Over-expression of DNA methyltransferase MET1 in Arabidopsis creates novel epi-alleles with heritable expression states

Samuel George Philip Brocklehurst

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

The University of Leeds

School of Biology

September 2017
Intellectual Property and Publication Statement.

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgment.

© 2017 The University of Leeds and Samuel George Philip Brocklehurst
Acknowledgements

Firstly, I would like to thank my project supervisor Professor Peter Meyer for his invaluable advice and support, especially for organizing the transcription data. I will always be grateful for the time you have committed to me and the project.

I also wish to thank Iris Heidmann and her team, including Suzan Out and Marieke Ykeme. Their willingness to guide me during the collaboration made gathering data at Enza Zaden infinitely easier. I would especially like to thank Iris for the excellent pictures she captured for this thesis and her support while at Enza.

To the Bioinformatician Dr. Ian Carr for compiling all the transcription data and spending the time and energy required to explain the process and what the data means. Thank you.

Thank you to my fellow lab members past and present. Dr. Michael Watson for his guidance and patience, especially during the early years of my Ph.D. Associate Professor. Valya Vassileva, Dr. Matthew Gentry, Dr. Mohd Ahmad Mokhtar, Elizabeth Hollwey, Susan Marcus and Najiha Mohamed, whose knowledge and friendship, made working much more enjoyable.

I owe a great deal of thanks to my mum and sister Lorraine and Megan Brocklehurst for their love and encouragement. Most of all, thank you to my loving and patient partner Louise Frankland whose support during this Ph.D. is greatly appreciated.
Abstract

DNA methylation marks and histone modifications are important factors involved in regulating gene expression and genome structure. By destabilizing these vital factors we can create novel epi-alleles that are transgenerational. To investigate the potential of destabilizing an epigenetic function, we over-expressed DNA METHYLTRANSFERASE1 (MET1) in both Arabidopsis and tomato. In Arabidopsis thaliana, MET1 controls maintenance of cytosine methylation at symmetrical CG positions. At certain densely methylated loci, loss of MET1 causes the loss of all cytosine methylation marks. Over-expression of either the catalytically active or inactive versions of MET1 in Arabidopsis stochastically generates new epi-alleles at loci encoding transposable elements, non-coding RNAs, and proteins, which mainly results in increased expression. These altered expression states can be transmitted to the next generation, without the need for increased MET1 concentration, but long-term stability differs for individual loci. Destabilizing epigenetic factors in tomato appears to be more sensitive, causing lethality when levels of MET1 are increased at certain stages of development. The over-expression of MET1, or other epigenetic factors, offers an alternative strategy to create novel epi-alleles, identify phenotypes under epigenetic control and determine which genes are epigenetically regulated.
Table of Contents

1 General Introduction ................................................................................................................. 12
  1.1 An Introduction to Epigenetics ........................................................................................... 12
  1.2 DNA Methylation .................................................................................................................. 14
  1.3 De novo DNA methylation ................................................................................................... 15
  1.4 Maintenance of DNA methylation ....................................................................................... 18
  1.5 DNA Demethylation ............................................................................................................. 19
  1.6 Histones .................................................................................................................................. 21
  1.7 The biological function of DNA methylation ...................................................................... 23
  1.8 The structure and novel functions of MET1 ....................................................................... 25
  1.9 Thesis objective .................................................................................................................... 27

2 Over-expression of DNA methyltransferase MET1 in Arabidopsis thaliana generates new epialleles with heritable expression states ............................................................................... 29
  2.1 Introduction .......................................................................................................................... 29
  2.2 Results .................................................................................................................................. 31
    2.2.1 Generating MET1 over-expression lines ....................................................................... 31
    2.2.2 Phenotypic analysis of the MET1 over-expression lines .................................................. 33
    2.2.3 Transcript analysis of the MET1 over-expression lines .................................................... 38
    2.2.4 Bisulphite analysis of MET1 over-expression lines ......................................................... 42
    2.2.5 Histone analysis of the MET1 over-expression lines ....................................................... 45
    2.2.6 Heritability of epigenetic changes ................................................................................... 48
    2.2.7 Protein analysis of over-expressing MET1 ..................................................................... 52
    2.2.8 Genetic analysis of AT5G34850 ...................................................................................... 54
    2.2.9 Investigating the phenotypic effects of the four candidate genes ................................... 55
  2.3 Discussion .............................................................................................................................. 58

3 Investigating the application of over-expressing MET1 in tomato ........................................... 65
  3.1 Introduction ............................................................................................................................ 65
  3.2 Results .................................................................................................................................. 67
    3.2.1 Determining the structure of tomato MET1 ..................................................................... 67
    3.2.2 Phenotypic analysis of the tomato ArMET1 over-expression lines ................................... 68
    3.2.3 Transcript analysis of MET1 over-expression in tomato ................................................... 73
  3.3 Discussion .............................................................................................................................. 78

4 General discussion .................................................................................................................... 80
  4.1 Over-expression of MET1 induces heritable epigenetic diversity ........................................ 80
  4.2 Investigating the tomato epigenome .................................................................................... 82
  4.3 Outlook and open questions ............................................................................................... 83
Materials and Methods .......................................................................................................................... 86

5.1 Materials .......................................................................................................................................... 86
5.1.1 Arabidopsis material ..................................................................................................................... 86
5.1.2 Solanum lycopersicum material ................................................................................................ 87
5.1.3 Bacterial strains .......................................................................................................................... 87

5.2 Primer list ......................................................................................................................................... 87

5.3 Construction of plasmids and plant transformation ........................................................................ 89
5.3.1 Over-expression Constructs ........................................................................................................ 89
5.3.2 Arabidopsis transformation by floral dip ................................................................................. 89
5.3.3 Leaf disc transformation of Solanum lycopersicum ................................................................... 90

5.4 Agrobacterium protocols ................................................................................................................ 91
5.4.1 Agrobacterium tumefaciens GV3101::pMP90 electro-competent cells ..................................... 91
5.4.2 Binary plasmid electroporation Agrobacterium ...................................................................... 91
5.4.3 Isolation of plasmid DNA from Agrobacterium tumefaciens GV3101::pMP90 ....................... 91

5.5 E. coli protocols .................................................................................................................................. 92
5.5.1 Preparation of chemically competent E. coli ............................................................................. 92
5.5.2 Heat-shock transformation ........................................................................................................ 93
5.5.3 Mini-prep isolation of plasmid DNA from E.coli .................................................................... 93

5.6 Phenotypic analysis ........................................................................................................................... 94
5.6.1 Arabidopsis phenotyping ............................................................................................................ 94
5.6.2 Tomato phenotyping .................................................................................................................. 94

5.7 DNA Protocol .................................................................................................................................... 95
5.7.1 Isolation of genomic DNA from Arabidopsis ............................................................................ 95
5.7.2 Restriction digest ....................................................................................................................... 95
5.7.3 Ligation reaction ......................................................................................................................... 95
5.7.4 Polymerase chain reaction (PCR) ............................................................................................. 96
5.7.5 Bisulphite analysis .................................................................................................................... 96
5.7.6 DNA sequencing ....................................................................................................................... 97
5.7.7 Data analysis ............................................................................................................................. 97

5.8 RNA analysis .................................................................................................................................... 98
5.8.1 Isolation of RNA from plants ..................................................................................................... 98
5.8.2 Semi-quantitative PCR ................................................................. 98
5.8.3 Quantitative RT-PCR assay .......................................................... 99
5.8.4 Sequencing and data analysis ...................................................... 99
5.9 Protein analysis ............................................................................. 102
  5.9.1 ChIP assay ............................................................................. 102
  5.9.2 Western blot assay .................................................................. 102
6 References ....................................................................................... 103
7 Supplementary data ......................................................................... 122
  Supplementary Table 1 ..................................................................... 139
  Supplementary Table 2 ..................................................................... 141
  Supplementary Table 3 ..................................................................... 147
  Supplementary Table 4 ..................................................................... 148
  Supplementary Table 5 ..................................................................... 160
  Supplementary Table 6 ..................................................................... 163
  Supplementary Figure 1 ................................................................... 164
  Supplementary Figure 2 ................................................................... 165
  Supplementary Figure 3 ................................................................... 166
### List of Abbreviations.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-aza</td>
<td>5-Azacytidine</td>
</tr>
<tr>
<td>A</td>
<td>Adenine</td>
</tr>
<tr>
<td>AGO</td>
<td>Argonaute</td>
</tr>
<tr>
<td>ArMET1</td>
<td>Arabidopsis Methyltransferase 1</td>
</tr>
<tr>
<td>BAH</td>
<td>Bromo-Adjacent homology</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>C</td>
<td>Cytosine</td>
</tr>
<tr>
<td>CaMV</td>
<td>Cauliflower Mosaic Virus</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary Deoxyribonucleic acid</td>
</tr>
<tr>
<td>CHD</td>
<td>Chromo-Domain</td>
</tr>
<tr>
<td>CLE3</td>
<td>Clavata 3</td>
</tr>
<tr>
<td>CLE9</td>
<td>Clavata 9</td>
</tr>
<tr>
<td>CMT2</td>
<td>Chromomethyltransferase 2</td>
</tr>
<tr>
<td>CMT3</td>
<td>Chromomethyltransferase 3</td>
</tr>
<tr>
<td>Col</td>
<td><em>Arabidopsis</em> Columbia</td>
</tr>
<tr>
<td>DCL</td>
<td>Dicer-Like</td>
</tr>
<tr>
<td>DCM</td>
<td>DNA Cytosine Methyltransferase</td>
</tr>
<tr>
<td>DDM1</td>
<td>Decrease in DNA Methylation 1</td>
</tr>
<tr>
<td>DME</td>
<td>Demeter</td>
</tr>
<tr>
<td>DML</td>
<td>Demeter-Like</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
</tbody>
</table>
DNMT1  DNA Methyltransferase 1
DNMT3A  DNA Methyltransferase 3 A
DNMT3B  DNA Methyltransferase 3 B
DNMT3L  DNA Methyltransferase 3-Like
DRM2  Domains Re-Arranged Methyltransferase 2
ds  double stranded
E.coli  Escherichia coli
EDTA  Ethylenediaminetetraacetic acid
FWA  Flowering Wageningen
G  Guanine
GUS  β-Glucuronidase
H3K4me3  Histone 3 lysine 4 trimethylation
H3K9me2  Histone 3 lysine 2 dimethylation
H3K27ac  Histone 3 lysine 27 acetylation
H3K27me3  Histone 3 lysine 27 trimethylation
H4ac  Histone 4 acetylation
HDA6  Histone Deacetylase 6
IAA  Isoamyl Alcohol
KTF1  Kow Domain-Containing Transcription Factor 1
LB  liquid lysogeny broth
LINE  Long Interspersed Nuclear Element
mC  5-methyl-cytosine
MEA  MEDEA
<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MET1</td>
<td>Methyltransferase 1</td>
</tr>
<tr>
<td>met1-1RE</td>
<td>MET1 Restored</td>
</tr>
<tr>
<td>METo</td>
<td>Methyltransferase 1 over-expressioning</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger Ribonucleic acid</td>
</tr>
<tr>
<td>MS</td>
<td>Murashige and Skoog</td>
</tr>
<tr>
<td>NASC</td>
<td>Nottingham Arabidopsis Stock Centre</td>
</tr>
<tr>
<td>nc</td>
<td>non-coding</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>PAP26</td>
<td>Purple Acid Phosphatase 26</td>
</tr>
<tr>
<td>PCNA</td>
<td>Proliferating Cell Nuclear Antigen</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PEV</td>
<td>position effect variegation</td>
</tr>
<tr>
<td>PGC</td>
<td>Primordial Germ Cells</td>
</tr>
<tr>
<td>pi</td>
<td>piwi-interacting</td>
</tr>
<tr>
<td>POLII</td>
<td>RNA Polymerase II</td>
</tr>
<tr>
<td>POLIV</td>
<td>RNA Polymerase IV</td>
</tr>
<tr>
<td>POLV</td>
<td>RNA Polymerase V</td>
</tr>
<tr>
<td>PTGS</td>
<td>Post-Transcriptional Gene Silencing</td>
</tr>
<tr>
<td>RdDM</td>
<td>RNA-directed DNA Methylation</td>
</tr>
<tr>
<td>rDNA</td>
<td>ribosomal DNA</td>
</tr>
<tr>
<td>RDR2</td>
<td>RNA-Dependent RNA Polymerase 2</td>
</tr>
<tr>
<td>RDR6</td>
<td>RNA-dependent RNA polymerase 6</td>
</tr>
</tbody>
</table>
RISC  RNA-Induced Silencing Complex
RNA  Ribonucleic acid
ROS  Repressor of Silencing
RPS  Repetitive Petunia Sequence
SAM  S-adenosyl methionine
SaM  Shoot apical Meristem
SDS  Sodium Dodecyl Sulphate
si  small interfering
SNP  Single Nucleotide Polymorphism
SP9D  Self-Pruning 9D
sq  semi-quantitative
SRA  Set and Ring Associated
sup  superman
SUVH4  Suppressor of Variegation 3-9 Homologue 4
T  Thymine
TAE  Tris-acetate-EDTA
T-DNA  Transfer-DNA
TE  Transposon
TGS  Transcriptional Gene Silencing
SlMET1  tomato Methyltransferase 1
VCN  Vegetative Cell Nucleus
WT  Wildtype
WUS  Wuschel
1 General Introduction

1.1 An Introduction to Epigentics

From the beginning when DNA was first identified as a distinct molecule by the Swiss physician Friedrich Miescher in 1869 (Dahm, 2005), to the discovery of the molecular structure and function of DNA in 1953 by Watson and Crick (Watson and Crick, 1953), DNA has established itself as the universal language of life. This led to the founding of the central dogma of molecular biology which describes the transfer of information from DNA to RNA to protein, giving rise to a wide variety of functions. This simplistic view led many scientists to focus solely on protein encoding genes, disregarding over 90% of the human genome leading to the term junk DNA which was believed to have no function or purpose. However, new findings have brought to light the importance of DNA structure, not just DNA sequence, in determining gene expression and development. The importance of genotype determining phenotypic change during development and the interaction between the two led to the coining of the term ‘epigenetics’ (Waddington, 1942). At the time, however, it was not yet fully understood the significance of epigenetic control and its wider implications in cell development. One of the most powerful metaphors used to illustrate cellular development is the epigenetic landscape (Fig 1.1) introduced by Conrad Waddington (Waddington, 1957).
Figure 1.1: Waddington’s classic model of an epigenetic landscape. At the top of the landscape is a ball which represents an undifferentiated totipotent cell, the troughs and peaks below portray the countless developmental pathways that the ball can traverse. The further the ball travels down a specific pathway, the more established it becomes in a particular developmental route until the cell becomes fully committed to its fate.

Over time the term epigenetics has been redefined as the study of internal factors and external stimuli leading to a heritable change in gene expression that is independent of alterations in the DNA sequence (Dupont et al., 2009). This can occur as post-transcriptional gene silencing (PTGS) which involves the cleavage of mRNAs directed by short interfering RNA (siRNAs), or as transcriptional gene silencing (TGS). The mechanism of TGS is primarily regulated by the structural organisation of the genome. In eukaryotes, the genomic structure is comprised of an eight histone protein wound by DNA twice to form a nucleosome. This, in turn, forms the fundamental repeating unit that structures the genome. Covalent marks such as acetylation and methylation at the N-terminal tails of histone proteins have the ability to alter the structural conformation of chromatin into a tightly packed and repressed heterochromatic state, or a more open transcriptionally active euchromatic state. Methylation marks are another important layer of epigenetic control which involves the transfer of a methyl group (-CH$_3$) from S-adenosyl methionine (SAM) to a cytosine residue. These epigenetic modifications must adhere to two fundamental rules. The alteration must be revisable in the sense that the original structure can be recovered; conversely, the modifications should be heritable so that the epi-alleles generated are stable following cell division, and even span multiple generations. The ability of epigenetic modifications to work independently or coordinate between each other is essential genome regulation. This includes position effect variegation (PEV) (Muller, 1930), paramutations (Brink, 1959), transgene silencing (Napoli et al., 1990; Meyer et al., 1992) and imprinting (Dechiara et al., 1991).
1.2 DNA Methylation

From Prokaryotes to Eukaryotic organisms the genetic code is highly conserved. However, the epigenetic strategies implemented to regulate the genome are often very different. For example, DNA methylation in bacteria occurs at nitrogen 6 of adenine (\(^6\text{m}A\)), carbon 5 of cytosine (\(^5\text{m}C\)) and in some cases nitrogen 4 of adenine (\(^4\text{m}A\)). DNA methylation in bacteria is used as a defence mechanism against foreign DNA. The methylated genome prevents cleavage from endonucleases, unlike unmethylated invading phage genomes which are readily degraded (Casadesus & Low, 2006).

In eukaryotes, the most prominent type of methylation is cytosine methylation, which in vertebrates accounts for nearly 100% of all methylated loci. In the mammalian system, methylation is found exclusively at the carbon 5 of cytosine, and this epigenetic modification is present throughout the whole genome. At gene promoters (Larsen et al., 1992) and origins of replication (Antequera & Bird, 1999) unmethylated regions can be found with higher than average CG content called CpG islands (Cooper et al., 1983) which function as a genomic platform for regulating transcription. In mammals, DNA methylation is located at cytosines in a CG sequence context, with the exception of embryonic stem cells, where methylation is found at cytosines in CA and CT sequence contexts (Ramsahoye et al., 2000).

In plants, DNA methylation can occur in CG, CHG and CHH sequence contexts (where H represents A, C or T). Establishment and maintenance of methylation in these three specific contexts requires the particular action of one of three methyltransferases; \textit{DOMAINS REARRANGED METHYLTRANSFERASE 2} (DRM2), \textit{METHYLTRANSFERASE 1} (MET1) and \textit{CHROMOMETHYLASE 3} (CMT3). DRM2 is a homologue of the mammalian DNA methyltransferase DNA METHYLTRANSFERASE 3 (DNMT3) and is responsible for de novo methylation in all sequence contexts (Cao et al., 2003). MET1 is a homolog of the mammalian DNA methyltransferase DNA METHYLTRANSFERASE 1 (DNMT1), that is responsible for maintenance of CG methylation during DNA replication (Jones et al., 2001). CMT3 is another maintenance methyltransferase that is plant specific and is required for CHG methylation (Lindroth et al., 2001). While these enzymes seem to have specific roles, there are studies that demonstrate their ability to coordinate or compete at specific
genomic regions. One example would be the maintenance of non-CG methylation which is carried out by both CMT3 and DRM2 via small interfering RNA (siRNA) (Cao et al., 2003). In a met1 mutant, methylation is lost in all three sequence contexts (Stroud et al., 2013). Though loss of CHG methylation is possibly due to the loss of chromatin marks required for CMT3 binding (Soppe et al., 2002; Tariq et al., 2003), the loss of CHH marks remain unclear, but may reflect a failed interaction between CMT3 and DRM2, or a direct role of MET1 in CHH methylation targeting.

1.3 De novo DNA methylation

*De novo* methylation is described as the addition of a methyl group to a cytosine base, that had been previously unmethylated. It was first identified from studies in tobacco when an RNA viroid containing a high level of secondary structures was shown to direct methylation to homologous transgenic sequences (Wassenegger et al., 1994). In mammals, there is an escalation of *de novo* methylation that occurs during two stages of development. During early embryogenesis, there is a universal removal of epigenetic markers shortly after fertilisation. This mechanism allows for a wave of *de novo* methylation to generate new methylation patterns (Reik et al., 2001), and silence one copy of the X-chromosome in females (Goto & Monk, 1998). During gametogenesis, there is a significant increase in the levels of *de novo* methylation which establishes DNA methylation at transposable elements (TE) and imprinted genes (Chen et al., 2003). The majority of the *de novo* methylation processes are catalysed by DNA METHYLTRANSFERASE 3A (DNMT3A) and DNA METHYLTRANSFERASE 3B (DNMT3B) which are guided by PIWI-like ARGONAUTE (AGO) proteins and PIWI-interacting RNA (piRNA). Although DNMT3A & B play different roles, they both have the ability to catalyse both CG and non-CG methylation. The major role of DNMT3A is the establishment of genomic imprinting, while DNMT3B serves to catalyse *de novo* methylation of repetitive centromeric DNA to increase the stability of the genome.
Unlike mammals, plants do not undergo an extensive reprogramming of DNA methylation but appear to have a stable methylation pattern. Plants have two distinct pathways which they can employ to target cytosine residues for methylation. The two pathways that control de novo methylation are the RNA-directed DNA methylation pathway (RdDM) and the independent pathway. DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), a homologue of the mammalian methyltransferase DNMT3A is highly regulated by the RdDM pathway which is responsible for de novo methylation. The RdDM pathway utilises siRNA in order to target de novo methylation in a sequence-specific manner to regulate various epigenetic activities. The generation of a diverse range of siRNAs for targeting requires a variety of plant-specific proteins and enzymes. Precursor dsRNA is generated by plant-specific RNA-DEPENDENT RNA-POLYMERASE 2 (RDR2) which syntheses the complementary strand of RNA POLYMERASE IV (PolIV) transcripts. Furthermore, dsRNA can be produced by the RNA POLYMERASE II (PolII) which produces transcripts derived from inverted repeats or over-lapping antagonistic regions. The generated dsDNA is recognised by DICER-LIKE (DCL) for cleavage into siRNAs. In total plants contain four DCL proteins, the first being DCL1 which is associated with the production of microRNAs (miRNA) (Henderson et al., 2006). DCL3 is the primary protein involved in cleavage of dsRNA into 21-24nt siRNAs, although DCL2 and DCL4 have some overlapping function (Xie et al., 2004). These 21-24nt, small RNA duplexes are in turn stabilised by ribose methylation at the 3’ terminal catalysed by the enzyme HEN1 methyltransferase (Yu et al., 2005). It is believed this modification acts as a two-fold function: firstly to protect the sRNAs from degradation by exonucleases, and secondly, it may help with recognition and loading onto ARGONAUTE (AGO) complexes (Yu et al., 2005). In Arabidopsis, there are ten AGO proteins present, of these, three (AGO4, AGO6 and AGO9) are known to associate with 24nt siRNAs and have a preference for sRNAs with a 5’ terminal adenosine (Mi et al., 2008; Havecker et al., 2010). The mature siRNA is loaded onto an AGO4- RNA-induced silencing complex (RISC) or AGO6-RISC (Li et al., 2006) which is facilitated by KOW DOMAIN-CONTAINING TRANSCRIPTION FACTOR 1 (KTF1) (He et al., 2009). The AGO4 complex is then targeted to a specific site through the use of the PolV scaffold complexes (Wierzbicki et al., 2009), which allows the AGO4 complex to direct DRM2 to the particular site (Wierzbicki et al., 2009), for targeted de novo methylation (Fig. 1.2).
Figure 1.2: RNA dependent DNA de novo methylation. Pol IV transcribes ssRNA that is copied into a dsRNA by RDR2. DCL3 processes the dsRNA into 24-nucleotide siRNAs. HEN 1 then methylates the 3' ends of the siRNA to be incorporated into the AGO4-RISC complex. Pol V transcribes a scaffold that base-pairs with the AGO4-bound siRNAs. The AGO4-RISC complex is recruited via Pol V to create a complex with KTF1, which binds the non-coding nascent PolV transcripts. AGO4 links DRM2, which catalyses de novo methylation of DNA.

Despite, RdDM pathway's ability to direct de novo methylation at all three-sequence context, there is an alternate path independent of siRNAs that can de novo methylate at CHH sites.
This alternative pathway requires CMT2 (Zemach et al., 2013). It is hypothesised that the targeting of CMT2 to CHH regions occurs through interaction with the dimethylation of lysine 9 on histone 3 (H3K9me2) via its chromodomain and not siRNA (Pikaard, 2013).

1.4 Maintenance of DNA methylation

Once methylation has been established, it must be maintained. In mammals, the responsibility of CG methylation maintenance is carried out by DNMT1. DNMT1’s primary function is the maintenance of DNA methylation lost during semi-conservative replication. To ensure the faithful duplication of methylation patterns, a range of histone modifications and proteins have evolved allowing DNMT1 to actively associate with the replication foci during S-phase (Fig 1.3). DNMT1 functions to restore hemimethylated DNA during semi-conservative replication to its fully methylated state (Kim et al., 2009). Studies have shown that the C-terminal catalytic domain of DNMT1 is recruited to the replication fork via PROLIFERATING CELL NUCLEAR ANTIGEN (PCNA) elements found at the replication foci (Chuang et al., 1997). Other chromatin-associated proteins, such as the multi-domain protein UHRF1 assist in the stable association of DNMT1 to hemimethylated DNA during replication. The PCNA/ DNMT1 interaction along with the activity of UHRF1 enhances DNMT1 affinity for hemimethylated DNA. This function allows DNMT1 to recognise and replicate the methylation pattern from the parent strand to the daughter strand as they are generated during semi-conservative replication (Mortusewicz et al., 2005).

Like DNMT1 the primary function of MET1 is the maintenance of methylation patterns during semi-conservative replication. The mechanism by which MET1 targets the replication foci is believed to be similar in manner to DNMT1 targeting (Fig 1.3). MET1 binds to hemimethylated DNA mediated by VARIANT IN METHYLATION (VIM) proteins. These contain the Set and Ring associated (SRA) domain that binds methylated Cytosine residues (Woo et al., 2008). The VIM1 protein also interacts with histone modifications (H2B, H3, H4, and HTR12) suggesting that the protein acts as a DNA methylation-histone interface. At heterochromatic regions, an SWI/SNF family chromatin remodelling factor, DECREASE IN DNA METHYLATION 1 (DDM1) allows access to the target area allowing MET1 to bind (Woo
et al., 2007). Once MET1 has attached, the parent DNA strand is utilised during semi-conservative replication to reproduce the methylation pattern onto the newly synthesised daughter strand.

Figure 1.3: Model is depicting the maintenance of CG methylation during replication in both plants and mammals. DNMT1 is believed to be recruited to the replication foci through UHRF1 which specifically interacts with hemimethylated DNA as well as with PCNA. After being recruited, DNMT1 functions to maintain methylation patterns by restoring the hemimethylated DNA to a fully methylated state. In plants, MET1 and the VIM family of SRA domain proteins, which are homologues of DNMT1 and UHRF1, respectively, are likely to function in a similar manner to maintain CG methylation patterns.

1.5 DNA Demethylation

Epigenetic modifications must be reversible, but for such an important role there is little documentation how active demethylation is carried out and the underlying mechanisms that are involved. There are two ways methylation can be removed; passive demethylation and active demethylation. Passive demethylation of genomic regions occurs via the loss of
function of maintenance enzymes, leading to a global loss of methylation following cell division. An example of passive DNA demethylation can be observed in MET1 mutants which possess a global loss of methylation in all three sequence contexts (Kankel et al., 2003). In plants, active demethylation removes methylated cytosines (5-mC) directly excised via the base excision repair process carried out by REPRESSOR OF SILENCING 1 (ROS1), DEMETER (DME), DEMETER like 2 (DML2) and DML3 (Penterman et al., 2007; Zhu, 2009). Each of the four enzymes can remove 5-mC in all sequence contexts, although preferential removal of specific substrates is debated (Agius et al., 2006; Gehring et al., 2006; Morales-Ruiz et al., 2006; Penterman et al., 2007). In addition to hydrolysing the glycosidic bond between the cytosine and sugar-phosphate backbone, they can also nick the DNA backbone at the abasic site via apurinic/apyrimidinic (AP) lyase activity (Bruner et al., 2000; Jiricny, 2002). During removal of 5-mC, these enzymes simultaneously glycosylate the 5-methylcytosine base and cleave the DNA backbone, generating a single-nucleotide gap that induces the base excision repair pathway, eventually refilling it with an unmethylated cytosine.

In vegetative cells, ROS1, DML2 and DML3 remove DNA methylation at the 5’ and 3’ ends of genes in a sequence-specific context (Penterman et al., 2007). ROS1 offsets the DNA methylation established by RdDM, to prevent hypermethylation and the spreading of DNA methylation via the self-reinforcement mechanisms that could lead to adverse gene silencing. DNA methylation and DNA methyltransferases may act as feedback mechanisms that play a role in establishing and regulating ROS1 demethylation (Mathieu et al., 2007; Williams et al., 2015). When DNA methylation levels are reduced in a met1 mutant or treated with DNA methylation inhibitors, ROS1 levels are down-regulated (Mathieu et al., 2007). However, it is still not fully understood how plant DNA glycosylases are targeted to specific sequences. One possibility is a RNA-based targeting system. It has been demonstrated that ROS1 co-localises with REPRESSOR OF SILENCING 3 (ROS3), which possesses RNA binding capacity (Zheng et al., 2008).

During gametogenesis, DME is responsible for maternal allele DNA demethylation in the endosperm that establishes gene imprinting (Gehring et al., 2006). Maternal DME also has the ability to change the endosperm methylation landscape completely, demethylating TEs and repeat sequences (Hsieh et al., 2009). In the central cell during female gametogenesis, the levels of MET1 are reduced (Jullien et al., 2008), and DME levels are increased (Choi et
al., 2002), culminating in the global loss of DNA methylation via both passive and active demethylation.

Although demethylation directly removes methylation modifications, it may also lead to modifications of the chromatin before or during the removal of 5-mC. A histone acetyltransferase INCREASED DNA METHYLATION 1 (IDM1) was identified as a regulator of DNA demethylation in Arabidopsis (Qian, W. et al., 2012). IDM1 recognises chromatic regions with CG methylation and low histone 3 lysine 4 (H3K4). ROS1 and IDM1 interact with each other and appear to function in the same pathway for DNA demethylation, albeit only a portion of ROS1 targets are regulated by IDM1. Currently, it is uncertain how ROS1 is recruited to the modified chromatin, but the discovery that there is interplay between IDM1 and ROS1 suggests that histone modifications and chromatin structure may play a role in active DNA demethylation. Maintenance of CHG methylation can be impaired by INCREASE IN BONSAI METHYLATION 1 (IBM1), which is a histone demethylase that prevents H3K9 methylation (Saze et al., 2008; Miura et al., 2009), further supporting this theory.

1.6 Histones

Chromatin is made up of the basic structural unit, the nucleosome, which consists of a 146-bp fragment of DNA that’s wrapped almost twice around a protein octamer which is comprised of a histone H3-H4 tetramer and two H2A-H2B dimers (Kornberg, 1974; Richmond et al, 1999; Wolffe and Guschin, 2000). Each histone protein possesses an N-terminal tail which allows for a variety of modifications such as methylation, acetylation, ubiquitination or phosphorylation of specific amino acids which can alter the conformational state of chromatin, essential for gene regulation. An example of a histone modification creating a conformational change is acetylation of lysine residues. One of the primary ‘active’ modifications is acetylation of the N-terminal tail which neutralises the positive charge on the ε amino group which acts as a binding site for the bromodomain-containing proteins, leading to a more open and active euchromatic state (Taverna et al., 2007).

Methylation of lysine residues are a more complex epigenetic mark, that can either silence transcription or activate chromatin domains, depending on which lysine residues are methylated and the degree of methylation. The overall charge of the histone tail is not
affected by lysine methylation, but an increase in hydrophobicity may allow intra- or intermolecular interactions between proteins or the modification alters the conformational structure allowing proteins to bind preferentially to the methylated domain (Du et al., 2012; Lanouette et al., 2014). Typically, histone H3K9 and H3K27 methylation are associated with repressed regions, whereas H3K4 methylation is associated with a more active open chromatin structure (Berger, 2007). In Arabidopsis, SET domain proteins are responsible for modulating lysine methylation, which can be classified into four distinct groups: SU(VAR)3–9, ENHANCER OF ZESTE (E(Z)) homologs, TRITHORAX (TRX) groups, and ABSENT, SMALL, OR HOMEOTIC DISCS 1 (ASH1) (Baumbusch et al., 2001).
Figure 1.4: The different conformational structures of chromatin. A) Euchromatic structure of histones with the active modification H3K4me3 at the N-terminal tail. B) Heterochromatic structure of histone with the repressive marks H3K9me2 and H3K27me3.

It has been well documented that there is a complex interplay between DNA methylation and the methylated H3K9 modification. In kyp/suvh4 mutants, H3K9me2 is reduced, leading to a loss of CHG methylation catalysed by CMT3 (Jackson et al., 2002). Transposable elements were also identified as common targets for both KYP/SUVH4 and CMT3 using expression profiling (Tran et al., 2005), and CHROMATIN IMMUNOPRECIPITATION (ChIP) analysis revealed a high correlation between H3K9me2 and CHG methylation (Bernatavichute et al., 2008). These results indicate that maintenance of non-CG methylation requires H3K9 methylation and is critical for preservation of genome stability and transcriptional repression. DNA methylation can also reinforce histone methylation in a positive feedback loop. In a met1 mutant where CG levels are diminished, a decrease in H3K9me2 is observed at 180bp centromeric repeats and transposable elements (Johnson et al., 2002; Tariq et al., 2003).

1.7 The biological function of DNA methylation

DNA methylation is essential for a number of cellular functions including gene expression, cell differentiation, regulation of transposable elements and even plant immunity (Choi et al., 2002; Chan et al., 2006; Agorio and Vera, 2007). DNA methylation present at gene promoters are commonly associated with transcriptional silencing, by directly obstructing transcription factors, and recruiting methyl-binding proteins that can modify the histone tail altering the chromatin structure. Methylation of DNA in all three-sequence context “dense methylation” is often found at TEs to prevent activation and mobilisation, which would disrupt the genome integrity by inserting into critical regions. One example would be the transposable element ATGP3, a class I TE in the gypsy family, which is not active in either a cmt3 or met1 mutant, but does occur in the ddm1 and met1 cmt3 double mutant.
(Tsukahara et al., 2009). This observation suggests a redundant function of CG and non-CG methylation in the transcriptional silencing of the TEs.

TEs should not be seen as just mobile deleterious mutagens, as they can also have positive regulatory roles. At regions with hypermethylated TEs, a lower meiotic recombination rate is observed compared to hypomethylated low copy number genes, which indicates that DNA methylation can influence the rate of recombination (Melamed-Bessudo & Levy, 2012). In the ddm1 hypomethylated mutant’s recombination rates were analysed in both euchromatin and heterochromatin. It was discovered that the rate of recombination between markers located in euchromatin increased, whereas recombination went unchanged between markers located in heterochromatin. Interestingly it is heterochromatic regions that are most affected by the loss of methylation in the ddm1 mutant; this suggests that DNA methylation may only act as a repressor of meiotic recombination at euchromatic regions where the chromatin is tightly packed (Melamed-Bessudo & Levy, 2012).

Genomic imprinting a mechanism by which genes are expressed in a parent-specific manner is also mediated by DNA methylation. Imprinting is present throughout the plant kingdom and occurs in the endosperm during seed development (Jahnke & Scholten, 2009). There are many examples of imprinted genes in plants including the Arabidopsis MEA, FIS2, FIE (Luo et al., 2000) and FWA (Kinoshita et al., 2004) genes, which are all expressed from the maternal alleles of the endosperm. FWA is an imprinted gene that is by default, silenced via methylation (Choi et al., 2002). Expression of FWA during endosperm development occurs via maternal-specific activation that is dependent on the removal of methylation by DME. However, the loss of MET1 activity induces ectopic FWA expression causing late flowering. This observation indicates that maintenance of endosperm-specific and parent of origin-specific FWA expression depends on MET1. If hypomethylation of FWA occurs in embryonic lineages, the fwa epigenetic mutation and its associated late-flowering phenotype can be stably inherited over many generations.

Methylation also plays a crucial role in plant immunity by regulating immune response genes. In Arabidopsis, crown gall tumours increased levels of MET1, DRM2, CMT3, and AGO4 transcripts were detected along with a global increase of methylation (Gohlke et al., 2013). These findings indicate that enhanced expression of these epigenetic factors may be part of a plant-induced defence response to prevent Agrobacterium-induced tumour
development. Six disease resistance genes carrying repeats in their vicinity were identified to be derepressed in *met1 nrpd2* mutants (Yu et al., 2013). One of these genes, *RESISTANCE METHYLATED GENE 1* (RMG1) possesses two repeats in its promoter, a distal repeat that is strongly methylated in all mC contexts and a proximal repeat that is unmethylated in the wildtype but hypermethylated in *ros1* mutants (Zhu et al., 2007). The hypermethylated repeat in the *ros1* mutant prevents transcriptional activation. This suggests that RMG1 is controlled by a dual and antagonistic mechanism. Basal expression of RMG1 is repressed via siRNA-directed DNA methylation, while active DNA demethylation maintains pathogen-triggered induction by regularly pruning DNA methylation at the boundaries of its proximal repeat, which may contain functional cis-regulatory elements (Deleris et al., 2016).

Currently, the significance of DNA methylation found within the gene bodies is poorly understood. DNA methylation found within the gene is predominantly located in exons especially those found within genes that are longer than average and functionally more important. This observation supports the hypothesis that body methylation may function as a marker allowing enhanced accuracy for splicing, preventing aberrant transcription (Takuno & Gaut, 2012). Since methylation of DNA is considered a repressive mark especially on single copy genes and TEs, it may play a role in the silencing of cryptic promoters found within the gene body (Zilberman et al., 2007).

### 1.8 The structure and novel functions of MET1

MET1 is structurally similar to that of the mammalian methyltransferase, DNMT1. Both MET1 and DNMT1 have a conserved methyltransferase domain but diverge at the N-terminus responsible for DNA targeting and regulation (Finnegan and Dennis, 1993). The N-terminal domain is connected to the C-terminus via a stretch of alternating glycine–lysine (GK) that acts as a nuclear import mechanism (Vanderkrol and Chua, 1991). Within the N-terminal of MET1, a slight hydrophobic region serves as a replication foci targeting sequence, assisting in MET1 localisation at the replication foci (Hermann et al., 2004). MET1 also contains two BAH domains, that have been demonstrated to interact with histone deacetylase HDA6 and coordinate the maintenance of TE silencing (Liu et al., 2012). The
methyltransferase domain of MET1 contains six highly conserved regions (Fig 1.5). Motifs I and X are responsible for the binding of the methyl donor SAM (Posfai et al., 1989). Motif IV contains an invariable Pro-Cys dipeptide that is the catalytic site of all known C-5 cytosine-specific DNA methylases. Motif VI is responsible for the binding of the methyltransferase domain to the targeted cytosine (Jeltsch, 2006), while motif VIII is believed to negate the negative charge of the DNA backbone via nonspecific association with cytosine residues. Motif IX interacts with the target recognition domain (TRD) located between motif VII and IX. The variable TRD is thought to be responsible for nearly all base-specific interactions with the 5′-GCGC-3′ target site (Lee et al., 2002).

![Figure 1.5: The structure of Arabidopsis MET1.](image)

In the literature, the function of MET1 is primarily discussed in the context of its maintenance of CG methylation marks. The role of MET1, however, is not strictly limited to maintenance of CG methylation. At some target loci, it was found that methylation lost from the body of an endogenous target gene in a met1 mutant, can be partially restored at CG sites when functional MET1 was re-introduced. The target also did not require passage through the germline to be re-methylated, suggesting MET1 may have de novo activity at CG
sequence contexts (Zubko et al., 2012). It has been observed that MET1 can also influence non-CG methylation. REPETITIVE PETUNIA SEQUENCE (RPS) is a repetitive hypermethylated DNA fragment from Petunia hybrida, that attracts DNA methylation in all sequence contexts when transferred into Arabidopsis thaliana. When the RPS was introduced into a met1 Arabidopsis mutant via a genetic cross, a reduction in both CG and non-CG methylation at the RPS was observed. Similarly, both CG and non-CG methylation were eliminated at the RPS when transferred into a drm2/cmt3 mutant. This suggests that MET1, DRM2 and CMT3 may coordinate with each other to establish methylation at the RPS (Singh et al., 2008).

The novel functions of MET1 further extends histone modification layer of epigenetic control. H3K9 methylation, which associates with transcriptionally repressed heterochromatin, is lost when CG methylation is completely removed in a met1 mutant (Tariq et al., 2003). Conversely, loss of non-CG methylation in a cmt3 mutant does not affect H3K9 methylation. When H3k9 methylation is reduced in a kyp suvh5 suvh6 mutant Arabidopsis, there was no significant change in CG methylation, indicating H3K9 methylation is not required for targeting CG methylation (Stroud et al., 2013). Therefore, MET1 and CG methylation may act as a scaffold to directly interact with H3K9 and initiate methylation at the N-terminal tail. The N-terminal domain of MET1 has also been shown to interact directly with the C-terminal domain of HISTONE DEACETYLASE 6 (HDA6) to cooperatively silence TEs and maintain heterochromatic gene silencing (Liu et al., 2012).

1.9 Thesis objective

MET1 does not exclusively act as a CG-specific maintenance methyltransferase; the literature has illustrated that at specific loci MET1 can interact with other methyltransferases to coordinate both CG and non-CG methylation. If MET1 does play a coordinating role, it should involve the direct interaction of MET1 with individual proteins or complexes. Any changes in MET1 concentration, reduction as well as increase, could affect the efficiency and stability of complex formation or interaction with different factors, and could potentially alter epigenetic states at dense methylation regions. Any effect that was induced by protein interaction of MET1 would not necessarily require an increase in MET1.
proteins with a functional catalytic activity. To test this model, we over-expressed catalytically active and inactive versions of the MET1 gene under the control of the 35S promoter. This was tested in both the model organism Arabidopsis and a commercially viable crop, Solanum lycopersicum (tomato). By over-expressing MET1 I hope to induce heritable epi-alleles at distinct loci with altered gene expression and epigenetic marks, to generate different phenotypes for commercial use.
2 Over-expression of DNA methyltransferase MET1 in Arabidopsis thaliana generates new epi-alleles with heritable expression states

2.1 Introduction

DNA methylation influences a number of important processes in plants, including DNA repair (Yao et al., 2012) transcription (Huettel et al., 2007), and recombination (Mirouze et al., 2012), with further implications for genome structure (Kato et al., 2003), plant development (Finnegan et al, 1996) and evolution (Lopez-Maury et al, 2008). The stable changes in expression created by DNA methylation has emerged as a significant factor in shaping phenotypic diversity (Becker et al., 2011; O’Malley and Ecker, 2012). DNA methylation patterns also respond and transform in response to environmental stimuli (Finnegan, 2002), indicating that DNA methylation may act as a molecular switch for evolutionary adaptation of plants to environmental change (Kou et al., 2011). Various biotic (Boyko et al., 2007), and abiotic stress conditions (Kovarik et al., 1997) have been shown to alter the DNA methylation profile, further supporting this model.

In Arabidopsis, maintenance of methylation at CG sequences is catalyzed by the maintenance methyltransferase MET1 (Kankel et al., 2003). To further understand the complex roles of MET1, MET1 knockdowns (Finnegan et al., 1996) and knockouts (Kankel et al., 2003) have been created, resulting in global DNA hypomethylation and developmental abnormalities. However, the effects of increasing MET1 levels and what effect that may have on the global DNA methylation profile have never been assessed in plants. The current dogma states that the sole role of MET1 is to maintain CG methylation if this was truly the case we would predict that over-expression of MET1 would result in continued maintenance of CG methylation. Nevertheless, recent evidence suggests that MET1 functions deviate from the classical CG maintenance model. In a met1 mutant, it has been observed that at certain loci the methylated histone mark H3K9 is lost, which results in the loss of CHG and CHH methylation marks (Stroud et al., 2013). A reduction in DNA methylation in both CHG and CHH sequences and the change in histone marks in the met1 mutant, highlights that MET1 has a more complex and diverse function than initially believed. Raising the question, what further roles and functions are MET1 responsible for?
Previous work had confirmed that MET1 does not exclusively act as a CG-specific maintenance methyltransferase. At particular loci, MET1 can act together with other methyltransferases as a central coordinator for both CG and non-CG methylation or ‘dense methylation’. Loss of a functional MET1 causes loss of cytosine methylation in all three sequence context at loci that are densely methylated (Singh et al., 2008). This coordinating role of MET1 has been observed in a number of Arabidopsis genes with dense DNA, further supporting this model (Watson et al., 2014). While CHH and CHG methylation at these loci are dependent on the chromomethylases CMT2 and CMT3, all methylation marks are lost in a met1 mutant, highlighting the important coordinating role MET1 plays. This indicates that DNA methylation patterns at particular loci are determined by the joint activity of several DNA methyltransferases that are fundamentally directed by MET1. This may be mediated by epigenetic complexes involving methyltransferases, methylcytosine-binding proteins (Woo et al., 2008), chromatin remodeling factors (Kakutani et al., 1996), and/or histone modifiers (Kelly et al., 2012). MET1 may interact directly and recruit these factors, while other factors are recruited indirectly via epigenetic marks which have been established by the MET1 complex, leading to a change in epigenetic state at some loci and even establishing novel dense methylation at others.

The role of MET1 co-ordinating dense methylation would involve direct interaction with individual proteins or complexes. Any changes in MET1 concentration, reduction as well as increase, could affect the efficiency and stability of complex formation or interaction with different factors, and could potentially alter epigenetic states at dense methylation regions. Any effect that was induced by protein interaction of MET1 would not necessarily require an increase in MET1 proteins with a functional catalytic activity. To test this model, we over-expressed catalytically active and inactive versions of the MET1 gene under the control of the 35S promoter. We find that at certain loci, over-expression of the MET1 transgene, can establish new epigenetic marks while removing/reducing the presence of previous marks and alter expression levels which are stably maintained over numerous generations. All of which is independent of the catalytic function of MET1 to methylate.
2.2 Results

2.2.1 Generating MET1 over-expression lines

To investigate the quantitative effects of increased MET1, it was first necessary to produce an over-expression construct. This was accomplished by inserting the MET1 cDNA sequence into the polylinker region of the plant transformation vector pGreen 0179, that contains the strong constitutive 35S promoter (Fig 2.1A). To analyze methylation-independent effects of overexpressing met1 in Arabidopsis, a second construct was created to over-express MET1 with a mutated active site loop, rendering the MET1 protein catalytically inactive (Fig 2.1B). To inhibit the catalytic function a point mutation in the catalytic region of MET1, GGPPCQGFSGMNRFN, introduced a Cystine/Serine replacement (Hsieh, 1999).

**Figure 2.1: Maps of the Active MET1 and Inactive MET1 constructs.** A) The MET1 wildtype cDNA sequence was inserted into the plant transformation vector 35S pGreen 0179, to produce the 35S MET1 over-expression construct. B) A mutation was introduced into the 35S MET1 construct to replace a cysteine with a serine codon within the active site. The position of the point mutation is highlighted by the purple marker.
The two transgenic constructs were transferred into *Arabidopsis*, and four transgenic lines were selected; A1 and A2 contained the catalytically active *MET1* cDNA, and I1 and I2 contained the catalytically inactive *MET1* cDNA. If a change in expression at particular loci was observed it was important to determine how stable these changes were if they were heritable over multiple generations or would revert to a wildtype expression state once the *MET1* over-expression (METo) transgene was lost. To investigate heritability, plants from the T2 generation of each line were selected which had retained the transgene (labeled ‘+’) and plants that had lost the transgene (labeled ‘−’) (Fig 2.2).

**Figure 2.2: Genotyping Arabidopsis over-expressing MET1.** Lines A1 and A2 express a catalytically active MET1 transgene, lines I1 and I2 express a catalytically inactive MET transgene. The larger bands at 1100 bp indicate the loss of the transgene; the primers were designed to span an intron generating a large band when amplifying wildtype MET1. The *Arabidopsis* plant that contains the transgene produces bands at 800 bp as the transgene contains a cDNA insert of MET1 containing no introns and 1100 bp (WT gene). In plants that contained the transgene, the MET1 cDNA was preferentially amplified producing a much brighter band than the WT band.

Before proceeding with any further analysis, the overall *MET1* expression levels were examined to ensure the transgene was not silenced (Fig 2.3). In the plants that had retained the transgene, *MET1* transcript levels were increased in all lines that possessed the METo
transgene, with the greatest increase in A2+ and I1+. In lines that had lost the transgene, MET1 transcript levels had been restored to wildtype levels.

Figure 2.3: Semi-quantitative PCR analysis of the overall MET1 transcript levels. cDNA was generated from seedlings 4 weeks after stratification. Lines A1 and A2 express a catalytically active MET1 transgene, lines I1 and I2 express a catalytically inactive MET transgene. (+) indicates MET1 transformants possessing the METo transgene. Lines derived from MET1 transformants, from which the transgene has been removed are labeled (-). Greater MET1 transcript levels were detected at 24 cycles in lines containing the METo transgene. Actin was used as a standardizing control to ensure equal concentration of input cDNA.

2.2.2 Phenotypic analysis of the MET1 over-expression lines

Once it was established that eight stable METo lines had been generated, four of which were homozygous for the METo transgene, and four lines that had lost the transgene, it was important to evaluate plant growth. This would allow the rapid screening and evaluation of the different lines and provide information on the plant status, enabling an insight into the potential mechanism underlying the phenotypic differences. The Arabidopsis met1 mutant has been previously well documented for its phenotypic differences, more specifically its delay in bolting (Ronemus et al., 1996). It was decided to exploit this phenotype as an indication of vegetative development (Fig 2.4). Plants that bolt later should have a larger number of rosette leaves upon bolting, due to greater resources and developmental time in the vegetative state (Kankel et al., 2003).
A

Number of basal rosette leaves at bolting

B

Control

A1+

A2+

A1

A2

I1+

I1

I2+

I2

met1

met1-1

met1-1 RE

34
Figure 2.4: Phenotypic analysis of bolting and late-stage development. METo lines were compared with a wildtype control, a MET1 mutant, met1-1, and a derivative of met1-1, met1-1 RE, with restored wildtype MET1 alleles. A) METo lines display a delay in bolting phenotype compared to both wildtype and mutated met1 Arabidopsis. Bolting time was analyzed by counting the number of basal rosette leaves upon bolting in long day conditions (Soppe et al., 2002). The parameter used to determine when bolting had occurred was defined, as the stem reaching a minimum of 1 cm in vertical height, for a basal rosette leaf to be counted in the study the leaf had to be at least 1 cm in length and 0.5 cm in width. The significance of a change from wildtype is indicated by asterisks (if present): * = P<0.05, ** = P<0.01 ***=P<0.005, calculated by Student’s two-tailed t-test. B) Image of wildtype, met1 mutant and METo Arabidopsis, taken eight weeks after stratification. The scale bar indicates 5 cm.

Seven of the METo lines displayed a significant delay in bolting compared to wildtype Arabidopsis, including lines that have lost the METo transgene. This implies that there are common target loci in the METo lines involved in bolting and these loci are stably altered, not requiring the presence of the METo transgene. The delayed bolting phenotype was also present in lines that over-expressed the catalytically inactive version of MET1 indicating that MET1 does not need its catalytic function to induce a phenotypic change. The catalytically inactive line I2+ displayed no significant change in bolting time but still possessed the METo transgene. This suggests that the delayed bolting observed in the majority of METo lines is a stochastic event that requires the over-expression of MET1 but is not always sufficient to cause a change. In the later stages of development, a wide range of phenotypic differences can be observed for the eight different METo lines even within their own independent line, highlighting the broad effect over-expressing MET1 has on healthy plant development and its stochastic nature. There are similar phenotypes between the met1-1 mutant and METo lines such as the irregular shoot structure that curls round, unable to support the entire plant. The root structure was also selected as a phenotypic marker (Fig 2.5). One study found that treating Arabidopsis seedlings with the DNA methylation inhibitor 5-azacytidine reduced the primary root length (Virdi et al., 2015), suggesting the phenotype is associated with the disruption of methylation.
Commented [CW3]: This axes is lateral roots per mm and so has units e.g. number of lateral roots (mm$^{-1}$ primary root).
Figure 2.5: Root phenotype analysis of METo lines. METo lines were compared with a wildtype control, a MET1 mutant, met1-1, and a derivative of met1-1, met1-1 RE, with restored wildtype MET1 alleles. A) Primary root length at four weeks of development. All METo lines display significantly reduced root length, although the severity of the phenotype varies between lines. B) Number of lateral roots greater than 2 mm per mm of primary root length, at four weeks of development. The significance of a change from wildtype is indicated by asterisks (if present): * = P<0.05, ** = P<0.01, *** = P<0.005, calculated by Student’s two-tailed t-test. C) Image of wildtype, met1-1 mutant and METo Arabidopsis, taken four weeks after stratification. The scale bar indicates 10mm.

Primary root length was significantly reduced in all METo lines and the met1-1 mutant compared to wildtype root length. The reduced primary root length is even maintained after the METo transgene has been lost. However, the severity of reduced primary root length varies between each line. The number of lateral root structures per mm of primary root length is increased in several METo lines, while a reduction in lateral roots is observed in the met1-1 mutant. The joint appearance of distinct phenotypes among different METo lines, including lines that have lost the transgene, suggests that these are due to heritable changes induced at common target loci. The changes caused at these common target loci seem not to require the catalytic function of MET1. Many of the phenotypic differences observed in the METo lines appear to be novel and independent of a mutated met1 Arabidopsis. The randomness of the induction of phenotypes in different lines and the lack of a correlation between phenotypes and transgene expression, suggests that the induction of heritable changes is a chance event and that increased MET1 levels are required but not always sufficient to induce the individual phenotypes.
2.2.3 Transcript analysis of the MET1 over-expression lines

Once it was confirmed that a change in phenotype could be caused by over-expressing MET1, transcript profiling was carried out. Due to the stochastic phenotypes observed by over-expressing MET1, it was important to reduce variation specifically resulting from secondary effects, such as the plant’s developmental stage (Ogneva et al., 2016), circadian cycle (Lim et al., 2014) and environmental stress (Secco et al., 2015). It was established that all molecular analysis would be carried out at 4 weeks to minimise variation and allow easier comparison of the different lines. Pools of ten, four-week-old seedlings for lines A1+, A1-, A2+ and A2- were selected for transcript profiling. Due to the ubiquitous nature of MET1, there was extensive expression variation among the METo lines, including the lines that had lost the transgene. The RNA sequencing data was meticulously analyzed, identifying genes that had a greater than 2.5-fold change in expression, then organized into three major categories; transposable elements (S1 Table), genes expressing non-coding transcripts (S3 Table) and coding genes (S5 Table). Many small RNAs were found to have a significant change in expression this included; microRNAs, Natural antisense transcript (NAT), ncRNA, rRNA, snoRNA, snRNA, and tRNA. MicroRNA, NATs, ncRNA, and rRNA were primarily upregulated in METo lines with and without the transgene. Conversely, snoRNA, snRNA, and tRNA were mainly down-regulated, highlighting the possibility of a common pathway responsible for regulating expression which is maintained by methylation. A large number of transposable elements including gypsy and copia like retrotransposons were upregulated in lines with and without the transgene. This suggests that by over-expressing MET1 we are disrupting dense methylation which typically represses transposon expression. Interestingly the domesticated TE (DTE) MUSTANG 8 (MUS8) was severely downregulated in the A1+ and A1- lines. In the mus7/mus8 double mutants, multiple developmental abnormalities occur including shortened stem structure and reduced sterility (Joly-Lopez et al., 2012), both phenotypes which are observed in multiple METo lines. Extensive expression changes were also found in protein-coding genes. Due to the large number of protein-coding genes that possessed greater than 2.5-fold change in expression, more stringent screening protocols were applied. Genes that maintained or increased in expression once the transgene was lost were selected. The shortened candidate list was then screened for the presence of MET1 dependent dense methylation, and six candidate genes were selected. The expression
levels of the six candidate genes and MET1 were tested further using qpcr for all eight METo lines along with a wildtype control, met1-1 mutant and a derivative of met1-1, met1-1 RE, with restored wildtype MET1 alleles (Fig 2.6).

**Figure 2.6: Comparison of MET1 expression levels.** The percentage of MET1 transcripts compared to Actin transcription. The analysis was carried in METo transformants with (+) and without the transgene (-). Lines A1 and A2 express a catalytically active MET1 transgene, lines I1 and I2 express a catalytically inactive MET transgene. Expression was also measured in the control line, met1-1, and met1-RE. In A1+ and I2+, MET1 expression is about 3-fold higher compared to the control. In A2+ and I1+, MET1 levels increase are about 15-fold compared to the control line.

In plants that had retained the transgene, MET1 transcript levels were ~3-fold greater in A1+ and I2+, and ~15-fold higher in A2+ and I1+. In lines that had lost the transgene, MET1 transcript levels had been restored to wildtype levels. The level of MET1 was also unaffected in a met1-1 mutant. This expression data confirms that excess MET1 is not required to maintain any of the phenotypes observed in the lines that have lost the METo transgene.
The difference in MET1 expression also verifies that the induction of phenotypes seen in the METo lines is a stochastic event that requires increased levels of MET1, but is not determined by the levels of MET1 transcript. QPCR was then used to confirm the change in expression levels for the six candidates identified in the transcript analysis (Fig 2.7)
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Gene Name</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G01345</td>
<td>N/A</td>
<td>Unknown expressed protein</td>
</tr>
<tr>
<td>AT3G27473</td>
<td>N/A</td>
<td>Cysteine/Histidine-rich C1 domain family protein</td>
</tr>
<tr>
<td>AT3G30720</td>
<td>QQS</td>
<td>Orphan gene that modulates carbon/nitrogen allocation</td>
</tr>
<tr>
<td>AT3G30820</td>
<td>ORF-1</td>
<td>RNA binding protein</td>
</tr>
<tr>
<td>AT4G25530</td>
<td>FWA</td>
<td>Homodomain-containing transcription factor for flowering</td>
</tr>
<tr>
<td>AT5G34850</td>
<td>PAP26</td>
<td>Purple acid phosphatase involved in phosphate scavenging</td>
</tr>
</tbody>
</table>

**Figure 2.7: qPCR analysis of six genes with dense methylation.** A) Analysis was carried in METo transformants with (+) and without the transgene (-). Lines A1 and A2 express a catalytically active MET1 transgene, lines I1 and I2 express a catalytically inactive MET transgene. Expression was also measured in the control line, met1-1, and met1-RE. The mean and the standard error are shown for three biological replicates each having three technical replicates for each line. Values on the y-axis represent the fold difference compared to the control line. B) Table of the six selected candidate genes containing gene ID, gene name and the function of the gene where possible.
Q-PCR analysis of the six target genes confirmed that a change in expression had occurred at these loci, and the presence of the catalytic region is not required to cause a change in expression. As already observed in the phenotype analysis, we detect common expression changes for individual genes in all or some METo lines. For most genes, similar changes in METo lines are also detected in the met1-1 mutant, while expression changes of AT4G35770 and AT5G34850 only occur in METo lines. In most METo lines that have lost the transgene, expression changes are conserved. In some cases; AT3G01345, A1, and AT3G38020, I2 an increase in expression are observed once the transgene is lost. Expression levels are predominantly increased in METo lines, but decreased expression can also occur. This is most evident for the gene AT5G34850, which is unchanged in the A2 lines but significantly reduced in all other METo lines.

2.2.4 Bisulphite analysis of MET1 over-expression lines

All the candidate genes selected for expression analysis contain MET1 dependent dense methylation located within or adjacent to its loci. To investigate if the change in expression corresponds to a change in dense methylation changes, bisulphite analysis was carried out for four of the target genes. The four candidate genes could be categorized into one of three distinct dense methylation categories (Fig 2.8A): Genic methylation, dense methylation present throughout the entire body of the gene, 5’ methylation, dense methylation that is located within the promoter region of the gene, but not present in the gene body. Regional methylation is dense methylation that is localized at a particular exon or intron. The four genes selected possessed one of these distinct methylation pattern; AT3G01345 (genic), AT3G27473 and AT5G34850 (5’), and AT3G30720 (regional) (Fig 2.8B). AT5G34850 was selected due to it being the only candidate gene down-regulated when MET1 is over-expressed (Fig 2.7). Due to highly repetitive nature of transposons and the low efficiency of bisulphite primers, a genomic region of AT5G34850 close to the dense methylation was selected. It was believed that AT5G34850 was silenced due to dense methylation spreading from the transposon upstream into the gene itself.

Commented [CW4]: Check formatting here
A

AT3G01345

GENIC METHYLATION

AT3G27473

5' METHYLATION

AT3G30720

REGIONAL METHYLATION

AT5G34850

5' METHYLATION
Figure 2.8: Bisulphite analysis of the METo lines. A) Genes with dense DNA methylation patterns in the genic region (AT3G01345), in the 5' region (AT3G27473), in the gene region (AT3G30720) and the down-regulated gene (AT5G34850). Boxes label sections that were analyzed by bisulphite sequencing (Figure 2.8B). DNA methylation patterns in the four genes with MET1-dependent dense methylation were extracted from http://genomes.mcdb.ucla.edu. B) DNA methylation analysis of AT3G01345, AT3G27473, AT3G30720, AT5G34850 regions in METo transformants (+) and in lines derived from METo transformants, from which the transgene has been removed (-). Line A expresses a catalytically active METo transgene; line I1 expresses a catalytically inactive METo transgene. Red denotes CG methylation, blue is CHG methylation, and green is CHH methylation.
Bisulphite sequencing analysis of the three candidate genes that were upregulated shows a reduction in methylation for all lines analyzed. In lines that have lost the METo transgene a decrease in methylation is maintained or even more severe, indicating that the METo transgene is not required to keep the hypomethylated state. For the A1 lines, a reduction in methylation correlates with an increase in expression. However, the I1 lines that also show a decline in methylation don't show a significant change in expression, highlighting a more complex mechanism determining the change in expression. Bisulphite analysis of AT5G34850 displayed no change in methylation for the region analyzed.

2.2.5 Histone analysis of the MET1 over-expression lines

As there was no direct correlation between a reduction in methylation and an increase in expression, other factors were also involved in determining expression change. The most likely candidate is histone modifications. There are numerous histone modifications that can alter the chromatin structure and regulate gene expression. Five histone marks were selected for further analysis, the repressive marks H3K9me2 and H3K27me3 and the active marks H3K4me3, H3K27ac and H4 acetylation (Fig 2.9).
Figure 2.9: Histone analysis of the METo lines. ChIP analysis of H3K9me2, H3K4me3, H4ac, H3K27me3 and H3K27ac for AT3G27473, AT3G01345, AT3G30720, and AT5G34850. The relative H3K9me2, H3K4me3, H4ac, H3K27me3 and H3K27ac levels were determined by ChIP assays and normalized via the input DNA. The mean and the standard error are shown for three biological replicates each having three technical replicates for each line. Values on the y-axis represent the relative fold enrichment of histone modification compared to the control line.
Analysis of AT3G01345, AT3G27473, AT3G30720, and AT5G34850 suggests some conclusions about the significance of different histone marks for expression changes. In AT5G34850 a reduction of H3K9me2 does not directly lead to a decrease in gene expression, observed in A2+ and A2-. A decrease in H3K4me3 and an increase in H3K27me3 correlates with a reduction in AT5G34850 expression in all lines. At AT5G34850 deacetylation of H3K27 is required before H3K27 methylation but is not sufficient to reduce expression, seen in line A2+ and A2-. Levels of H3K4me3 need to be maintained or increased to enhance gene expression, a reduction in H3K4me3 directly leads to silencing which is seen in line A1- for gene AT3G27473. Significant increases of H3K4me3 can cause increased levels of acetylation, but the increase does not proportionately correlate. While MET1 expression may induce H4 acetylation and H3K9me2 changes, changes in expression appear to depend on changes in H3K4me3 levels. Over-expressing MET1 allows the disruption of dense methylation and alter multiple histone marks at particular loci, leading to complete restructuring of the chromatin environment. Studies have shown that a change to the chromatic structure can become stably maintained over multiple generations (Silveira et al., 2013).
2.2.6 Heritability of epigenetic changes

Once it was established that an epigenetic change could be induced by over-expressing MET1, and change was still maintained without the presence of the METo transgene, it was important to determine if the epigenetic changes observed spanned multiple generations. We proceeded to self-fertilize the initial eight METo lines already analyzed, and carried out both phenotypic and molecular analysis. The easiest phenotypic marker used for determining heritability was primary root length and density of lateral root growth (Fig 2.10).
Seven of the eight, next-generation METo lines still possessed significantly shorter root length, regardless of METo transgene presence. However, the primary root length of A2 returned to a wildtype length after only one generation. This could be due to the lack of excess MET1 and reduced phenotypic severity in the third generation, allowing the A2- lines to immediately recover after one generation. Though the reduction in primary root length is still significant in the fourth generation, it is not as severe as the third generation of METo lines. Compared to the third generation which shows a significant increase in lateral root density in some lines, the fourth generation appears more sporadic. Lines A1+, A2+, A2- and I1+ show a significant reduction in lateral root density in the fourth generation, compared to the previous generation which has no significant change or an increase in lateral root density. The substantial increase in lateral root density observed in I2+ in the third generation is restored to a wildtype density in the next generation. Only A1- and I2- have maintained their increased lateral root density after one generation. Lateral root density appears to be more irregular, varying from one generation to the next, making it a poor phenotypic marker. However, primary root length is stably maintained after one generation without the presence of the METo transgene. This makes primary root length a good hereditary marker. Once it was confirmed that a change in phenotype is heritable over a generation, it was critical to determine if changes in gene expression are still maintained in the next generation. Expression analysis of MET1 and the six candidate genes were carried out in the fourth generation of METo lines (Fig 2.11).
Initial analysis
Next generation analysis
Figure 2.11: qRT-PCR analysis of MET1 and the six different candidate genes for the 3rd and 4th generation. A) MET1 expression levels as a percentage compared to actin expression for both 3rd and 4th generation METo Arabidopsis. B) Expression analysis of the six candidate genes compared to the wildtype control for both 3rd and 4th generation METo Arabidopsis. The mean and the standard error are shown for three biological replicates each having three technical replicates for each line. Values on the y-axis represent the fold difference compared to the control line.

In lines that still retained the METo transgene in the 4th generation, the levels of MET1 transcript was maintained in A1+, I1+, and I2+ lines compared to the 3rd generation, though at greater variation. This could be due to silencing of the METo transgene in some of the plants within the line. This observation is supported by the complete restoration of native MET1 transcript levels in the fourth generation of A2+, which had the strongest MET1 expression in the previous generation. Lines that didn’t possess the METo transgene in the 3rd generation still maintained innate MET1 transcript levels in the next generation. There appears to be no correlation between the level of MET1 and the increase or reduction of lateral root density. The heritability of expression in the six candidate genes can be categorized into three distinct groups. Genes that have been restored to wildtype in some lines and reduced in expression for others after one generation, AT3G01345, AT3G27473, and AT4G25530. Genes that show a reduction in expression change for all lines in the fourth generation, AT3G38020, and AT5G34850. And finally, genes that have retained or reduced expression change after one generation, AT3G30720. When the level of MET1 transcript is compared to the heritability of each gene, there appears to be no definite correlation. Line A2+ which possessed high MET1 transcript levels in the third generation but native MET1 concentration in the fourth generation, displays varying expression heritability for each gene, demonstrating how over-expressing MET1 is a stochastic event.
2.2.7 Protein analysis of over-expressing MET1

Some genes with altered expression in MET1 over-expression lines have also been reported to be affected in met1 mutants. This implies that increased levels of MET1 transcript may generate co-suppression or protein degradation effects that would resemble a met1 mutant. To confirm this is not the case, and increased MET1 transcripts lead to an increase in MET1 protein, Western blot was carried out. Unfortunately, there was no MET1 Arabidopsis antibody available at the time. To overcome this problem, a FLAG-tagged MET1 over-expression construct was created (S2 Fig) which would allow the analysis of the MET1 protein using an anti-FLAG antibody(Fig 2.12).
Figure 2.12: Protein analysis of over-expressed FLAG-tagged MET1. A) Western blot was carried out for FLAG-tagged MET1 transformants and wildtype control. The expected protein size of the FLAG-tagged MET1 is 176 kDa. Actin (40 kDa) was used as an internal control to determine the concentration of protein extract. B) Semi-quantitative PCR of control line and three biologically different over-expressing MET1 FLAG-tagged lines. Primers were used that only amplified MET1 cDNA with a FLAG tag. Three technical replicates were carried out for each line at 26 cycles. Actin was used as an internal reference to ensure the cDNA analyzed was of the same concentration.

Western blot identified that the MET1 FLAG-tagged protein was present only in the over-expression lines, with an expected protein band of 170 kDa. Secondary bands were observed in all lines at approximately 50 kDa indicating non-specific binding of the FLAG tag antibody. Semi-quantitative PCR of the control line and MET1 FLAG-tagged over-expression lines confirms that increased MET1 transcripts are translated, and doesn’t cause co-suppression or protein degradation. Although changes in expression of individual genes are similar to that of a met1 mutant, the mechanism behind the epigenetic change is distinctly different and doesn’t involve the loss or reduction of the MET1 protein.
2.2.8 Genetic analysis of AT5G34850

Bisulphite analysis of the silenced gene AT5G34850 (Fig 2.8B), did not demonstrate any change in DNA methylation. However, dense methylation was present in a transposable region upstream of the gene (Fig 2.8A). Analysis of this region turned out to be more complicated than initially expected and PCR analysis of this locus was performed (Fig 2.13).
Figure 2.13: Mapping AT5G34850. A) The region of AT5G34850 mapped using four different primer pairs (Pp1-Pp4). B) PCR Mapping analysis of AT5G34850 region using the four different primer pairs. PCR analysis was carried out in MET1 transformants (+) and in lines derived from MET1 transformants, from which the transgene has been removed (-). Lines A expresses a catalytically active MET1 transgene; line I1 expresses a catalytically inactive MET transgene. Actin was used as an internal reference to confirm the DNA analyzed was of similar concentration.

PCR-analysis of the locus revealed that the upstream region of the gene, which contains multiple repetitive elements, had been deleted or rearranged in all six lines, in which the gene had been silenced. Moreover, a central region of AT5G34850 could not be amplified, in lines A1+ and A1- suggesting an extensive rearrangement of the locus. Highlighting the possibility of disrupted dense methylation at repetitive regions leading to transposition and gene disruption.

2.2.9 Investigating the phenotypic effects of over-expressing MET1.

Over-expression of MET1 caused distinct changes in both phenotype and gene expression. Although we have identified many different phenotypic changes, few have been directly correlated with a change in gene expression. One gene that has been identified is AT3G30820 (FWA) which is well documented to cause a delay in bolting when it is expressed (Koornneef et al., 1991). In the METO lines, up-regulation of FWA directly correlates with late bolting. METo lines also display reduced primary root length which appears to correlate with the reduction of AT5G34850. However, line A2 that possess increased AT5G34850 levels, still have shorter root length, though not as severe as the remaining METo lines. Indicating that a reduction in AT5G34850 maybe partially responsible for the reduced primary root length, but other factors are involved. Although two genes have been identified to cause a change in phenotype, the remaining candidate genes have not been linked to a change in phenotype. To determine if the up-regulated candidate genes contribute to the observed METo phenotypes, constructs were created that either over-
expressed AT3G01345, AT3G27473, or AT3G30720 (S1 Fig). AT5G34850 was also further examined. AT5G34850 encoded a *PURPLE ACID PHOSPHATASE 26* (PAP26) which cleaves inorganic phosphate and transports it to the vacuole. As phosphate is involved in many important plant processes (Bolan *et al.*, 2003), silencing of PAP26 may be responsible for other phenotypic abnormalities. Phenotypic analysis was then carried out for each of the over-expression transformants, along with a pap26 mutant (Fig 2.14).

**A**

Number of basal rosette leaves at bolting

![Bar graph showing number of basal rosette leaves at bolting](image)

**B**

Primary root length

![Bar graph showing primary root length](image)
Figure 2.14: Phenotypic analysis of the candidate genes. A) Bolting analysis of the over-expressed AT3G01345, AT3G27473, and AT3G30720 lines compared to wildtype control and pap26 mutant line. Bolting time was analyzed by counting the number of basal rosette leaves upon bolting in long day conditions (Soppe et al., 2002). The parameter used to determine when bolting had occurred was defined, as the stem reaching a minimum of 1 cm in vertical height, for a basal rosette leaf to be counted in the study the leaf had to be at least 1 cm in length and 0.5 cm in width. B) Root phenotype analysis of over-expressed AT3G01345, AT3G27473 and AT3G30720 lines compared to wildtype control and pap26 mutant line, at four weeks of development. The significance of a change from wildtype is indicated by asterisks (if present): * = P<0.05, ** = P<0.01 *** = P<0.005, calculated by Student’s two-tailed t-test.

Over-expression of AT3G01345 and AT3G307020 caused no significant change in primary root length compared to the control line. However, there was a significant reduction in lines that over-expressed AT3G27473. Although there was a significant decrease in the primary root length, it was not as significant as that observed in the METO lines (Fig 2.5A), indicating that an increase in AT3G27473 could contributes to shorter primary root length. A reduction in primary root length was even greater in the pap26 mutant which had an average primary root length comparable to that of the METO lines (Fig 2.5A), suggesting that a reduction of PAP26 is likely to plays a role in reducing the primary root length for METO lines. No substantial change in bolting was observed in either of the AT3G01345 over-expression lines or the pap26 mutant. Lines that over-expressed AT3G27473 or AT3G30720 did take longer to bolt, but the severity of the delay was not as significant as that observed in the METO lines (Fig 2.4A). Furthermore, METO lines that don’t have increased AT3G27473 or AT3G30720 expression (Fig 2.7) still take longer to bolt. This illustrates the complexity of over-expressing MET1 and the numerous effects it has on gene expression. Although genes have been identified in the METO lines to cause a primary phenotypic change, such as FWA, many genes that change in expression may enhance or diminish this phenotypic change. Along with the stochastic nature of increased MET1 and gene expression, directly correlating a change in gene expression with a phenotypic change becomes increasingly challenging.
2.3 Discussion

It has been extensively documented (Vongs et al., 1993) (Finnegan et al., 1996) that MET1 is exclusively responsible for the maintenance of cytosine methylation in a CG-specific context. Recently an alternative function of MET1 has been identified affecting specific loci with dense methylation in CG and non-CG contexts. At these loci, elimination of MET1 activity does not only cause loss of CG methylation but the loss of methylation marks in all sequence contexts. For some loci, this can result in heritable loss of dense methylation patterns creating novel epi-alleles and states of expression (Watson et al., 2014). MET1 dependent dense methylation at many loci is independent of de novo methylation and other components of the RdDM pathway. Instead, dense methylation at these loci requires the nucleosome remodeler DDM1, with CHH methylation being controlled by CMT2 and CHG methylation by CMT3 (Watson et al., 2014).

The coordinating role of MET1 for dense methylation, illustrated by the loss of CG and non-CG marks in met1 mutants, could be based on the MET1 protein facilitating the access of CMT2 and CMT3 to dense methylation targets if MET1 is an essential component of a multi-protein complex that also contains CMT2 and/or CMT3. Alternatively, dense methylation could be mediated by MET1-controlled CG-methylation or by other epigenetic marks established by CG-methylation, which may be required to recruit CMT2 and CMT3. This could involve interaction of MET1 with histone regulators like HDA6, for which direct binding to MET1 has been demonstrated (Liu et al., 2012) and which has been proposed to recruit MET1 to particular target loci as the initial step in establishing subsequent non-CG methylation (To et al., 2011). As MET1 may interact with other epigenetic factors to form a stable complex, it may be sensitive to changes in MET1 concentration, leading to a disruption in complex formation. To test this, high levels of catalytically active and inactive MET1 proteins were introduced to the plant system.

Over-expression of MET1 led to numerous phenotypic variations, some of which were common in multiple METo lines. One example is the delayed bolting phenotype observed in seven of the eight METo lines. FWA is the gene primarily responsible for this delayed bolting and is commonly silenced in wildtype Arabidopsis. Although DNA methylation of FWA has been thoroughly examined, little is understood about the molecular basis of the late-
flowering phenotype (Ikeda et al., 2007). In the METo lines increased expression of FWA correlates with a delay in bolting. Silencing of FWA is mediated by transposable-element-derived tandem repeats in the promoter region which are densely methylated (Lippman et al., 2004). In lines that have lost the METo transgene, FWA activation is still retained, which suggests MET1 over-expression can induce heritable activation. In contrast, FWA allele activated in a met1-1 mutant was efficiently remethylated and re-silenced upon restoration of the MET1 function (Kankel et al., 2003). As the met1-1 allele encodes a MET1 protein with a single amino acid substitution, it is possible that some of the induced phenotypes in the met1-1 mutants are generated by changes in protein structure and interaction, which may produce similar effects as an increase in MET1 concentration.

Reduced primary root length which is present in the METo lines, is also observed in Arabidopsis seedlings treated with the DNA methylation inhibitor 5-azacytidine (Virdi et al., 2015) suggesting the phenotype is associated with cytosine hypomethylation. Among the METo lines, leaf shape, flower structure, and floral organ identity were not significantly altered. Many of these phenotypes, however, have been reported in either the ddm1 mutant (Kakutani et al., 1996) or MET1 antisense lines (Finnegan et al., 1996), but the delay in bolting resembles phenotypes observed in some mutants associated with DNA methylation pathways. Both the HDA6 mutant axe1-5 and HDA6 RNAi lines display late-flowering phenotypes (Wu et al., 2008). Plants with altered MET1 functions show a range of flowering time effects. In both met1-1 and met1-3 mutants, a consistent delay in flowering is observed (Kankel et al., 2003)(Saze et al., 2003). Demethylation of DNA via 5-azacytidine (5-azaC) treatment or via expression of a MET1 antisense gene causes early flowering, with the promotion of flowering being directly proportional to the decrease in methylation in MET1 antisense lines (Finnegan et al., 1998).

Delayed bolting phenotype and reduced primary root length are common phenotypes present in multiple METo lines, highlighting the possibility of common target loci. Though there are common targets, there was no direct correlation between the increased levels of MET1 and the severity of the phenotypes. The randomness of these induced phenotypes in different lines and the lack of a correlation between phenotypic severity and transgene expression levels suggests that the induction of heritable changes is a chance event and that
increased MET1 levels are required but not always sufficient to induce the individual phenotypes.

To identify potential target loci, transcript profiling was carried out for the catalytically active MET1 over-expression lines. In each line, the majority of genes with altered expression show an increase. Applying a cut-off of a log2-fold change of 2.5, increased expression levels were observed in 644 genes in A1+, 565 genes in A1-, 22 in A2+ and 37 in A2-. Reduced expression was found in 240 genes in A1+, 77 genes in A1-, 0 genes in A2+ and 85 genes A2-. Genes with altered expression were organised into three categories; transposable elements (S2 Table), genes expressing non-coding transcripts (S4 Table) and coding genes (S6 Table).

The majority of genes encoding transposable elements are up-regulated (S1 Table). Silencing of the TE populations depends highly on methylation and small RNAs. Most TEs contain and are silenced via the presence of both CG and non-CG methylation (Cokus et al., 2008). Indeed, a TE in the gypsy family, ATGP3, remains silent in single mutants of met1 or cmt3 but activates in a met1 cmt3 double mutant (Tsukahara et al., 2009). This suggests a redundant function of CG and non-CG methylation in the transcriptional silencing of TEs. Within the METo lines, numerous TEs are upregulated, implying a loss of methylation in all three sequence contexts. Among the many TEs upregulated a large subset of these are CACTA-like transposable elements. All CACTA elements carry short sequence motif repeats called subterminal repeats (STRs) in their subterminal regions. It is believed that methylation at these repeat motifs may prevent binding, impairing efficient excision/transposition (Miura et al., 2001). In met1 mutants increased transposition and transcription of CAC1 a CACTA element is observed (Vicient, 2010; Park et al., 2014). This observation correlates with increased CACTA-like transcripts present in the METo lines. It is uncertain if transposition occurs in the METo lines, but given the increase in transcription, it is more than likely.

A number of retrotransposons are also upregulated in the METo lines, which are categorized into either Copia-like or gypsy-like elements. The gypsy-like elements can be further split into those that possess the Athila element and those that are unspecified. Research into Athelia elements identified they contain transcriptional silencing information (TSI) that are released in met1 mutants (Kanno et al., 2005). These TSI or a similar element may also be
present in other TEs, allowing targeted methylation at specific loci. This may explain why we observed common upregulation for TEs that contain TSI-like elements in the METo lines. Some TEs activated in MET1 over-expression lines also deviate in their heritability levels. While, for example, Athila elements that are activated in met1 mutants are efficiently silenced again after re-introduction of a MET1 transgene copy (Catoni et al., 2017), two-thirds of all Athila elements activated in MET1 over-expression lines, retain this status after removal of the MET1 transgene.

Small RNAs have emerged as key regulators of gene expression, genome stability, and defense against foreign genetic elements (D’Ario et al., 2017). When MET1 is over-expressed a number of these small RNAs change in expression (S3 Table), these changes appear to have a common underlying regulator for each class of small RNA. A group of micro RNAs that all code for precursors for miR854 are up-regulated in response to increased levels of MET1, which is heritably maintained once the METo transgene is lost. These precursors are all equally up-regulated in both lines indicating a mutual regulator for these miRNAs. Interestingly all miR854 precursors are located within the ATHILA retrotransposon family (Arteaga-Vazquez et al., 2006) which may explain the common response to increased levels of MET1. Transposable elements are frequently methylated at all sequence context which has been shown to be heritably disrupted when MET1 levels are significantly increased. With the removal of dense methylation, transcription at these loci can occur leading to increased miR854 expression. Interestingly these miR854 precursors are not located within the same retrotransposons indicating a common regulator of dense methylation at particular loci.

snoRNAs are an ancient class of small non-coding RNAs present in all eukaryotes and a subset of archaea that carry out a fundamental role in the modification and processing of ribosomal RNA. Within the METo lines a large number of these snoRNAs are heritably downregulated. Unfortunately little is known about the role of snoRNAs in plants. There are two main classes of snoRNAs those that direct 2′-O-methylation of the ribose (KissLaszlo et al., 1996) and a group that guides pseudouridination of rRNAs, snRNAs and other RNA targets (Ganot et al., 1997). Many of the snoRNAs identified by the transcript analysis are not annotated. However, two snoRNAs were identified. One of the snoRNAs, SNO30 belongs to the class that methylates ribose, while SNO111 carries out pseudouridylation. Little else is
known about these snoRNAs, but there appears to be no dense methylation present at these loci. The large number of snoRNAs that are heritably upregulated indicates a common epigenetic mechanism that controls expression.

Due to the large number of coding genes with altered expression, it was important to differentiate between potential primary and secondary targets of MET1-based epigenetic modifications. Using a methylome genome browser 31 primary target candidate genes with heritable dense methylation were identified (S6 Table). These genes were grouped into three categories, based on the presence of dense methylation in the promoter or 5’ region (upstream), in the gene region (genic) or in the genomic region into which the gene is embedded (region). Several of the genes listed have been shown to be sensitive to DNA methylation changes. The gene responsible for delayed flowering, FWA, is up-regulated in METo lines and under the control of MET1 (Kinoshita et al., 2004). The up-regulated gene AT4G03950, which encodes a nucleotide/sugar transporter family protein, is activated in a ddm1-2 mutant (Lippman et al., 2004). AT3G30720, Qua-Quine Starch (QQS), which is up-regulated in METo lines, is embedded within a TE-rich region and its expression levels are increased in met1, ddc (ddm1/ddm2/cmt3), ddm1 and in the RNA-DEPENDENT RNA POLYMERASE 2 mutant rdr2. QQS expression levels correlate negatively with the DNA methylation level of repeated sequences located within the 5’end of the gene and can be inherited for several generations (Silveira et al., 2013). Two genes are directly regulated by DNA methylation. The up-regulated gene AT3G50770, calmodulin-like 41 (CML41,) contains transposon promoter insertions (Baev et al., 2010). Its increased expression, in response to elevated temperature, correlates with reduced promoter DNA methylation (Naydenov et al., 2015). The down-regulated gene AT3G18610, nucleolin like 2 (NOR2), is involved in epigenetic regulation, as its disruption induces rDNA hypermethylation (Durut et al., 2014).

Six genes were selected for further analysis for both expression changes and epigenetic features. Similar to the observed phenotypes, expression changes of the six analyzed genes occur independently of expression levels, catalytic activity or conservation of the MET1 transgene. Within individual lines, expression changes occur stochastically and with different intensity, inducing an increase in expression for all genes except AT5G34850, which displays a significant reduction in expression in six out of eight MET1-overexpression lines. In most MET1-overexpression lines that have lost the transgene, expression changes
were conserved. Bisulphite sequencing analysis was carried out for four of the target genes, and a reduction or loss of dense methylation marks for three of these genes was identified, independent of the expression levels of the three activated genes in different lines. This suggests that MET1 overexpression induced heritable hypomethylation at these loci, which in some cases was not sufficient to increase gene expression. The analysis of the silenced gene AT5G34850 revealed that the upstream region of the gene may have been deleted or rearranged in the lines that had been silenced. This along with the large number of transposable elements upregulated in the MET0 lines suggests transposition activity maybe occurring on a genome-wide scale.

Analysis of the five different histone marks (H3K9me2, H3K4me3, H4ac, H3K27me3, and H3K27ac), for the four target genes, revealed the complex nature of histone modifications. Among the histone marks tested, Acetylation and H3K4me3 levels show the most significant changes. While there was no consistent correlation between expression changes and individual H3K4me3 marks, some locus-specific correlations were detectable. Increased H3K4me3 levels correlated in all MET1 overexpression lines with enhanced AT3G27473 expression, and in seven out of eight MET1 overexpression lines with enhanced expression of AT3G01345. In the six lines with reduced expression of AT5G34850 H3K4me3 levels are also significantly reduced. However, it is unclear if silencing of AT5G34850 is the consequence of H3K4me3 reduction or of the loss of upstream regions that are required for gene expression. It is also unclear if H3K4me3 reduction is linked to DNA rearrangements or expression changes. Acetylation of either H4 or H3K27 positively correlates with H3K4me3 similar to what is seen in transposable elements in met1 and hda6 mutants (Liu et al., 2012).

Expression analysis identified loci for which the presence of the MET1 transgene was not required to maintain expression changes. This suggests that for individual loci, altered gene expression can be inherited without the continuous presence of increased MET1 levels. Conversely, lines that have maintained the MET1 transgene and enhanced MET1 levels may continuously induce new novel epigenetic changes. To investigate this hypothesis and to test the long-term stability of MET1-induced expression changes, the expression profiles of six genes in the T3 and T4 generation were compared. In most lines, expression change in genes observed in the T3 generation, were also detectable in the T4 generation, although at lower levels. A comparison of the four lines that had lost the MET1 transgene suggests
locus-specific differences in the efficiency of maintaining expression levels, with altered states being preserved for AT3G30720 but reduced for AT5G34850. This corresponds to previous reports about locus-specific differences in the maintenance of epigenetic changes (Mirouze et al., 2012). The stable epigenetic state of AT3G30720 confirms reports about a ddm1-derived hypomethylated epiallele of AT3G30720 that was inherited for at least eight generations (Silveira et al., 2013). In some lines, enhanced expression levels are higher in T4 lines that have retained the MET1 transgene, supporting the hypothesis that epigenetic changes can be continuously induced in lines that have maintained increased MET1 expression.

The phenotypic investigation into the upregulated genes AT3G01345, AT3G27473, and AT3G30720, along with silenced gene AT5G34859, demonstrated the complex nature of phenotypic change and genetic cause. Increased expression of AT3G27473 induced both a delay in bolting and reduced primary root length. The severity of these phenotypes was not comparable to that seen in the METo lines. However, with the increase of FWA, which also causes late bolting, the phenotype may become compounded explaining the severe bolting delay observed in some METo lines. Similarly, the reduced root length seen in the pop26 mutant may also be enhanced by the increased expression of AT3G27473. Little is known about AT3G27473 except that it is a Cysteine/Histidine-rich C1 domain family protein and that it is upregulated in an ibm1 mutant, which displays no significant phenotypic abnormalities (Duque & Chua, 2003), although a number of genes have been identified to cause a phenotypic difference many remain uncharacterized.

Our data shows that MET1 over-expression can induce epigenetic changes, with enhanced MET1 expression levels being required but not always sufficient to cause epigenetic change. There is no direct correlation between the level of increased MET1 expression and the efficiency of the induction of epigenetic changes. This implies that MET1 proteins do not act as a transcription factor or like any other concentration-dependent gene regulator. MET1 over-expression behaves stochastically but not randomly as it induces similar changes in epigenetic and expression states at specific target loci in different MET1-overexpression lines. While the mechanisms involved in MET1 over-expression remain unclear, our data show that MET1 over-expression offers a new strategy to induce variants with novel combinations of epi-alleles.
3 Investigating the application of over-expressing MET1 in tomato

3.1 Introduction

Mankind has been improving crops for thousands of years (Doebley et al., 2006), but recent developments in breeding strategies have given rise to many different crop varieties with greater yield and survivability. Varieties with desirable phenotypes must be self-fertilized multiple times to produce a pure line, which can then be crossed and used for breeding (Rommens et al., 2007). With the advancement of technologies such as marker-assisted selection (MAS), plant breeders can identify single nucleotide polymorphisms (SNPs) associated with a trait (Tester & Langridge, 2010). This is particularly useful for traits that don’t display a visible phenotype or when multiple genes are required for the desired characteristic. However, traditional crop breeding does have its limitations. The crossing of traits can only be carried out between plants that can sexually mate with each other, and even if they are successfully crossed unwanted traits may be introduced. Classical breeding strategies also rely on changes in the genome, though there is growing evidence that stable unwanted traits may be introduced. Classical breeding strategies also rely on changes in the genome, though there is growing evidence that stable changes in gene expression can occur without altering the DNA sequence (Watson et al., 2014). The molecular mechanisms that contribute to this epigenetic phenomenon are DNA methylation and histone modifications.

DNA methylation is the most frequent modification in plants, commonly acting as a transcriptional repressor, either by directly obstructing transcriptional proteins or by serving as a target for specific proteins which signal chromatin condensation (Klose & Bird, 2006). Changes in DNA methylation at individual loci can cause heritable changes in gene expression (epi-mutant), leading to epigenetic variation. These epigenetic changes are present throughout the plant kingdom. One example is the late flowering phenotype seen in Arabidopsis, which is caused by an epi-mutant, fwa, due to hypomethylation at direct repeats within the 5’ region of the gene (Soppe et al., 2000). The rice epi-mutant, Epi-d1, has shortened vegetative branch shoots, caused by hypermethylation at the gene promoter of dwarf1 (Miura et al., 2009). In tomato, hypermethylation of the colourless non-ripening locus (Cnr) inhibits fruit ripening (Fraser et al., 2001). Histone modifications also play a fundamental role in epigenetic variation by altering the chromatin structure, allowing or
preventing transcriptional activators or repressors access to loci. A mutation in HDA6, a histone deacetylase, causes a delay in flowering in Arabidopsis. HDA6 prevents the chromatin forming an open structure, preventing transcription factors access to FLC, a gene which inhibits flowering (Yu et al., 2011). In tomatoes, the histone deacetylase, SlHDA1, plays a significant role in fruit ripening by negatively regulating carotenoids, the chemical responsible for the red pigmentation (Guo et al., 2017).

All studied plants use DNA methylation (Lane et al., 2014), but not necessarily in the same way (Gent et al., 2013)(Takuno & Gaut, 2013). Even the distribution of methylation varies among different plant species. In Arabidopsis, 22-30% of cytosines are methylated in a CG context, in comparison to 6-9% for CHG and 1.5-4% for CHH sites, giving an overall methylation level of 5% (Cokus et al., 2008). Other important crop species contain higher levels of methylation, for example, rice, which has 14-18% of its genome methylated (Zemach et al., 2010), and tomato, which has 22-24% of its genome methylated (Du et al., 2012). 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). Even the way the plant responds to a disruption in the epigenetic machinery varies among plant species. A loss of function for met1 in Arabidopsis causes developmental abnormalities such as delayed bolting and abnormal floral development (Kankel et al., 2003)(Finnegan et al., 1996).

There is also no phenotypic change to Arabidopsis in either a cmt3 mutant (Lindroth et al., 2001) or a drm2 mutant (Cao & Jacobsen, 2002). However, a mutation in met1 is lethal in rice at the seedling stage (Hu et al., 2014), cmt3 mutants have reduced fertility and dwarf phenotypes (Cheng et al., 2015) and drm2 mutants are sterile and have developmental abnormalities (Moritoh et al., 2012). In tomato, an RNA Polymerase V mutation, one of the components of the RdDM pathway, causes lethality (Gouil & Baulcombe, 2016) and transformants containing a MET1 RNAi construct directed against the tomato MET1 gene could not be produced, suggesting seedling lethality (Watson, 2013). In tomato plants, it appears that accurate epigenetic regulation is essential for plant development. This makes understanding how different epigenetic mechanisms function in tomato, challenging to uncover.
Little is known about the role methylation plays in tomato development, except for its importance in fruit ripening. 5-azacytidine which removes methylation, induces early fruit ripening in the tomato, via the demethylation of the CNR gene promoter. This allows the binding of the transcription factor RIN (Ripening Inhibitor) and subsequent gene expression of CNR and other fruit ripening genes (Zhong et al, 2013). Although investigation into tomato methylation has been limited it still possesses many of the methyltransferase homologues found in Arabidopsis (Cao et al., 2014). MET1 is structurally conserved but its precise function and its role in DNA methylation in tomato is still not clearly understood. It has been suggested that MET1 is essential for tomato development (Watson, M. R., 2013) making studying a met1 mutant difficult. In the previous chapter, it has been demonstrated that over-expressing MET1 in Arabidopsis can disrupt dense methylation and cause epigenetic change. By applying the same strategy to tomato, we will be able to determine if MET1 functions in a similar manner in tomato as it does in Arabidopsis. If dense methylation is disrupted, we can investigate if it plays a larger role in gene regulation and stability and, if this disruption or alteration generates novel epi-alleles, whether they are heritably maintained in the next generation.

3.2 Results

3.2.1 Determining the structure of tomato MET1

Silencing of the tomato MET1 (SiMET) appears to be detrimental to the development of the plant (Watson, 2013). To investigate the function of MET1, the strategy to over-express the protein was chosen. Analysis of the SiMET protein was carried out, revealing that the structure and feature of the methyltransferase in tomato was essentially the same in Arabidopsis, consistent with both factors having similar functions in methylating DNA (Fig 3.1).
Figure 3.1: Predicted domains of Methyltransferases in *Solanum lycopersicum* and *Arabidopsis thaliana*. Both proteins are almost structurally identical possessing two DNMT1-RFD domains and BAH domain, and a single DNA methylase domain. Structural analysis was carried out using NCBI Conserved Domain Search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

As both methyltransferases are fundamentally identical, the MET1 over-expression constructs (Fig 1.1) used in the *Arabidopsis* study were used to study tomato as well. The previous study confirmed that by altering the concentration of MET1 in *Arabidopsis*, the coordination of dense methylation was disrupted. As the structure of the methyltransferases is the same, it would be safe to assume they carry out the same function in tomato, including the coordination of dense methylation. If this is correct, by over-expressing either the catalytic active or inactive version of the *Arabidopsis* MET1 the epigenetic state could be altered in tomato. The two transgenic constructs were transferred into tomato, and six transgenic lines were selected; tA1+, tA2+, tA3+ and tA4+ containing the catalytically active *Arabidopsis* MET1 cDNA (*ArMET1*), and tI1+ and tI2+ contained the catalytically inactive *ArMET1*. Due to the long developmental time of tomato, only lines that possessed the transgene were analysed. Positive transformants were selected using the same primers designed for genotyping the *Arabidopsis* METo lines. The positive transformants were then allowed to self-fertilize, and their offspring were analysed.

3.2.2 Phenotypic analysis of the tomato ArMET1 over-expression lines

Over-expression of MET1 in *Arabidopsis* caused numerous phenotypes such as delayed bolting and shorter root growth. To determine if increased levels of *ArMET1* caused similar developmental abnormalities in tomato, extensive phenotypic analysis was carried out. Disrupting the epigenetic machinery has been shown to be detrimental to tomato [Gouil & Baulcombe, 2016; Watson, 2013]. However, original tomato transformants containing the METo transgene fully developed and produced fruit. This indicates that over-expression of
ArMET1 may not critically disrupt the epigenome. Although offspring were produced, increased levels of MET1 may still cause adverse developmental effects in the following generation. To ascertain if this was the case, germination for each METo line was assessed (Fig 3.2).

Figure 3.2: The germination of METo tomato lines. Line tA expresses a catalytically active METo transgene; line tI expresses a catalytically inactive METo transgene. Germination was determined by the emergence of the radicle after four weeks of development. The significance of a change from wildtype is indicated by asterisks (if present): * = P <0.05, ** = P<0.01 ***=P<0.005, calculated by Student’s two-tailed t-test.

Germination for three of the catalytically active METo lines was not significantly changed compared to the control lines, with the exception of line tA2+ which displayed increased germination. In comparison, both catalytically inactive METo lines have severely reduced germination with only 33% of seedlings germinating in line tI2+. Interestingly, germination was only severely affected in the catalytically inactive lines suggesting that that the mutated ArMET1 may have a more dramatic effect on tomato development. To confirm if this is correct, further phenotypic analysis was carried out. Previous study of METo Arabidopsis
found that root length was reduced in all METo lines. To determine if similar phenotypes were observed in tomato, analysis of the root length and stem was carried out at four weeks of development (Fig3.3).
Figure 3.3: Phenotype analysis of tomato METo lines. A) Images of METo lines and wildtype control taken four weeks after stratification. The scale bar for shoot images indicates 3cm. B) The average primary root length at four weeks of development, comparing wildtype control and tMETo lines. B) The average shoot length at four weeks of development comparing wildtype control and tMETo lines. The significance of a change from wildtype is indicated by asterisks (if present): * = P < 0.05, ** = P<0.01 ***=P<0.005, calculated by Student’s two-tailed t-test.

Compared to the control tomato line there was no significant change in the primary root length for all METo lines analysed, indicating that over-expression of ArMET1 responds differently in tomato root tissue than Arabidopsis root tissue. There was also no significant change in stem length for the tA lines. However, there is a significant reduction in the stem length for the catalytically inactive METo lines. Curiously the reduced stem length phenotype and reduction in germination are only present for the tI lines. This further supports the hypothesis that the catalytically inactive version of ArMET1 has a much more significant effect on tomato development. During the phenotypic analysis, a number of
developmental abnormalities were observed (Fig 3.4); such as the degradation of chlorophyll and the acute curling of cotyledons and true leaves (Fig 3.4A). In one tomato, roots developed on the stem (Fig 3.4B). Unfortunately, these phenotypic abnormalities seldom appeared making it impossible to draw any significant conclusions. One phenotype that did appear a number of times in all METo lines was the abnormal development of the shoot apical meristem (SaM) or “blind” phenotype (Fig 3.4C-D). To determine if this was a significant phenotypic event, the number of blind tomatoes were determined for each METo line along with the control line and a line that had undergone plant transformation containing a GUS expression construct (Fig 3.4E).
Figure 3.4: Phenotypic analysis of abnormal tomato development present in tMETo lines. A) Image of diminished chlorophyll and acute leaf curling taken at four weeks of development for a catalytically active METo line. B) Image of abherant root growth on the stem of a catalytically inactive METo tomato. C) Image was taken of blind phenotype present in a catalytically active METo tomato. D) Image was taken of blind phenotype found in a catalytically inactive METo tomato. E) The percentage of tomato plants that possess the blind phenotype at four weeks of development.

The blind phenotype was not present in any of the wildtype control plants. However, a number of blind plants were observed for each of the METo lines, independent of the catalytic activity of MET1. The percentage of blind plants was greater in the catalytically inactive METo lines which occurred in up to 10% of the plants analysed. The GUS transformant line also produced a similar number of plants possessing the blind phenotype, compared to the METo lines. This indicates the blind phenotype observed in the METo lines is, in fact, an artefact of plant transformation and not a causal effect of over-expressing MET1. Further investigation into the later stages of METo tomato development found no significant phenotypic abnormalities. Though numerous phenotypic analyses had been carried out, a characteristic phenotype that occurred across all METo lines could not be detected.

3.2.3 Transcript analysis of MET1 over-expression in tomato

The detection of no common phenotype in the METo lines implies that either the over-expression of MET1 causes no phenotypic difference due to its inability to disrupt dense methylation in tomato, or the METo transgene is silenced in tomato lines and over-expression of MET1 does not occur. To investigate if the METo transgene is silenced in tomato, qRT-PCR was used to measure the level of MET1 transcripts (Fig 3.5). As the METo transgene over-expresses the Arabidopsis homologue of MET1, both ArMET1 and SIMET1 were analysed. Previous analysis identified that the blind phenotype observed in the METo lines might be linked to plants that have undergone a transformation. Plant transformation
has been shown to cause significant changes in the expression of many genes (Veena et al., 2003) and is likely to be associated with many epigenetic changes. As the blind phenotype appears in the next generation of METo lines and occurs stochastically, it strongly suggests the phenotype is epigenetically linked. To study this possibility, MET1 transcripts were also examined in blind plants for both tA and tI plants (Fig 3.5)
For both $S\text{IMET}1$ and $A\text{rMET}1$ the values on the y-axis represent the log2 fold difference compared to the wildtype $S\text{IMET}1$.

qRT-PCR determined that $A\text{rMET}1$ was not expressed in any of the catalytically active METo lines which resolves why no phenotypic abnormalities were observed. Conversely, $A\text{rMET}1$ expression is detected in one of the catalytically inactive lines albeit with significant variation. This extreme difference in $A\text{rMET}1$ expression is likely caused by the silencing of $A\text{rMET}1$ in some of the plants analysed. No increase in $A\text{rMET}1$ was detected for either of the blind tomatoes analysed. Interestingly an increase of $S\text{IMET}1$ is detected for both the tI lines and plants that are blind. It may be possible that an increase in $S\text{IMET}1$ is observed in response to a disruption of the epigenome, the blind phenotype has been linked to epigenetic change, and increased levels of MET1 have also been shown to alter the epigenome. No increase in $S\text{IMET}1$ was detected in the tA lines, but they still possess the METo transgene, suggesting the transgene had been directly silenced. As no significant phenotypic marker was identified for the METo lines, and the blind phenotype is likely linked to an epigenetic change. It was decided to investigate if there were any genes with altered expression in the blind lines that are epigenetically regulated and have dense methylation. Three genes were identified that may play a role in causing the blind phenotype; $S\text{ELF-PRUNING 9D (SP9D)}$ which causes abnormal SaM development when silenced (Thouet et al., 2008), a homologue of the Arabidopsis floral repressor ($C\text{EN1.1}$) (Cao et al., 2016), over-expression causes delayed flowering and floral defects (Yoo et al., 2010), and $W\text{USCHEL (WUS)}$ required for shoot and floral meristem integrity (Xu et al., 2015). All three genes selected are involved in regulating correct SaM development and possess dense methylation within or adjacent to the gene (S3 Fig), $C\text{LE3}, C\text{LE9}$ and $B\text{lind}$ were also selected as they play a role in correct SaM development but didn’t possess any dense methylation (Fig 3.6).
Figure 3.6: qRT-PCR analysis of the six genes selected for further analysis. Transcription was measured for six genes believed to play a role in the blind phenotype. Both increases and decreases in gene expression can be seen across different lines, with the exception of WUS, whose expression is reduced in all lines. Lines tA expressing a catalytically active MET1 transgene, lines tI expressing a catalytically inactive MET transgene. Tomato lines possessing the blind phenotype are denoted with a "b". The analysis was carried out at 4 weeks after stratification and compared to the wildtype control. The mean and the standard
error are shown for three biological replicates each having three technical replicates for each line. Values on the y-axis represent the log2 fold difference compared to the control line.

Analysis of all six genes found no significant correlation between increased SIMET1 and gene expression. WUS was downregulated in all the METo lines including those that were blind which may imply that it is a common target in plants that have undergone transformation but that it is not responsible for the blind phenotype. Expression of CEN1.1 was upregulated in all three of the t1 lines which had increased SIMET1 levels, but not in tAb+. The Blind gene was down-regulated in the tA+ and tAb+ lines but cannot be directly linked to the blind phenotype as the gene is downregulated in lines that are not blind as well. The remaining genes displayed no common expression pattern and were highly variable across multiple lines.
3.3 Discussion

In the previous chapter, it was discovered that over-expression of MET1 could be used as a strategy to induce new epigenetic variants with novel epi-alleles. MET1 is structurally and functionally similar in many important crops species including; wheat (Thomas et al., 2014), maize (Steward et al., 2000), rice (Teerawanichpan et al., 2004) and tomato (Cao et al., 2014), making over-expression of MET1 in crops ideal. It was decided first to test this strategy in tomato plants, as the tomato MET1 is structurally and functionally similar to Arabidopsis MET1. Both the catalytically active and inactive METo construct was introduced in tomato and tested. Unfortunately, no significant phenotypic difference was observed in any of the over-expression lines except for a severe reduction in germination for lines that over-expressed the catalytically inactive form of MET1. Transcript analysis of MET1 revealed that ArMET1 was completely silenced in all the catalytically active METo lines. Expression of the catalytically inactive ArMET1 was detected, but due to the significant variation in transcription, stable expression of ArMET1 could not be confirmed. As no significant expression of ArMET1 could be detected, and germination in the catalytically inactive lines were significantly lower than both the wildtype control and the catalytically active lines, suggests excess catalytically inactive MET1 could be detrimental to tomato development. Over-expression of the catalytically active may also be detrimental but due to transgene silencing in the previous generation expression of ArMET1 was silenced, preventing any developmental abnormalities in further generations. Disruption of the epigenome and MET1 has been shown to be detrimental to tomato development (Watson, 2013) supporting our observation that increased levels of MET1 causes tomato termination. Interfering with MET1 in tomato may cause such a severe response due to only possessing one copy of MET1. In comparison, Arabidopsis has three additional homologs of MET1 (MET2a, MET2b and MET3) which are structurally similar (Cao et al., 2014), allowing for redundancy. Although over-expression of MET1 appears to be difficult in tomato, many other crop species have multiple copies of MET1 such as maize (Qian et al., 2014) and wheat (Thomas et al., 2014), making them prime candidates for over-expression of MET1.
Although \textit{ArMET1} is silenced in tomato, it was identified that \textit{SIMET1} levels increased in lines that possess the catalytically inactive METo transgene and those with the blind phenotype. The stochastic nature of the blind phenotype which appears in less than 10\% of tomato plants, spanning at least two generations, implies an epigenetic disturbance possibly brought about by plant transformation. As increased levels of \textit{MET1} were observed in lines that were epigenetically compromised, there may be a link between \textit{MET1} levels and the disruption of the epigenome. To determine if this was the case, six genes were selected all of which play a role in SaM development. Three of the selected genes were also densely methylated, as increased levels of \textit{MET1} have been shown to disrupt dense methylation. Analysis of the six genes could not confirm any direct link between the increased levels of \textit{MET1} and abnormal development of the SaM causing the blind phenotype. One of the target genes, \textit{WUS}, which causes abnormal SaM development in \textit{Arabidopsis} when downregulated (Laux et al., 1996), is also downregulated in all METo lines. However, downregulation of the \textit{WUS} gene was observed in tomatoes that didn’t possess the blind phenotype, indicating downregulation of \textit{WUS} does not directly cause abnormal SaM development.

Despite not being able to over-express \textit{MET1} in tomato we have identified that the stress of plant transformation disrupts the epigenome generating a stochastic phenotype. Further investigation into the epigenetic differences in plant transformants will be critical to determine common loci that are altered in response. We have also shown that \textit{MET1} is upregulated in response to possible epigenetic disruption (Fig 3.5) making it an attractive candidate to follow in lines that have undergone transformation stress. Over-expressing \textit{MET1} in tomato has also highlighted the importance of different expression strategies, such as targeted over-expression in certain tissue types or at distinct \textit{MET1} target loci. Spatial and temporal over-expression of \textit{MET1} will also offer the opportunity to test if \textit{MET1} target loci alter their susceptibility to \textit{MET1} over-expression in different tissues and identify developmental stages that are particularly sensitive to the induction of epigenetic switches.
4 General discussion

Variation in gene expression and phenotypes in plants can be induced by different epigenetic states. Our study has shown that by temporarily increasing the level of DNA methyltransferase MET1 we can cause heritable epigenetic changes at specific loci. This provides a new strategy to generate novel epi-alleles, and identify common epigenetic target loci and phenotypes. MET1 over-expression serves as a proof-of-concept study that should stimulate a wider application of over-expressing epigenetic regulator genes to examine the significance and targets of epigenetic regulation in different species.

4.1 Over-expression of MET1 induces heritable epigenetic diversity

The coordinating role of MET1 for dense methylation, illustrated by the loss of CG and non-CG marks in met1 mutants (Singh et al., 2008), may be facilitated by a MET1 multi-protein complex which guides CMT2 and CMT3 to dense methylation targets. Alternatively, MET1 could interact with histone regulators like HDA6, for which direct binding to MET1 has been demonstrated (Liu et al., 2012) and which has been proposed to recruit MET1 to particular target loci as the initial step in establishing subsequent non-CG methylation (To et al., 2011).

As the MET1 complex may involve direct interaction with other epigenetic factors, it should be sensitive to changes in MET1 concentration. To test this, high levels of catalytically active and inactive MET1 protein were introduced. The observation that both approaches can cause expression changes and hypomethylation of dense methylation loci resembles disruption caused by the imbalance of multi-protein complexes induced by over-expression of individual complex partners (Sopko et al., 2006). Stoichiometric imbalances can sequester complex partners and disrupt the multiprotein complex. One of the earliest examples demonstrating this effect is the over-expression of either histone H2A-H2B or histone H3-H4 gene pairs in yeast, which causes aberrant chromosome segregation (Meeks-Wagner and Hartwell, 1986) and alters transcription due to disturbance of the histone octamer (Clarkadams et al., 1988).
We observe that increased MET1 expression is required but not always sufficient to induce novel epi-alleles. The efficiency of which these epigenetic changes are caused is not directly correlated with the level of MET1. This implies that the epigenetic changes occur in a stochastic manner but with defined probability for individual loci, similar to the effects of position-effect-variegation (Elgin and Reuter, 2013). This explains why not all transformants display the same phenotypic changes and why particular phenotypes occur more frequently than others. This provides us with a pool of new epi-variants that can be used to link phenotypes to ectopically expressed epi-alleles.

The changes in histone marks that accompany expression changes in METo lines suggest a possible involvement of HDA6 or a related histone modifier. Similar effects are observed in transposable elements activated in met1 and hda6 mutants, which also show increased H4 acetylation and H3K4 methylation levels (Liu et al., 2012). So, at some loci, increased levels of MET1 may interfere with the targeting functions of HDA6, causing the observed histone acetylation increases, stimulating hypomethylation and H3K4 methylation, leading to increased expression.

MET1 over-expression does not just copy the effects induced in a met1 mutant. Some, but not all, phenotypes and genes whose expression are altered in METo lines, are not observed in a met1 mutant. Expression changes of common target genes are also reversed when the met1 mutant is restored to a wildtype background, but retained if the METo transgene is out crossed. Although expression change is preserved over multiple generations when the METo transgene is out crossed, the efficiency with which it is maintained varies between target loci. Further enhancement of gene expression is also observed in lines that have retained the METo transgene, suggesting epigenetic changes can be continuously induced in lines that have maintained increased MET1 expression.
4.2 Investigating the tomato epigenome

As tomato MET1 is structurally and functionally similar to Arabidopsis MET1 possessing the same conserved domains (Fig3.1), both the catalytically active and inactive METo constructs were introduced in tomato. However, unlike Arabidopsis which displayed a number of different phenotypic variants, tomato METo lines did not. Transcript analysis confirmed that ArMET1 was silenced in all the METo lines. As only tomatoes with silenced ArMET1 are observed, increased levels of MET1 may be detrimental to plant development. However, unlike the attempted silencing of MET1 which couldn’t produce tomato lines transformed with a MET1 inverted repeat, tomatoes can be created containing the METo transgene. This suggests that enhanced MET1 levels may only be detrimental at a specific stage of development, allowing silencing of the transgene.

Although ArMET1 is silenced in the METo lines, enhanced levels of SlMET1 are observed in catalytically inactive METo lines. The increased levels of SlMET1 in the METo lines could imply that MET1 is involved in transgene silencing, and increased levels of MET1 are required to establish this state. Enhanced levels of SlMET1 are also observed in plants that are blind, but transcript analysis of six candidate genes involved in correct SaM development could not determine a correlation. Nevertheless, transcript analysis identified that WUS, a gene which possesses dense methylation, is down-regulated in every METo line, highlighting it as a possible epi-allele that is targeted during plant transformation.
4.3 Outlook and open questions

4.3.1 Linking phenotypic change to alternatively expressed epi-alleles

Over-expression of MET1 has identified a number of novel epi-alleles that induce distinct phenotypes. FWA is upregulated in METo lines and causes delayed bolting. PAP26 is downregulated and plays a significant role in reducing primary root length, and AT3G27473 contributes to both delayed bolting and reduced primary root length when upregulated. Although we have determined three epi-alleles that induce phenotypic change, transcript analysis has identified 31 protein-coding genes and many more small RNAs that may play a role in phenotypic development. Altering these genes may also increase or reduce certain stress tolerance. AT3G30775 (ERD5) for example is involved in pathogen response (Fabro et al., 2016). Increased expression of ERD5 is observed in A1 lines, which may enhance their responsiveness to infections. We ascertained that a reduction of PAP26 correlates with shorter primary root length. PAP26 does not appear to cause root shortening directly, but a recent study found that Pi-deprived plants accelerate the degradation of AUX/IAA proteins (Perez-Torres et al., 2008), which may disrupt root development. To see if this is the case, fluorescent auxin reporters can be transformed into METo lines, to visualize the transportation and distribution of auxin in the roots.

4.3.2 Using an inducible system to over-express MET1

The inability to over-express MET1 in tomato led to the speculation that increased levels of MET1 are lethal to tomato. To test this theory, the METo construct could be modified with an inducible system allowing over-expression of MET1 at different stages of development. The ability to generate tomatoes containing the METo transgene suggests that increased levels of MET1 can be tolerated at later stages of development and lethality occurs during the earlier stages of development. One stage that may be especially sensitive to altered levels of MET1 is embryogenesis. Genes that specify embryo cell identity are incorrectly
expressed, and auxin hormone gradients are not properly formed in abnormal met1 embryos (Xiao et al., 2006), leading to abnormal development and reduced seed viability. Different tissue types may also be more susceptible to epigenetic change. A reduction in root length is observed in all METo lines, yet the development of the flower appeared unaffected. Using a tissue-specific promoter, such as the root-specific promoter HPX1 (Park et al., 2013), will allow us to test this susceptibility, and if the phenotypic changes observed are caused by an epigenetic disruption during embryogenesis or occur during root growth. We may also be able to induce epigenetic changes at distinct MET1 targets via CRISPR dCas9-MET1 fusions construct allowing precise induction of epi-alleles.

4.3.3 Investigating the protein interactions of MET1

Our findings support the idea that MET1 is part of a multiple protein complex that regulates dense methylation. However, we have yet to determine what proteins MET1 directly interacts with. Candidates include CMT2 and CMT3, which have been proposed to play a role in establishing dense methylation (Singh et al., 2008), and HDA6 which has been shown to directly bind to MET1 (Liu et al., 2012) and recruit it to particular target loci as the initial step in establishing subsequent non-CG methylation (To et al., 2011). Using a FLAG-tagged MET1 and tandem affinity purification (TAP), proteins and complexes can be isolated that directly interact with MET1. This can also be used to investigate if the mutation in the catalytically inactive MET1 causes a novel conformational change preventing interaction with individual proteins or allows new interactions. Chromatin tandem affinity purification Sequencing (chTAP-seq) can also be implemented to map the genome-wide binding of the MET1 complex (Soleimani et al., 2013), this may identify other target regions that don’t have altered expression levels when grown in normal conditions, but change in response to environmental stress.
4.3.4 The function of the MET1 homologs

Over-expressing MET1 in Arabidopsis causes numerous phenotypic abnormalities, yet in tomato it appears to be lethal. This variation in susceptibility to epigenetic disruption may be due to the different homologs of MET1 found in Arabidopsis (Cao et al., 2014). The structural similarity of both MET2a and MET2b compared to MET1 may imply some redundancy in function. Very little is known about the roles of the MET1 homologs except that MET2a and MET2b are expressed in the mature ovules (Jullien et al., 2012). To determine if there is any redundancy in function a knockout line of all four METs (MET2, MET2a, MET2b, MET3) could be generated to see if the loss of all four METs are lethal. Investigation of the individual MET knockouts should also be studied to determine if the MET1 homologs are active at different stages of development or different tissue types and if they are part of a mechanism that can offset the loss of MET1.

4.3.5 Epigenetic changes induced by transformation

The investigation into the recurrence of the tomato blind phenotype implied it was likely an epigenetic event caused by the stress of plant transformation. It has been documented that plant transformation causes significant expression changes in many genes (Veena et al., 2003), but little is known about the epigenetic effect. Whole transcript and methylome analysis should be carried out for plants that have undergone transformation to determine common target loci along with phenotypic analysis to determine the frequency at which abnormal phenotypes occur. We believe to have identified one common epi-allele, WUS, which is downregulated in all transformed lines and has dense methylation adjacent to the gene. The epigenetic effect of transforming crops will become more important as GM crops increase in popularity.
5 Materials and Methods

5.1 Materials

5.1.1 Arabidopsis material

All Arabidopsis analyzed possessed a Columbia background. Control Arabidopsis plants were derived from non-transgenic seeds raised from a transformation experiment where seeds were cultured on selection-free media. The Arabidopsis met1-1 mutant was provided by Dr. Ortrun Mittelsten Scheid (GMI, Vienna, Austria) and genotyped according to (Singh et al., 2008). MET1 levels were restored (Met1-RE), by self-pollinating a plant derived from a cross between the met1-1 mutant and a wildtype line, and selected a line homozygous for the wildtype MET1 alleles. Homozygous atpap26 T-DNA insertion mutants (Salk_152821) were obtained from the Nottingham Arabidopsis Stock Centre (http://arabidopsis.info) and genotyped using the T-DNA left-border and gene-specific primers.

T1 METo transformants A1, A2, I1, and I2, were selected on hygromycin medium and self-pollinated. T2 progeny plants of each line were grown without selections and were genotyped. To differentiate between transformants that had retained or lost the MET1 transgene, respectively, primers were designed annealing either side of an intron of the MET1 gene. These primers amplify part of the endogenous MET1 gene yielding a 1161bp fragment, while amplification of a part of the MET1 cDNA transgene without the intron produces a 786bp fragment. Plant with (+) and without (-) the transgene was isolated and selfed. T3 seeds of these plant were placed on hygromycin selection to confirm that the transgene had been lost in (-) plants and to identify (+) lines that were homozygous for the transgene. One (-) plant and one (+) plant, homozygous for the transgene, were selected for each line. The T1 AT3G01345, AT3G27473, and AT3G30720 over-expression lines were selected on hygromycin medium and genotyped using the corresponding primers.
5.1.2 Solanum Lycopersicum material

All Solanum Lycopersicum analyzed possessed the Moneyberg background. All tomato analysis was carried out in T1 METo transformants. A1, A2, A3, A4, I1, and I2 were genotyped using the Arabidopsis MET1 primers to generate one band at 786bp.

5.1.3 Bacterial strains

Plasmid cloning was carried out using Escherichia coli Dh5α (New England Biolabs). Plant transformations were performed using Agrobacterium tumefaciens GV3101::pMP90 (Hellens et al., 2000).

5.2 Primer list

<table>
<thead>
<tr>
<th></th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G01345</td>
<td>TGTCCCGGGCCAATTTGACATTTTAA</td>
<td>TGTACTAGTTTTTTTTTTGTCAAAACAGATC</td>
</tr>
<tr>
<td>AT3G27473</td>
<td>TGTCCGGGATGACCTTTAGGCTAGAAG</td>
<td>TGTACTAGTCAAATGCCCATGGAGGGCA</td>
</tr>
<tr>
<td>AT3G30720</td>
<td>TGTCGGGCGTCAGAAGAAGGCTCGTCTTC</td>
<td>TGTACTAGTTGCTAGTTGAATGGGCGATTA</td>
</tr>
<tr>
<td>Genotyping</td>
<td></td>
<td></td>
</tr>
<tr>
<td>METo</td>
<td>TCCAATCACCGTGAGACAC</td>
<td>TCATAGCTCTAGACATCTTGCTTG</td>
</tr>
<tr>
<td>met1-1</td>
<td>CTCTTTAGTAGAAAGTGGCATG</td>
<td>GTAAAGCTCTAGTACAGGCTGTC</td>
</tr>
<tr>
<td>Atpap-26</td>
<td>ATTGCTGAAACTTAAAGCGGG</td>
<td>ATTTTGCGGATTTCGGGAAC</td>
</tr>
<tr>
<td>AT3G01345</td>
<td>TGCTTTGAGACGTGGTTGAACG</td>
<td>CAGCATTAGTTAGCGCGCCA</td>
</tr>
<tr>
<td>AT3G27473</td>
<td>TGCTTTAGAGACGTGGTTGAACG</td>
<td>GAGCTCATTGGGTAGATAGC</td>
</tr>
<tr>
<td>AT3G30720</td>
<td>TGCTTTGAGACGTGGTTGAACG</td>
<td>GACACAAACACCGACAGG</td>
</tr>
<tr>
<td>MET1 FLAG</td>
<td>TCCAATCACCGTGAGACAC</td>
<td>TCAGTGTTACTTTGCTAGTC</td>
</tr>
<tr>
<td>Semi-qPCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>analysis</td>
<td>AT3G49160 (METI)</td>
<td>GGGGTTACCGCTGGTTTGGATAGACACG</td>
</tr>
<tr>
<td></td>
<td>GGGTACCGCTGGTTTGGATAGACACG</td>
<td>GGGGTTACCGCTGGTTTGGATAGACACG</td>
</tr>
<tr>
<td>Gene</td>
<td>Sequence 1</td>
<td>Sequence 2</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>AT3G01345</td>
<td>GCGGCTAAACCTAATGCTGC</td>
<td>CAGACCTGGGCGCTTAGAGGA</td>
</tr>
<tr>
<td>AT3G27473</td>
<td>CCCAAACTGTATTTACCAAAAC</td>
<td>GTACGATCGCAAGCATCACA</td>
</tr>
<tr>
<td>AT3G0720</td>
<td>TTCTTCACAGATGAAGACCAA</td>
<td>ATTTTGAGCCTTGCGACACC</td>
</tr>
<tr>
<td>AT1G07920</td>
<td>GCGGTGTATTTGAGAGTTCG</td>
<td>GTCAAGAGCCCTAAGGAGAG</td>
</tr>
<tr>
<td><strong>qPCR analysis</strong> Arabidopsis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G01345</td>
<td>TTGCTGCCACACCAAGTATG</td>
<td>ACCAGCCCAAACAGAAGT</td>
</tr>
<tr>
<td>AT3G27473</td>
<td>GCGGCAATGCTAGCCTCGCA</td>
<td>TGAAGCAACGCTGACATCCC</td>
</tr>
<tr>
<td>AT3G0720</td>
<td>AAACCTCCTTTCGATCTGCGA</td>
<td>ATGGGTGACCGTGAAGTCTTG</td>
</tr>
<tr>
<td>AT3G30820</td>
<td>CAGAGCATCTTGCTGACCTG</td>
<td>TGGTGGCAGAGAAAGTAGG</td>
</tr>
<tr>
<td>AT4G25530</td>
<td>TCCCATGACTTGCTGACTCTG</td>
<td>CAGTGTGACCCAATTGCCCTG</td>
</tr>
<tr>
<td>AT5G4850</td>
<td>TCAGATTTGAGACGGGAAATCC</td>
<td>TGGCTGTTGGTCCCGTAAACCTCC</td>
</tr>
<tr>
<td>AT5G49160</td>
<td>AGACCTCCGAAGAAGAACAGA</td>
<td>CTACAGGTGATTGAGCCGA</td>
</tr>
<tr>
<td>AT3G18780</td>
<td>CGGTATGGTGAAGGCTGAGAT</td>
<td>ACAACGAGGCGCTGAAACAA</td>
</tr>
<tr>
<td><strong>qPCR analysis</strong> Tomato</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WUS</td>
<td>GCGGCAACTACCCCTTTCTTG</td>
<td>TAAAGCAGAGTTACCCCTTGG</td>
</tr>
<tr>
<td>CLE9</td>
<td>CAAATCAACGACAATCTCTCT</td>
<td>CCTGCCATCTGCGTTATTT</td>
</tr>
<tr>
<td>CLE3</td>
<td>CTTGCTGAGATTTTAAGAAGCTG</td>
<td>GAATGCTCTTTGGTATCATC</td>
</tr>
<tr>
<td>CEN1.1</td>
<td>GACCGCTATGCTCCAGGTC</td>
<td>TGGCTGAGCTTCCCTTG</td>
</tr>
<tr>
<td>SP9D</td>
<td>GGTGAGCTATGAGACCCCAAAG</td>
<td>CAGCGGTTTCAGTTGAGT</td>
</tr>
<tr>
<td>Blind</td>
<td>TCTCAGCAGCCAACAAACAC</td>
<td>GAAACACTGCACTFCCCAA</td>
</tr>
<tr>
<td>SMT1</td>
<td>CGGCTGTTGGTGTTATAT</td>
<td>GATGCAAAGTCCCCTGAGG</td>
</tr>
<tr>
<td>elf3-E</td>
<td>GAGCGATGGATGGTGAATCTT</td>
<td>TGTACGTTGCGTCCAGAAAG</td>
</tr>
<tr>
<td><strong>Bisulphite Sequencing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G01345</td>
<td>GTCTGGATAYAAGAAGAAGATG</td>
<td>ATAAACCACTACAAAAATTT</td>
</tr>
<tr>
<td>AT3G27473</td>
<td>ATAAAATATTAGTATTAGAG</td>
<td>ATCTCRAATCAATTTTCACAC</td>
</tr>
<tr>
<td>AT3G0720</td>
<td>GAGATATGGYTTTTGATTGTTT</td>
<td>TCTTTTTTTCRATCTTCACAT</td>
</tr>
<tr>
<td>AT5G34850</td>
<td>GAATTGTGATTTAAAYTAGAATGAG</td>
<td>CAAACTTTTTTTCTTACCAAACTACTTC</td>
</tr>
<tr>
<td><strong>ChIP analysis</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G01345</td>
<td>CGAGGCGAACAAGCCTTCAAAC</td>
<td>SAGAGCAGAACAAGGAGAGAT</td>
</tr>
<tr>
<td>AT3G27473</td>
<td>ATCCACCAACGCGATGACCTT</td>
<td>GAGAACCACTACACAGCAGA</td>
</tr>
<tr>
<td>AT3G0720</td>
<td>AGGTTCATTTTGCCTCAACT</td>
<td>GCCCGACCAAATGATAGACC</td>
</tr>
<tr>
<td>AT5G34850</td>
<td>TGGTACACCTGGATGAACCTG</td>
<td>TGGTAAGCTCCCTTGGAGCAAAC</td>
</tr>
</tbody>
</table>

88
5.3 Construction of plasmids and plant transformation

5.3.1 Over-expression Constructs

The METo and MET1 FLAG-tagged plasmids were constructed by Michael Watson (Watson, 2013). To create the AT3G01345, AT3G27473 and AT3G30720 over-expression constructs the MET1 gene was excised from the METo construct using SpeI and Xmal. The target genes were then amplified with primers containing the corresponding site and inserted into the over-expression construct.

5.3.2 Arabidopsis transformation by floral dip

Arabidopsis transformation was carried out by floral dip (Clough and Bent, 1998). 500 ml of LB media (10 g/l bacto-tryptone; 5 g/l bacto-yeast extract; 10 g/l NaCl) containing the appropriate antibiotics and bacteria were grown at 28°C until OD600 1.0 was reached. Cells were pelleted using centrifugation and re-suspended to an OD60 0.8 in 5% sucrose; 0.05% Silwet-L77. Wildtype plants were grown at 25°C, under long day conditions till they had appropriately matured at which point clipping of the primary bolt was carried out to induce lateral bolt formation. The plants were inverted into the re-suspended culture for 1 minute and placed into sealed bags for 24 hours to encourage infiltration. This process was repeated one week after the original floral dip. Seeds were harvested by bagging the matured plants. Positive transformants were identified by growing seeds on MS20 medium (4.405 g/l Murashige and Skoog plus vitamins; 20 g/l Sucrose; 0.55% agar; pH 5.8) containing the appropriate antibiotics, at 25°C, 16/8 hour day/light conditions for 2 weeks. Seeds were
sterilized by washing in 70% ethanol for 2 minutes, soaking in 30% bleach (4.8% active hypochlorite) for 10 minutes and washing 5 times with sterilized water.

5.3.3 Leaf disc transformation of Solanum lycopersicum

Leaf disc transformation of Solanum lycopersicum was carried out at the premises of ENZA ZADEN, Enkhuizen, The Netherlands (supervised by Iris Heidmann). Seeds of Moneyberg were sown onto MSB530 medium (Murashige and Skoog salts, B5 vitamins, Duchefa M0231; 30g/l Sucrose; 0.8% agar; pH 5.8) and germinated at 25°C, 16/8 hour day/light conditions for 10 days (until cotyledons expanded). Cotyledons were cut into 0.5cm pieces, placed onto solid co-cultivation medium (4.405g/l MSB5; 3% glucose; 0.8% agar; 200mg/l KH2PO4; 0.2mg/l 2; 4D, 0.1mg/l Kinetin; 0.1mg/l indole-3-acetic acid; 46.8 mM Acetosyringone; pH 5.8) and pre-cultured overnight. During this time Agrobacterium containing the required clone was grown in YEB media (5g/l Yeast extract; 5g/l Beef extract; 20g/l Sucrose; pH 7.2; 2.5mM MgSO4, with the appropriate antibiotics. The Agrobacterium was washed with liquid co-cultivation medium (4.405g/l MSB5; 3% glucose; 200mg/l KH2PO4; 0.2mg/l 2; 4D, 0.1mg/l Kinetin; 0.1mg/l indole-3-acetic acid; 46.8 mM Acetosyringone; pH 5.8), diluted to a density of OD600 0.4 and poured over the explants. After one hour the explants were briefly dried, transferred onto fresh co-cultivation medium and incubated at 25°C for 76 hours under dim light conditions. The explants were transferred to selective medium (4.405g/l MSB5; 3% glucose; 0.8% agar; 2mg/l Zeatin; 500mg/l Cefotaxime; selective antibiotic; pH 5.8) for regeneration.
5.4 Agrobacterium protocols

5.4.1 Agrobacterium tumefaciens GV3101::pMP90 electro-competent cells

Agrobacterium was grown in 500 ml of liquid lysogeny broth (LB) (10 g/l bacto-tryptone; 5 g/l bacto-yeast extract; 10 g/l NaCl) and the appropriate antibiotics at 28°C, gently agitated. When an OD600 of 0.8 was reached the cells were pelleted and re-suspended in ice cold, sterile water. This procedure was repeated 3 times with a final re-suspension in 10% glycerol. The final re-suspension was made into stock aliquots, then frozen using liquid nitrogen and stored at -80°C.

5.4.2 Binary plasmid electroporation Agrobacterium

The plasmid construct pGreenII0029 were co-transferred with pSoup into Agrobacterium. pSoup is a helper plasmid that provides the replicase function for the pSa replication origin of pGreen. A pre-chilled 1 mm cuvette was loaded with 10–50 ng of plasmid construct, 10 ng of pSoup and 50 µl of electrocompetent cells of Agrobacterium. The cuvette was transferred onto a BioRAD Gene Pulser cell-porator using the following parameters: C = 25 µF, R = 400 Ω, 5 ms delay, and pulsed at V = 1.8 kV. Immediately after electroporation the cells were mixed with 950 µl LB in a 15 ml tube and incubated at 28°C for 4 hours with gentle agitation. The transformation mix was spread on LB plates containing 50 µg/ml Kanamycin, 50 µg/ml Gentamycin, and 12 µg/ml Tetracycline prior to incubation at 29°C for 3-4 days.

5.4.3 Isolation of plasmid DNA from Agrobacterium tumefaciens GV3101::pMP90

Mini-prep isolation of plasmid DNA from Agrobacterium was carried out using a modified alkaline lysis method (Wang, 2006). Individual colonies were grown in 10ml of liquid lysogeny broth (LB) (10 g/l bacto-tryptone; 5 g/l bacto-yeast extract; 10 g/l NaCl) supplemented with the required antibiotics for 48 hours at 28°C, and cultures were gently
shaken. The culture was pelleted by centrifugation, and the supernatant was discarded. The cells were re-suspended in 100µl of solution 1 (50mM glucose; 25mM Tris-HCl, pH8.0; 10mM EDTA, pH8.0; 4mg/ml lysozyme) and incubated at room temperature for 30mins, 200 µl of solution 2 (0.2M NaOH; 1%SDS) was added and mixed. Finally, 150 µl of solution 3 (3M KAc, pH 5.5) was added and mixed thoroughly. The solution was then centrifuged at 12,000 x g for 5 min at 4°C, and the supernatant was transferred to a new tube containing phenol:chloroform:IAA. The suspension was centrifuged at 12,000 x g for 1 minute, and the upper layer was transferred to a new tube. The isolated DNA was precipitated with an equal volume of Isopropanol and centrifuged at 12,000 x g for 10 min at 4°C to pellet the DNA. The pellet was washed with 70% ethanol and allowed to air dry. The DNA was dissolved in 30-50 µl of sterile ddH2O.

5.5 E. coli protocols

5.5.1 Preparation of chemically competent E. coli cell

E.coli competent cells were made according to (Sambrook et al., 1989). A glycerol stock of the E. coli strain DH5α was plated on LB (10g/l bacto-tryptone; 5g/l bacto-yeast extract; 10g/l NaCl 0.8% agar) and incubated overnight at 37°C. A single colony was selected and inoculated in 2 ml of LB broth and incubated overnight at 37°C with agitation. 1ml of the overnight culture was added to 500 ml of LB broth in a 2000 ml Erlenmeyer flask, followed by incubation at 37°C with agitation until the OD600 reached 0.3–0.4. The culture was cooled on ice for 10 min and divided into two sterile round-bottom centrifuge tubes. The cells were collected by centrifugation at 5,000 x g for 10 min at 4°C and the supernatant discarded. The pelleted cells were kept on ice and gently resuspended in 100 ml of 100 mM ice cold, sterile MgCl₂. The cells were collected again by centrifugation. The supernatant was discarded, followed by cells re-suspension in 20 ml of ice-cold 100 mM CaCl₂. An additional 180 ml of ice-cold 100 mM CaCl₂ was added. This suspension was kept on ice for 20 min, and the cells were collected via centrifugation. The supernatant was discarded, and the cells were re-suspended in the 4ml volume of ice cold, sterile 85 mM CaCl₂ and 15% of glycerol...
(w/v). The suspension was aliquoted into 1.5 ml tubes and frozen in liquid nitrogen prior to storage in -80°C freezer.

5.5.2 Heat-shock transformation

10–50 ng of plasmid DNA or 10 µl of ligation reaction were added to 100 µl of thawed competent cells on ice and gently mixed. The suspension was incubated for 30 minutes on ice prior to heat shock treatment at 42°C for 90 seconds and transferred immediately back on ice for 2 min. Then, 900 µl of LB was added and incubated at 37°C with agitation for 1 hour. Next, 100 µl of the culture was spread on warmed LB plates containing the appropriate antibiotic for correct transformation selection, followed by overnight incubation at 37°C.

5.5.3 Mini-prep isolation of plasmid DNA from E.coli

Mini-prep isolation of plasmid DNA from E.coli was carried out using a modified alkaline lysis method (Sambrook et al., 1989). Individual colonies were grown in 2ml of liquid LB with the required antibiotics for 17 hours at 37°C. 1 ml of the overnight culture was transferred to a 1.5 ml tube and centrifuged for 5 min at max speed to pellet the cells. The supernatant was removed and resuspended in 100 µl of solution I (50 mM glucose; 25 mM Tris-HCl, pH 8; and 10 mM EDTA, pH 8) and thoroughly mixed. 200 µl of solution II (0.2 M NaOH and 1% SDS) and 150 µl of solution III (5 M Potassium Acetate; pH 5.5, adjusted with Glacial Acetic Acid) was added and placed on ice for 10mins. The tubes were then centrifuged at 12,000 x g for 5 min at 4°C. The supernatants were transferred into new tubes with an equal volume of ice-cold Isopropanol. The suspension was centrifuged at 12,000 x g for 10 min at 4°C to pellet the precipitated plasmid DNA. The supernatant was discarded, and the pellet was washed with 200 µl of 70% Ethanol and allowed to air-dry. The DNA was re-suspended using sterile distilled H₂O and RNase A (20mg/l).
5.6 Phenotypic analysis

5.6.1 *Arabidopsis* phenotyping

Seeds were sterilized by washing in 70% ethanol for 2 minutes, then soaked in 30% bleach (4.8% active hypochlorite) for 10 minutes and washed 3 times with sterilized water. Sterilised seeds were sown on MS15 medium (4.405g/l Murashige and Skoog plus vitamins; 15g/l Sucrose; 1% agar; pH 5.8) and germinated under long day conditions (25°C, 16/8 hour day/light). For the bolting analysis, 24 seedlings for each line were transferred to soil after two weeks and grown under long day conditions. Once the primary bolt reached 1cm in height from the base of the plant, leaves above 1cm in length were counted. Therefore, flowering time is measured as the total number of leaves before flowering, described by (Soppe *et al*., 2000). For the root analysis, 30 seedlings for each line were transferred 120mm square Petri-dishes containing MS15 (1% agar). Each plate contained 10 seedlings and was grown in a vertical position under long day conditions (25°C, 16/8 hour day/light). After four weeks of development, root images were captured using a flat-bed scanner at 800ppi (HP Scanjet G3110) and analyzed using ImageJ (Schneider *et al*., 2012). Lateral roots were only counted if they were bigger than 2mm. The ratio was calculated by dividing the length of the primary root (mm) by the number of lateral roots.

5.6.2 *Tomato* phenotyping

Seeds were washed with 99% ethanol, soaked in 25% bleach (4% active hypochlorite) for 20 min and rinsed three times with sterilized water. Seeds were then sewn onto MSB530 medium (Murashige and Skoog salts, B5 vitamins, Duchefa M0231; 30 g/l Sucrose; 0.8% agar; pH 5.8) and grown under long day conditions (25°C, 16/8 hour day/light). For the phenotypic analysis, 100 seeds were sown for each line with each line was analyzed at 4 weeks of development. Germination was calculated by counting the number of seedlings that had emerged from the seed. The stem and root length was measured by removing each tomato plant from the media and using a ruler to measure their lengths. Blind plants were determined via abnormal growth at the shoot apical meristem.
5.7 DNA Protocol

5.7.1 Isolation of genomic DNA from *Arabidopsis*

Isolation of plant genomic DNA was carried out using a modified Vejlupkova and Fowler protocol (Vejlupkova & Fowler, 2003). Plant tissue was frozen in liquid nitrogen and ground, 560 µl of extraction buffer (200 mM NaCl; 18 mM NaHSO₃; 200 mM Tris-HCl, pH 8.0; 0.07 mM EDTA) and 50 µl of 5% sarkosyl was added and mixed. The suspension was mixed and incubated at 65° for 1 hour. An equal volume of phenol:chloroform: Isoamyl alcohol (IAA) (12:12:1) was added then centrifuged at 12,000 x g for 10 min at 4°C. The upper phase was transferred to a new Eppendorf tube and the phenol:chloroform:IAA extraction was repeated. The DNA was precipitated using 300µl of isopropanol and pelleted by centrifugation at 12,000 x g for 10 min at 4°C. The supernatant was discarded and washed with 70% ethanol. The DNA was re-pelleted and the supernatant removed and allowed to air-dry. The DNA was re-suspended using sterile distilled H₂O and RNase A (20 mg/l).

5.7.2 Restriction digest

Digestion reactions were made to a final volume of 50µl. A standard digestion reaction consisted of approximately 1µg of DNA, the 1x concentration of appropriate digestion buffer, 5-10 units of restriction endonuclease enzyme and 0.1mg/ml Bovine Serum Albumin (BSA) when required. The reaction was incubated at the optimum temperature for 2 hours.

5.7.3 Ligation reaction

Ligation reactions consisted of 1 x ligase buffer, 1 U/µl of T4 DNA Ligase (Promega M180A) and a 3:1 ratio of insert to vector. The final reaction was incubated at 16°C overnight.
5.7.4 Polymerase chain reaction (PCR)

PCR for genotyping was carried out using MyTaq DNA polymerase (Bioline). The reaction was made according to manufacturer’s instructions which consisted of 0.3 μl of MyTaq DNA polymerase, the 1x concentration of red buffer, 10 μM of both forward and reverse primers, 250 ng of DNA template and H2O to make the final volume to 50 μl. The reaction was placed into a thermocycler with the following settings: initial denaturation at 95°C for 5 min, 25–29 cycles of denaturation at 15°C for 30 sec, annealing temperature (Ta, according to the primer annealing temperature) for 15 sec and extension at 72°C for 10 sec/kb and followed by final extension at 72°C for 5 min.

PCR for plasmid construction was carried out using the high fidelity polymerase Phusion (Finnzymes). The reaction was made according to manufacturer’s instructions which consisted of 0.5 μl of Phusion High-Fidelity DNA Polymerase at 1 U/μl, 10 μl of 5x Phusion HF Buffer, 10 mM dNTPs, 10 μM of both forward and reverse primer, and 250 ng of template DNA was added with H2O to a final volume of 50 μl. The reaction was placed into a thermocycler with the following setting: initial denaturation at 98°C for 5 min, 25–29 cycles of denaturation at 98°C for 30 sec, annealing (Ta, according to the primer annealing temperature) for 30 sec and extension at 72°C for 30 sec/kb and followed by final extension at 72°C for 5 min. A negative control was run together with the DNA template replaced with H2O.

5.7.5 Bisulphite analysis

Genomic DNA was isolated from, three replica samples, each contained ten pooled four-week-old seedlings of the T3 generation and subjected to bisulfite treatment using an EZ DNA Methylation-lightning kit (Zymo Research) according to the manufacturer’s instructions. Regions containing dense methylation for AT3G01345 (Chr3: 129684..129860 - 177 bp), AT3G27473 (Chr3: 10171884..10172090 - 207 bp), AT3G30720 (Chr3: 12348994..12349109 - 116 bp) and AT5G34850 (Chr5: 13111304..13111574 – 271bp) were amplified by primers listed S8 Table. For each line, 10 clones were sequenced, and sequences were exported into
the BioEdit program (Hall, T. A., 1999). Aligned sequences were saved in FASTA format and analyzed by the CyMATE program (Hetzl et al., 2007).

5.7.6 DNA sequencing

Approximately 100 ng/µl isolated plasmid DNA was sent to Beckman Genomics for sequencing using the appropriate primers. Sequencing reads were aligned using the Clustal function in Bioedit 7.0.9.0 (Higo et al., 1999).

5.7.7 Data analysis

The ThaleMine platform https://apps.araport.org/thalemine/begin.do was used to extract the annotation for extracted genes. DNA methylation patterns for Arabidopsis were extracted from the Neomorph platform http://neomorph.salk.edu/epigenome/epigenome.html to identify genes with dense DNA methylation patterns. The tomato DNA methylation patterns were extracted using the tomato epigenome database http://ted.bti.cornell.edu/cgi-bin/epigenome/home.cgi to determine genes with dense DNA methylation patterns.
5.8 RNA analysis

5.8.1 Isolation of RNA from plants

Total plant RNA was isolated from each line, with for each line having replica samples, each contained ten pooled four-week-old seedlings of the T3 generation, and performed as described by (Stam et al., 2000). 750µl of RNA extraction buffer (100mM Tris-HCL, pH8.5; 100mM NaCl; 20mM EDTA; 1% sarcosyl) was added to 0.5g of plant tissue ground in liquid nitrogen. The suspension was mixed and equal volume of phenol:chloroform:IAA (12:12:1) was added then centrifuged at 12,000 x g for 10 min at 4°C. The top phase was transferred to a new Eppendorf and phenol:chloroform:IAA extraction was repeated. Precipitation of the RNA was performed using Isopropanol, 4M LiCl, and 3M, pH7.0 NaAc. Extractions were quantified using the nano-drop spectrophotometer ND-1000 (Thermo Scientific). DNA was removed using the TURBO DNase kit (Ambion Applied Biosystems) according to the manufacturer’s instructions.

5.8.2 Semi-quantitative PCR

cDNA was generated using SuperScript™ II. The reaction was made according to manufacturer’s instructions which consisted of 10µg of isolated RNA, 50 µM oligo(dT)20 and 10 mM dNTP mix. The reaction was incubated at 65°C for 5 min and then placed on ice for at least 1 min. The 1x concentration of First stand buffer, 0.2M of DTT and 40 units/µl of RNaseOUT was added to each reaction. The mix was incubated at 25°C for 2 minutes at which point 200 units of SuperScript™ II was added to a final volume of 20 µl. The mix was incubated for a further 10 min, then 42°C for 50 min then final 70°C for 15 min. The solution was diluted with 20 µl of ddH2O. Random primers were used as non-coding RNA would be analyzed in the expression analysis. 1µl of the diluted cDNA solution was added to a standard MyTaq reaction. The reaction was placed into a thermocycler with the following settings: initial denaturation at 95°C for 5 min, 25–29 cycles of denaturation at 15°C for 30 sec, annealing temperature (Ta, according to the primer annealing temperature) for 15 sec
and extension at 72°C for 10 sec/kb, when the thermocycler had performed 20 cycles the reaction was held at 72°C while 6 µl was removed from the total reaction, this was repeated twice more every three cycles then ran on a Agarose gel. Using Elongation Factor 1α which is ubiquitously expressed the reactions exponential phase could be determined and used to standardize each reaction for expression analysis.

5.8.3 Quantitative RT-PCR assay

Gene expression was analyzed using SsoFast EvaGreen Supermix (BioRad) on the Fluidigm Biomark 96.96 Dynamic Array according to the manufacturer’s protocol. Data analysis was carried out utilizing the Fluidigm Gene Expression Analysis software using ACTIN 2 (AT3G18780) as the reference gene.

5.8.4 Sequencing and data analysis

Next-generation sequencing libraries were created from RNA using the TruSeq Stranded total RNA kit (Illumina) which removes rRNA and cleaves the remaining RNA allowing the ligation of random hexamers, and synthesis of cDNA for further analysis. Sequencing was carried out on a HiSeq 2500 to generate 50 bp single-end sequence data. Data analysis was carried out Dr. Ian M Carr (Leeds University, Leeds, England). The data were aligned to the Arabidopsis genome (TAIR website [https://www.arabidopsis.org]) using the STAR aligner (Dobin et al., 2013). Reads mapping to each transcript were determined using the R package rsubRead (Liao et al., 2013) and pairwise comparisons between the wildtype sample and each of the modified samples were performed using the R package DESeq2 (Love et al., 2014) to identify transcripts whose expression varied markedly between the control and experimental sample for each condition. Reads were used to calculate the mean value of read mapping to a transcript in all sample in the analysis (base Mean), the
change in expression between the control sample and the test sample given as a Log to the base 2 value (log2FoldChange), the standard error of variation for the log2FoldChange values in the analysis (lfcSE = log fold change Standard Error), the Wald statistic; the log2FoldChange divided by lfcSE, the probability the result is real; the log2FoldChange divided by lfcSE, compared to a standard Normal distribution to generate a two-tailed p-value (pvalue) and the pvalue adjusted for multiple testing using the Benjamini-Hochberg test (Padj).
Raw data were submitted to the short-read archive of NCBI BioProject database under SubmissionID SUB2885208, BioProject ID PRJNA395995 for the following Datasets:

<table>
<thead>
<tr>
<th>Accession</th>
<th>Sample Name</th>
<th>Organism</th>
<th>Tax ID</th>
<th>BioProject</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAMN07419160</td>
<td>WT_1</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419161</td>
<td>WT_2</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419162</td>
<td>WT_3</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419163</td>
<td>A1+_1</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419164</td>
<td>A1+_2</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419165</td>
<td>A1+_3</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419166</td>
<td>A1_-1</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419167</td>
<td>A1_-2</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419168</td>
<td>A1_-3</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419169</td>
<td>A2+_1</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419170</td>
<td>A2+_2</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419171</td>
<td>A2+_3</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419172</td>
<td>A2_-1</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419173</td>
<td>A2_-2</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
<tr>
<td>SAMN07419174</td>
<td>A2_-3</td>
<td>Arabidopsis thaliana</td>
<td>3702</td>
<td>PRJNA395995</td>
</tr>
</tbody>
</table>
5.9 Protein analysis

5.9.1 ChIP assay

28-day-old seedlings were harvested and cross-linked with 1% formaldehyde. Chromatin was extracted using the ChromaFlash Plant Chromatin Extraction Kit (Epigentek) and sheared to 200-500bp fragments using a Bioruptor (Diagenode). ChIP was carried out using the EpiQuik Plant ChIP Kit (Epigentek). Input samples and immunoprecipitated samples were analyzed using SsoFast EvaGreen Supermix (BioRad) on the Fluidigm Biomark 96.96 Dynamic Array according to the manufacturer’s protocol. ChIP-qPCR results were first normalized with input sample. Relative enrichment was then calculated via the enrichment of the signal (the antibody of interest) compared to the enrichment of the noise (negative control). Antibodies used for ChIP: anti-acetyl-histone H4K5K8K12K16 (06-866; Millipore), H3K4me3 (07-473, Millipore), H3K9me3 (07-442, Millipore), normal rabbit IgG (12-370, Millipore).

5.9.2 Western blot assay

Protein was isolated from 28-day-old seedlings using the P-PER plant protein extraction kit (Pierce) in accordance with the manufacturer’s instructions. Protein was then denatured by suspending the extract in equal volume laemmli buffer (Bio-Rad) and heated at 95°C for 5 minutes. The samples were resolved on 12% Mini-PROTEAN TGX Stain-Free Protein Gels (Bio-Rad) and transferred to Trans-Blot Turbo Mini Nitrocellulose Transfer Packs using the Trans-Blot Turbo Transfer (Bio-Rad). The nitrocellulose membranes were blocked with 5% (w/v) nonfat dry milk in Tris-buffered saline (TBS; 20 mM Tris and 137 mM NaCl, pH 7.7) with 0.1% (v/v) Tween-20. The Anti-FLA (ab197345, Abcam) and Anti-Actin (ab197345, Abcam) were diluted 1/1000 in TBS-Tween/5% milk solution and incubated overnight at 4°C. Blots were washed with TBS-Tween 6 times for 10 min each. Blots were further incubated with horseradish peroxidase–conjugated anti-rabbit IgG (ab6721, abcam) at 1/2000 dilution and then washed as above. Blots were covered with SuperSignal™ West Pico PLUS Chemiluminescent Substrate (Thermo Scientific) for 20 minutes; bound antibodies were visualized on a G:BOX Chemi XX6 (Syngene).
6 References


March, pp. 1330–1344.


Hsieh, C. L. (1999) In Vivo Activity of Murine de Novo Methyltransferases, Dnmt3a and


Johnson, L. M., Cao, X. F. & Jacobsen, S. E. (2002) Interplay between Two Epigenetic Marks:


Kinoshita, T., Miura, A., Choi, Y. H., Kinoshita, Y., Cao, X. F., Jacobsen, S. E., Fischer, R. L. & 

*TRENDS IN BIOCHEMICAL SCIENCES*, 31 (2) February, pp. 89–97.

Analysis of Late Flowering Mutants in Arabidopsis-Threadiana. *MOLECULAR & GENERAL 
GENETICS*, 229 (1) September, pp. 57–66.

184 (4139), pp. 868–871.

Heritable Alteration in DNA Methylation Induced by Nitrogen-Deficiency Stress 
Accompanies Enhanced Tolerance by Progenies to the Stress in Rice (Oryza Sativa L.). 

Heterochromatic Loci in Response to Osmotic Stress. *THEORETICAL AND APPLIED 

Lane, A. K., Niederhuth, C. E., Ji, L. & Schmitz, R. J. (2014) pENCODE: A Plant Encyclopedia of 
DNA Elements. vol. 48. pp. 49–70.

Protein Lysine Methylation. *MOLECULAR SYSTEMS BIOLOGY*, 10 (4) April.


87–96.


Reik, W., Dean, W. & Walter, J. (2001) Epigenetic Reprogramming in Mammalian Development. SCIENCE, 293 (5532) August, pp. 1089–1093.


Virdi, K. S., Laurie, J. D., Xu, Y.-Z., Yu, J., Shao, M.-R., Sanchez, R., Kundariya, H., Wang, D.,


### Supplementary data

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Assembly</th>
<th>Segment length</th>
<th>Segment start</th>
<th>Segment end</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT2G13160</td>
<td>1 - 18431</td>
<td>4.487655</td>
<td>3755560</td>
<td>5.052103</td>
</tr>
<tr>
<td>AT2G12300</td>
<td>1 - 39382</td>
<td>5.342745</td>
<td>7597004</td>
<td>8.345745</td>
</tr>
<tr>
<td>AT2G06670</td>
<td>1 - 24949</td>
<td>4.194542</td>
<td>6792000</td>
<td>9.176748</td>
</tr>
<tr>
<td>AT1G40118</td>
<td>1 - 152013</td>
<td>9.245220</td>
<td>5120307</td>
<td>14.232504</td>
</tr>
<tr>
<td>AT2G12308</td>
<td>1 - 103877</td>
<td>3.733609</td>
<td>1225000</td>
<td>5.678609</td>
</tr>
<tr>
<td>AT2G13110</td>
<td>1 - 1026954</td>
<td>1.950432</td>
<td>1078000</td>
<td>12.95432</td>
</tr>
<tr>
<td>AT2G22200</td>
<td>1 - 363259</td>
<td>5.105428</td>
<td>2150000</td>
<td>7.215428</td>
</tr>
<tr>
<td>AT2G13118</td>
<td>1 - 1572130</td>
<td>2.574621</td>
<td>1302507</td>
<td>2.574621</td>
</tr>
<tr>
<td>AT2G13116</td>
<td>1 - 2117200</td>
<td>4.274210</td>
<td>1202507</td>
<td>4.274210</td>
</tr>
<tr>
<td>AT4G01290</td>
<td>1 - 2310055</td>
<td>4.075802</td>
<td>5025000</td>
<td>8.075802</td>
</tr>
<tr>
<td>AT2G13119</td>
<td>1 - 2081687</td>
<td>3.655409</td>
<td>1575000</td>
<td>7.577609</td>
</tr>
<tr>
<td>AT2G12309</td>
<td>1 - 2927160</td>
<td>5.979875</td>
<td>6525000</td>
<td>11.979875</td>
</tr>
<tr>
<td>AT2G22209</td>
<td>1 - 4180272</td>
<td>1.428328</td>
<td>5120400</td>
<td>1.428328</td>
</tr>
<tr>
<td>AT2G13120</td>
<td>1 - 3859104</td>
<td>2.085391</td>
<td>5825000</td>
<td>4.085391</td>
</tr>
<tr>
<td>AT2G12308</td>
<td>1 - 2773195</td>
<td>2.3773195</td>
<td>920000</td>
<td>7523195</td>
</tr>
<tr>
<td>AT2G13118</td>
<td>1 - 2955028</td>
<td>1.3295503</td>
<td>1300000</td>
<td>1.3295503</td>
</tr>
<tr>
<td>AT4G01290</td>
<td>1 - 5951085</td>
<td>0.5951086</td>
<td>7000000</td>
<td>0.5951086</td>
</tr>
<tr>
<td>AT2G12308</td>
<td>1 - 3773195</td>
<td>2.3773195</td>
<td>920000</td>
<td>7523195</td>
</tr>
<tr>
<td>AT2G13119</td>
<td>1 - 2081687</td>
<td>3.655409</td>
<td>1575000</td>
<td>7.577609</td>
</tr>
<tr>
<td>AT2G12309</td>
<td>1 - 2927160</td>
<td>5.979875</td>
<td>6525000</td>
<td>11.979875</td>
</tr>
<tr>
<td>AT2G22209</td>
<td>1 - 4180272</td>
<td>1.428328</td>
<td>5120400</td>
<td>1.428328</td>
</tr>
<tr>
<td>AT2G13120</td>
<td>1 - 3859104</td>
<td>2.085391</td>
<td>5825000</td>
<td>4.085391</td>
</tr>
<tr>
<td>AT2G13120</td>
<td>1 - 3859104</td>
<td>2.085391</td>
<td>5825000</td>
<td>4.085391</td>
</tr>
<tr>
<td>AT2G13118</td>
<td>1 - 2955028</td>
<td>1.3295503</td>
<td>1300000</td>
<td>1.3295503</td>
</tr>
<tr>
<td>AT4G01290</td>
<td>1 - 5951085</td>
<td>0.5951086</td>
<td>7000000</td>
<td>0.5951086</td>
</tr>
<tr>
<td>AT2G12308</td>
<td>1 - 3773195</td>
<td>2.3773195</td>
<td>920000</td>
<td>7523195</td>
</tr>
<tr>
<td>AT2G13119</td>
<td>1 - 2081687</td>
<td>3.655409</td>
<td>1575000</td>
<td>7.577609</td>
</tr>
<tr>
<td>AT2G12309</td>
<td>1 - 2927160</td>
<td>5.979875</td>
<td>6525000</td>
<td>11.979875</td>
</tr>
<tr>
<td>AT2G22209</td>
<td>1 - 4180272</td>
<td>1.428328</td>
<td>5120400</td>
<td>1.428328</td>
</tr>
<tr>
<td>AT2G13120</td>
<td>1 - 3859104</td>
<td>2.085391</td>
<td>5825000</td>
<td>4.085391</td>
</tr>
<tr>
<td>AT2G13310</td>
<td>AT4G07518</td>
<td>AT3G46487</td>
<td>AT2G14970</td>
<td>AT2G14230</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>237.40606</td>
<td>142.49938</td>
<td>24.237105</td>
<td>20.103713</td>
<td>178.89096</td>
</tr>
<tr>
<td>17.084336</td>
<td>7.713349</td>
<td>20.295202</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>14.260127</td>
<td>3.213349</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>15.112134</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

| 0.47346802 | 0.333787318 | 171871458 | 71641     | -         |
| 0.411031186 | 0.375058759 | 0.499871218 | 0.3999233 | -         |
| 0.325619928 | 0.171952347 | 0.333787318 | 0.13          | -         |
| 0.13347447  | 0.066387318 | 0.13347447  | 0.066387318 | -         |
| 0.03347447  | 0.01673714  | 0.03347447  | 0.01673714  | -         |

<p>| 0.940325156 | 0.325619928 | 0.333787318 | 171871458 | 71641     |
| 0.871745884 | 0.694589061 | 0.325619928 | 0.333787318 | -         |
| 0.534325156 | 0.252639854 | 0.325619928 | 0.333787318 | -         |
| 0.252639854 | 0.126319928 | 0.325619928 | 0.333787318 | -         |
| 0.126319928 | 0.06319928  | 0.325619928 | 0.333787318 | -         |</p>
<table>
<thead>
<tr>
<th>AT1G35600</th>
<th>AT5G36655</th>
<th>AT5G30450</th>
<th>AT5G29408</th>
<th>AT4G08598</th>
<th>AT4G07518</th>
</tr>
</thead>
<tbody>
<tr>
<td>23.558115</td>
<td>7.6375554</td>
<td>33.347057</td>
<td>3.5397797</td>
<td>8.8468137</td>
<td>4.2318645</td>
</tr>
<tr>
<td>8.4934548</td>
<td>29.121055</td>
<td>320.78694</td>
<td>63.36628</td>
<td>385.48988</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>201.58489</td>
<td>5.35287065</td>
<td>1.162837</td>
<td>0.9549149</td>
<td>7.1886687</td>
<td>8.328826</td>
</tr>
<tr>
<td>2.2826924</td>
<td>6.2762677</td>
<td>5.807358</td>
<td>4.4629371</td>
<td>6.105016</td>
<td>5.3078482</td>
</tr>
<tr>
<td>3.1274708</td>
<td>7.38991287</td>
<td>5.3562094</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AT1G35600</th>
<th>AT5G36655</th>
<th>AT5G30450</th>
<th>AT5G29408</th>
<th>AT4G08598</th>
<th>AT4G07518</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1.228419184</td>
<td>3.063565776</td>
<td>1.156916239</td>
<td>4.218748586</td>
<td>0.637906993</td>
<td>0.627890677</td>
</tr>
<tr>
<td>2.525671883</td>
<td>0.937100426</td>
<td>6.423606911</td>
<td>5.09407699</td>
<td>0.392471801</td>
<td></td>
</tr>
<tr>
<td>0.478972938</td>
<td>0.58218025</td>
<td>0.169278724</td>
<td>0.400963157</td>
<td>0.312398996</td>
<td></td>
</tr>
<tr>
<td>0.472E</td>
<td>6.64E</td>
<td>1.10E</td>
<td>1.35E</td>
<td>9.44E</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>24</td>
<td>06</td>
<td>06</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>8.29E</td>
<td>1.62E</td>
<td>1.03E</td>
<td>9.05E</td>
<td>0.000591073</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5.192160237</td>
<td>7.367754976</td>
<td>2.797162804</td>
<td>3.480233365</td>
<td>4.72E</td>
<td></td>
</tr>
<tr>
<td>7.29246369</td>
<td>9.919469095</td>
<td>6.952811471</td>
<td>7.260839943</td>
<td>6.079018602</td>
<td></td>
</tr>
<tr>
<td>5.09407699</td>
<td>0.392471801</td>
<td>0.58218025</td>
<td>0.169278724</td>
<td>0.400963157</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>34</td>
<td>05</td>
<td>48</td>
<td>07</td>
<td></td>
</tr>
</tbody>
</table>

**like transposase family (Tnp2/En/Spm)**

**like transposase family (Tnp1/En/Spm)**
<table>
<thead>
<tr>
<th>AT2G16000</th>
<th>AT2G16000</th>
<th>AT2G11950</th>
<th>AT2G07550</th>
<th>AT2G07400</th>
<th>AT1G30180</th>
<th>AT5G30470</th>
<th>AT5G28926</th>
<th>A</th>
<th>45085</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.158340814</td>
<td>6.720861493</td>
<td>3.983552009</td>
<td>1.359</td>
<td>1.397860822</td>
<td>2.832287813</td>
<td>2.919275487</td>
<td>1.421169185</td>
<td>3.28011116</td>
<td>3.966917784</td>
</tr>
<tr>
<td>0.575387679</td>
<td>0.143833593</td>
<td>0.172507351</td>
<td>0.590626951</td>
<td>0.180244527</td>
<td>0.590101718</td>
<td>0.159267496</td>
<td>0.488407911</td>
<td>0.480776707</td>
<td>0.175857278</td>
</tr>
<tr>
<td>8.95E-2</td>
<td>2.86E-1</td>
<td>1.48E-1</td>
<td>3.37E-1</td>
<td>5.88E-1</td>
<td>3.33E-1</td>
<td>2.04E-1</td>
<td>9.86E-1</td>
<td>4.35E-2</td>
<td>2.28E-1</td>
</tr>
<tr>
<td>7.10E-1</td>
<td>4.34E-1</td>
<td>4.01E-1</td>
<td>3.37E-1</td>
<td>5.88E-1</td>
<td>3.33E-1</td>
<td>2.04E-1</td>
<td>9.86E-1</td>
<td>4.35E-2</td>
<td>2.28E-1</td>
</tr>
<tr>
<td>11</td>
<td>07</td>
<td>09</td>
<td>08</td>
<td>05</td>
<td>06</td>
<td>10</td>
<td>16</td>
<td>08</td>
<td>22</td>
</tr>
<tr>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT1G38450</td>
<td>4.2826734213</td>
<td>0.5071870229</td>
<td>6.46084346</td>
<td>1.34-30</td>
<td>3.56视线</td>
<td>0.08视线</td>
<td>6.70视线</td>
<td>0.56视线</td>
<td>1.28视线</td>
</tr>
<tr>
<td>AT1G38360</td>
<td>0.06视线</td>
<td>4.3视线</td>
<td>0.56视线</td>
<td>1.2视线</td>
<td>1.1视线</td>
<td>1.2视线</td>
<td>1.1视线</td>
<td>1.2视线</td>
<td>1.1视线</td>
</tr>
<tr>
<td>AT1G38330</td>
<td>0.06视线</td>
<td>4.3视线</td>
<td>0.56视线</td>
<td>1.2视线</td>
<td>1.1视线</td>
<td>1.2视线</td>
<td>1.1视线</td>
<td>1.2视线</td>
<td>1.1视线</td>
</tr>
<tr>
<td>AT1G36795</td>
<td>60视线</td>
<td>7.41视线</td>
<td>3.2视线</td>
<td>83.1视线</td>
<td>38.4视线</td>
<td>67视线</td>
<td>13视线</td>
<td>17视线</td>
<td>40视线</td>
</tr>
<tr>
<td>AT1G36540</td>
<td>6视线</td>
<td>7.4视线</td>
<td>3.2视线</td>
<td>83.1视线</td>
<td>38.4视线</td>
<td>67视线</td>
<td>13视线</td>
<td>17视线</td>
<td>40视线</td>
</tr>
<tr>
<td>AT5G35052</td>
<td>6视线</td>
<td>7.4视线</td>
<td>3.2视线</td>
<td>83.1视线</td>
<td>38.4视线</td>
<td>67视线</td>
<td>13视线</td>
<td>17视线</td>
<td>40视线</td>
</tr>
<tr>
<td>AT4G06682</td>
<td>6视线</td>
<td>7.4视线</td>
<td>3.2视线</td>
<td>83.1视线</td>
<td>38.4视线</td>
<td>67视线</td>
<td>13视线</td>
<td>17视线</td>
<td>40视线</td>
</tr>
<tr>
<td>AT4G04426</td>
<td>6视线</td>
<td>7.4视线</td>
<td>3.2视线</td>
<td>83.1视线</td>
<td>38.4视线</td>
<td>67视线</td>
<td>13视线</td>
<td>17视线</td>
<td>40视线</td>
</tr>
<tr>
<td>AT3G43867</td>
<td>6视线</td>
<td>7.4视线</td>
<td>3.2视线</td>
<td>83.1视线</td>
<td>38.4视线</td>
<td>67视线</td>
<td>13视线</td>
<td>17视线</td>
<td>40视线</td>
</tr>
<tr>
<td>AT2G16670</td>
<td>6视线</td>
<td>7.4视线</td>
<td>3.2视线</td>
<td>83.1视线</td>
<td>38.4视线</td>
<td>67视线</td>
<td>13视线</td>
<td>17视线</td>
<td>40视线</td>
</tr>
</tbody>
</table>

### Table 1: Summary of Gene Expression Data

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Description</th>
<th>Expression Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G38450</td>
<td>Gene 1</td>
<td>4.2826734213</td>
</tr>
<tr>
<td>AT1G38360</td>
<td>Gene 2</td>
<td>0.06</td>
</tr>
<tr>
<td>AT1G38330</td>
<td>Gene 3</td>
<td>0.06</td>
</tr>
<tr>
<td>AT1G36795</td>
<td>Gene 4</td>
<td>60</td>
</tr>
<tr>
<td>AT1G36540</td>
<td>Gene 5</td>
<td>6</td>
</tr>
<tr>
<td>AT5G35052</td>
<td>Gene 6</td>
<td>6</td>
</tr>
<tr>
<td>AT4G06682</td>
<td>Gene 7</td>
<td>6</td>
</tr>
<tr>
<td>AT4G04426</td>
<td>Gene 8</td>
<td>6</td>
</tr>
<tr>
<td>AT3G43867</td>
<td>Gene 9</td>
<td>6</td>
</tr>
<tr>
<td>AT2G16670</td>
<td>Gene 10</td>
<td>6</td>
</tr>
</tbody>
</table>

### Table 2: Gene Expression Comparison

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Description</th>
<th>Expression Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G38450</td>
<td>Gene 1</td>
<td>4.2826734213</td>
</tr>
<tr>
<td>AT1G38360</td>
<td>Gene 2</td>
<td>0.06</td>
</tr>
<tr>
<td>AT1G38330</td>
<td>Gene 3</td>
<td>0.06</td>
</tr>
<tr>
<td>AT1G36795</td>
<td>Gene 4</td>
<td>60</td>
</tr>
<tr>
<td>AT1G36540</td>
<td>Gene 5</td>
<td>6</td>
</tr>
<tr>
<td>AT5G35052</td>
<td>Gene 6</td>
<td>6</td>
</tr>
<tr>
<td>AT4G06682</td>
<td>Gene 7</td>
<td>6</td>
</tr>
<tr>
<td>AT4G04426</td>
<td>Gene 8</td>
<td>6</td>
</tr>
<tr>
<td>AT3G43867</td>
<td>Gene 9</td>
<td>6</td>
</tr>
<tr>
<td>AT2G16670</td>
<td>Gene 10</td>
<td>6</td>
</tr>
<tr>
<td>Gene 1</td>
<td>Gene 2</td>
<td>Gene 3</td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>AT3G32880</td>
<td>AT3G30713</td>
<td>AT2G11120</td>
</tr>
<tr>
<td>13.892168</td>
<td>17.894267</td>
<td>13.399389</td>
</tr>
<tr>
<td>16.86-06</td>
<td>1.99519.08</td>
<td>16.86-06</td>
</tr>
<tr>
<td>16.81-27</td>
<td>1.99519.08</td>
<td>16.81-27</td>
</tr>
<tr>
<td>AT3G29792</td>
<td>AT2G01026</td>
<td>AT4G05587</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
</tr>
</tbody>
</table>

### Table 1

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 3

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 5

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 6

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 7

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 8

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 9

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
</tbody>
</table>

### Table 10

<table>
<thead>
<tr>
<th>AT3G29792</th>
<th>AT2G01026</th>
<th>AT4G05587</th>
<th>AT5G35057</th>
<th>AT5G33050</th>
<th>AT5G31804</th>
<th>AT5G31087</th>
<th>12.929668</th>
<th>289.25982</th>
<th>45.83491</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1925654</td>
<td>2.6802759</td>
<td>1.2966643</td>
<td>1.8745843</td>
<td>0.5540613</td>
<td>0.3656152</td>
<td>0.6913011</td>
<td>3.0184450</td>
<td>3.3582332</td>
<td>1.5016574</td>
</tr>
<tr>
<td>AT2G15940</td>
<td>AT2G15160</td>
<td>AT2G14950</td>
<td>AT2G06760</td>
<td>AT2G05700</td>
<td>AT1G42110</td>
<td>AT5G31981</td>
<td>AT5G29562</td>
<td>AT5G29562</td>
<td>AT4G22415</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>91</td>
<td>0.489637743</td>
<td>0.160127077</td>
<td>0.509740879</td>
<td>0.129966326</td>
<td>0.489691757</td>
<td>0.473502323</td>
<td>0.428162796</td>
<td>0.511678809</td>
<td>0.588150227</td>
</tr>
<tr>
<td>8.64E</td>
<td>9.02E</td>
<td>6.79E</td>
<td>1.75E</td>
<td>2.08E</td>
<td>1.48E</td>
<td>1.14E</td>
<td>4.25E</td>
<td>7.24E</td>
<td>1.10E</td>
</tr>
<tr>
<td>130</td>
<td>1.101575979</td>
<td>0.116146416</td>
<td>1.101575979</td>
<td>0.694589061</td>
<td>0.653468127</td>
<td>0.653468127</td>
<td>1.718867729</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td>AT1G41680</td>
<td>AT1G40310</td>
<td>AT1G40135</td>
<td>AT1G40121</td>
<td>AT1G40121</td>
<td>AT1G40115</td>
<td>AT1G40072</td>
<td>AT1G38194</td>
<td>AT1G36763</td>
<td>AT5G38365</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>1.045789</td>
<td>0.000380</td>
<td>0.008410</td>
<td>0.108258</td>
<td>2.956015</td>
<td>0.356359</td>
<td>0.000024</td>
<td>0.000042</td>
<td>0.000042</td>
<td>0.000042</td>
</tr>
<tr>
<td>3.200000</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

| 1711.4608  | 39.766456  | 95.651336  | 12.384131  | 16.547026  | 5.1674818  | 11.102231  | 2.1844454  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  | 0.000000  |

**Mutator**

---

**transposable element gene.**

**LTR retrotransposon family (LINE)**
<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
<th>AT2G23710</th>
<th>AT2G23480</th>
<th>AT2G22710</th>
<th>AT2G15800</th>
<th>AT2G13320</th>
<th>AT2G13050</th>
<th>AT2G12910</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT2G23710</td>
<td>transposable element gene.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G23480</td>
<td>transposable element gene.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G22710</td>
<td>transposable element gene.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G15800</td>
<td>transposable element gene.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G13320</td>
<td>transposable element gene.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G13050</td>
<td>transposable element gene.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT2G12910</td>
<td>transposable element gene.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Additional Notes:**
- The table above lists genes with related descriptions.
- Each gene is identified with its accession number.
- The descriptions provide information on their category.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>16.191992</th>
<th>14.4</th>
<th>26.74328</th>
<th>18.11128</th>
<th>112.83328</th>
<th>38.368603</th>
<th>77.285393</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G33050</td>
<td>414654</td>
<td>1.7595</td>
<td>3.0679</td>
<td>4.1303</td>
<td>1.3921</td>
<td>0.495392</td>
<td>0.57841</td>
<td>0.442929</td>
</tr>
<tr>
<td>AT3G40120</td>
<td>2532941</td>
<td>2.4980</td>
<td>4.7971</td>
<td>5.5765</td>
<td>9.4503</td>
<td>8.82864</td>
<td>5.986553</td>
<td>2.765478</td>
</tr>
<tr>
<td>AT3G41000</td>
<td>3603701</td>
<td>1.2493</td>
<td>1.8493</td>
<td>3.0665</td>
<td>1.2380</td>
<td>0.319029</td>
<td>0.389194</td>
<td>0.274744</td>
</tr>
<tr>
<td>AT3G42100</td>
<td>3827060</td>
<td>1.0413</td>
<td>1.6413</td>
<td>2.8587</td>
<td>1.0785</td>
<td>0.279029</td>
<td>0.349194</td>
<td>0.234744</td>
</tr>
<tr>
<td>AT3G43980</td>
<td>2545580</td>
<td>1.1495</td>
<td>1.7495</td>
<td>3.2865</td>
<td>1.2185</td>
<td>0.319029</td>
<td>0.389194</td>
<td>0.274744</td>
</tr>
<tr>
<td>AT3G44000</td>
<td>3642941</td>
<td>1.2493</td>
<td>1.8493</td>
<td>3.0665</td>
<td>1.2380</td>
<td>0.319029</td>
<td>0.389194</td>
<td>0.274744</td>
</tr>
<tr>
<td>AT3G44120</td>
<td>38527060</td>
<td>1.0413</td>
<td>1.6413</td>
<td>2.8587</td>
<td>1.0785</td>
<td>0.279029</td>
<td>0.349194</td>
<td>0.234744</td>
</tr>
<tr>
<td>AT3G45120</td>
<td>4043600</td>
<td>1.2493</td>
<td>1.8493</td>
<td>3.0665</td>
<td>1.2380</td>
<td>0.319029</td>
<td>0.389194</td>
<td>0.274744</td>
</tr>
<tr>
<td>AT3G46100</td>
<td>4249270</td>
<td>1.0413</td>
<td>1.6413</td>
<td>2.8587</td>
<td>1.0785</td>
<td>0.279029</td>
<td>0.349194</td>
<td>0.234744</td>
</tr>
</tbody>
</table>

**Note:** The table contains gene accession numbers and various numerical data points, which may represent expression levels, functional categories, or other biological measurements. The context of the data requires additional information about the specific biological system or experiment to provide a detailed interpretation.
<table>
<thead>
<tr>
<th>Gene</th>
<th>AT5G30762</th>
<th>AT5G28870</th>
<th>AT5G27180</th>
<th>AT4G28960</th>
<th>AT4G09425</th>
<th>AT4G09370</th>
<th>AT4G08710</th>
<th>AT4G07939</th>
<th>AT4G07570</th>
<th>AT4G07520</th>
<th>AT4G07485</th>
<th>AT4G06736</th>
<th>AT4G06720</th>
<th>AT4G06544</th>
<th>AT4G06544</th>
<th>AT4G06533</th>
<th>AT4G06517</th>
<th>AT4G06517</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.650685782</td>
<td>6.215060583</td>
<td>4.030065422</td>
<td>2.728630574</td>
<td>1.875969921</td>
<td>3.72463792</td>
<td>0.906307289</td>
<td>5.559277784</td>
<td>9.233090089</td>
<td>3.125373588</td>
<td>0.863025863</td>
<td>1.15706901</td>
<td>2.555797231</td>
<td>4.2660467</td>
<td>14979365</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.142671814</td>
<td>0.531278785</td>
<td>0.579188404</td>
<td>0.513944674</td>
<td>0.421259382</td>
<td>0.590921074</td>
<td>0.526332181</td>
<td>0.583023552</td>
<td>0.512400757</td>
<td>0.38397787</td>
<td>0.180083665</td>
<td>0.491399001</td>
<td>0.403848484</td>
<td>1.65E5</td>
<td>3.90E4</td>
<td>14979365</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.34E3</td>
<td>3.01E3</td>
<td>1.92E3</td>
<td>4.06E3</td>
<td>1.23E3</td>
<td>3.93E3</td>
<td>2.00E3</td>
<td>1.17E3</td>
<td>5.84E3</td>
<td>7.78E3</td>
<td>1.55E4</td>
<td>9.90E3</td>
<td>9.08E3</td>
<td>2.88E3</td>
<td>1.18E3</td>
<td>5.43E3</td>
<td>1.65E4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**transposable element gene.**
| AT5G33391 | 4.052257715 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | 1.936048031 | A1+ | transposable element gene. |
| AT5G33255 | 1.993407045 | 0.990885637 | 4.052257715 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | A1+ | transposable element gene. |
| AT5G32511 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32434 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32241 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32107 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G30852 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |

| AT5G33391 | 4.052257715 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | 1.936048031 | A1+ | transposable element gene. |
| AT5G33255 | 1.993407045 | 0.990885637 | 4.052257715 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | A1+ | transposable element gene. |
| AT5G32511 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32434 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32241 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32107 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G30852 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |

| AT5G33391 | 4.052257715 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | 1.936048031 | A1+ | transposable element gene. |
| AT5G33255 | 1.993407045 | 0.990885637 | 4.052257715 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | A1+ | transposable element gene. |
| AT5G32511 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32434 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32241 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G32107 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
| AT5G30852 | 1.934557862 | 6.270129482 | 0.000000000 | 0.990885637 | 4.056250578 | 1.934557862 | 6.270129482 | 0.000000000 | A1+ | transposable element gene. |
**Supplementary Table 1**: List of transposable elements with at least log2 fold change of 2.5 in at least one of the four lines A1+, A1, A2+, A1.

<table>
<thead>
<tr>
<th>Transposable Element</th>
<th>Gene Name</th>
<th>Line A1+</th>
<th>Line A1</th>
<th>Line A2+</th>
<th>Line A1</th>
<th>Log2 Fold Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G05306</td>
<td>18.837320</td>
<td>1.07138885</td>
<td>1.16337811</td>
<td>6.54386684</td>
<td>0.012-11</td>
<td>5.66E-09</td>
</tr>
<tr>
<td>AT3G04906</td>
<td>14.13815</td>
<td>3.959424198</td>
<td>3.959381705</td>
<td>7.200975011</td>
<td>0.231-12</td>
<td>3.77E-11</td>
</tr>
<tr>
<td>AT3G05930</td>
<td>0.0219043</td>
<td>2.665155815</td>
<td>5.314281615</td>
<td>6.795439355</td>
<td>1.62-06</td>
<td>1.44E-05</td>
</tr>
<tr>
<td>AT3G05910</td>
<td>7.1123975</td>
<td>5.082525746</td>
<td>5.818249079</td>
<td>9.871647388</td>
<td>6.13E-07</td>
<td>7.60E-06</td>
</tr>
<tr>
<td>AT3G05905</td>
<td>0.6736985</td>
<td>0.599392562</td>
<td>0.137263358</td>
<td>0.946154301</td>
<td>1.75E-05</td>
<td>3.96E-03</td>
</tr>
<tr>
<td>AT1G50723</td>
<td>0.0441916</td>
<td>1.046217427</td>
<td>1.534818108</td>
<td>0.530855171</td>
<td>1.75E-07</td>
<td>1.08E-06</td>
</tr>
<tr>
<td>AT1G50780</td>
<td>0.000677</td>
<td>1.01774704</td>
<td>0.978379777</td>
<td>0.950682712</td>
<td>0.0009828</td>
<td>0.000234976</td>
</tr>
<tr>
<td>AT4G04890</td>
<td>6.717875</td>
<td>5.121353548</td>
<td>0.383765458</td>
<td>5.540486866</td>
<td>1.70E-08</td>
<td>5.52E-07</td>
</tr>
<tr>
<td>AT5G04305</td>
<td>19.863705</td>
<td>1.07168865</td>
<td>1.66731831</td>
<td>6.943946614</td>
<td>0.012-11</td>
<td>6.38E-09</td>
</tr>
<tr>
<td>AT5G02177</td>
<td>0.0445528</td>
<td>1.078620342</td>
<td>0.945674565</td>
<td>0.948147592</td>
<td>1.01E-07</td>
<td>1.38E-09</td>
</tr>
<tr>
<td>AT5G01355</td>
<td>12.535980</td>
<td>7.108067159</td>
<td>0.908574746</td>
<td>5.943917487</td>
<td>1.06E-12</td>
<td>0.44E-13</td>
</tr>
<tr>
<td>AT5G01354</td>
<td>7.1385666</td>
<td>1.230824056</td>
<td>0.899675195</td>
<td>0.558691541</td>
<td>1.75E-07</td>
<td>1.56E-06</td>
</tr>
<tr>
<td>AT1G51510</td>
<td>0.036035</td>
<td>3.007713627</td>
<td>0.895711555</td>
<td>0.970313888</td>
<td>0.66E-07</td>
<td>6.21E-06</td>
</tr>
<tr>
<td>AT5G02173</td>
<td>12.707371</td>
<td>3.643534543</td>
<td>1.056766763</td>
<td>6.907521063</td>
<td>3.01E-12</td>
<td>5.18E-10</td>
</tr>
<tr>
<td>AT5G02082</td>
<td>10.414294</td>
<td>5.250484767</td>
<td>0.430753175</td>
<td>3.290951393</td>
<td>1.32E-03</td>
<td>1.81E-03</td>
</tr>
<tr>
<td>AT5G01987</td>
<td>14.881048</td>
<td>1.281771074</td>
<td>1.281815184</td>
<td>1.7933849</td>
<td>4.75E-07</td>
<td>2.13E-09</td>
</tr>
<tr>
<td>AT5G02085</td>
<td>12.436503</td>
<td>0.460143759</td>
<td>0.441293597</td>
<td>0.68E-17</td>
<td>1.82E-15</td>
<td>0.40E48721</td>
</tr>
<tr>
<td>AT5G02086</td>
<td>14.449411</td>
<td>3.31043950</td>
<td>0.460118317</td>
<td>7.909392506</td>
<td>1.75E-12</td>
<td>4.14E-11</td>
</tr>
<tr>
<td>AT5G02071</td>
<td>12.18168</td>
<td>0.650458482</td>
<td>0.478224641</td>
<td>0.582025678</td>
<td>0.36E-06</td>
<td>4.35E-08</td>
</tr>
<tr>
<td>AT5G01873</td>
<td>11.819740</td>
<td>1.124531355</td>
<td>1.057753813</td>
<td>6.93486666</td>
<td>1.75E-11</td>
<td>0.05E-09</td>
</tr>
<tr>
<td>AT5G01872</td>
<td>0.0383983</td>
<td>0.847927619</td>
<td>0.282987215</td>
<td>0.699153815</td>
<td>1.57E-07</td>
<td>2.77E-06</td>
</tr>
<tr>
<td>AT5G01870</td>
<td>1.100397</td>
<td>0.814573695</td>
<td>0.22207465</td>
<td>0.576798368</td>
<td>0.72E-05</td>
<td>1.71E-04</td>
</tr>
<tr>
<td>AT5G01868</td>
<td>0.043656</td>
<td>0.792617469</td>
<td>0.08672204</td>
<td>0.305672468</td>
<td>0.00679088</td>
<td>0.000825489</td>
</tr>
<tr>
<td>AT5G01867</td>
<td>1.101576</td>
<td>0.859109509</td>
<td>0.30793458</td>
<td>0.974143542</td>
<td>0.016732</td>
<td>0.07E45721</td>
</tr>
<tr>
<td>AT5G01865</td>
<td>1.101576</td>
<td>0.859109509</td>
<td>0.30793458</td>
<td>0.974143542</td>
<td>0.016732</td>
<td>0.07E45721</td>
</tr>
<tr>
<td>AT5G01866</td>
<td>1.101576</td>
<td>0.859109509</td>
<td>0.30793458</td>
<td>0.974143542</td>
<td>0.016732</td>
<td>0.07E45721</td>
</tr>
<tr>
<td>AT5G01864</td>
<td>1.101576</td>
<td>0.859109509</td>
<td>0.30793458</td>
<td>0.974143542</td>
<td>0.016732</td>
<td>0.07E45721</td>
</tr>
</tbody>
</table>

--

**Note**: The table lists transposable elements with at least log2 fold change of 2.5 in at least one of the four lines A1+, A1, A2+, A1. The log2 fold change is calculated as the ratio of expression levels between the transgenic line and the wild type. Positive fold changes indicate an increase in expression, while negative fold changes indicate a decrease. The elements are categorized based on their type and function.
<table>
<thead>
<tr>
<th>Transposable elements</th>
<th>No of genes</th>
<th>Genes with heritable changes</th>
<th>Percentage heritable changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CACTA-like transposase family (En/Spm)</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CACTA-like transposase family (Ptta/En/Spm)</td>
<td>59</td>
<td>39</td>
<td>66.1</td>
</tr>
<tr>
<td>CACTA-like transposase family (Tnp1/En/Spm)</td>
<td>18</td>
<td>3</td>
<td>16.7</td>
</tr>
<tr>
<td>CACTA-like transposase family (Tnp2/En/Spm)</td>
<td>32</td>
<td>7</td>
<td>21.9</td>
</tr>
<tr>
<td>CACTA-like transposase family, putative</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>copia-like retrotranspon family (Ty1-Copia-element)</td>
<td>19</td>
<td>7</td>
<td>36.8</td>
</tr>
<tr>
<td>gypsy-like retrotranspon family (Athila)</td>
<td>59</td>
<td>39</td>
<td>66.1</td>
</tr>
<tr>
<td>gypsy-like retrotranspon pseudogene (Athila)</td>
<td>4</td>
<td>4</td>
<td>100</td>
</tr>
<tr>
<td>gypsy-like retrotranspon genes and pseudogenes (Athila)</td>
<td>63</td>
<td>43</td>
<td>68.3</td>
</tr>
<tr>
<td>gypsy-like retrotranspon family (Ty3-element)</td>
<td>26</td>
<td>16</td>
<td>61.5</td>
</tr>
<tr>
<td>hAT-like transposase family (hobo/Ac/Tam3)</td>
<td>12</td>
<td>9</td>
<td>75</td>
</tr>
<tr>
<td>Mutator-like transposase family</td>
<td>24</td>
<td>16</td>
<td>66.7</td>
</tr>
<tr>
<td>non-LTR retrotranspon family (LINE)</td>
<td>11</td>
<td>8</td>
<td>72.7</td>
</tr>
<tr>
<td>transposable element gene</td>
<td>64</td>
<td>37</td>
<td>57.8</td>
</tr>
<tr>
<td>transposable element gene; pseudogene, hypothetical protein</td>
<td>86</td>
<td>69</td>
<td>80.2</td>
</tr>
<tr>
<td>AT1G05913</td>
<td>AT5G03195</td>
<td>AT5G03195</td>
<td>AT3G01055</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>4.8081274</td>
<td>4.8920416</td>
<td>4.8489135</td>
<td>4.8749734</td>
</tr>
<tr>
<td>4.5327085</td>
<td>4.5327085</td>
<td>4.5327085</td>
<td>4.5327085</td>
</tr>
<tr>
<td>4.3551469</td>
<td>4.3551469</td>
<td>4.3551469</td>
<td>4.3551469</td>
</tr>
<tr>
<td>4.0628368</td>
<td>4.0628368</td>
<td>4.0628368</td>
<td>4.0628368</td>
</tr>
<tr>
<td>3.7381287</td>
<td>3.7381287</td>
<td>3.7381287</td>
<td>3.7381287</td>
</tr>
<tr>
<td>3.4114728</td>
<td>3.4114728</td>
<td>3.4114728</td>
<td>3.4114728</td>
</tr>
<tr>
<td>2.8415537</td>
<td>2.8415537</td>
<td>2.8415537</td>
<td>2.8415537</td>
</tr>
</tbody>
</table>

**Table 2: Summary of transposable elements with altered transcript levels and their heritability rates.** Data were compiled for different categories of transposable elements (S1 table) that showed at least log2-fold changes of +/- 2.5 in line A1+ compared to wildtype. For each gene, the values in A1+ and A1- were compared to score the heritability of expression changes.
<table>
<thead>
<tr>
<th>Gene1</th>
<th>Gene2</th>
<th>Gene3</th>
<th>Gene4</th>
<th>Gene5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT4G15242</td>
<td>AT4G05025</td>
<td>AT3G09855</td>
<td>AT3G06955</td>
<td>AT3G06715</td>
</tr>
<tr>
<td>69.574642</td>
<td>33.731313</td>
<td>21.767391</td>
<td>88.600286</td>
<td>24.620587</td>
</tr>
<tr>
<td>40.076715</td>
<td>31.925181</td>
<td>157.56276</td>
<td>36.282964</td>
<td>10.843168</td>
</tr>
</tbody>
</table>

**Table: Gene Expression Studies**

<table>
<thead>
<tr>
<th>Gene1</th>
<th>Gene2</th>
<th>Gene3</th>
<th>Gene4</th>
<th>Gene5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G06715</td>
<td>AT1G06321</td>
<td>AT4G19651</td>
<td>AT4G09450</td>
<td>AT1G17487</td>
</tr>
<tr>
<td>0.768408</td>
<td>0.94245294</td>
<td>0.29342174</td>
<td>0.40114857</td>
<td>0.39867511</td>
</tr>
</tbody>
</table>

**Table: Additional Gene Expression Data**

<table>
<thead>
<tr>
<th>Gene1</th>
<th>Gene2</th>
<th>Gene3</th>
<th>Gene4</th>
<th>Gene5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G06321</td>
<td>AT4G19651</td>
<td>AT4G09450</td>
<td>AT1G17487</td>
<td>AT2G10230</td>
</tr>
<tr>
<td>2.925224</td>
<td>4.14364914</td>
<td>3.79399475</td>
<td>5.923057</td>
<td>3.294327</td>
</tr>
</tbody>
</table>

**Table: Further Gene Expression Measurements**

<table>
<thead>
<tr>
<th>Gene1</th>
<th>Gene2</th>
<th>Gene3</th>
<th>Gene4</th>
<th>Gene5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G06321</td>
<td>AT4G19651</td>
<td>AT4G09450</td>
<td>AT1G17487</td>
<td>AT2G10230</td>
</tr>
<tr>
<td>7.558966</td>
<td>6.45052134</td>
<td>1.87039716</td>
<td>5.38024211</td>
<td>0.70437554</td>
</tr>
</tbody>
</table>

**Table: Additional Gene Expression Results**

<table>
<thead>
<tr>
<th>Gene1</th>
<th>Gene2</th>
<th>Gene3</th>
<th>Gene4</th>
<th>Gene5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G06321</td>
<td>AT4G19651</td>
<td>AT4G09450</td>
<td>AT1G17487</td>
<td>AT2G10230</td>
</tr>
<tr>
<td>5.709246</td>
<td>8.25446021</td>
<td>1.4623804</td>
<td>0.74812207</td>
<td>0.54738703</td>
</tr>
</tbody>
</table>

**Table: Gene Expression Analysis**

<table>
<thead>
<tr>
<th>Gene1</th>
<th>Gene2</th>
<th>Gene3</th>
<th>Gene4</th>
<th>Gene5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G06321</td>
<td>AT4G19651</td>
<td>AT4G09450</td>
<td>AT1G17487</td>
<td>AT2G10230</td>
</tr>
<tr>
<td>0.360704315</td>
<td>0.389840635</td>
<td>0.210388554</td>
<td>0.391418267</td>
<td>0.550896091</td>
</tr>
</tbody>
</table>

**Table: Additional Gene Expression Values**

<table>
<thead>
<tr>
<th>Gene1</th>
<th>Gene2</th>
<th>Gene3</th>
<th>Gene4</th>
<th>Gene5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G06321</td>
<td>AT4G19651</td>
<td>AT4G09450</td>
<td>AT1G17487</td>
<td>AT2G10230</td>
</tr>
<tr>
<td>0.360704315</td>
<td>0.389840635</td>
<td>0.210388554</td>
<td>0.391418267</td>
<td>0.550896091</td>
</tr>
</tbody>
</table>

**Table: Further Gene Expression Data**

<table>
<thead>
<tr>
<th>Gene1</th>
<th>Gene2</th>
<th>Gene3</th>
<th>Gene4</th>
<th>Gene5</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G06321</td>
<td>AT4G19651</td>
<td>AT4G09450</td>
<td>AT1G17487</td>
<td>AT2G10230</td>
</tr>
<tr>
<td>0.360704315</td>
<td>0.389840635</td>
<td>0.210388554</td>
<td>0.391418267</td>
<td>0.550896091</td>
</tr>
</tbody>
</table>

**Table: Gene Expression Summary**
<table>
<thead>
<tr>
<th>AT1G12013</th>
<th>AT1G07897</th>
<th>AT1G07593</th>
<th>AT1G05917</th>
<th>AT1G05247</th>
<th>AT1G04527</th>
<th>AT3G4176</th>
</tr>
</thead>
<tbody>
<tr>
<td>49.197211</td>
<td>37.315144</td>
<td>129.85936</td>
<td>27.692177</td>
<td>461.63003</td>
<td>537.78415</td>
<td>28.6164</td>
</tr>
<tr>
<td>2.3974089</td>
<td>2.566643</td>
<td>2.4537107</td>
<td>2.4221971</td>
<td>3.3889002</td>
<td>3.3688507</td>
<td>0.4494234</td>
</tr>
<tr>
<td>0.356426406</td>
<td>0.279585449</td>
<td>0.26587131</td>
<td>0.189447222</td>
<td>0.298007517</td>
<td>0.3688507</td>
<td>0.463000314</td>
</tr>
<tr>
<td>3.80E-14</td>
<td>1.49E-03</td>
<td>2.60E-03</td>
<td>9.18E-04</td>
<td>4.30E-04</td>
<td>1.93E-04</td>
<td>7.93E-05</td>
</tr>
<tr>
<td>5.10E-14</td>
<td>2.76E-04</td>
<td>9.77E-05</td>
<td>4.76E-04</td>
<td>2.71E-04</td>
<td>4.70E-04</td>
<td>2.90E-04</td>
</tr>
<tr>
<td>61E-14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>snoRNA</th>
<th>snoRNA</th>
<th>snoRNA</th>
<th>snoRNA</th>
<th>snoRNA</th>
<th>snoRNA</th>
<th>snoRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>box snoRNA (snoR111)</td>
<td>box snoRNA (snoR111)</td>
<td>box snoRNA (snoR111)</td>
<td>box snoRNA (snoR111)</td>
<td>box snoRNA (snoR111)</td>
<td>box snoRNA (snoR111)</td>
<td>box snoRNA (snoR111)</td>
</tr>
</tbody>
</table>

| A1 | A1 | A1 | A1 | A2 | A1 | A1 |

<table>
<thead>
<tr>
<th>AT1G12013</th>
<th>AT1G07897</th>
<th>AT1G07593</th>
<th>AT1G05917</th>
<th>AT1G05247</th>
<th>AT1G04527</th>
<th>AT3G4176</th>
</tr>
</thead>
<tbody>
<tr>
<td>49.197211</td>
<td>37.315144</td>
<td>129.85936</td>
<td>27.692177</td>
<td>461.63003</td>
<td>537.78415</td>
<td>28.6164</td>
</tr>
<tr>
<td>2.3974089</td>
<td>2.566643</td>
<td>2.4537107</td>
<td>2.4221971</td>
<td>3.3889002</td>
<td>3.3688507</td>
<td>0.4494234</td>
</tr>
<tr>
<td>0.356426406</td>
<td>0.279585449</td>
<td>0.26587131</td>
<td>0.189447222</td>
<td>0.298007517</td>
<td>0.3688507</td>
<td>0.463000314</td>
</tr>
<tr>
<td>3.80E-14</td>
<td>1.49E-03</td>
<td>2.60E-03</td>
<td>9.18E-04</td>
<td>4.30E-04</td>
<td>1.93E-04</td>
<td>7.93E-05</td>
</tr>
<tr>
<td>5.10E-14</td>
<td>2.76E-04</td>
<td>9.77E-05</td>
<td>4.76E-04</td>
<td>2.71E-04</td>
<td>4.70E-04</td>
<td>2.90E-04</td>
</tr>
<tr>
<td>61E-14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Gene ID</td>
<td>AT3G14735</td>
<td>AT3G0533</td>
<td>AT3G02445</td>
<td>AT3G02445</td>
<td>AT3G01155</td>
<td>AT2G35747</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>----------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>1</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
</tr>
<tr>
<td>2</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
</tr>
<tr>
<td>3</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
</tr>
</tbody>
</table>

**Table:**

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>AT3G14735</th>
<th>AT3G0533</th>
<th>AT3G02445</th>
<th>AT3G02445</th>
<th>AT3G01155</th>
<th>AT2G35747</th>
<th>AT2G3</th>
<th>AT2G09705</th>
<th>AT2G08900</th>
<th>AT2G07605</th>
<th>AT2G07605</th>
<th>AT1G75163</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
<tr>
<td>2</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
<tr>
<td>3</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
</tbody>
</table>

**Table:**

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>AT3G14735</th>
<th>AT3G0533</th>
<th>AT3G02445</th>
<th>AT3G02445</th>
<th>AT3G01155</th>
<th>AT2G35747</th>
<th>AT2G3</th>
<th>AT2G09705</th>
<th>AT2G08900</th>
<th>AT2G07605</th>
<th>AT2G07605</th>
<th>AT1G75163</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
<tr>
<td>2</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
<tr>
<td>3</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
</tbody>
</table>

**Table:**

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>AT3G14735</th>
<th>AT3G0533</th>
<th>AT3G02445</th>
<th>AT3G02445</th>
<th>AT3G01155</th>
<th>AT2G35747</th>
<th>AT2G3</th>
<th>AT2G09705</th>
<th>AT2G08900</th>
<th>AT2G07605</th>
<th>AT2G07605</th>
<th>AT1G75163</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
<tr>
<td>2</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
<tr>
<td>3</td>
<td>0.290928</td>
<td>0.629258</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
<td>1.010583</td>
<td>0.98647</td>
<td>0.594072</td>
<td>0.618069</td>
<td>0.312611</td>
<td>0.250062</td>
<td>0.590583</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Expression Value</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G28305</td>
<td>0.9692475</td>
<td>1.30322792</td>
<td>0.24788076</td>
<td>15.770077</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G09515</td>
<td>9.00700049</td>
<td>1.60419654</td>
<td>0.23430325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G08515</td>
<td>2.07500725</td>
<td>1.24930787</td>
<td>0.75300325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G07105</td>
<td>0.01409937</td>
<td>0.23486644</td>
<td>0.18034325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G07105</td>
<td>9.00700049</td>
<td>1.60419654</td>
<td>0.23430325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G02665</td>
<td>2.07500725</td>
<td>1.24930787</td>
<td>0.75300325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G02655</td>
<td>0.01409937</td>
<td>0.23486644</td>
<td>0.18034325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G01785</td>
<td>9.00700049</td>
<td>1.60419654</td>
<td>0.23430325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G00750</td>
<td>2.07500725</td>
<td>1.24930787</td>
<td>0.75300325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G39366</td>
<td>7.30462472</td>
<td>0.90000890</td>
<td>1.23842897</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G39361</td>
<td>9.00700049</td>
<td>1.60419654</td>
<td>0.23430325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G15258</td>
<td>2.07500725</td>
<td>1.24930787</td>
<td>0.75300325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G06250</td>
<td>9.00700049</td>
<td>1.60419654</td>
<td>0.23430325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G47348</td>
<td>9.00700049</td>
<td>1.60419654</td>
<td>0.23430325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G47358</td>
<td>2.07500725</td>
<td>1.24930787</td>
<td>0.75300325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G21805</td>
<td>7.30462472</td>
<td>0.90000890</td>
<td>1.23842897</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G39175</td>
<td>2.07500725</td>
<td>1.24930787</td>
<td>0.75300325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G39185</td>
<td>9.00700049</td>
<td>1.60419654</td>
<td>0.23430325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G39195</td>
<td>2.07500725</td>
<td>1.24930787</td>
<td>0.75300325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G39205</td>
<td>9.00700049</td>
<td>1.60419654</td>
<td>0.23430325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G39215</td>
<td>2.07500725</td>
<td>1.24930787</td>
<td>0.75300325</td>
<td>13.193353</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The table above represents gene expression data with gene IDs and corresponding expression values. Each gene ID is associated with a specific gene expression level. The expressions are quantified in arbitrary units.
**Supplementary Table 3:** List of non-coding RNAs with at least log2 fold-change differences (negative log2 fold-change) or decreases (positive log2 fold-change) of 2.5 in at least one of the four lines A1+, A1-, A2+ or A2-.

| AT2G07759  | AT2G07759  | AT2G07754  | AT2G07754  | AT5G40395  | AT5G04465  | AT5G04465  | AT4G04185  | AT4G04185  | AT1G05853  | AT1G04263  | AT5G66564  | AT5G66564  | AT5G51174  | AT5G51174  | AT5G51174  | AT5G51174  | AT5G13225  | AT5G13225  | AT5G13225  | AT5G13225  |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| 137.2616 | 228.85452 | 2.6722019 | 2.6265867 | 2.5558965 | 1.9429112 | 3.1853596 | 4.3088034 | 6.825543535 | 8.76E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E |
| 0.334587185 | 0.353380812 | 0.278126205 | 0.250223738 | 0.282282904 | 0.402518134 | 0.298422822 | 0.264606008 | 6.825543535 | 8.76E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E | 3.43E |

**Note:** The table lists non-coding RNAs with at least log2 fold-change differences (negative log2 fold-change) or decreases (positive log2 fold-change) of 2.5 in at least one of the four lines A1+, A1-, A2+ or A2-.
Supplementary Table 4: Summary of non-coding RNAs with altered transcript levels and their heritability rates. Data were compiled for different categories of genes expressing non-coding RNAs (S3 Table) that showed at least log2-fold changes of +/- 2.5 in line A1+ compared to wildtype. For each gene, the values in A1+ and A1- were compared to score the heritability of expression changes.
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Strain</th>
<th>Esr</th>
<th>Nsp</th>
<th>Dsp</th>
<th>Msp</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT1G21400</td>
<td>18.83978</td>
<td>0.31820</td>
<td>0.76166</td>
<td>1.28728</td>
<td>2.36864</td>
<td>BUL, ABF1, Epithelial cell adhesion molecule, resistance to oxidative stress, playing a role in stress response and cytokinesis.</td>
</tr>
<tr>
<td>AT1G20310</td>
<td>13.6841</td>
<td>9.49059</td>
<td>1.47E+06</td>
<td>1.14E+06</td>
<td>1.36E+06</td>
<td>Auxin response factor, Auxin/indole-3-acetic acid response factor, role in auxin signaling.</td>
</tr>
<tr>
<td>AT1G01060</td>
<td>3.14829</td>
<td>2.78398</td>
<td>3.36205</td>
<td>2.45810</td>
<td>2.09505</td>
<td>Functions in auxin signaling and response.</td>
</tr>
</tbody>
</table>

**Notes:**
- **Esr:** Effective signal ratio
- **Nsp:** Normal signal ratio
- **Dsp:** Differential signal ratio
- **Msp:** Mean signal ratio
- **Description:** Detailed function and role in various biological processes.
<table>
<thead>
<tr>
<th>AT1G65610</th>
<th>AT1G65481</th>
<th>AT1G61800</th>
<th>AT1G59920</th>
<th>AT1G57650</th>
<th>AT1G55390</th>
<th>AT1G51850</th>
<th>AT1G42980</th>
<th>AT1G33760</th>
<th>AT1G32080</th>
<th>AT1G31095</th>
<th>AT1G24260</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.08165</td>
<td>52.363097</td>
<td>74.177788</td>
<td>25.507512</td>
<td>18.541913</td>
<td>76.765807</td>
<td>481.38601</td>
<td>83.406609</td>
<td>24.791551</td>
<td>10.52504</td>
<td>20.558726</td>
<td>394.42426</td>
</tr>
<tr>
<td>2.800766342</td>
<td>3.998164215</td>
<td>-</td>
<td>2.656145851</td>
<td>1.00523313</td>
<td>-</td>
<td>-</td>
<td>2.527846059</td>
<td>-</td>
<td>-</td>
<td>4.268422494</td>
<td>-</td>
</tr>
<tr>
<td>0.26759391</td>
<td>0.40154293</td>
<td>0.24956056</td>
<td>0</td>
<td>0.48959777</td>
<td>0.53315789</td>
<td>0.54177664</td>
<td>0.36499634</td>
<td>0.56591766</td>
<td>0.56376718</td>
<td>0.3839564</td>
<td>0.56376718</td>
</tr>
</tbody>
</table>

---

**Notes:**
- Various proteins are listed, each with a specific function and involvement in different processes.
- The table includes activities such as kinase activity, binding, translation, and specific binding.
- Some proteins are involved in binding events with other proteins or entities.
- There are genetic expressions and interactions related to various processes, including defense and development.

IDA, INFLORESCENCE DEFICIENT IN ABSCISSION, Putative membrane lipoprotein, encodes a small protein of 77 amino acids. casm
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Expression</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT2G15890</td>
<td>80.35867</td>
<td></td>
</tr>
<tr>
<td>AT2G11778</td>
<td>150.86473</td>
<td></td>
</tr>
<tr>
<td>AT2G07240</td>
<td>6.4377869</td>
<td></td>
</tr>
<tr>
<td>AT1G76410</td>
<td>28.995233</td>
<td></td>
</tr>
<tr>
<td>AT1G69930</td>
<td>164.05302</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>16.97418</td>
<td></td>
</tr>
<tr>
<td>AT1G05560</td>
<td>133.91134</td>
<td></td>
</tr>
<tr>
<td>AT1G76410</td>
<td>27.70764</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.18026689</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.42963267</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.15548841</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.47770972</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.17380294</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.52298166</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.52590128</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.37756237</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>04</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>09</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>09</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>04</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>1.80E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>2.57E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>2.43E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>3.58E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>7.62E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>9.17E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>2.91E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>2.65E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>9.73E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>9.51E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>3.36E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>3.73E+01</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>10.3562651</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.71097342</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.64014872</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>8.24916071</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.5990688</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>4.04846158</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.28456251</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.01236438</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>10.3562651</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.71097342</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.64014872</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>8.24916071</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.5990688</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>4.04846158</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.28456251</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.01236438</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>10.3562651</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.71097342</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.64014872</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>8.24916071</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.5990688</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>4.04846158</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.28456251</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.01236438</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>10.3562651</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.71097342</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.64014872</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>8.24916071</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.5990688</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>4.04846158</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.28456251</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.01236438</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>10.3562651</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>5.71097342</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0.64014872</td>
<td></td>
</tr>
<tr>
<td>AT1G62987</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

**Function mutations are deletion in the process of ethylene independent floral organ abscission. Although the flowers have a normal appearing abscission zone, the floral organs do not abscise. The peptide appears to be secreted and may function as a ligand. Arabidopsis 35S:GUS lines completely overexpressing HA exhibit earlier abscission of floral organs, showing that the abscission zones are responsive to GA after the opening of the flowers. In addition, epidermal abscission was observed at the base of the pedicel, branches of the inflorescence, and cauline leaves. The osmotic values also declined prematurely.**
<table>
<thead>
<tr>
<th>Accession</th>
<th>AT2G36790</th>
<th>AT2G35930</th>
<th>AT2G34130</th>
<th>AT2G33830</th>
<th>AT2G33830</th>
<th>AT2G25460</th>
<th>AT2G19800</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>36.425532</td>
<td>93.085353</td>
<td>178.91183</td>
<td>71.059387</td>
<td>1938.0129</td>
<td>17.897106</td>
<td>1.783897</td>
</tr>
<tr>
<td>Value</td>
<td>0.788484544</td>
<td>3.343311505</td>
<td>-</td>
<td>-</td>
<td>2.696036325</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Value</td>
<td>0.56146439</td>
<td>0.17873414</td>
<td>0.18708031</td>
<td>0.47375406</td>
<td>0.2114287</td>
<td>3.542659</td>
<td>11.7707</td>
</tr>
<tr>
<td>Value</td>
<td>14</td>
<td>09</td>
<td>42</td>
<td>32</td>
<td>25</td>
<td>08</td>
<td>27</td>
</tr>
<tr>
<td>Value</td>
<td>1.48E+09</td>
<td>9.50E+09</td>
<td>3.28E+09</td>
<td>6.01E+09</td>
<td>6.07596656</td>
<td>4.88274439</td>
<td>12.366197</td>
</tr>
<tr>
<td>Value</td>
<td>0.15392552</td>
<td>5235</td>
<td>1.66E+09</td>
<td>3.28E+09</td>
<td>6.13896187</td>
<td>10.5590142</td>
<td>12.023298</td>
</tr>
</tbody>
</table>

**Regulatory Type**

- **AT2G36790**: conserved within the At1g19630 family.
- **AT2G35930**: conserved within the At1g19630 family.
- **AT2G34130**: conserved within the At1g19630 family.
- **AT2G33830**: conserved within the At1g19630 family.

**Regulatory Mechanisms**

- **AT2G36790**: interacts with At1g07194 via a phi interaction.
- **AT2G35930**: interacts with At1g07194 via a phi interaction.
- **AT2G34130**: interacts with At1g07194 via a phi interaction.
- **AT2G33830**: interacts with At1g07194 via a phi interaction.

**Expression Data**

- **AT2G36790**: expressed in the late embryogenesis abiotic stress (LEA) gene family.
- **AT2G35930**: expressed in the LEA gene family.
- **AT2G34130**: expressed in the LEA gene family.
- **AT2G33830**: expressed in the LEA gene family.

**Genetic Analysis**

- **AT2G36790**: regulates a subset of late embryogenesis abiotic stress, including decreased sensitivity to ABA inhibition of germination and altered expression of some ABA response regulators.
- **AT2G35930**: regulates the circadian oscillations of its own transcript. Gene expression is induced by cold.
- **AT2G34130**: appears to promote osmotic opening and reduce tolerance under salt and dehydration stress conditions, but promotes osmotic closing and thereby increases stress tolerance under conditions of cold tolerance.
- **AT2G33830**: encodes a putative transcription factor.

**Functional Analysis**

- **AT2G36790**: encodes a putative transcription factor.
- **AT2G35930**: encodes a putative transcription factor.
- **AT2G34130**: encodes a putative transcription factor.
- **AT2G33830**: encodes a putative transcription factor.

**Additional Information**

- **AT2G36790**: involved in the regulation of ABA response, including decreased sensitivity to ABA inhibition of germination.
- **AT2G35930**: involved in the regulation of ABA response, including decreased sensitivity to ABA inhibition of germination.
- **AT2G34130**: involved in the regulation of ABA response, including decreased sensitivity to ABA inhibition of germination.
- **AT2G33830**: involved in the regulation of ABA response, including decreased sensitivity to ABA inhibition of germination.

**References**

### Table 1: Functional Classification of Arabidopsis Genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G16860</td>
<td>SHYH, SHY5, RING/U binding; Calmodulin ATPase activity, coupled to transmembrane movement of substances, glutathione S-transferase family protein</td>
</tr>
<tr>
<td>AT3G15440</td>
<td>ERF, ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 18 members in this subfamily including ATIS-1, ATIS-2, and ATIS-3.</td>
</tr>
<tr>
<td>AT3G14735</td>
<td>Small auxin upregulated RNA (SAUR) is a small family of auxin-inducible genes that encode small polypeptides (SAUR). The SAUR proteins are involved in auxin-mediated growth and development, including root growth, lateral root formation, and floral organ development.</td>
</tr>
<tr>
<td>AT3G14450</td>
<td>Small auxin upregulated RNA 48, SAUR48, SAUR family,eshire of oxygen in the root tip. Exhibits apolar plasma membrane localization in the root cap and polar localization in tissues above and is involved in root hair degeneration.</td>
</tr>
<tr>
<td>AT3G13600</td>
<td>Small auxin upregulated RNA 3, SAUR3, multidrug resistance subfamily 3, multidrug resistance protein 7 (MRP7); MRP7, multidrug resistance protein 7, auxiliary output; MRP3; FUNCTIONS IN: chlorophyll catabolite transmembrane transporter activity, resistance P (PXR) subfamily B (PXR); functions in the basipetal redirection of auxin from the root tip. Exhibits apolar plasma membrane localization in the root cap and polar localization in tissues above and is involved in root hair degeneration.</td>
</tr>
<tr>
<td>AT3G10986</td>
<td>Small auxin upregulated RNA 1, SAUR1, multidrug resistance subfamily 2, nucleolin like 2 (NUC2); NUC2, nucleolin-like protein; SMALL AUXIN UPREGULATED RNA 2 (SAUR2), SAUR, small auxin upregulated RNA, a small family of auxin-inducible genes that encode small polypeptides. The SAUR proteins are involved in auxin-mediated growth and development, including root growth, lateral root formation, and floral organ development.</td>
</tr>
<tr>
<td>AT3G06890</td>
<td>GRF, GLYCOPROTEIN 4, Encodes an auxin efflux transmembrane transporter that is a member of the multidrug resistance superfamily. The gene is involved in auxin transport and auxin-mediated cellular response.</td>
</tr>
<tr>
<td>AT3G05320</td>
<td>GRF, GLYCOPROTEIN 4, Encodes an auxin efflux transmembrane transporter that is a member of the multidrug resistance superfamily. The gene is involved in auxin transport and auxin-mediated cellular response.</td>
</tr>
<tr>
<td>AT3G01760</td>
<td>GRF, GLYCOPROTEIN 4, Encodes an auxin efflux transmembrane transporter that is a member of the multidrug resistance superfamily. The gene is involved in auxin transport and auxin-mediated cellular response.</td>
</tr>
<tr>
<td>AT3G01345</td>
<td>GRF, GLYCOPROTEIN 4, Encodes an auxin efflux transmembrane transporter that is a member of the multidrug resistance superfamily. The gene is involved in auxin transport and auxin-mediated cellular response.</td>
</tr>
</tbody>
</table>

### Table 2: Expression Analysis

<table>
<thead>
<tr>
<th>Gene</th>
<th>Expression Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G16860</td>
<td>2.764927</td>
</tr>
<tr>
<td>AT3G15440</td>
<td>0.491451</td>
</tr>
<tr>
<td>AT3G14735</td>
<td>1.787649</td>
</tr>
<tr>
<td>AT3G14450</td>
<td>1.848187</td>
</tr>
<tr>
<td>AT3G13600</td>
<td>3.566355</td>
</tr>
<tr>
<td>AT3G10986</td>
<td>3.852767</td>
</tr>
<tr>
<td>AT3G06890</td>
<td>2.479528</td>
</tr>
<tr>
<td>AT3G05320</td>
<td>4.166408</td>
</tr>
<tr>
<td>AT3G01760</td>
<td>1.254629</td>
</tr>
<tr>
<td>AT3G01345</td>
<td>3.766379</td>
</tr>
</tbody>
</table>

### Table 3: Functional Category:

<table>
<thead>
<tr>
<th>Gene</th>
<th>Functional Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G16860</td>
<td>SHYH, SHY5, RING/U binding; Calmodulin ATPase activity, coupled to transmembrane movement of substances, glutathione S-transferase family protein</td>
</tr>
<tr>
<td>AT3G15440</td>
<td>ERF, ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 18 members in this subfamily including ATIS-1, ATIS-2, and ATIS-3.</td>
</tr>
<tr>
<td>AT3G14735</td>
<td>Small auxin upregulated RNA (SAUR) is a small family of auxin-inducible genes that encode small polypeptides (SAUR). The SAUR proteins are involved in auxin-mediated growth and development, including root growth, lateral root formation, and floral organ development.</td>
</tr>
<tr>
<td>AT3G14450</td>
<td>Small auxin upregulated RNA 48, SAUR48, SAUR family,eshire of oxygen in the root tip. Exhibits apolar plasma membrane localization in the root cap and polar localization in tissues above and is involved in root hair degeneration.</td>
</tr>
<tr>
<td>AT3G13600</td>
<td>Small auxin upregulated RNA 3, SAUR3, multidrug resistance subfamily 3, multidrug resistance protein 7 (MRP7); MRP7, multidrug resistance protein 7, auxiliary output; MRP3; FUNCTIONS IN: chlorophyll catabolite transmembrane transporter activity, resistance P (PXR) subfamily B (PXR); functions in the basipetal redirection of auxin from the root tip. Exhibits apolar plasma membrane localization in the root cap and polar localization in tissues above and is involved in root hair degeneration.</td>
</tr>
<tr>
<td>AT3G10986</td>
<td>Small auxin upregulated RNA 1, SAUR1, multidrug resistance subfamily 2, nucleolin like 2 (NUC2); NUC2, nucleolin-like protein; SMALL AUXIN UPREGULATED RNA 2 (SAUR2), SAUR, small auxin upregulated RNA, a small family of auxin-inducible genes that encode small polypeptides. The SAUR proteins are involved in auxin-mediated growth and development, including root growth, lateral root formation, and floral organ development.</td>
</tr>
<tr>
<td>AT3G06890</td>
<td>GRF, GLYCOPROTEIN 4, Encodes an auxin efflux transmembrane transporter that is a member of the multidrug resistance superfamily. The gene is involved in auxin transport and auxin-mediated cellular response.</td>
</tr>
<tr>
<td>AT3G05320</td>
<td>GRF, GLYCOPROTEIN 4, Encodes an auxin efflux transmembrane transporter that is a member of the multidrug resistance superfamily. The gene is involved in auxin transport and auxin-mediated cellular response.</td>
</tr>
<tr>
<td>AT3G01760</td>
<td>GRF, GLYCOPROTEIN 4, Encodes an auxin efflux transmembrane transporter that is a member of the multidrug resistance superfamily. The gene is involved in auxin transport and auxin-mediated cellular response.</td>
</tr>
<tr>
<td>AT3G01345</td>
<td>GRF, GLYCOPROTEIN 4, Encodes an auxin efflux transmembrane transporter that is a member of the multidrug resistance superfamily. The gene is involved in auxin transport and auxin-mediated cellular response.</td>
</tr>
</tbody>
</table>

### Notes:
- The data presented in this document was derived from various Arabidopsis gene expression studies and functional analysis experiments. The genes listed are involved in various aspects of plant growth and development, including auxin response, transcription factor activity, and membrane transport functions.
- The functional categories assigned to each gene are based on their known or predicted roles in plant biology.
- The expression values represent fold changes in gene expression relative to control conditions, as determined by quantitative real-time PCR (qRT-PCR) experiments.
<table>
<thead>
<tr>
<th>GenBank Accession</th>
<th>Description</th>
<th>Function</th>
<th>Involved in</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G44350</td>
<td>Major facilitator superfamily protein</td>
<td>Functions in: transporter activity</td>
<td>Inducible element to oxidative stress</td>
</tr>
<tr>
<td>AT3G44265</td>
<td>EARLY RESPONSIVE TO DEHYDRATION 5 (ERD5)</td>
<td>Contains InterPro DOMAINS: Proline dehydrogenase (IPR002872), Galactose oxidase/kelch repeat superfamily protein</td>
<td>Target gene of MIR2111</td>
</tr>
<tr>
<td>AT3G44070</td>
<td>Proline</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G42</td>
<td>Granulin repeat cysteine protease family protein</td>
<td>Functions in: cysteine galactosidase family 35 protein</td>
<td></td>
</tr>
<tr>
<td>AT3G31910</td>
<td>like protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G30770</td>
<td>tRNA ligase/ nucleotide binding protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G3016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G20340</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G19390</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: The table above lists some of the genes and their associated functions. The complete table can be found in the original document.
<table>
<thead>
<tr>
<th>AT4G16250</th>
<th>AT4G15530</th>
<th>AT4G15530</th>
<th>AT4G14130</th>
<th>AT4G13540</th>
<th>AT4G13540</th>
<th>AT4G09430</th>
<th>AT4G03950</th>
<th>AT3G61060</th>
<th>AT3G56380</th>
<th>AT3G53910</th>
</tr>
</thead>
<tbody>
<tr>
<td>27.77997</td>
<td>2054.2744</td>
<td>8997.8</td>
<td>489</td>
<td>67</td>
<td>236.04993</td>
<td>23.708583</td>
<td>55.835697</td>
<td>31.875399</td>
<td>.594726</td>
<td>2.697492529</td>
</tr>
<tr>
<td>2.3433641</td>
<td>3.69538578</td>
<td>2.67426255</td>
<td>4.9760323</td>
<td>3.51273244</td>
<td>5.2215664</td>
<td>3.70054128</td>
<td>.8417949</td>
<td>1.58E-06</td>
<td>6.74E-04</td>
<td>4.60E-06</td>
</tr>
<tr>
<td>0.27529796</td>
<td>0.49436674</td>
<td>0.21595521</td>
<td>0.29577578</td>
<td>0.35138355</td>
<td>0.39260152</td>
<td>0.39260152</td>
<td>0.39260152</td>
<td>0.2722502</td>
<td>0.2722502</td>
<td>0.2722502</td>
</tr>
<tr>
<td>7.681268</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1.58E+06</td>
<td>6.74E+06</td>
<td>4.60E+06</td>
<td>4.10E+06</td>
<td>4.40E+06</td>
<td>4.40E+06</td>
<td>4.40E+06</td>
<td>4.40E+06</td>
<td>4.40E+06</td>
<td>4.40E+06</td>
<td>4.40E+06</td>
</tr>
<tr>
<td>156</td>
<td>156</td>
<td>156</td>
<td>156</td>
<td>156</td>
<td>156</td>
<td>156</td>
<td>156</td>
<td>156</td>
<td>156</td>
<td>156</td>
</tr>
</tbody>
</table>

**Note:** The values represent different biological contexts, possibly involving gene expression levels or protein interactions. The specific biological context is not clear from the data alone.
| AT4G34410     | 0.200882 | 0.0083891 | 0.008897 | 1.3239 | 2.36-52 | 0.0000974 | 1 | 0.0000795 | 1.243959 | 0.0705759 |
| AT4G30430     | 3.778486 | 1.266666 | 1.266666 | 1.7925 | 2.20-16 | 0.0000461 | 1 | 0.0000298 | 1.403191 | 0.0705759 |
| AT4G3028      | 1.113342 | 0.3262549 | 0.3262549 | 1.7925 | 2.20-16 | 0.0000461 | 1 | 0.0000298 | 1.403191 | 0.0705759 |
| AT4G29200     | 0.220760 | 0.220760 | 0.220760 | 1.7925 | 2.20-16 | 0.0000461 | 1 | 0.0000298 | 1.403191 | 0.0705759 |
| AT4G27654     | 1.113342 | 0.3262549 | 0.3262549 | 1.7925 | 2.20-16 | 0.0000461 | 1 | 0.0000298 | 1.403191 | 0.0705759 |
| AT4G25580     | 1.98945362 | 0.787481005 | - | - | - | 0.172789 | 1.190453 | 0.0705759 |
| AT4G18150     | 0.787481005 | - | - | - | - | 0.172789 | 1.190453 | 0.0705759 |
| AT4G24950     | 0.787481005 | - | - | - | - | 0.172789 | 1.190453 | 0.0705759 |
| AT4G24145     | 0.787481005 | - | - | - | - | 0.172789 | 1.190453 | 0.0705759 |
| AT4G24145     | 0.787481005 | - | - | - | - | 0.172789 | 1.190453 | 0.0705759 |
| AT4G24145     | 0.787481005 | - | - | - | - | 0.172789 | 1.190453 | 0.0705759 |
| AT4G24145     | 0.787481005 | - | - | - | - | 0.172789 | 1.190453 | 0.0705759 |
| AT4G24145     | 0.787481005 | - | - | - | - | 0.172789 | 1.190453 | 0.0705759 |
| AT4G24145     | 0.787481005 | - | - | - | - | 0.172789 | 1.190453 | 0.0705759 |

**Table:**

<table>
<thead>
<tr>
<th>Genes and Functions</th>
<th>Expression Data</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOT3, DEFECTIVELY ORGANIZED TRIBUTARIES 3 (DOT3)</td>
<td>Encodes a protein with an N terminal Homeodomain specific SCR region</td>
<td>Controls nitrate removal from the xylem sap. Mediates cations translocation.</td>
</tr>
<tr>
<td>DOT3, DEFECTIVELY ORGANIZED TRIBUTARIES 3 (DOT3)</td>
<td>Encodes a protein with an N terminal Homeodomain specific SCR region</td>
<td>Controls nitrate removal from the xylem sap. Mediates cations translocation.</td>
</tr>
<tr>
<td>DOT3, DEFECTIVELY ORGANIZED TRIBUTARIES 3 (DOT3)</td>
<td>Encodes a protein with an N terminal Homeodomain specific SCR region</td>
<td>Controls nitrate removal from the xylem sap. Mediates cations translocation.</td>
</tr>
</tbody>
</table>

**Legend:**

- **DOT3:** DEFECTIVELY ORGANIZED TRIBUTARIES 3 (DOT3)
- **N:** Nitrate transporter 3 (NIT3)
- **NIT4:** Nitrate transporter 4 (NIT4)
- **NIT5:** Nitrate transporter 5 (NIT5)

**Note:**

- The table above summarizes expression data for various genes involved in nitrate transport and metabolism. The genes listed are DOT3, which plays a crucial role in controlling nitrate removal from the xylem sap and mediating cation translocation.

children.
<table>
<thead>
<tr>
<th>Gene</th>
<th>AT5G28810</th>
<th>AT5G28235</th>
<th>AT5G24080</th>
<th>AT5G23240</th>
<th>AT5G22680</th>
<th>AT5G15360</th>
<th>AT5G13930</th>
<th>AT5G10250</th>
</tr>
</thead>
<tbody>
<tr>
<td>104.34</td>
<td>116.68</td>
<td>15.78</td>
<td>26.01</td>
<td>131.79</td>
<td>48.62</td>
<td>22.58</td>
<td>103.42</td>
<td>73.34</td>
</tr>
<tr>
<td>129.35</td>
<td>27.09</td>
<td>3.09</td>
<td>0.81</td>
<td>3.47</td>
<td>3.47</td>
<td>2.79</td>
<td>4.21</td>
<td>1.81</td>
</tr>
<tr>
<td>0.36</td>
<td>0.17</td>
<td>0.14</td>
<td>0.58</td>
<td>0.53</td>
<td>0.48</td>
<td>0.42</td>
<td>0.23</td>
<td>0.90</td>
</tr>
<tr>
<td>9.99</td>
<td>4.84</td>
<td>9.00</td>
<td>8.20</td>
<td>6.15</td>
<td>8.49</td>
<td>14.99</td>
<td>4.84</td>
<td>7.74</td>
</tr>
<tr>
<td>6.45E</td>
<td>1.97E</td>
<td>2.31E</td>
<td>4.71E</td>
<td>3.52E</td>
<td>3.59E</td>
<td>5.86E</td>
<td>2.23E</td>
<td>6.81E</td>
</tr>
<tr>
<td>-4</td>
<td>-3</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
<td>-4</td>
</tr>
</tbody>
</table>

**Note:** The table above represents gene expression levels or other relevant data. The exact nature of the data and its interpretation would depend on the specific context of the experiment or study.
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Description</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT5G67450</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G64810</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G61910</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G61270</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G49660</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G48900</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G47850</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G47240</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G42900</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G41730</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G36950</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G56870</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G56870</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G47240</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G42900</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G41730</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G36950</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G56870</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G56870</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G47240</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G42900</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
<tr>
<td>AT5G41730</td>
<td>Unidentified protein, hypothetical protein</td>
<td>Putative role in protein metabolism.</td>
</tr>
</tbody>
</table>

**Table Notes:**
- The table above lists genes and their putative functions in various biological processes.
- For a comprehensive understanding, review the original source for more detailed information.

**Additional Information:**
- **Gene Expression:**
  - Expression data for each gene shows varying levels of expression from 0 to 100.
  - Different colors indicate different expression levels.
- **Function:**
  - Functions are broadly categorized into protein metabolism, disease resistance, and other biological activities.
  - Each function has a brief description indicating its role in cellular processes.
**Supplementary Table 5**: List of coding genes at least $\log_2$-fold increases (negative $\log_2$-fold change) or decreases (positive $\log_2$-fold change) of 2.5 in at least one of the four lines A1+, A1, A2+ or A2-.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>line</th>
<th>log2-fold change</th>
<th>pvalue</th>
<th>Location of dense C methylation</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT2G34130</td>
<td>A1+</td>
<td>-6.343</td>
<td>7.37E-42</td>
<td>genic</td>
<td>MEE19 maternal effect embryo arrest 19; hypothetical protein</td>
</tr>
<tr>
<td></td>
<td>A1-</td>
<td>-1.855</td>
<td>0.000285</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A2+</td>
<td>-2.970</td>
<td>5.17E-62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G01345</td>
<td>A1+</td>
<td>-7.076</td>
<td>4.86E-76</td>
<td>genic</td>
<td>Expressed protein</td>
</tr>
<tr>
<td></td>
<td>A1-</td>
<td>-6.829</td>
<td>5.84E-86</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>A2+</td>
<td>-0.883</td>
<td>6.86E-08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G21570</td>
<td>A1+</td>
<td>-3.451</td>
<td>2.47E-09</td>
<td>genic</td>
<td>proline-rich nuclear receptor coactivator</td>
</tr>
<tr>
<td></td>
<td>A1-</td>
<td>-1.949</td>
<td>0.000119</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G24542</td>
<td>A1+</td>
<td>-5.551</td>
<td>8.93E-28</td>
<td>genic</td>
<td>Beta-galactosidase related protein</td>
</tr>
<tr>
<td></td>
<td>A1-</td>
<td>-4.889</td>
<td>3.90E-26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G53910</td>
<td>A1+</td>
<td>-4.682</td>
<td>3.88E-18</td>
<td>genic</td>
<td>Malate dehydrogenase-like protein</td>
</tr>
<tr>
<td></td>
<td>A1-</td>
<td>-3.701</td>
<td>8.64E-14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G18150</td>
<td>A1+</td>
<td>-5.660</td>
<td>7.99E-34</td>
<td>genic</td>
<td>Serine/Threonine-kinase, putative</td>
</tr>
<tr>
<td></td>
<td>A1-</td>
<td>-5.388</td>
<td>3.45E-37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Class</td>
<td>Expression</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G15360</td>
<td>genic</td>
<td>-4.214</td>
<td>Transmembrane protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>6.81E-14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>-6.021</td>
<td>7.80E-51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G26270</td>
<td>genic</td>
<td>-3.470</td>
<td>Transmembrane protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>2.04E-09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>-5.411</td>
<td>4.71E-64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G35375</td>
<td>genic</td>
<td>-3.379</td>
<td>Transmembrane protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>4.08E-09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>-2.398</td>
<td>2.47E-06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G01080</td>
<td>upstream/</td>
<td>-2.550</td>
<td>Beta-galactosidase related protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>genic</td>
<td>1.39E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>-4.218</td>
<td>5.14E-21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G27473</td>
<td>upstream</td>
<td>-2.984</td>
<td>Cysteine/Histidine-rich C1 domain family protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.32E-09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>-1.797</td>
<td>0.000464</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT3G30775</td>
<td>upstream</td>
<td>-1.205</td>
<td>EARLY RESPONSIVE TO DEHYDRATION 5 (ERD5); Encodes a proline oxidase, its</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.18E-08</td>
<td>mRNA expression induced by high levels of Al and by osmotic stress. The</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>promoter contains an L-proline-inducible element.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G09430</td>
<td>upstream</td>
<td>-2.531</td>
<td>Disease resistance protein (TIR-NBS-LRR class) family; with Natural antisense</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.12E-06</td>
<td>mRNA expression induced by high levels of Al and by osmotic stress. The</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>promoter contains an L-proline-inducible element.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G25530</td>
<td>upstream</td>
<td>-10.316</td>
<td>FLOWERING WAGENINGEN, FWA, HDG6, HOMEODOMAIN GLABROUS6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.26E-118</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>-3.858</td>
<td>3.10E-14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2+</td>
<td>-2.503</td>
<td>7.00E-44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G23240</td>
<td>upstream</td>
<td>-2.881</td>
<td>ATDJC17, DJC75, DNAJ PROTEIN C76, DNAJ heat shock N-terminal domain-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.78E-30</td>
<td></td>
<td>containing protein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1-</td>
<td>-2.544</td>
<td>1.21E-20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2-</td>
<td>-3.046</td>
<td>1.10E-41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GenBank Accession</td>
<td>Gene Name</td>
<td>Start</td>
<td>End</td>
<td>Function</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------</td>
<td>-------</td>
<td>-------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>AT5G24240</td>
<td></td>
<td></td>
<td></td>
<td>upstream</td>
<td>Phosphoinositide 4-kinase PI4Kc3, Overexpression mutants display late-flowering phenotype.</td>
</tr>
<tr>
<td>AT2G06904</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>Nucleic acid / zinc ion binding protein</td>
</tr>
<tr>
<td>AT2G07240</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>Cysteine-type peptidase</td>
</tr>
<tr>
<td>AT2G11778</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>Transmembrane protein</td>
</tr>
<tr>
<td>AT3G28193</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>Transmembrane protein</td>
</tr>
<tr>
<td>AT3G30720</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>QOS qua-quine starch</td>
</tr>
<tr>
<td>AT3G30770</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>Eukaryotic aspartyl protease family protein</td>
</tr>
<tr>
<td>AT3G31910</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>Ulp1 protease family protein (DUF1985)</td>
</tr>
<tr>
<td>AT3G42723</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>ATP binding / aminoacyl-tRNA ligase/ nucleotide binding protein</td>
</tr>
<tr>
<td>AT3G44070</td>
<td></td>
<td></td>
<td></td>
<td>region</td>
<td>Glycosyl hydrolase family 35 protein</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Line</td>
<td>Log2 Fold Change</td>
<td>p-value</td>
<td>Feature</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-----------------</td>
<td>---------------</td>
<td>-------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>AT3G44265</td>
<td>A1+</td>
<td>-6.376</td>
<td>7.87E-46</td>
<td>region</td>
<td>Beta-galactosidase-like protein</td>
</tr>
<tr>
<td>A1-</td>
<td></td>
<td>-5.866</td>
<td>2.31E-47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2+</td>
<td></td>
<td>-0.549</td>
<td>8.62E-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G03950</td>
<td>A1+</td>
<td>-3.246</td>
<td>2.27E-08</td>
<td>region</td>
<td>Nucleotide/sugar transporter family protein</td>
</tr>
<tr>
<td>A1-</td>
<td></td>
<td>-3.250</td>
<td>4.93E-11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT5G45570</td>
<td>A1+</td>
<td>-4.759</td>
<td>5.05E-17</td>
<td>region</td>
<td>Ulp1 protease family protein</td>
</tr>
<tr>
<td>A1-</td>
<td></td>
<td>-3.487</td>
<td>4.75E-13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2+</td>
<td></td>
<td>-1.290</td>
<td>8.65E-14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Reduced transcript levels**

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Line</th>
<th>Log2 Fold Change</th>
<th>p-value</th>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT5G34850</td>
<td>A1+</td>
<td>7.956</td>
<td>1.09E-105</td>
<td>upstream</td>
<td>Purple acid phosphatase 26</td>
</tr>
<tr>
<td>A1-</td>
<td></td>
<td>7.971</td>
<td>6.56E-111</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Antagonistic transcript level changes in A1+ and A1-**

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Line</th>
<th>Log2 Fold Change</th>
<th>p-value</th>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT3G50770</td>
<td>A1+</td>
<td>2.816</td>
<td>3.50E-15</td>
<td>upstream</td>
<td>CML41, calmodulin-like 41 FUNCTIONS IN: calcium ion binding</td>
</tr>
<tr>
<td>A1-</td>
<td></td>
<td>-0.948</td>
<td>0.00018459</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AT4G00130</td>
<td>A1+</td>
<td>2.452</td>
<td>6.56E-08</td>
<td>region</td>
<td>DNA-binding storekeeper protein-related transcriptional regulator</td>
</tr>
</tbody>
</table>

**Supplementary Table 6**: List of all coding genes with heritably increased (negative log2-fold change) or reduced (positive log2-fold change) transcript levels in the A1 lines with dense cytosine methylation in all three sequence contexts (CG, CHG, CHH).
Supplementary Figure 1: Map and expression analysis of the three over-expression lines. A) AT3G01345 was inserted into the plant transformation vector 35S pGreen 0179, to produce the 35S AT3G01345 over-expression construct. B) AT3G27473 was inserted into the plant transformation vector 35S pGreen 0179, to produce the 35S AT3G27473 over-expression construct. C) AT3G30720 was inserted into the plant transformation vector 35S pGreen
0179, to produce the 35S AT3G01345 over-expression construct. D) cDNA was generated from seedlings 4 weeks after stratification. Greater transcript levels were detected for all over-expression lines (AT3G01345, AT3G27473, AT3G30720). EF1α was used as a standardizing control to ensure equal concentration of input cDNA.

**Supplementary Figure 2: Maps of the FLAG-tagged MET1.** A triple FLAG sequence was inserted at the 3' end of the catalytically active MET1 construct.
Supplementary Figure 3: Bisulphite analysis of CEN1.1, WUS, SP9D. Methylation data from tomato epigenome database. Yellow bars indicate the presence of CG methylation, blue indicates the presence of CHG methylation and red indicate the presence of methylation in a CHH context. Two Genes, CEN1.1 and WUS, possess dense DNA methylation adjacent to the gene. SP9D contains dense methylation within the gene. Methylation profiles in the three genes were extracted from http://epigenome.genetics.uga.edu/PlantMethylome/.