7.2.1.C. Isolation of plasmid DNA from *Agrobacterium tumefaciens*

Miniprep of *Agrobacterium tumefaciens* cells was carried out using the method from (Wang, 2006) with some modifications. Individual colonies were grown overnight at 28°C in 3mL of LB broth (10g/L bactotryptone; 5g/L bacto-yeast extract; 10g/L NaCl) with the appropriate antibiotics. 1mL of this culture was pelleted, and resuspended

in 100µL of Solution I (50mM glucose; 25mM Tris HCl pH 8; 10mM EDTA pH 8, 4mg/mL lysozyme). 200µL of Solution II (0.2M NaOH, 1% SDS) and 150µL Solution III (3M CH_3COOK pH 8) were added. Phenol:chloroform extraction was carried out, and DNA was precipitated using isopropanol. The pellet was washed with 70% ethanol and resuspended in 50µL of dH2O + RNase (20mg/L).

7.2.1.D. Preparation and Transformation of Chemically Competent *Escherichia coli* Cells

500µL of fresh stationary culture was transferred to 50mL of sterile LB broth (10g/L bactotryptone; 5g/L bacto-yeast extract; 10g/L NaCl) and incubated at 37°C with agitation at 200rpm until an OD_{600} of 0.3-0.4 was reached. The culture was chilled, pelleted and resuspended in TSS buffer (100mM MgCl2; 10% PEG; 5% DMSO; LB broth pH 6.5) to an OD_{600} of 1.0 per 100µL. The resuspension was dispensed in 100µL aliquots, snap-frozen in liquid nitrogen and stored at -80°C. 10-50ng of plasmid DNA were added to 100µL of freshly thawed chemically competent *E. coli* cells and transformation by heat shock was performed.

7.2.1.E. Preparation and Transformation of Electrocompetent *Agrobacterium tumefaciens* cells

Preparation and transformation of electrocompetent *Agrobacterium* cells was carried out according to (Mersereau et al., 1990; Shen and Forde, 1989). 500mL of sterile LB broth (10g/L bactotryptone; 5g/L bacto-yeast extract; 10g/L NaCl) was inoculated with 5mL of a fresh saturated culture of *Agrobacterium* and incubated at 28°C until OD_{600} 0.5-0.8 was reached. Cells were pelleted three times, and resuspended in ice-cold sterile H_2O. Cells were then pelleted for a fourth time and resuspended in 2.5mL of 10% ice-cold sterile glycerol. This was aliquot in 50µL volumes. DNA for transformation was added to aliquots and transformed by electroporation at 1.8V.

7.2.1.F. Polymerase chain reaction (PCR)

PCR for genotyping and for expression analysis was carried out using MyTaq polymerase (Bioline) according to the manufacturer's instructions. PCR for cloning was carried out using Herculase II Fusion DNA polymerase (Agilent) according to the manufacturer's instructions.

7.2.1.G. Agarose gel electrophoresis of DNA

DNA was separated using agarose gels of 0.7-3% concentration, depending on the desired DNA fragment size. Ethidium bromide was added to the liquid agarose gel at a concentration of 1µg/ml prior to pouring. Electrophoresis was carried out in 1 x TAE buffer and the DNA was visualised on a UV trans-illuminator. DNA size was determined using a 1kb+ DNA ladder (Invitrogen).

7.2.1.H. Ploidy analysis using flow cytometry

Ploidy of tomato samples was analysed using the Cell Lab Quanta SC (Beckman Coulter). Samples were chopped with a razor blade in LB01 buffer (15mM Tris HCl, 2mM EDTA, 0.5mM spermine.4HCl, 80mM KCl, 20mM NaCl, 0.1% Triton X-100) to release nuclei. DAPI (4'6-diamidino-2-phenylindole, 3µM) was added to the lysis buffer to stain nuclei. Samples were loaded onto the Cell Lab Quanta SC flow cytometer (Beckman Coulter) in a 96 well plate format using a multi-platform loader.

7.2.1.I. Construction of plasmids

Desired fragments were amplified using PCR and purified from an agarose gel. The fragment was excised from the gel under long wave UV light. DNA was purified from the fragment using the Q-spin gel extraction/PCR purification kit (Geneflow#0030) following the manufacturer's instructions. DNA was eluted in dH2O with a final volume of 10-50µL. Plasmids were digested with the appropriate restriction enzymes according to the manufacturer’s instructions (New England Biolabs). Digested

fragments were dephosphorylated in a 50µL volume using 1µL of calf intestinal alkaline phosphatase (Promega), 1xCIAP reaction buffer and up to 300ng of DNA. The reaction was incubated at 37°C for 60-90 minutes before purification with the Q-spin purification kit. DNA ligation was carried out in a 10µL volume using 0.5µL DNA ligase (Promega) following the manufacturer's instructions. Vector:insert ratios of 1:3, 1:7 and 3:1 were used and the ligation reaction was incubated overnight at 4°C.

7.2.1.I.1. 35S::TET1

A GCG codon was added after the ATG start codon of *TET1* to encourage translation in plants using PCR. A primer containing the ATG start codon and a GCG codon as a mismatch was used to amplify the first kilobase of the *TET1* gene, and the altered version was inserted into TET1 pENTR using *BamHI* and *SfiI*. An *SfiI* restriction site was then inserted into the multiple cloning site of the pGreen 0029 35S vector using two complementary oligonucleotides. These contained an *SfiI* restriction site and were designed to anneal in such a way as to leave sticky ends which could anneal into a digested *HindIII* site and *EcoRI* site. pGreen 0029 35S contains a *HindIII* site and an *EcoRI* site in the multiple cloning site between the promoter and the *nos* terminator. The pGreen 0029 35S vector was digested with *HindIII* and *EcoRI* and the *SfiI* site was inserted in this manner. *TET1* pENTR was then digested with *SfiI* and transferred into pGreen 0029 35S.

7.2.1.I.2. 35S::CEN1.1

The *CEN1.1* gene of *Solanum lycopersicum* (1066bp) was amplified from tomato DNA using primers which added a *HindIII* site at the 5' end and an *EcoRI* site to the 3' end. pGreen 0029 35S contains a *HindIII* site and an *EcoRI* site in the multiple cloning site between the promoter and the *nos* terminator. The *CEN1.1* amplicon was transferred into this vector using these enzymes.

7.2.1.I.3. 35S::SP9DcDNA

The cDNA of the *SP9D* gene of *S. lycopersicum* (509bp) was amplified from tomato cDNA using primers which added a *HindIII* site at the 5' end and an *EcoRI* site to the 3' end. The *SP9D* cDNA amplicon was transferred into pGreen 0029 35S using these enzymes.

7.2.1.I.4. 35S::CLV3

The *CLV3* gene of *S. lycopersicum* (599bp) was amplified from tomato DNA using primers which added a *HindIII* site at the 5' end and an *EcoRI* site to the 3' end. The *CLV3* amplicon was transferred into pGreen 0029 35S using these enzymes.

7.2.1.I.5. 35S::CLE9

The *CLE9* gene of *S. lycopersicum* (430bp) was amplified from tomato DNA using primers which added a *HindIII* site at the 5' end and an *EcoRI* site to the 3' end. The *CLE9* amplicon was transferred into pGreen 0029 35S using these enzymes.

7.2.1.I.6. 35S::WUS

The *WUS* gene of *S. lycopersicum* (1239bp) was amplified from tomato DNA using primers which added a *HindIII* site at the 5' end and an *EcoRI* site to the 3' end. The *WUS* amplicon was transferred into pGreen 0029 35S using these enzymes.

7.2.1.I.7. *35S::CEN1.1* inverted repeat

A 313bp region of the *CEN1.1* gene of *S. lycopersicum* was amplified from tomato DNA using two primer sets, one which added a *ClaI* site at the 5' end and an *XbaI* site to the 3' end and another set which added *KpnI* and *XhoI* sites. The *35S::MET1ir* plasmid was digested to remove *MET1* regions using *ClaI/XbaI* and *KpnI/XhoI* and the *CEN1.1* sense and antisense regions were transferred in their place. The inverted repeat

amplicon was cut out with *NotI* and transferred into pGreen 0029 35S using this enzyme.

7.2.1.I.8. 35S::WUS:GR

The glucocorticoid tag was amplified from a plasmid provided by Dr Barry Causier (B. Davies lab, Centre for Plant Sciences, University of Leeds, LS2 9JT), using primers which added a BamHI site to either end. The WUS gene of S. lycopersicum was amplified without the termination codon using primers which added a HindIII site to the 5' end and a BamHI site to the 3' end. The WUS amplicon was transformed into pGreen 0029 35S using these enzymes. pGreen 0029 35S::WUS was then cut with BamHI and the glucocorticoid tag was transferred to the vector.

7.2.1.J. DNA sequencing

The DNA sequence of plasmids was confirmed using sequencing by Beckman Genomics using the appropriate primer.

7.2.1.K. Bisulfite sequencing

Bisulfite treatment was carried out on denatured DNA in dH_2O using the EZ DNA Methylation-Lightning Kit (Zymo Research) according to the manufacturer's instructions. DNA regions of interest were amplified from bisulfite-treated DNA using PCR with MyTaq DNA Polymerase (Bioline).

Amplicons were A-tailed by MyTaq DNA polymerase (Bioline) and then ligated into the pGEM T-easy vector system (Promega) following the manufacturer's instructions. Clones containing the fragment of interest were identified and sequenced by Beckman Genomics using the universal SP6 primer. Sequence reads were aligned using the BioEdit program (T.Hall) and identification of methylated cytosines was carried out using the online CYMATE tool (Hetzl et al., 2007). Data was analysed using the program SequenceFileConverter (J. Royle).

7.2.1.K.1. Oxidative Bisulfite Sequencing

Oxidation of DNA was carried out as described in (Booth et al., 2012). Genomic DNA was denatured by incubation at 37°C for 30 minutes in 0.05M NaOH followed by snap cooling on ice. The sample was oxidised by addition of 1µL of KRuO4, followed by incubation on ice for 1 hour, with occasional vortexing. The reaction was purified using an Illustra MicroSpin G-50 column (GE Healthcare). This was followed by bisulfite treatment.

7.2.1.K.2. Reduced Bisulfite Sequencing

Reduction of DNA for reduced bisulfite sequencing was carried out as described in (Booth et al., 2014). 5µL of an aqueous 1M $NaBH_4$ solution was added to 15µL of DNA in water. The reaction was incubated in the dark at room temperature for 1 hour, with frequent intervals to vortex the reaction and remove built up gas. CH3COONa (750mM, pH 5) was added to quench the reaction. Once the reaction was quenched, bisulfite treatment was carried out.

7.2.1.L. Isolation of SAM DNA for bisulfite sequencing

FFPE blocks of tomato shoot apices were made according to (Vitha et al., 2000). Three week old tomato shoot apices were fixed in formalin for one hour before progressive dehydration in ethanol, and infiltration with wax. Samples were sectioned using a rotary microtome (Microm) and subsequently dewaxed and rehydrated using decreasing ethanol concentrations. Sections were stained using 0.05% toluidine blue and meristem cells were isolated using a dissection microscope (Leica). DNA from meristem cells was recovered using overnight digestion with Proteinase K followed by phenol-chloroform purification.

7.2.1.M. Hydroxymethylcytosine DNA immunoprecipitation

3µg of starting DNA was fragmented to a size of 200-400bp using a sonicator (Diogenode Bioruptor) and an aliquot was run on a gel to confirm. Samples were immunoprecipitated using the hMeDIP Kit (Active Motif). The antibody was digested with Proteinase K to release the DNA, which was then purified using the Chromatin IP DNA Purification Kit (Active Motif). DNA was subsequently analysed using quantitative RT-PCR. Enrichment of hmC-containing DNA was confirmed using the provided control DNA and primers.

7.2.2. RNA analysis

7.2.2.A. Isolation of RNA from plants

RNA was isolated using the method described in (Stam et al., 2000). 750µL of RNA Extraction Buffer (100mM Tris HCl pH8; 100mM NaCl; 20mM EDTA pH 8; 1% sarcosyl) was added to ground tissue. Two phenol:chloroform:IAA (24:24:1) extractions were performed. Nucleic acids were precipitated with 325µL isopropanol and 325µL high salt buffer (0.8M sodium citrate; 1.2M NaCl). An overnight precipitation with 4M lithium chloride was performed, followed by a final precipitation using 40µL 3M NaOAc and 1mL 99% ethanol. DNA was removed using the TURBO DNase kit (Ambion applied biosystems) following the manufacturer's instructions.

7.2.2.B. cDNA synthesis

cDNA was synthesised using SuperscriptII Reverse Transcriptase (Invitrogen), following the manufacturer's instructions and using random primers.

7.2.2.C. Quantitative PCR

Quantitative RT-PCR analysis was performed with the Biomark HD system (Fluidigm) using SsoFast Eva Green Supermix (Bio-Rad) on a 96.96 dynamic array chip according

to the manufacturer's instructions. Data was analysed using the manufacturer's software (Fluidigm RT PCR Analysis) using the $2^{-\Delta\Delta Ct}$ method.

7.2.3. Microscopy

7.2.3.A. Tissue fixation and sectioning

Tissue was fixed using Technovit 7100 (Hereaus Kulzer, Germany). Tissue was fixed in ethanol:acetic acid (3:1) overnight, before being dehydrated progressively using 70% - 100% ethanol. Samples were infiltrated with Technovit 7100 using ethanol:Solution A (Technovit 7100, Hardener I) with an increasing proportions of Solution A over 4 steps. Polymerisation was carried out using Solution B (Solution A, hardener II). Sections 0.6µM thick were generated using a rotary microtome (Zeiss HM340E) and stained using 0.05% toluidine blue (Merck, Germany). Sections were mounted in Euparal (Roth, Germany).
Sections were visualised using an IX70 microscope (Olympus) and CellSense software (Olympus).

7.2.4. Plant transformation and tissue culture

7.2.4.A. Transformation by floral dip of *Arabidopsis thaliana*

Floral dip transformation was carried out according to (Clough and Bent, 1998). Wild type plants were grown for 4 weeks under long day conditions (16 hours of light/8 hours of dark, 25°C). 100mL of sterile LB broth (10g/L bactotryptone; 5g/L bacto-yeast extract; 10g/L NaCl) was inoculated with 1mL of saturated *Agrobacterium* containing the construct and grown to OD_{600} 0.8. Cells were pelleted and resuspended in an aqueous solution of 5% sucrose and 0.05% Silwet L-77 to OD_{600} 0.8. Plants were inverted into the solution for one minute. Seeds were harvested and sterilised using 70% ethanol for 1 minute, followed by 30% bleach for 10 minutes and 3 washes in sterile H_2O. Successful transformants were identified by growth on MS20 medium

(4.4g/L Murashige and Skoog salts + B5 vitamins, 20g/L sucrose, 0.8% agar, pH 5.8) containing the appropriate antibiotic.

7.2.4.B. Leaf disc transformation of *Solanum lycopersicum*

Leaf disc transformation was carried out at the premises of Enza Zaden, Enkhuizen, the Netherlands under the supervision of Iris Heidmann and Suzan Out according to (Rai et al., 2012) with some modifications. Tomato seeds of the EZCBT1 variety were sterilised using 100% ethanol for 1 minute, followed by 1% bleach for 10 minutes and 3 washes in sterile H_2O. Seeds were sown onto MSB530 (4.4g/L Murashige and Skoog salts + B5 vitamins, 30g/L sucrose, 0.8% agar, pH 5.8) and grown for 8-10 days. Cotyledons were removed, cut into 0.5cm pieces and placed onto KMSC medium (4.4g/L Murashige and Skoog salts + B5 vitamins, 30g/L sucrose, 0.8% agar, pH 5.8, 200mg/L KH_2PO_4, 0.2mg/L 2,4-D, 0.1 mg/L kinetin, 0.1 mg/L IAA, 10mg/L acetosyringone) overnight in dim light. *Agrobacterium* containing the construct was grown overnight in YEB (5g/L yeast extract, 5g/L beef extract, 20g/L sucrose, pH 7.2) with the appropriate antibiotics. The culture was washed and diluted to a density of 0.4 OD_{600} and poured over the explants. This was incubated for two hours before explants were removed to fresh KMSC plates and incubated for 72 hours in dim light. Explants were then moved to selection medium I for regeneration (4.4g/L Murashige and Skoog salts + B5 vitamins, 30g/L sucrose, 0.8% agar, pH 5.8, 2mg/L zeatin, 0.1mg/L IAA, 500mg/L cefotaxime, 100mg/L kanamycin). When calli were visible, explants were transferred to selection medium II (4.4g/L Murashige and Skoog salts + B5 vitamins, 30g/L sucrose, 0.8% agar, pH 5.8, 0.2mg/L zeatin, 500mg/L cefotaxime, 100mg/L kanamycin). Calli which regenerated shoot meristems were transferred to rooting medium (4.4g/L Murashige and Skoog salts + B5 vitamins, 10g/L sucrose, 0.8% agar, pH 5.8, 100mg/L cefotaxime, 50mg/L kanamycin).

8. References

Agius, F., Kapoor, A., and Zhu, J.-K. (2006). Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. Proc. Natl. Acad. Sci. U. S. A. *103*, 11796–11801.

Amaya, I., Ratcliffe, O.J., and Bradley, D.J. (1999). Expression of CENTRORADIALIS (CEN) and CEN-like genes in tobacco reveals a conserved mechanism controlling phase change in diverse species. Plant Cell *11*, 1405–1417.

Amouroux, R., Nashun, B., Shirane, K., Nakagawa, S., Hill, P.W.S., D'Souza, Z., Nakayama, M., Matsuda, M., Turp, A., Ndjetehe, E., et al. (2016). De novo DNA methylation drives 5hmC accumulation in mouse zygotes. Nat. Cell Biol. *18*, 225–233.

Bachman, M., Uribe-Lewis, S., Yang, X., Williams, M., Murrell, A., and Balasubramanian, S. (2014). 5-Hydroxymethylcytosine is a predominantly stable DNA modification. Nat. Chem. *6*, 1049–1055.

Bachman, M., Uribe-Lewis, S., Yang, X., Burgess, H.E., Iurlaro, M., Reik, W., Murrell, A., and Balasubramanian, S. (2015). 5-Formylcytosine can be a stable DNA modification in mammals. Nat. Chem. Biol. *11*, 555–557.

Banfield, M.J., and Brady, R.L. (2000). The structure of Antirrhinum centroradialis protein (CEN) suggests a role as a kinase regulator1. J. Mol. Biol. *297*, 1159–1170.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

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

Bartkuhn, M., and Renkawitz, R. (2008). Long range chromatin interactions involved in gene regulation. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1783*, 2161–2166.

Bestor, T.H. (1992). Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. EMBO J. *11*, 2611–2617.

Betsuyaku, S., Takahashi, F., Kinoshita, A., Miwa, H., Shinozaki, K., Fukuda, H., and Sawa, S. (2011). Mitogen-activated protein kinase regulated by the CLAVATA receptors contributes to shoot apical meristem homeostasis. Plant Cell Physiol. *52*, 14–29.

Bewick, A.J., Ji, L., Niederhuth, C.E., Willing, E.-M., Hofmeister, B.T., Shi, X., Wang, L., Lu, Z., Rohr, N.A., Hartwig, B., et al. (2016). On the origin and evolutionary consequences of gene body DNA methylation. Proc. Natl. Acad. Sci. *113*, 9111–9116.

Bönisch, C., and Hake, S.B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? Nucleic Acids Res. *40*, 10719–10741.

Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W., and Balasubramanian, S. (2012). Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science *336*, 934–937.

Booth, M.J., Marsico, G., Bachman, M., Beraldi, D., and Balasubramanian, S. (2014). Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. Nat. Chem. *6*, 435–440.

Bouyer, D., Kramdi, A., Kassam, M., Heese, M., Schnittger, A., Roudier, F., and Colot, V. (2017). DNA methylation dynamics during early plant life. Genome Biol. *18*, 179.

Bradley, D., Carpenter, R., Copsey, L., Vincent, C., Rothstein, S., and Coen, E. (1996). Control of inflorescence architecture in Antirrhinum. Nature *379*, 791–797.

Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M., and Simon, R. (2000). Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by CLV3 activity. Science *289*, 617–619.

Brooks, S.C., Fischer, R.L., Huh, J.H., and Eichman, B.F. (2014). 5-Methylcytosine recognition by Arabidopsis thaliana DNA glycosylases DEMETER and DML3. Biochemistry (Mosc.) *53*, 2525–2532.

Byrne, M.E., Simorowski, J., and Martienssen, R.A. (2002). ASYMMETRIC LEAVES1 reveals knox gene redundancy in Arabidopsis. Development *129*, 1957–1965.

Cao, X., and Jacobsen, S.E. (2002a). Locus-specific control of asymmetric and CpNpG methylation by the DRM and CMT3 methyltransferase genes. Proc. Natl. Acad. Sci. *99*, 16491–16498.

Cao, X., and Jacobsen, S.E. (2002b). Role of the Arabidopsis DRM methyltransferases in de novo DNA methylation and gene silencing. Curr. Biol. *12*, 1138–1144.

Cao, K., Cui, L., Zhou, X., Ye, L., Zou, Z., and Deng, S. (2016). Four tomato FLOWERING LOCUS T-like proteins act antagonistically to regulate floral initiation. Front. Plant Sci. *6*.

Cao, X., He, Z., Guo, L., and Liu, X. (2015). Epigenetic mechanisms are critical for the regulation of WUSCHEL expression in floral meristems. Plant Physiol. *168*, 1189–1196.

Carmel-Goren, L., Liu, Y.S., Lifschitz, E., and Zamir, D. (2003). The SELF-PRUNING gene family in tomato. Plant Mol. Biol. *52*, 1215–1222.

Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. Nat. Rev. Genet. *10*, 295–304.

Chan, S.W.-L., Henderson, I.R., and Jacobsen, S.E. (2005). Gardening the genome: DNA methylation in Arabidopsis thaliana. Nat. Rev. Genet. *6*, 351–360.

Chardon, F., and Damerval, C. (2005). Phylogenomic analysis of the PEBP gene family in cereals. J. Mol. Evol. *61*, 579–590.

Chen, C.-C., Wang, K.-Y., and Shen, C.-K.J. (2012). The Mammalian de novo DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases♦. J. Biol. Chem. *287*, 33116–33121.

Chen, J.-G., Ullah, H., Young, J.C., Sussman, M.R., and Jones, A.M. (2001). ABP1 is required for organized cell elongation and division in Arabidopsis embryogenesis. Genes Dev. *15*, 902–911.

Chen, L., Chen, K., Lavery, L.A., Baker, S.A., Shaw, C.A., Li, W., and Zoghbi, H.Y. (2015). MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. Proc. Natl. Acad. Sci. U. S. A. *112*, 5509–5514.

Cheng, C., Tarutani, Y., Miyao, A., Ito, T., Yamazaki, M., Sakai, H., Fukai, E., and Hirochika, H. (2015). Loss of function mutations in the rice chromomethylase OsCMT3a cause a burst of transposition. Plant J. *83*, 1069–1081.

Chiang, P.K., Gordon, R.K., Tal, J., Zeng, G.C., Doctor, B.P., Pardhasaradhi, K., and McCann, P.P. (1996). S-Adenosylmethionine and methylation. FASEB J. *10*, 471–480.

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

Chu, H., Qian, Q., Liang, W., Yin, C., Tan, H., Yao, X., Yuan, Z., Yang, J., Huang, H., Luo, D., et al. (2006). The FLORAL ORGAN NUMBER4 gene encoding a putative

ortholog of Arabidopsis CLAVATA3 regulates apical meristem size in rice. Plant Physiol. *142*, 1039–1052.

Clark, S.E., Running, M.P., and Meyerowitz, E.M. (1993). CLAVATA1, a regulator of meristem and flower development in Arabidopsis. Development *119*, 397–418.

Clark, S.M., Running, M.P., and Meyerowitz, E.M. (1995). CLAVATA3 is a specific regulator of shoot and floral meristem development affecting the same processes as CLAVATA1. Development *121*, 2057–2067.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J. *16*, 735–743.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature *452*, 215–219.

Corbit, K.C., Trakul, N., Eves, E.M., Diaz, B., Marshall, M., and Rosner, M.R. (2003). Activation of Raf-1 signaling by Protein Kinase C through a mechanism involving Raf Kinase Inhibitory Protein. J. Biol. Chem. *278*, 13061–13068.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., et al. (2011). Thymine DNA Glycosylase Is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67–79.

Davis, T.L., Yang, G.J., McCarrey, J.R., and Bartolomei, M.S. (2000). The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. Hum. Mol. Genet. *9*, 2885–2894.

Dawlaty, M.M., Ganz, K., Powell, B.E., Hu, Y.-C., Markoulaki, S., Cheng, A.W., Gao, Q., Kim, J., Choi, S.-W., Page, D.C., et al. (2011). Tet1 is dispensable for maintaining

pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell *9*, 166–175.

Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M., et al. (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev. Cell *24*, 310–323.

Dawlaty, M.M., Breiling, A., Le, T., Barrasa, M.I., Raddatz, G., Gao, Q., Powell, B.E., Cheng, A.W., Faull, K.F., Lyko, F., et al. (2014). Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. Dev. Cell *29*, 102–111.

DeChiara, T.M., Robertson, E.J., and Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. Cell *64*, 849–859.

Delatte, B., Deplus, R., and Fuks, F. (2014). Playing TETris with DNA modifications. EMBO J. *33*, 1198–1211.

Delhommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Massé, A., Kosmider, O., Le Couedic, J.-P., Robert, F., Alberdi, A., et al. (2009). Mutation in TET2 in myeloid cancers. N. Engl. J. Med. *360*, 2289–2301.

Dowen, R.H., Pelizzola, M., Schmitz, R.J., Lister, R., Dowen, J.M., Nery, J.R., Dixon, J.E., and Ecker, J.R. (2012). Widespread dynamic DNA methylation in response to biotic stress. Proc. Natl. Acad. Sci. *109*, E2183–E2191.

Ebbs, M.L., and Bender, J. (2006). Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase. Plant Cell *18*, 1166–1176.

Ellul, P., Garcia-Sogo, B., Pineda, B., Ríos, G., Roig, L.A., and Moreno, V. (2003). The ploidy level of transgenic plants in Agrobacterium-mediated transformation of tomato cotyledons (Lycopersicon esculentum Mill.) is genotype and procedure dependent [corrected]. TAG Theor. Appl. Genet. Theor. Angew. Genet. *106*, 231–238.

Erdmann, R.M., Souza, A.L., Clish, C.B., and Gehring, M. (2014). 5-hydroxymethylcytosine is not present in appreciable quantities in Arabidopsis DNA. G3 GenesGenomesGenetics *5*, 1–8.

Feng, S., Cokus, S.J., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. Proc. Natl. Acad. Sci. U. S. A. *107*, 8689–8694.

Ferguson-Smith, A.C. (2011). Genomic imprinting: the emergence of an epigenetic paradigm. Nat. Rev. Genet. *12*, 565–575.

Finnegan, E.J., Peacock, W.J., and Dennis, E.S. (1996). Reduced DNA methylation in Arabidopsis thaliana results in abnormal plant development. Proc. Natl. Acad. Sci. *93*, 8449–8454.

Gallois, J.-L., Woodward, C., Reddy, G.V., and Sablowski, R. (2002). Combined SHOOT MERISTEMLESS and WUSCHEL trigger ectopic organogenesis in Arabidopsis. Development *129*, 3207–3217.

Gehring, M., and Satyaki, P.R. (2017). Endosperm and imprinting, inextricably linked. Plant Physiol. *173*, 143–154.

Gehring, M., Reik, W., and Henikoff, S. (2009). DNA demethylation by DNA repair. Trends Genet. *25*, 82–90.

Gheysen, G., Herman, L., Breyne, P., Gielen, J., Van Montagu, M., and Depicker, A. (1990). Cloning and sequence analysis of truncated T-DNA inserts from Nicotiana tabacum. Gene *94*, 155–163.

Golan, I., Dominguez, P.G., Konrad, Z., Shkolnik-Inbar, D., Carrari, F., and Bar-Zvi, D. (2014). Tomato ABSCISIC ACID STRESS RIPENING (ASR) gene family revisited. PLoS ONE *9*.

González, R.M., Ricardi, M.M., and Iusem, N.D. (2013). Epigenetic marks in an adaptive water stress-responsive gene in tomato roots under normal and drought conditions. Epigenetics *8*, 864–872.

Gouil, Q., and Baulcombe, D.C. (2016). DNA methylation signatures of the plant chromomethyltransferases. PLOS Genet. *12*, e1006526.

Gouil, Q., Novák, O., and Baulcombe, D.C. (2016). SLTAB2 is the paramutated SULFUREA locus in tomato. J. Exp. Bot. *67*, 2655–2664.

Gregory, D.J., Mikhaylova, L., and Fedulov, A.V. (2012). Selective DNA demethylation by fusion of TDG with a sequence-specific DNA-binding domain. Epigenetics *7*, 344–349.

Gu, T.-P., Guo, F., Yang, H., Wu, H.-P., Xu, G.-F., Liu, W., Xie, Z.-G., Shi, L., He, X., Jin, S., et al. (2011). The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. Nature *477*, 606–610.

Guo, J.U., Su, Y., Zhong, C., Ming, G., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423–434.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mech. Dev. *117*, 15–23.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303–1307.

Hellens, R., Mullineaux, P., and Klee, H. (2000a). Technical Focus:A guide to Agrobacterium binary Ti vectors. Trends Plant Sci. *5*, 446–451.

Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S., and Mullineaux, P.M. (2000b). pGreen: a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant Mol. Biol. *42*, 819–832.

Hengst, U., Albrecht, H., Hess, D., and Monard, D. (2001). The phosphatidylethanolamine-binding protein is the prototype of a novel family of serine protease inhibitors. J. Biol. Chem. *276*, 535–540.

Hetzl, J., Foerster, A.M., Raidl, G., and Scheid, O.M. (2007). CyMATE: a new tool for methylation analysis of plant genomic DNA after bisulphite sequencing. Plant J. *51*, 526–536.

Hicklenton, P.R., Newman, S.M., and Davies, L.J. (1993). Growth and flowering of Gypsophila paniculata L. “Bristol Fairy” and “Bridal Veil” in relation to temperature and photosynthetic photon flux. Sci. Hortic. *53*, 319–331.

Hollwey, E., Watson, M., and Meyer, P. (2016). Expression of the C-terminal domain of mammalian TET3 DNA dioxygenase in Arabidopsis thaliana induces heritable methylation changes at rDNA loci. Adv. Biosci. Biotechnol. *07*, 243–250.

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. Proc. Natl. Acad. Sci. 201410761.

Huang, R., Ding, Q., Xiang, Y., Gu, T., and Li, Y. (2016). Comparative analysis of DNA methyltransferase gene family in fungi: a focus on basidiomycota. Front. Plant Sci. *7*.

Huang, Y., Chavez, L., Chang, X., Wang, X., Pastor, W.A., Kang, J., Zepeda-Martínez, J.A., Pape, U.J., Jacobsen, S.E., Peters, B., et al. (2014). Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. Proc. Natl. Acad. Sci. *111*, 1361–1366.

Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.-F., Uzawa, R., Rodrigues, J.A., Zemach, A., Chumak, N., Machlicova, A., Nishimura, T., et al. (2012). Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes. Science *337*, 1360–1364.

Inoue, A., and Zhang, Y. (2011). Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science *334*, 194.

Inoue, S., and Oishi, M. (2005). Effects of methylation of non-CpG sequence in the promoter region on the expression of human synaptotagmin XI (syt11). Gene *348*, 123–134.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129–1133.

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

Iurlaro, M., Ficz, G., Oxley, D., Raiber, E.-A., Bachman, M., Booth, M.J., Andrews, S., Balasubramanian, S., and Reik, W. (2013). A screen for hydroxymethylcytosine and formylcytosine binding proteins suggests functions in transcription and chromatin regulation. Genome Biol. *14*, R119.

Iurlaro, M., McInroy, G.R., Burgess, H.E., Dean, W., Raiber, E.-A., Bachman, M., Beraldi, D., Balasubramanian, S., and Reik, W. (2016). In vivo genome-wide profiling reveals a tissue-specific role for 5-formylcytosine. Genome Biol. *17*, 141.

Iyer, L.M., Tahiliani, M., Rao, A., and Aravind, L. (2009). Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. Cell Cycle Georget. Tex *8*, 1698–1710.

Jang, H., Shin, H., Eichman, B.F., and Huh, J.H. (2014). Excision of 5-hydroxymethylcytosine by DEMETER family DNA glycosylases. Biochem. Biophys. Res. Commun. *446*, 1067–1072.

Jeon, J., Choi, J., Lee, G.-W., Park, S.-Y., Huh, A., Dean, R.A., and Lee, Y.-H. (2015). Genome-wide profiling of DNA methylation provides insights into epigenetic regulation of fungal development in a plant pathogenic fungus, Magnaporthe oryzae. Sci. Rep. *5*, 8567.

Ji, L., Jordan, W.T., Shi, X., Hu, L., He, C., and Schmitz, R.J. (2017). TET-mediated epimutagenesis of the Arabidopsis thaliana methylome. bioRxiv.

Jiang, H., Doerge, R. w., and Gelvin, S.B. (2003). Transfer of T-DNA and Vir proteins to plant cells by Agrobacterium tumefaciens induces expression of host genes involved in mediating transformation and suppresses host defense gene expression. Plant J. *35*, 219–236.

Jiang, K., Liberatore, K.L., Park, S.J., Alvarez, J.P., and Lippman, Z.B. (2013). Tomato yield heterosis is triggered by a dosage sensitivity of the florigen pathway that fine-tunes shoot architecture. PLOS Genet *9*, e1004043.

Jin, S.-G., Zhang, Z.-M., Dunwell, T.L., Harter, M.R., Wu, X., Johnson, J., Li, Z., Liu, J., Szabó, P.E., Lu, Q., et al. (2016). Tet3 reads 5-carboxylcytosine through its CXXC domain and is a potential guardian against neurodegeneration. Cell Rep. *14*, 493–505.

de Jonge, J. (2013). Shoot apical meristem arrest in brassica and tomato. PhD Thesis. Wageningen University.

de Jonge, J., Kodde, J., Severing, E.I., Bonnema, G., Angenent, G.C., Immink, R.G.H., and Groot, S.P.C. (2016). Low temperature affects stem cell maintenance in Brassica oleracea seedlings. Front. Plant Sci. *7*.

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

Kakutani, T., Jeddeloh, J.A., Flowers, S.K., Munakata, K., and Richards, E.J. (1996). Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc. Natl. Acad. Sci. U. S. A. *93*, 12406–12411.

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

Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M., and Matzke, A.J.M. (2005). Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat. Genet. *37*, 761–765.

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

Kawakatsu, T., Nery, J.R., Castanon, R., and Ecker, J.R. (2017). Dynamic DNA methylation reconfiguration during seed development and germination. Genome Biol. *18*, 171.

Kayes, J.M., and Clark, S.E. (1998). CLAVATA2, a regulator of meristem and organ development in Arabidopsis. Dev. Camb. Engl. *125*, 3843–3851.

Kim, J., Kim, J.H., Richards, E.J., Chung, K.M., and Woo, H.R. (2014). Arabidopsis VIM proteins regulate epigenetic silencing by modulating DNA methylation and histone modification in cooperation with MET1. Mol. Plant *7*, 1470–1485.

Kinoshita, A., Betsuyaku, S., Osakabe, Y., Mizuno, S., Nagawa, S., Stahl, Y., Simon, R., Yamaguchi-Shinozaki, K., Fukuda, H., and Sawa, S. (2010). RPK2 is an essential receptor-like kinase that transmits the CLV3 signal in Arabidopsis. Development *137*, 3911–3920.

Kinoshita, T., Miura, A., Choi, Y., Kinoshita, Y., Cao, X., Jacobsen, S.E., Fischer, R.L., and Kakutani, T. (2004). One-way control of *FWA* imprinting in *Arabidopsis* endosperm by DNA methylation. Science *303*, 521–523.

Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem. Sci. *31*, 89–97.

Koh, K.P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C.A., Mostoslavsky, G., et al. (2011). Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. Cell Stem Cell *8*, 200–213.

Kondo, T., Sawa, S., Kinoshita, A., Mizuno, S., Kakimoto, T., Fukuda, H., and Sakagami, Y. (2006). A plant peptide encoded by *CLV3* identified by in situ MALDI-TOF MS analysis. Science *313*, 845–848.

Kornberg, R.D. (1977). Structure of chromatin. Annu. Rev. Biochem. *46*, 931–954.

Kou, H.P., Li, Y., Song, X.X., Ou, X.F., Xing, S.C., Ma, J., Von Wettstein, D., and Liu, B. (2011). Heritable alteration in DNA methylation induced by nitrogen-deficiency stress accompanies enhanced tolerance by progenies to the stress in rice (Oryza sativa L.). J. Plant Physiol. *168*, 1685–1693.

Kovarik, A., Koukalová, B., Bezdeˇk, M., and Opatrn´, Z. (1997). Hypermethylation of tobacco heterochromatic loci in response to osmotic stress. Theor. Appl. Genet. *95*, 301–306.

Kovarík, A., Koukalová, B., Lim, K.Y., Matyásek, R., Lichtenstein, C.P., Leitch, A.R., and Bezdek, M. (2000). Comparative analysis of DNA methylation in tobacco heterochromatic sequences. Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol. *8*, 527–541.

Lai, Y.-S., Zhang, X., Zhang, W., Shen, D., Wang, H., Xia, Y., Qiu, Y., Song, J., Wang, C., and Li, X. (2017). The association of changes in DNA methylation with temperature-dependent sex determination in cucumber. J. Exp. Bot.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. *11*, 204–220.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.

Li, H., Qi, M., Sun, M., Liu, Y., Liu, Y., Xu, T., Li, Y., and Li, T. (2017). Tomato transcription factor SlWUS plays an important role in tomato flower and locule development. Front. Plant Sci. *8*.

Li, Q., Wang, X., Sun, H., Zeng, J., Cao, Z., Li, Y., and Qian, W. (2015a). Regulation of active DNA demethylation by a methyl-CpG-binding domain protein in Arabidopsis thaliana. PLOS Genet. *11*, e1005210.

Li, W., Liu, H., Cheng, Z.J., Su, Y.H., Han, H.N., Zhang, Y., and Zhang, X.S. (2011a). DNA methylation and histone modifications regulate de novo shoot regeneration in Arabidopsis by modulating WUSCHEL expression and auxin signaling. PLoS Genet. *7*.

Li, Y., Córdoba-Cañero, D., Qian, W., Zhu, X., Tang, K., Zhang, H., Ariza, R.R., Roldán-Arjona, T., and Zhu, J.-K. (2015b). An AP endonuclease functions in active DNA demethylation and gene imprinting in Arabidopsis. PLoS Genet. *11*.

Li, Y., Duan, C.-G., Zhu, X., Qian, W., and Zhu, J.-K. (2015c). A DNA ligase required for active DNA demethylation and genomic imprinting in Arabidopsis. Cell Res. *25*, 757–760.

Li, Z., Cai, X., Cai, C.-L., Wang, J., Zhang, W., Petersen, B.E., Yang, F.-C., and Xu, M. (2011b). Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. Blood *118*, 4509–4518.

Lian, H., Li, W.-B., and Jin, W.-L. (2016). The emerging insights into catalytic or non-catalytic roles of TET proteins in tumors and neural development. Oncotarget *7*, 64512–64525.

Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C., et al. (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat. Genet. *47*, 469–478.

Lifschitz, E., Eviatar, T., Rozman, A., Shalit, A., Goldshmidt, A., Amsellem, Z., Alvarez, J.P., and Eshed, Y. (2006). The tomato FT ortholog triggers systemic signals that regulate growth and flowering and substitute for diverse environmental stimuli. Proc. Natl. Acad. Sci. *103*, 6398–6403.

Lim, A.S.P., Srivastava, G.P., Yu, L., Chibnik, L.B., Xu, J., Buchman, A.S., Schneider, J.A., Myers, A.J., Bennett, D.A., and Jager, P.L.D. (2014). 24-hour rhythms of DNA methylation and their relation with rhythms of RNA expression in the human dorsolateral prefrontal cortex. PLOS Genet. *10*, e1004792.

Lindroth, A.M., Cao, X., Jackson, J.P., Zilberman, D., McCallum, C.M., Henikoff, S., and Jacobsen, S.E. (2001). Requirement of CHROMOMETHYLASE3 for maintenance of CpXpG methylation. Science *292*, 2077–2080.

Lindroth, A.M., Shultis, D., Jasencakova, Z., Fuchs, J., Johnson, L., Schubert, D., Patnaik, D., Pradhan, S., Goodrich, J., Schubert, I., et al. (2004). Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. EMBO J. *23*, 4146–4155.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature *462*, 315–322.

Liu, R., How-Kit, A., Stammitti, L., Teyssier, E., Rolin, D., Mortain-Bertrand, A., Halle, S., Liu, M., Kong, J., Wu, C., et al. (2015). A DEMETER-like DNA demethylase governs tomato fruit ripening. Proc. Natl. Acad. Sci. *112*, 10804–10809.

Maeder, M.L., Angstman, J.F., Richardson, M.E., Linder, S.J., Cascio, V.M., Tsai, S.Q., Ho, Q.H., Sander, J.D., Reyon, D., Bernstein, B.E., et al. (2013). Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. Nat. Biotechnol. *31*, 1137–1142.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. J. Biol. Chem. *286*, 35334–35338.

Manning, K., Tör, M., Poole, M., Hong, Y., Thompson, A.J., King, G.J., Giovannoni, J.J., and Seymour, G.B. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat. Genet. *38*, 948–952.

Mao, L., Begum, D., Chuang, H., Budiman, M.A., Szymkowiak, E.J., Irish, E.E., and Wing, R.A. (2000). JOINTLESS is a MADS-box gene controlling tomato flower abscission zone development. Nature *406*, 910–913.

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

Matzke, M.A., Kanno, T., and Matzke, A.J.M. (2015). RNA-Directed DNA Methylation: The Evolution of a Complex Epigenetic Pathway in Flowering Plants. Annu. Rev. Plant Biol. *66*, 243–267.

Mayer, K.F., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., and Laux, T. (1998). Role of WUSCHEL in regulating stem cell fate in the Arabidopsis shoot meristem. Cell *95*, 805–815.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Embryogenesis: Demethylation of the zygotic paternal genome. Nature *403*, 501–502.

McCabe, M.S., Mohapatra, U.B., Debnath, S.C., Power, J.B., and Davey, M.R. (1999a). Integration, expression and inheritance of two linked T-DNA marker genes in transgenic lettuce. Mol. Breed. *5*, 329–344.

McCabe, M.S., Schepers, F., van der Arend, A., Mohapatra, U., de Laat, A.M., Power, J.B., and Davey, M.R. (1999b). Increased stable inheritance of herbicide resistance in transgenic lettuce carrying a petE promoter-bar gene compared with a CaMV 35S-bar gene. TAG Theor. Appl. Genet. Theor. Angew. Genet. *99*, 587–592.

McGarry, R.C., and Ayre, B.G. (2012). Manipulating plant architecture with members of the CETS gene family. Plant Sci. *188–189*, 71–81.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343–349.

Mersereau, M., Pazour, G.J., and Das, A. (1990). Efficient transformation of Agrobacterium tumefaciens by electroporation. Gene *90*, 149–151.

Messeguer, R., Ganal, M.W., Steffens, J.C., and Tanksley, S.D. (1991). Characterization of the level, target sites and inheritance of cytosine methylation in tomato nuclear DNA. Plant Mol. Biol. *16*, 753–770.

Messerschmidt, D.M., Knowles, B.B., and Solter, D. (2014). DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev. *28*, 812–828.

Meyer, P. (2015). Epigenetic variation and environmental change. J. Exp. Bot. *66*, 3541–3548.

Meyer, P., and Heidmann, I. (1994). Epigenetic variants of a transgenic petunia line show hypermethylation in transgene DNA: an indication for specific recognition of foreign DNA in transgenic plants. Mol. Gen. Genet. MGG *243*, 390–399.

Meyer, P., Linn, F., Heidmann, I., Meyer, H., Niedenhof, I., and Saedler, H. (1992). Endogenous and environmental factors influence 35S promoter methylation of a maize A1 gene construct in transgenic petunia and its colour phenotype. Mol. Gen. Genet. MGG *231*, 345–352.

Mimida, N., Goto, K., Kobayashi, Y., Araki, T., Ahn, J.H., Weigel, D., Murata, M., Motoyoshi, F., and Sakamoto, W. (2001). Functional divergence of the TFL1-like gene family in Arabidopsis revealed by characterization of a novel homologue. Genes Cells *6*, 327–336.

Mishra, P.K., Baum, M., and Carbon, J. (2011). DNA methylation regulates phenotype-dependent transcriptional activity in Candida albicans. Proc. Natl. Acad. Sci. U. S. A. *108*, 11965–11970.

Molinero-Rosales, N., Latorre, A., Jamilena, M., and Lozano, R. (2003). SINGLE FLOWER TRUSS regulates the transition and maintenance of flowering in tomato. Planta *218*, 427–434.

Morales-Ruiz, T., Ortega-Galisteo, A.P., Ponferrada-Marín, M.I., Martínez-Macías, M.I., Ariza, R.R., and Roldán-Arjona, T. (2006). DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. Proc. Natl. Acad. Sci. *103*, 6853–6858.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. J. Biol. Chem. *279*, 52353–52360.

Morita, S., Noguchi, H., Horii, T., Nakabayashi, K., Kimura, M., Okamura, K., Sakai, A., Nakashima, H., Hata, K., Nakashima, K., et al. (2016). Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions. Nat. Biotechnol. *34*, 1060–1065.

Moritoh, S., Eun, C.-H., Ono, A., Asao, H., Okano, Y., Yamaguchi, K., Shimatani, Z., Koizumi, A., and Terada, R. (2012). Targeted disruption of an orthologue of DOMAINS REARRANGED METHYLASE 2, OsDRM2, impairs the growth of rice plants by abnormal DNA methylation. Plant J. *71*, 85–98.

Motorin, Y., Lyko, F., and Helm, M. (2010). 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. Nucleic Acids Res. *38*, 1415–1430.

Müller, R., Borghi, L., Kwiatkowska, D., Laufs, P., and Simon, R. (2006). Dynamic and compensatory responses of Arabidopsis shoot and floral meristems to CLV3 signaling. Plant Cell *18*, 1188–1198.

Müller, R., Bleckmann, A., and Simon, R. (2008). The receptor kinase CORYNE of Arabidopsis transmits the stem cell–limiting signal CLAVATA3 independently of CLAVATA1. Plant Cell *20*, 934–946.

Müller, U., Bauer, C., Siegl, M., Rottach, A., and Leonhardt, H. (2014). TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. Nucleic Acids Res. *42*, 8592–8604.

Nabel, C.S., Jia, H., Ye, Y., Shen, L., Goldschmidt, H.L., Stivers, J.T., Zhang, Y., and Kohli, R.M. (2012). AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. Nat. Chem. Biol. *8*, 751–758.

Nakano, Y., Steward, N., Sekine, M., Kusano, T., and Sano, H. (2000). A tobacco NtMET1 cDNA encoding a DNA methyltransferase: molecular characterization and abnormal phenotypes of transgenic tobacco plants. Plant Cell Physiol. *41*, 448–457.

Navarro, C., Abelenda, J.A., Cruz-Oró, E., Cuéllar, C.A., Tamaki, S., Silva, J., Shimamoto, K., and Prat, S. (2011). Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. Nature *478*, 119–122.

Neri, F., Incarnato, D., Krepelova, A., Rapelli, S., Pagnani, A., Zecchina, R., Parlato, C., and Oliviero, S. (2013). Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. Genome Biol. *14*, R91.

Niederhuth, C.E., Bewick, A.J., Ji, L., Alabady, M.S., Kim, K.D., Li, Q., Rohr, N.A., Rambani, A., Burke, J.M., Udall, J.A., et al. (2016). Widespread natural variation of DNA methylation within angiosperms. Genome Biol. *17*.

Nilsson, E. (1950). Some Experiments with Tetraploid Tomatoes. Hereditas *36*, 181–204.

Ogawa, M., Shinohara, H., Sakagami, Y., and Matsubayashi, Y. (2008). *Arabidopsis* CLV3 peptide directly binds CLV1 ectodomain. Science *319*, 294–294.

Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

Ortega-Galisteo, A.P., Morales-Ruiz, T., Ariza, R.R., and Roldán-Arjona, T. (2008). Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. Plant Mol. Biol. *67*, 671–681.

Parrilla-Doblas, J.T., Ariza, R.R., and Roldán-Arjona, T. (2017). Targeted DNA demethylation in human cells by fusion of a plant 5-methylcytosine DNA glycosylase to a sequence-specific DNA binding domain. Epigenetics *12*, 296–303.

Penterman, J., Zilberman, D., Huh, J.H., Ballinger, T., Henikoff, S., and Fischer, R.L. (2007a). DNA demethylation in the Arabidopsis genome. Proc. Natl. Acad. Sci. *104*, 6752–6757.

Penterman, J., Uzawa, R., and Fischer, R.L. (2007b). Genetic interactions between DNA demethylation and methylation in Arabidopsis. Plant Physiol. *145*, 1549–1557.

Piccolo, F.M., and Fisher, A.G. (2014). Getting rid of DNA methylation. Trends Cell Biol. *24*, 136–143.

Pnueli, L., Carmel-Goren, L., Hareven, D., Gutfinger, T., Alvarez, J., Ganal, M., Zamir, D., and Lifschitz, E. (1998). The SELF-PRUNING gene of tomato regulates vegetative to reproductive switching of sympodial meristems and is the ortholog of CEN and TFL1. Development *125*, 1979–1989.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature *463*, 1101–1105.

Praça, M.M., Carvalho, C.R., and Clarindo, W.R. (2009). A practical and reliable procedure for in vitro induction of tetraploid tomato. Sci. Hortic. *122*, 501–505.

Pröls, F., and Meyer, P. (1992). The methylation patterns of chromosomal integration regions influence gene activity of transferred DNA in Petunia hybrida. Plant J. *2*, 465–475.

Qian, W., Miki, D., Zhang, H., Liu, Y., Zhang, X., Tang, K., Kan, Y., La, H., Li, X., Li, S., et al. (2012). A histone acetyltransferase regulates active DNA demethylation in Arabidopsis. Science *336*, 1445–1448.

Quinet, M., Dubois, C., Goffin, M.-C., Chao, J., Dielen, V., Batoko, H., Boutry, M., and Kinet, J.-M. (2006). Characterization of tomato (Solanum lycopersicum L.)

mutants affected in their flowering time and in the morphogenesis of their reproductive structure. J. Exp. Bot. *57*, 1381–1390.

Quinet, M., Kinet, J.-M., and Lutts, S. (2011). Flowering response of the uniflora:blind:self-pruning and jointless:uniflora:self-pruning tomato (Solanum lycopersicum) triple mutants. Physiol. Plant. *141*, 166–176.

Rai, G.K., Rai, N.P., Kumar, S., Yadav, A., Rathaur, S., and Singh, M. (2012). Effects of explant age, germination medium, pre-culture parameters, inoculation medium, pH, washing medium, and selection regime on Agrobacterium-mediated transformation of tomato. Vitro Cell. Dev. Biol. - Plant *48*, 565–578.

Raissig, M.T., Bemer, M., Baroux, C., and Grossniklaus, U. (2013). Genomic imprinting in the Arabidopsis embryo is partly regulated by PRC2. PLOS Genet. *9*, e1003862.

Rajakumara, E., Nakarakanti, N.K., Nivya, M.A., and Satish, M. (2016). Mechanistic insights into the recognition of 5-methylcytosine oxidation derivatives by the SUVH5 SRA domain. Sci. Rep. *6*, 20161.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc. Natl. Acad. Sci. U. S. A. *97*, 5237–5242.

Ratcliffe, O.J., Amaya, I., Vincent, C.A., Rothstein, S., Carpenter, R., Coen, E.S., and Bradley, D.J. (1998). A common mechanism controls the life cycle and architecture of plants. Development *125*, 1609–1615.

Ratel, D., Ravanat, J.-L., Berger, F., and Wion, D. (2006). N6-methyladenine: the other methylated base of DNA. Bioessays *28*, 309–315.

Regulski, M., Lu, Z., Kendall, J., Donoghue, M.T.A., Reinders, J., Llaca, V., Deschamps, S., Smith, A., Levy, D., McCombie, W.R., et al. (2013). The maize

methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. Genome Res. gr.153510.112.

Reik, W., and Walter, J. (2001). Genomic imprinting: parental influence on the genome. Nat. Rev. Genet. *2*, 21–32.

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

Reinhardt, D., Frenz, M., Mandel, T., and Kuhlemeier, C. (2003). Microsurgical and laser ablation analysis of interactions between the zones and layers of the tomato shoot apical meristem. Development *130*, 4073–4083.

Rigal, M., Kevei, Z., Pélissier, T., and Mathieu, O. (2012). DNA methylation in an intron of the IBM1 histone demethylase gene stabilizes chromatin modification patterns. EMBO J. *31*, 2981–2993.

Riggs, A.D., Martienssen, R., and Russo, V. (1996). Introduction. In Epigenetic mechanisms of gene regulation, (Cold Spring Harbor: Cold Spring Harbor Laboratory Press), pp. 1–4.

Róis, A.S., Rodríguez López, C.M., Cortinhas, A., Erben, M., Espírito-Santo, D., Wilkinson, M.J., and Caperta, A.D. (2013). Epigenetic rather than genetic factors may explain phenotypic divergence between coastal populations of diploid and tetraploid Limonium spp. (Plumbaginaceae) in Portugal. BMC Plant Biol. *13*, 205.

Rojo, E., Sharma, V.K., Kovaleva, V., Raikhel, N.V., and Fletcher, J.C. (2002). CLV3 is localized to the extracellular space, where it activates the Arabidopsis CLAVATA stem cell signaling pathway. Plant Cell *14*, 969–977.

Rougeulle, C., Glatt, H., and Lalande, M. (1997). The Angelman syndrome candidate gene, UBE3AIE6-AP, is imprinted in brain. Nat. Genet. *17*, 14–15.

Rougier, N., Bourc'his, D., Gomes, D.M., Niveleau, A., Plachot, M., Pàldi, A., and Viegas-Péquignot, E. (1998). Chromosome methylation patterns during mammalian preimplantation development. Genes Dev. *12*, 2108–2113.

Salter, P.J. (1957). Blindness in Early Summer Cauliflower. Nature *180*, 1056–1056.

Sambrook, J., Fritsch, E., and Maniatis, T. (1989). Molecular cloning: a laboratory manual (Cold Spring Harbour Laboratory Press).

SanMiguel, P., Tikhonov, A., Jin, Y.K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z., et al. (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science *274*, 765–768.

Santo, T.E., Pereira, R.J., and Leitão, J.M. (2017). The pea (Pisum sativum L.) rogue paramutation is accompanied by alterations in the methylation pattern of specific genomic sequences. Epigenomes *1*, 6.

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

Schiesser, S., Hackner, B., Pfaffeneder, T., Müller, M., Hagemeier, C., Truss, M., and Carell, T. (2012). Mechanism and stem-cell activity of 5-carboxycytosine decarboxylation determined by isotope tracing. Angew. Chem. Int. Ed. *51*, 6516–6520.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F.X., Jürgens, G., and Laux, T. (2000). The stem cell population of Arabidopsis shoot meristems is maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. Cell *100*, 635–644.

Schuermann, D., Weber, A.R., and Schär, P. (2016). Active DNA demethylation by DNA repair: Facts and uncertainties. DNA Repair *44*, 92–102.

Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W., and Reik, W. (2012). The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol. Cell *48*, 849–862.

Shalit, A., Rozman, A., Goldshmidt, A., Alvarez, J.P., Bowman, J.L., Eshed, Y., and Lifschitz, E. (2009). The flowering hormone florigen functions as a general systemic regulator of growth and termination. Proc. Natl. Acad. Sci. *106*, 8392–8397.

Shemer, O., Landau, U., Candela, H., Zemach, A., and Eshed Williams, L. (2015). Competency for shoot regeneration from Arabidopsis root explants is regulated by DNA methylation. Plant Sci. *238*, 251–261.

Shen, W.J., and Forde, B.G. (1989). Efficient transformation of Agrobacterium spp. by high voltage electroporation. Nucleic Acids Res. *17*, 8385.

Shirane, K., Toh, H., Kobayashi, H., Miura, F., Chiba, H., Ito, T., Kono, T., and Sasaki, H. (2013). Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. PLoS Genet. *9*.

Singh, A., Zubko, E., and Meyer, P. (2008). Cooperative activity of DNA methyltransferases for maintenance of symmetrical and non-symmetrical cytosine methylation in Arabidopsis thaliana. Plant J. *56*, 814–823.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. Nat. Rev. Genet. *8*, 272–285.

Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., and Lucchesi, J.C. (2000). The Drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. Mol. Cell. Biol. *20*, 312–318.

Somssich, M., Je, B.I., Simon, R., and Jackson, D. (2016). CLAVATA-WUSCHEL signaling in the shoot meristem. Development *143*, 3238–3248.

Song, J., Rechkoblit, O., Bestor, T.H., and Patel, D.J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. Science *331*, 1036–1040.

Song, S.-K., Lee, M.M., and Clark, S.E. (2006). POL and PLL1 phosphatases are CLAVATA1 signaling intermediates required for Arabidopsis shoot and floral stem cells. Development *133*, 4691–4698.

Soppe, W.J.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T., Koornneef, M., and Peeters, A.J.M. (2000). The late flowering phenotype of fwa mutants is caused by gain-of-function epigenetic alleles of a homeodomain gene. Mol. Cell *6*, 791–802.

Soyk, S., Müller, N.A., Park, S.J., Schmalenbach, I., Jiang, K., Hayama, R., Zhang, L., Van Eck, J., Jiménez-Gómez, J.M., and Lippman, Z.B. (2017). Variation in the flowering gene SELF PRUNING 5G promotes day-neutrality and early yield in tomato. Nat. Genet. *49*, 162–168.

Spofford, J.B. (1967). Single-locus modification of position-effect variegation in Drosophila melanogaster. I. white variegation. Genetics *57*, 751–766.

Spruijt, C.G., Gnerlich, F., Smits, A.H., Pfaffeneder, T., Jansen, P.W.T.C., Bauer, C., Münzel, M., Wagner, M., Müller, M., Khan, F., et al. (2013). Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. Cell *152*, 1146–1159.

Stam, M., De Bruin, R., Van Blokland, R., Van Der Hoorn, R.A.L., Mol, J.N.M., and Kooter, J.M. (2000). Distinct features of post-transcriptional gene silencing by antisense transgenes in single copy and inverted T-DNA repeat loci. Plant J. *21*, 27–42.

Steward, N., Ito, M., Yamaguchi, Y., Koizumi, N., and Sano, H. (2002). Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. J. Biol. Chem. *277*, 37741–37746.

Stroud, H., Greenberg, M.V.C., Feng, S., Bernatavichute, Y.V., and Jacobsen, S.E. (2013). Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. Cell *152*, 352–364.

Stroud, H., Do, T., Du, J., Zhong, X., Feng, S., Johnson, L., Patel, D.J., and Jacobsen, S.E. (2014). Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. Nat. Struct. Mol. Biol. *21*, 64–72.

Sunilkumar, G., Mohr, L., Lopata-Finch, E., Emani, C., and Rathore, K.S. (2002). Developmental and tissue-specific expression of CaMV 35S promoter in cotton as revealed by GFP. Plant Mol. Biol. *50*, 463–479.

Suzaki, T., Yoshida, A., and Hirano, H.-Y. (2008). Functional diversification of CLAVATA3-related CLE proteins in meristem maintenance in rice. Plant Cell Online *20*, 2049–2058.

Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat. Rev. Genet. *9*, 465–476.

Szymkowiak, E.J., and Irish, E.E. (2006). JOINTLESS suppresses sympodial identity in inflorescence meristems of tomato. Planta *223*, 646–658.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930–935.

Takuno, S., and Gaut, B.S. (2012). Body-methylated genes in Arabidopsis thaliana are functionally important and evolve slowly. Mol. Biol. Evol. *29*, 219–227.

Takuno, S., Ran, J.-H., and Gaut, B.S. (2016). Evolutionary patterns of genic DNA methylation vary across land plants. Nat. Plants *2*, 15222.

Tal, L., Friedlander, G., Gilboa, N.S., Unger, T., Gilad, S., and Eshed, Y. (2017). Coordination of meristem doming and the floral transition by late termination, a Kelch repeat protein. Plant Cell tpc.00030.2017.

Tan, L., and Shi, Y.G. (2012). Tet family proteins and 5-hydroxymethylcytosine in development and disease. Development *139*, 1895–1902.

Tang, K., Lang, Z., Zhang, H., and Zhu, J.-K. (2016). The DNA demethylase ROS1 targets genomic regions with distinct chromatin modifications. Nat. Plants *2*, 16169.

Tremblay, K.D., Duran, K.L., and Bartolomei, M.S. (1997). A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. Mol. Cell. Biol. *17*, 4322–4329.

Valinluck, V., and Sowers, L.C. (2007). Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. Cancer Res. *67*, 946–950.

Vassileva, V., Hollwey, E., Todorov, D., and Meyer, P. (2016). Leaf epidermal profiling as a phenotyping tool for DNA methylation mutants. Genet. Plant Physiol. 3–13.

Vejlupkova, Z., and Fowler, J. (2003). Maize DNA preps for undergraduate students: a robust method for PCR genotyping. Maize Genet. Coop. Newsl. 24–25.

Veluchamy, A., Lin, X., Maumus, F., Rivarola, M., Bhavsar, J., Creasy, T., O'Brien, K., Sengamalay, N.A., Tallon, L.J., Smith, A.D., et al. (2013). Insights into the role of DNA methylation in diatoms by genome-wide profiling in Phaeodactylum tricornutum. Nat. Commun. *4*.

Vitha, S., Baluška, F., Jasik, J., Volkmann, D., and Barlow, P.W. (2000). Steedman's wax for F-actin visualization. In Actin: a dynamic framework for multiple plant cell functions, (Springer, Dordrecht), pp. 619–636.

Vrebalov, J., Ruezinsky, D., Padmanabhan, V., White, R., Medrano, D., Drake, R., Schuch, W., and Giovannoni, J. (2002). A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. Science *296*, 343–346.

Waddington, C.H. (1942). The Epigenotype. Endeavour *41*, 18–20.

Wang (2006). Agrobacterium Protocols - Volume I | Kan Wang | Springer.

Wang, C., Dong, X., Jin, D., Zhao, Y., Xie, S., Li, X., He, X., Lang, Z., Lai, J., Zhu, J.-K., et al. (2015). Methyl-CpG-binding domain protein MBD7 is required for active DNA demethylation in Arabidopsis. Plant Physiol. *167*, 905–914.

Wang, X., Wang, X., Ren, J., Ma, Y., and Yin, J. (2012). Characterization of tomato transcription factor WUSCHEL and functional study in Arabidopsis. J. Integr. Agric. *11*, 1257–1265.

Wassenegger, M., Heimes, S., Riedel, L., and Sänger, H.L. (1994). RNA-directed de novo methylation of genomic sequences in plants. Cell *76*, 567–576.

Watanabe, Y. (2011). Overview of Plant RNAi. In RNAi and Plant Gene Function Analysis, (Humana Press), pp. 1–11.

Watson, M.R. (2013). Heritable epigenetic variation of DNA methylation targets in plants. PhD Thesis. University of Leeds.

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

Weber, A.R., Krawczyk, C., Robertson, A.B., Kuśnierczyk, A., Vågbø, C.B., Schuermann, D., Klungland, A., and Schär, P. (2016). Biochemical reconstitution of

TET1–TDG–BER-dependent active DNA demethylation reveals a highly coordinated mechanism. Nat. Commun. *7*, 10806.

Werner, T., Motyka, V., Laucou, V., Smets, R., Onckelen, H.V., and Schmülling, T. (2003). Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. Plant Cell *15*, 2532–2550.

Wetzstein, H.Y., and Vavrina, C.S. (2002). Morphological evaluation of apical meristem decline in greenhouse-grown tomato transplants and the effect of mineral nutrition on Its occurrence. J. Am. Soc. Hortic. Sci. *127*, 635–638.

Williams, B.P., Pignatta, D., Henikoff, S., and Gehring, M. (2015). Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat. PLoS Genet. *11*.

Williams, R.W., Wilson, J.M., and Meyerowitz, E.M. (1997). A possible role for kinase-associated protein phosphatase in the Arabidopsis CLAVATA1 signaling pathway. Proc. Natl. Acad. Sci. *94*, 10467–10472.

Wilson, G.G., and Murray, N.E. (1991). Restriction and modification systems. Annu. Rev. Genet. *25*, 585–627.

Witte, C.-P., Le, Q.H., Bureau, T., and Kumar, A. (2001). Terminal-repeat retrotransposons in miniature (TRIM) are involved in restructuring plant genomes. Proc. Natl. Acad. Sci. U. S. A. *98*, 13778–13783.

Wu, T.P., Wang, T., Seetin, M.G., Lai, Y., Zhu, S., Lin, K., Liu, Y., Byrum, S.D., Mackintosh, S.G., Zhong, M., et al. (2016). DNA methylation on N6-adenine in mammalian embryonic stem cells. Nature *532*, 329–333.

Xu, C., Liberatore, K.L., MacAlister, C.A., Huang, Z., Chu, Y.-H., Jiang, K., Brooks, C., Ogawa-Ohnishi, M., Xiong, G., Pauly, M., et al. (2015). A cascade of

arabinosyltransferases controls shoot meristem size in tomato. Nat. Genet. *47*, 784–792.

Xu, X., Tao, Y., Gao, X., Zhang, L., Li, X., Zou, W., Ruan, K., Wang, F., Xu, G., and Hu, R. (2016). A CRISPR-based approach for targeted DNA demethylation. Cell Discov. *2*, celldisc20169.

Xu, Y., Xu, C., Kato, A., Tempel, W., Abreu, J.G., Bian, C., Hu, Y., Hu, D., Zhao, B., Cerovina, T., et al. (2012). Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. Cell *151*, 1200–1213.

Yaari, R., Noy-Malka, C., Wiedemann, G., Gershovitz, N.A., Reski, R., Katz, A., and Ohad, N. (2015). DNA METHYLTRANSFERASE 1 is involved in mCG and mCCG DNA methylation and is essential for sporophyte development in Physcomitrella patens. Plant Mol. Biol. *88*, 387–400.

Yamagata, Y., Szabó, P., Szüts, D., Bacquet, C., Aránvi, T., and Paldi, A. (2012). Rapid turnover of DNA methylation in human cells. Epigenetics *7*, 141–145.

Yamamuro, C., Miki, D., Zheng, Z., Ma, J., Wang, J., Yang, Z., Dong, J., and Zhu, J.-K. (2014). Overproduction of stomatal lineage cells in Arabidopsis mutants defective in active DNA demethylation. Nat. Commun. *5*, 4062.

Yao, Q., Song, C.-X., He, C., Kumaran, D., and Dunn, J.J. (2012). Heterologous expression and purification of Arabidopsis thaliana VIM1 protein: In vitro evidence for its inability to recognize hydroxymethylcytosine, a rare base in Arabidopsis DNA. Protein Expr. Purif. *83*, 104–111.

Yang, H., Chang, F., You, C., Cui, J., Zhu, G., Wang, L., Zheng, Y., Qi, J., and Ma, H. (2015). Whole-genome DNA methylation patterns and complex associations with gene structure and expression during flower development in Arabidopsis. Plant J. *81*, 268–281.

Yoo, S.J., Chung, K.S., Jung, S.H., Yoo, S.Y., Lee, J.S., and Ahn, J.H. (2010). BROTHER OF FT AND TFL1 (BFT) has TFL1-like activity and functions redundantly with TFL1 in inflorescence meristem development in Arabidopsis. Plant J. Cell Mol. Biol. *63*, 241–253.

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

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

Zeng, Y., Yao, B., Shin, J., Lin, L., Kim, N., Song, Q., Liu, S., Su, Y., Guo, J.U., Huang, L., et al. (2016). Lin28A Binds Active Promoters and Recruits Tet1 to Regulate Gene Expression. Mol. Cell *61*, 153–160.

Zhang, L., Eugeni, E.E., Parthun, M.R., and Freitas, M.A. (2003). Identification of novel histone post-translational modifications by peptide mass fingerprinting. Chromosoma *112*, 77–86.

Zhang, M., Kimatu, J.N., Xu, K., and Liu, B. (2010). DNA cytosine methylation in plant development. J. Genet. Genomics *37*, 1–12.

Zhang, Q., Zhao, K., Shen, Q., Han, Y., Gu, Y., Li, X., Zhao, D., Liu, Y., Wang, C., Zhang, X., et al. (2015). Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. Nature *525*, 389–393.

Zhang, R.-R., Cui, Q.-Y., Murai, K., Lim, Y.C., Smith, Z.D., Jin, S., Ye, P., Rosa, L., Lee, Y.K., Wu, H.-P., et al. (2013). Tet1 regulates adult hippocampal neurogenesis and cognition. Cell Stem Cell *13*, 237–245.

Zhang, Y., Yang, S., Song, Y., and Wang, J. (2014). Genome-wide characterization, expression and functional analysis of CLV3/ESR gene family in tomato. BMC Genomics *15*, 827.

Zhao, Y., Xie, S., Li, X., Wang, C., Chen, Z., Lai, J., and Gong, Z. (2014). REPRESSOR OF SILENCING5 encodes a member of the small heat shock protein family and is required for DNA demethylation in Arabidopsis. Plant Cell *26*, 2660–2675.

Zhao, Z., Andersen, S.U., Ljung, K., Dolezal, K., Miotk, A., Schultheiss, S.J., and Lohmann, J.U. (2010). Hormonal control of the shoot stem-cell niche. Nature *465*, 1089–1092.

Zheng, X., Pontes, O., Zhu, J., Miki, D., Zhang, F., Li, W.-X., Iida, K., Kapoor, A., Pikaard, C.S., and Zhu, J.-K. (2008). ROS3 is an RNA-binding protein required for DNA demethylation in Arabidopsis. Nature *455*, 1259–1262.

Zhong, S., Fei, Z., Chen, Y.-R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., et al. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nat. Biotechnol. *31*, 154–159.

Zhu, B., Zheng, Y., Angliker, H., Schwarz, S., Thiry, S., Siegmann, M., and Jost, J.-P. (2000a). 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. Nucleic Acids Res. *28*, 4157–4165.

Zhu, B., Zheng, Y., Hess, D., Angliker, H., Schwarz, S., Siegmann, M., Thiry, S., and Jost, J.-P. (2000b). 5-Methylcytosine-DNA glycosylase activity is present in a cloned G/T mismatch DNA glycosylase associated with the chicken embryo DNA demethylation complex. Proc. Natl. Acad. Sci. *97*, 5135–5139.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. *39*, 61–69.

Ziller, M.J., Müller, F., Liao, J., Zhang, Y., Gu, H., Bock, C., Boyle, P., Epstein, C.B., Bernstein, B.E., Lengauer, T., et al. (2011). Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. PLoS Genet. *7*.

9. Appendix

Figure 9.1

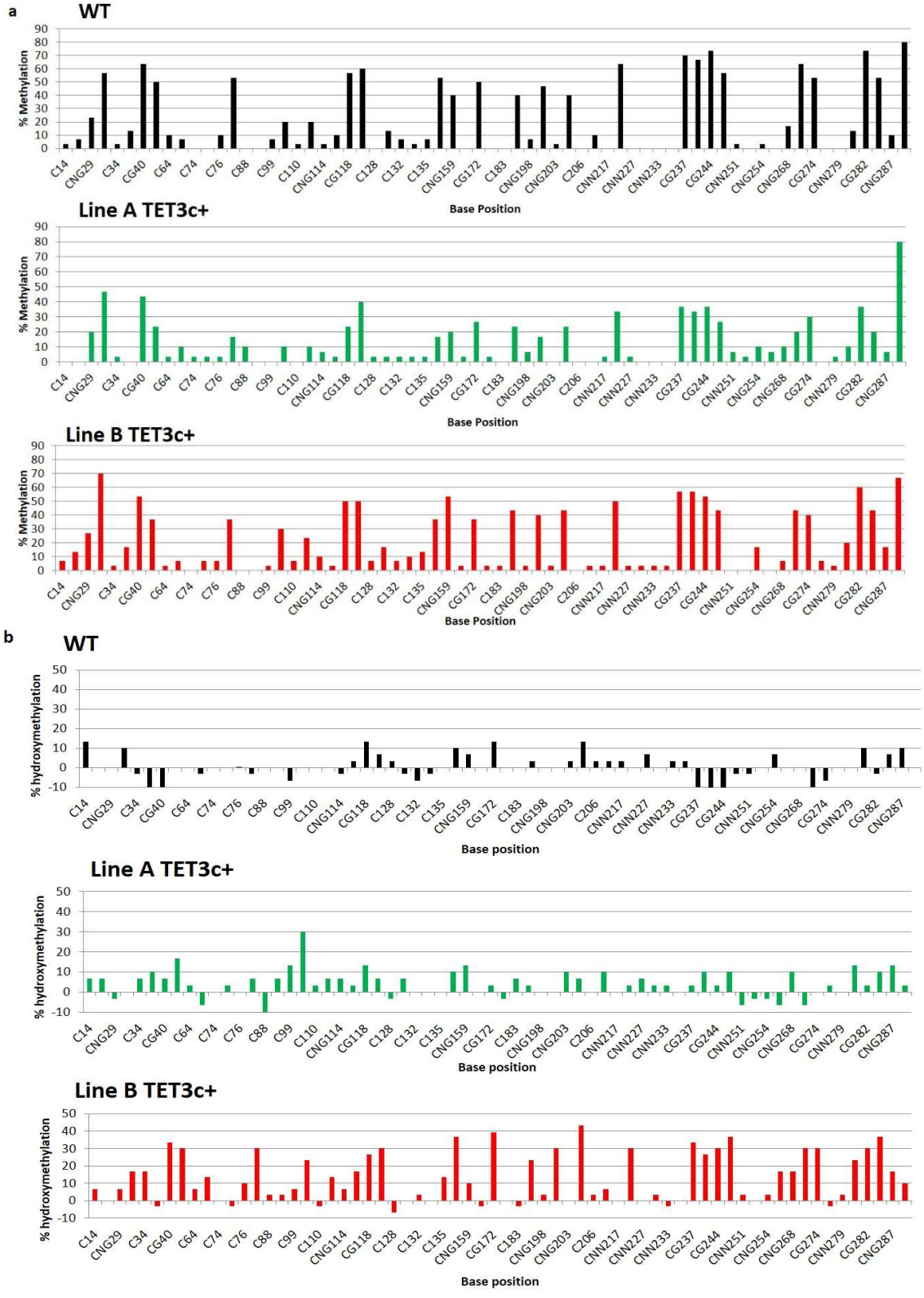


Figure 9.1. Methylation levels in a region of the 18S rDNA of Arabidopsis were altered in TET3c plants – shown site by site

(a) *Levels of 5-methylcytosine were reduced in Arabidopsis plants containing TET3c in Line A (Line A TET3c+) but did not change in Line B (Line A TET3c+). The % of 5-methylcytosine was calculated as the number of methylated clones divided by the total number of clones at each site. Methylation here has been shown on a site by site basis.*

(b) *5-hydroxymethylcytosine could be detected in Line A and Line B plants containing TET3c. These figures were calculated by subtracting the methylation level at each site in oxidative bisulfite sequencing from the methylation level at each site in standard bisulfite sequencing. This can result in negative values, which were used as an indication of the level of technical error.*

Figure 9.2

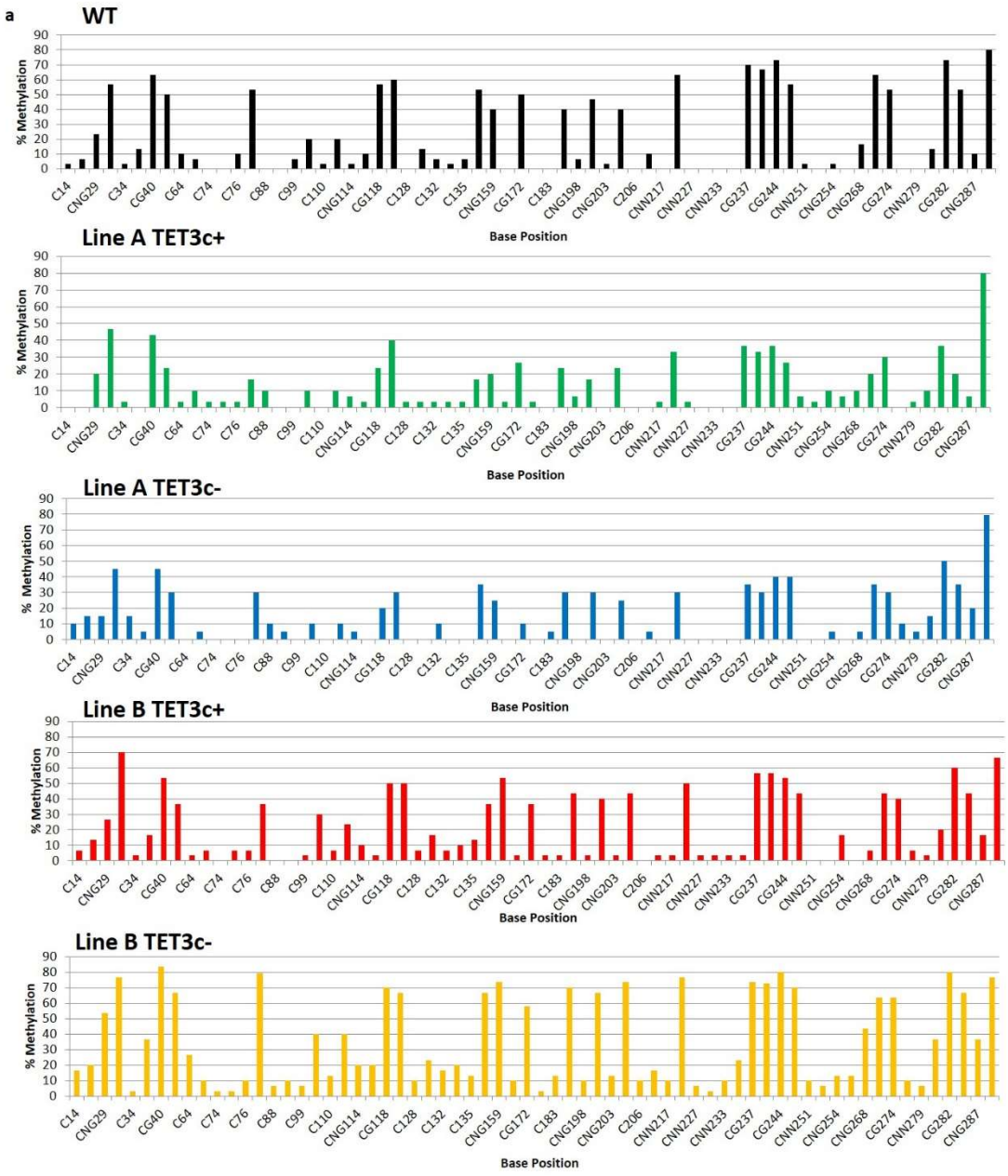


Figure 9.2. Methylation changes observed in Line A TET3c+ were heritable – shown on a site by site basis

(a) *The reduction in methylation seen in Line A TET3c+ was still present after the transgene had been segregated away (Line A TET3c-). In Line B, an increase of methylation was seen after the TET3c transgene was no longer present (Line B TET3c-). The % of 5-methylcytosine was calculated as the number of methylated clones divided by the total number of clones at each site. Methylation here has been shown on a site by site basis.*

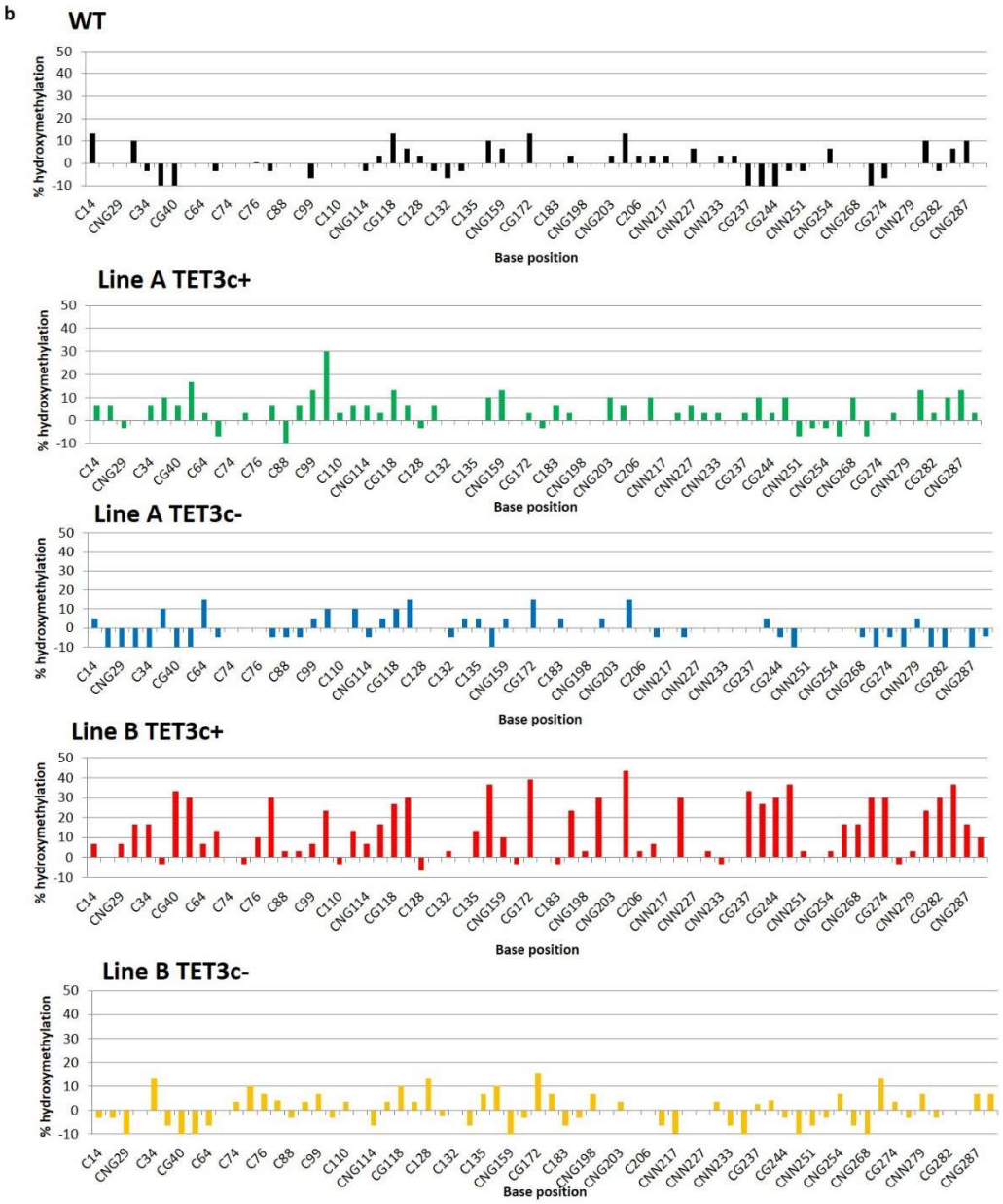


Figure 9.2. Methylation changes observed in Line A TET3c+ were heritable – shown on a site by site basis

(b) 5-hydroxymethylcytosine was not present in Line A and Line B plants after the TET3c transgene had been segregated away (Line A TET3c- and Line B TET3c-). These figures were calculated by subtracting the methylation level at each site in oxidative bisulfite sequencing from the methylation level at each site in standard bisulfite sequencing. This can result in negative values, which were used as an indication of the level of technical error.

Figure 9.3

Figure 9.3. Triple glycosylase mutants (ros1-3;dml2-1;dml3-1) with TET3c (rdd TET3c) and without (rdd) could still be demethylated by TET3c and showed increased levels of 5-fC – shown on a site by site basis

TET3c caused demethylation in rdd plants, to similar levels as that seen in Line A TET3c+. The % of 5-methylcytosine was calculated as the number of methylated clones divided by the total number of clones at each site. Methylation here has been shown on a site by site basis.

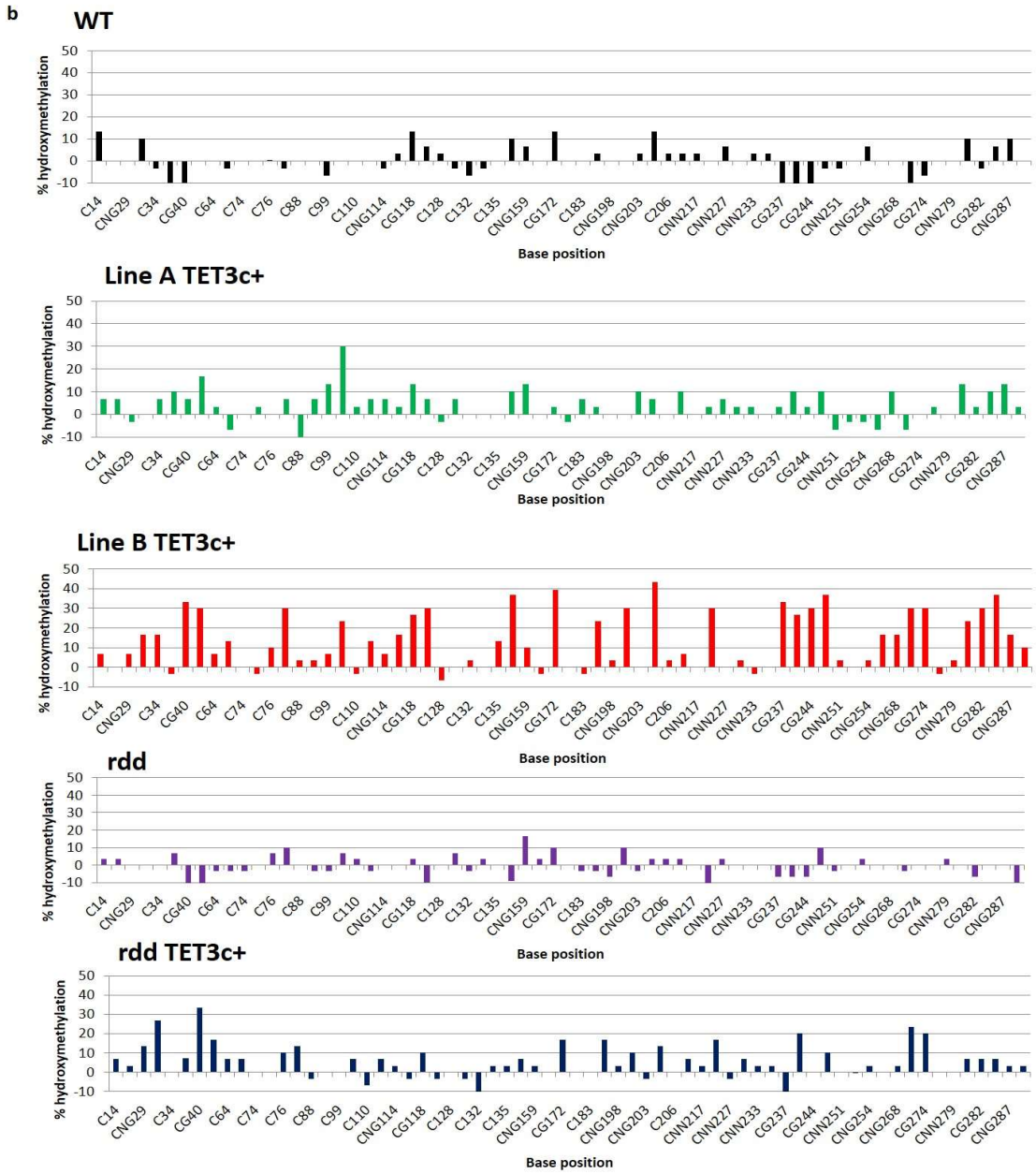


Figure 9.3. Triple glycosylase mutants (ros1-3;dml2-1;dml3-1) with TET3c (rdd TET3c) and without (rdd) could still be demethylated by TET3c and showed increased levels of 5-fC – shown on a site by site basis

(b) *Levels of 5-hydroxymethylcytosine were not increased in rdd plants in comparison to TET3c+ plants in the wild type background, suggesting that these glycosylases do not excise 5-hydroxymethylcytosine. These figures were calculated by subtracting the methylation level at each site in oxidative bisulfite sequencing from the methylation level at each site in standard bisulfite sequencing. This can result in negative values, which were used as an indication of the level of technical error.*

c

Figure 9.3. Triple glycosylase mutants (ros1-3;dml2-1;dml3-1) with TET3c (rdd TET3c) and without (rdd) could still be demethylated by TET3c and showed increased levels of 5-fC – shown on a site by site basis

(c) *5-formylcytosine was absent in all plants except rdd TET3c+ plants, where it was present only in the CG context. These figures were calculated by subtracting the methylation level at each site in standard bisulfite sequencing from the methylation level at each site in reduced bisulfite sequencing. This can result in negative values, which were used as an indication of the level of technical error.*

Figure 9.4

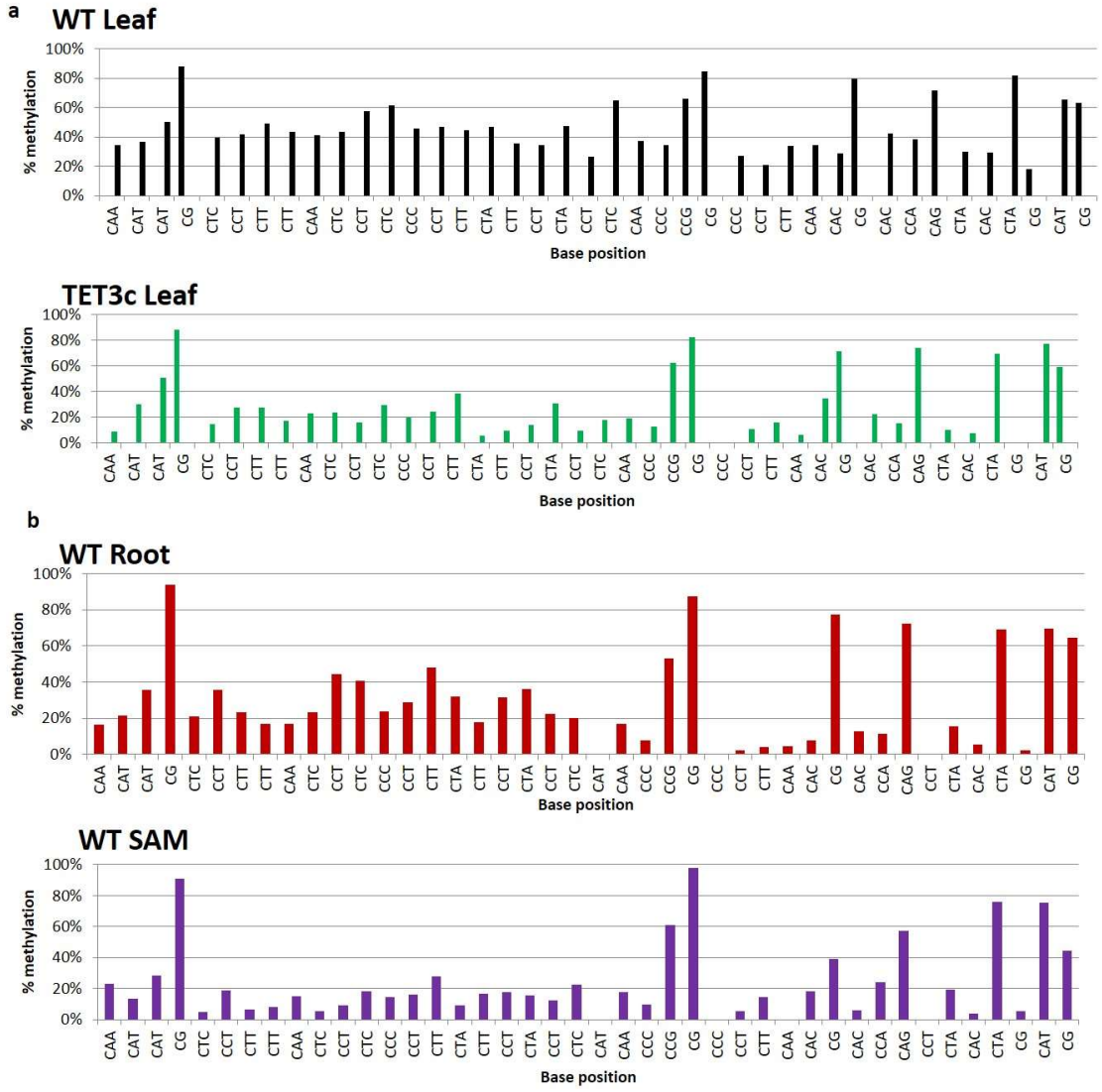


Figure 9.4. CEN1.1 was hypomethylated in TET3c lines and in the root and shoot apex – shown on a site by site basis

(a) Levels of 5-methylcytosine were reduced in a 206bp region upstream of the CEN1.1 transcriptional start site in TET3c tomato leaves compared to wild type tomato leaves.

(b) Levels of 5-methylcytosine were reduced in the same 206bp region upstream of the CEN1.1 transcriptional start site in both of the wild type tomato tissues which express the CEN1.1 gene (root and shoot apical meristem (SAM)).

The % of 5-methylcytosine was calculated as the number of methylated clones divided by the total number of clones at each site. Methylation here has been shown on a site by site basis. The first CAA site corresponds to -1837 in Figure 3.4c and the final CG site to -1631.

Figure 9.5

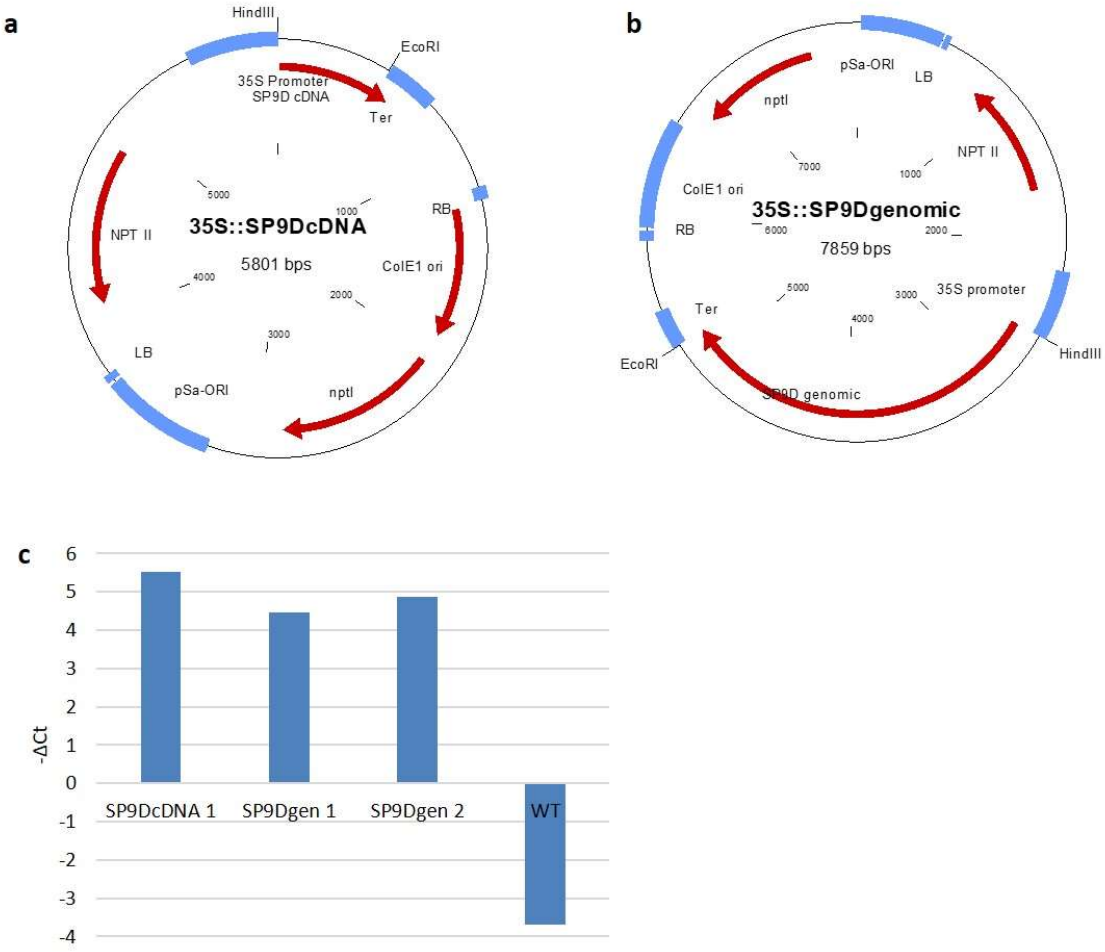


Figure 9.5. Additional constructs were made to investigate expression of the SP9D gene and expression of the transgene was confirmed

(a) Map of the plant expression vector pGreen 0029 35S containing the cDNA encoding the full length SP9D gene.

(b) Map of the plant expression vector pGreen 0029 35S containing the genomic DNA of the full length SP9D gene.

(c) T0 plants which contained the 35S::SP9Dgenomic or 35S::SP9DcDNA transgenes were shown to express it more strongly than the constitutively expressed eIF3-E using quantitative RT-PCR of cDNA. -ΔCt is calculated by subtracting $Ct_{eIF3\text{-}E}$ *from* Ct_{SP9D} *to control for sample concentration and calculate expression of a gene in comparison to eIF3-E.*

Figure 9.6

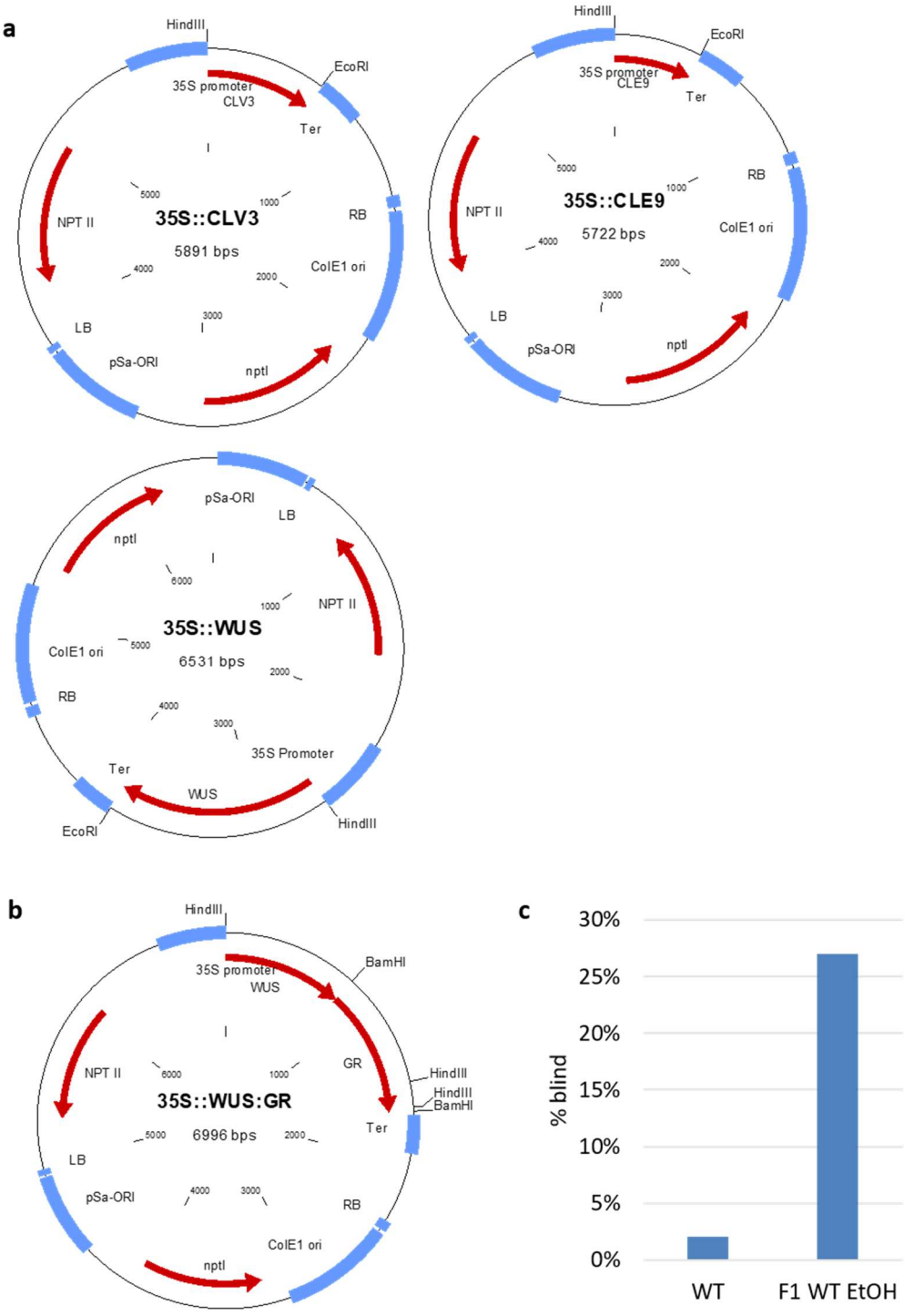

a
HindIII
EcoRI
35S promoter
CLV3
Ter
NPT II
35S::CLV3
5891 bps
RB
ColE1 ori
LB
pSa-ORI
nptI
HindIII
EcoRI
35S promoter
CLE9
Ter
RB
NPT II
35S::CLE9
5722 bps
ColE1 ori
LB
pSa-ORI
nptI
pSa-ORI
LB
nptI
NPT II
35S::WUS
6531 bps
ColE1 ori
RB
35S Promoter
Ter
WUS
EcoRI
HindIII
b
HindIII
BamHI
35S promoter
WUS
NPT II
GR
HindIII
HindIII
BamHI
35S::WUS:GR
6996 bps
Ter
LB
pSa-ORI
RB
ColE1 ori
nptI
c
30%
25%
20%
15%
10%
5%
0%
% blind
WT
F1 WT EtOH

Figure 9.6. Constructs were made to overexpress CLV3, CLE9 and WUS

(a) Maps of the plant expression vector pGreen 0029 35S containing the full length genomic DNA of CLV3, CLE9 and WUS, respectively.

(b) Map of the plant expression vector pGreen 0029 35S containing the genomic DNA of WUS fused to a glucocorticoid tag, allowing inducible control of the activity of the WUS protein.

(c) Offspring of wild type tomato plants exposed to ethanol vapour (F1 WT EtOH) showed an increased frequency of the blind phenotype in comparison to the unexposed wild type plants.