

DESCRIPTION OF A PUTATIVE P-EAMET FROM *ARABIDOPSIS THALIANA*

**Phylogenetic, Biochemical and Structural Comparisons Show a Putative  
Phosphobase *N*-methyltransferase from *Arabidopsis thaliana* to be a Homolog of  
Phosphoethanolamine *N*-methyltransferase**

By

Mitchell MacLeod, BSc

A Thesis

Submitted to the School of Graduate Studies

In Partial Fulfilment of the Requirements

For the Degree

Master of Science

McMaster University

2010

© Copyright by Mitchell MacLeod, January 2010



Master's Thesis – M. MacLeod

McMaster – Biology

Master of Science (2010)

(Biology)

McMaster University

Hamilton, ON

TITLE: Phylogenetic, biochemical and structural comparisons show a putative phosphobase *N*-methyltransferase from *Arabidopsis thaliana* to be a homolog of phosphoethanolamine *N*-methyltransferase

AUTHOR: Mitchell J.R. MacLeod

SUPERVISOR: Dr. Elizabeth A. Weretilnyk

CO-SUPERVISOR: Dr. Robin Cameron

NUMBER OF PAGES: xii, 142

**ABSTRACT**

Phosphocholine is a precursor of free choline and phosphatidylcholine, the most abundant phospholipid in non-plastid plant membranes (Moore, 1976). Phosphocholine synthesis in several plants studied to date occurs via a phosphobase route using water-soluble intermediates (Hanson and Rhodes, 1983). In this pathway, three sequential *N*-methylations of phosphoethanolamine catalyzed by the enzyme phosphoethanolamine-*N*-methyltransferase (P-EAMeT) produces phosphocholine.

*Arabidopsis* has three P-EAMeT-like enzymes annotated as *S*-adenosylmethionine (SAM)-dependent *N*-methyltransferases (NMTs). These three proteins share greater than 84% identity at the amino acid level. Two of the genes encoding these enzymes, *NMT1* and *NMT2*, have been cloned from *Arabidopsis* and the proteins they encode shown to catalyze the three reactions required to convert phosphoethanolamine into phosphocholine (Bolognese *et al.*, 2000; Begora *et al.*, 2010). The objective of this work is to determine if the third gene, *NMT3*, also encodes a phosphobase methyltransferase.

Phylogenetic analyses of the *Arabidopsis* NMT proteins and related plant sequences suggest that a full-length NMT3 should methylate all three phosphobases to synthesize phosphocholine. Moreover, homology modelling predicts a highly conserved tertiary structure for NMT1 and NMT3 as well as spinach and wheat P-EAMeT enzymes. A cDNA encoding the N-terminal half of NMT3 was cloned and used to assay for phosphobase activity. *In vitro* assays showed this portion of NMT3 to methylate phosphoethanolamine, confirming that NMT3 is a P-EAMeT enzyme.

A mutant deficient in transcripts associated with *NMT3* shows an early-flowering phenotype but no apparent alterations in phospholipid membrane composition relative to wild-type plants. Together these results suggest that *NMT3* expression is not required for phosphatidylcholine production but somehow a deficiency in this gene product accelerates the transition to reproductive development. Future work is needed to clone the full length *NMT3* and further characterization of *NMT3* deficient plants may offer insight into the role of *NMT3* in plant development.

## **ACKNOWLEDGEMENTS**

I first must thank Dr. Elizabeth Weretilnyk for her wisdom and encouragement over the past few years. Without your many contributions and support this work would not have been possible.

I would also like to acknowledge other students from the lab who have helped me along the way. First and foremost, Mike BeGora, who has helped guide my research since my time as a summer student and knows every protocol inside and out. Also, to Alicia DiBattista, Ashley Tattersall, Sam Bremner, Kristy Shulist, Kim Carruthers and John Mansbridge for all your help. It has been a pleasure working with you all.

Lastly, I would like to thank Dr. Peter Summers for always being there to lend a hand and Dr. Robin Cameron for her many insights as my co-supervisor. Also, I extend special thanks to Dr. Cameron and Dr. Brian Golding for their valuable input as members of my thesis committee.

**TABLE OF CONTENTS**

<b>Literature Review</b>	<b>1</b>
Choline synthesis	1
Uses of choline	5
Phosphatidylcholine synthesis in plants	8
Phosphatidylcholine synthesis in yeast	11
The P-EAMeT family	12
Spinach P-EAMeTs	12
Wheat P-EAMeTs	14
Non-plant P-EAMeTs	16
Arabidopsis P-EAMeTs	16
<b>Materials and Methods</b>	<b>20</b>
Plant material and growth	20
Salt stress experiment	22
Extraction and analysis of nucleic acids	23
Arabidopsis <i>NMT3</i> clone preparation	31
Radioassays	34
Phospholipid profiling	36
Construction of phylogenetic relationships among selected NMTs	37
Homology modelling of selected NMTs	41
Analysis of Arabidopsis <i>NMT</i> gene expression	41

<b>Results</b>	<b>43</b>
<b>Sequence and structure analysis of NMT proteins</b>	<b>43</b>
Phylogenetic analysis of NMT-like Proteins	43
Detailed alignment analysis of NMT amino acid sequences	51
Comparison of tertiary protein structures predicted by homology modelling	54
<b>Organization and expression of genes encoding Arabidopsis NMTs</b>	<b>66</b>
Transcription factor binding site prediction	66
Expression of Arabidopsis <i>NMT</i> genes	69
Differential expression of <i>NMT</i> genes in Arabidopsis	69
Salt stress experiment	72
Early flowering mutant phenotype	79
Phospholipid profiling	82
Cloning <i>NMT3</i>	82
Assaying NMT3 activity	111
<b>Discussion</b>	<b>115</b>
<b>Conclusions</b>	<b>127</b>
<b>References</b>	<b>131</b>

## LIST OF FIGURES AND TABLES

### Figures

Figure 1. Metabolism involving choline and phosphatidylcholine showing pathway intermediates and regulation by feed-back mechanisms involving phosphocholine	2
Figure 2. Consensus phylogenetic tree constructed from full length amino acid sequences of 23 known/putative SAM-dependent methyltransferases	44
Figure 3. Phylogenetic tree constructed from the N-terminal half sequences of 30 similar SAM-dependent methyltransferases	47
Figure 4. Phylogenetic tree constructed from the C-terminal half sequences of 36 similar SAM-dependent methyltransferases	49
Figure 5. Deduced amino acid sequence alignment of full length Arabidopsis, spinach and wheat NMT proteins.	52
Figure 6. Alignment of deduced amino acid sequences for SBD1 and SBD2 of Arabidopsis, spinach and wheat NMT proteins	56
Figure 7. Protein modelling of five plant NMT proteins binding SAH	58
Figure 8. $\alpha$ -carbon backbones of Arabidopsis NMT protein models	60
Figure 9. Comparison of Arabidopsis NMT protein models binding SAH	62
Figure 10. Protein modelling of five plant NMT proteins binding SAM	64
Figure 11. RT-PCR of <i>NMT3</i> transcript from WT and SALK_062703 leaf RNA	70
Figure 12. Expression of <i>NMT</i> genes in various developmental tissues of Arabidopsis	73
Figure 13. Growth of Arabidopsis seedlings on defined media in response to NaCl	75
Figure 14. Developmental phenotype of early flowering <i>nmt3</i> plants	80

Figure 15. Phospholipid profiles of WT and <i>nmt3</i> leaf extracts	83
Figure 16. Isolation of a partial <i>NMT3</i> cDNA template	85
Figure 17. Isolation of <i>NMT3</i> N and C-terminal halves amplified from full-length <i>NMT3</i> cDNA template	88
Figure 18. Strategy for cloning <i>NMT3</i> from N and C-terminal halves	90
Figure 19. N-terminal half DNA sequence and translated amino acid sequence	92
Figure 20. C-terminal half DNA sequence and translated amino acid sequence	94
Figure 21. C-terminal half of <i>NMT3</i> amplified from an 1800 bp RACE generated cDNA template	98
Figure 22. Predicted RNA secondary structure of the transcripts encoding Arabidopsis <i>NMT</i> proteins	100
Figure 23. Amplification of a full length <i>NMT3</i> product by RT-PCR	102
Figure 24. Approach to cloning <i>NMT3</i> and insertion into pET-30a(+) expression vector	104
Figure 25. DNA sequence of full length <i>NMT3</i> ORF inserted into pET-30a(+) vector	106
Figure 26. DNA sequence of N-terminal half portion of <i>NMT3</i> inserted into pET-30a(+) vector	109
Figure 27. Cartoon depicting regions of <i>NMT3</i> cDNA sub-cloned into the pET-30a(+) protein expression vector	113

## Tables

Table 1. Legend of protein sequences included in phylogenetic trees	39
Table 2. Potential transcription factor binding sites upstream of Arabidopsis <i>NMT</i> genes	67
Table 3. Response of Arabidopsis to NaCl exposure on defined media	77



**Appendix**

## Appendix A. Primer information

129

**ABBREVIATIONS**

AG	Agamous transcription factor
BADH	Betaine aldehyde dehydrogenase enzyme
CBF1/2/3	C-repeat binding transcription factor
CDP/CTP	Cytidine di/tri phosphate
CK	Choline kinase enzyme
CMO	Choline monooxygenase enzyme
CTT	Cytidyltransferase enzyme
d	Day(s)
EA	Ethanolamine
EFM	Early-flowering mutant phenotype
DEA	Diethanolamine
DREB1a/2a	Dehydration response element binding transcription factor
ESI/MS/MS	Electron spray ionization/mass spectrometer/mass spectrometer
GB	Glycine betaine
GT1/2	Guanidine/thymine binding transcription factor
h	Hours(s)
Hmz	Homozygous
IPTG	Isopropyl- $\beta$ -D-thio-galactoside
MEA	Methylethanolamine

MS	Murashige & Skoog
NMT	<i>N</i> -methyltransferase
ORF	Open reading frame
PA	Phosphatidic acid
P-Cho	Phosphocholine
P-DEA	Phosphodimethylethanolamine
P-EA	Phosphoethanolamine
P-EAMeT	Phosphoethanolamine- <i>N</i> -methyltransferase
PLMT	Phospholipid- <i>N</i> -methyltransferase
P-MEA	Phosphomethylethanolamine
Ptd-Cho	Phosphatidylcholine
Ptd-DEA	Phosphatidyl dimethylethanolamine
Ptd-EA	Phosphatidylethanolamine
Ptd-MEA	Phosphatidylmethylethanolamine
RACE	Rapid amplification of cDNA ends
RAV	Related to ABI3/VP1 transcription factor
RNAi	RNA interference
SAH	<i>S</i> -adenosylhomocysteine
SAM	<i>S</i> -adenosylmethionine
SBD	SAM binding domain
T-DNA	Transfer DNA

TFBS	Transcription factor binding site
uORF	Upstream open reading frame
UTR	Untranslated region
WRKY18	WRKY-motif binding transcription factor
WT	Wild type

## LITERATURE REVIEW

### Choline synthesis

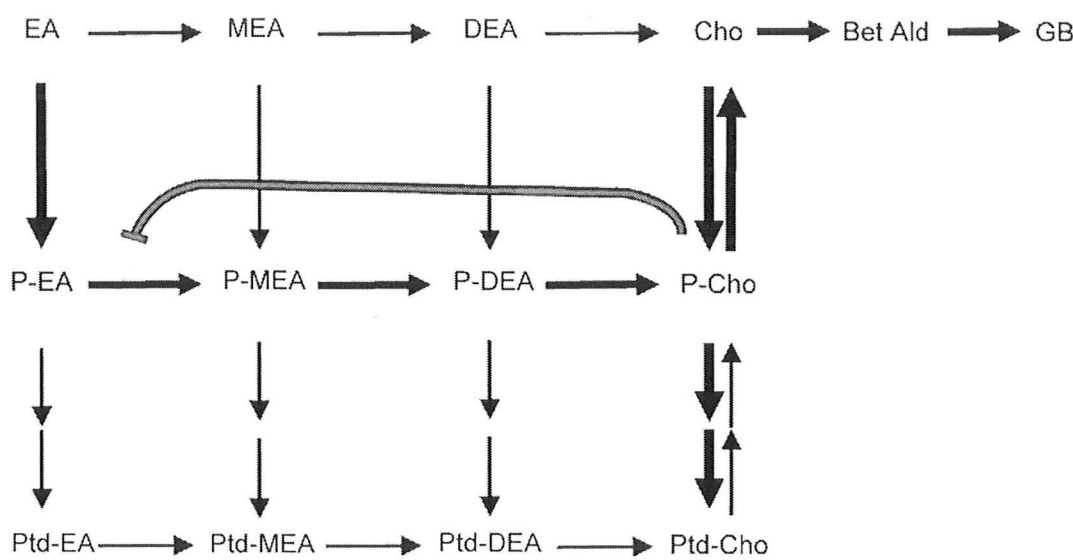
In plants studied to date, choline can be produced via free base, phosphobase or phosphatidylbase intermediates (Prud'homme and Moore, 1992; Hanson and Rhodes, 1983; Datko and Mudd, 1988a). Choline can also be recycled by release from the headgroup of the membrane lipid phosphatidylcholine (Hitz *et al.*, 1981).

The phosphobase intermediates likely constitute the main route of choline production in *Arabidopsis* (Bolognese *et al.*, 2000). This pathway has received comparatively greater attention in two members of the *Chenopodiaceae* family *Beta vulgaris* (sugar beet) and *Spinacia oleracea* (spinach) (Hanson and Rhodes, 1983; Summers and Weretilnyk, 1993). These chenopods have the ability to synthesize greater amounts of choline. Choline can subsequently be used to synthesize the compatible solute glycine betaine (GB) in response to water stress (Hanson and Nelson, 1978).

Ethanolamine (EA) enters the phosphobase pathway via the decarboxylation of serine (Rontein *et al.*, 2001). EA is rapidly phosphorylated by ethanolamine kinase (Delwiche and Bregoff, 1958) to produce phosphoethanolamine (P-EA). P-EA is sequentially *N*-methylated first to phosphomethylethanolamine (P-MEA) then phosphodimethylethanolamine (P-DEA) and finally phosphocholine (P-Cho). The three methylation reactions that carry P-EA to P-Cho can be catalyzed by a single enzyme phosphoethanolamine-*N*-methyltransferase (P-EAMeT) as shown in Fig. 1 (Nuccio *et al.*, 2000). P-Cho is converted to free choline for use in GB synthesis through the action of P-Cho phosphatase (Summers and Weretilnyk, 1993).

Figure 1 – Metabolism involving choline and phosphatidylcholine showing plant and animal pathway intermediates and regulation by feed-back mechanisms involving phosphocholine.

EA, ethanolamine; MEA, methylethanolamine; DEA, dimethylethanolamine; Cho, choline; Bet Ald, betaine aldehyde; GB, glycine betaine; P-EA, phosphoethanolamine; P-MEA, phosphomethylethanolamine; P-DEA, phosphodimethylethanolamine, P-Cho, phosphocholine; Ptd-EA, phosphatidylethanolamine; Ptd-MEA, phosphatidylmethylethanolamine, Ptd-DEA, phosphatidyl dimethylethanolamine; Ptd-Cho, phosphatidylcholine. P-Cho can feedback inhibit this methylation pathway (Smith *et al.*, 2000; Mudd and Datko, 1989).



In addition to the phosphobase route, sequential methylations can take place at the phosphatidylbase level from the point of Ptd-MEA ( $\text{Ptd-MEA} \rightarrow \text{Ptd-DEA} \rightarrow \text{Ptd-Cho}$ ) and there is one report of synthesis at the free base level ( $\text{EA} \rightarrow \text{MEA} \rightarrow \text{DEA} \rightarrow \text{choline}$ ) (Datko and Mudd, 1988a; Prud'homme and Moore, 1992). The phosphatidylbase pathway is the dominant pathway in mammals as well as yeast (Greenberg *et al.*, 1983). Datko and Mudd (1988a) have shown that some methylation occurs at the phosphatidyl level in soybean and carrot. For either the phosphatidyl or phosphobase routes, the committing step for the synthesis of choline is the methylation of P-EA by P-EAMeT (Mudd and Datko, 1986, Datko and Mudd, 1988a). For example, in soybean tissue culture P-EA is initially methylated to P-MEA by P-EAMeT. P-MEA is then converted to Ptd-MEA, via the intermediate cytidinediphospho (CDP)-MEA, and from that point on the methylations proceed predominantly at the phosphatidylbase level ( $\text{Ptd-MEA} \rightarrow \text{Ptd-DEA} \rightarrow \text{Ptd-Cho}$ ; Datko and Mudd, 1986). In carrot cell cultures, P-EA is converted to P-MEA but subsequent methylation reactions proceed along both the phosphobase and phosphatidylbase route (Datko and Mudd, 1988a). *Lemma* typifies the route found for spinach plants where the main route of methylation is via the phosphobase intermediates P-MEA and P-DEA to P-Cho (Datko and Mudd, 1988b). Soybean and carrot obtain the majority of their free choline from Ptd-Cho (Datko and Mudd, 1988a).

$^{14}\text{C}$ -Labelled ethanolamine incubated with *Lemma* fronds was rapidly converted to P-EA and Ptd-EA but Ptd-EA is not methylated to Ptd-MEA (Mudd and Datko, 1986). A recent report confirms that the ability to methylate  $\text{Ptd-EA} \rightarrow \text{Ptd-MEA}$  has not been found in plants studied to date (Keogh *et al.*, 2009). Thus, P-EA serves as a substrate for P-Cho

production indirectly via P-MEA and directly for synthesis of the phospholipid Ptd-EA.

As mentioned above, methylation at the free base level has been demonstrated in castor bean endosperm (Prud'homme and Moore, 1992). Coughlan and Wyn-Jones (1982) originally suggested that choline was produced mainly via the free bases in spinach but this was later shown not to be the case and that the phosphobase route is the dominant pathway for the production of P-Cho in spinach (Summers and Weretilnyk, 1993). The free base pathway has not been identified as a major contributor to the P-Cho, choline or Ptd-Cho pool using radiotracer analysis performed using *Lemna*, soybean or carrot plants (Datko and Mudd, 1988b).

### **Uses of choline**

One of the most studied uses of choline in plants is its role in the synthesis of GB (Hanson and Nelson, 1978). GB is a quaternary ammonium compound that acts as a compatible solute and osmoprotectant in plant cells (Hanson and Nelson, 1978; Hitz and Hanson, 1980; Wyn Jones and Storey, 1978). GB is accumulated by many species in response to water deficits and salinity including sugar beet and spinach (Hanson and Rhodes, 1983; Coughlan and Wyn-Jones, 1982). In plants accumulating this organic solute, GB synthesis occurs primarily in leaves (Hanson and Rhodes, 1983). In spinach leaves choline is produced primarily via the phosphobase methylation pathway and the choline produced is transported into the chloroplast (Weigel *et al.*, 1986). In the chloroplast the enzyme choline monooxygenase (CMO) converts choline to betaine aldehyde that is then converted into GB via the action of betaine aldehyde dehydrogenase



(BADH) (Weretilnyk and Hanson, 1990).

In spinach, P-Cho is hydrolyzed by P-Cho phosphatase to yield free choline (Summers and Weretilnyk, 1993). When plants are exposed to saline conditions, spinach leaves are able to accumulate up to three times more GB than control plants (Summers and Weretilnyk, 1993). This increase in GB synthesis is made possible, in part, by the upregulation of P-EAMeT, ethanolamine kinase, BADH and CMO. Ethanolamine kinase produces the P-EA that serves as the initial substrate for P-EAMeT in the synthesis of P-Cho (Summers and Weretilnyk, 1993). Spinach P-EAMeT is also regulated by light as the plant must be exposed to both light and salt to observe the salt-responsive increase in activity (Weretilnyk *et al.*, 1995). In spinach it has been shown that the enzyme P-Cho phosphatase is more active than either choline kinase (choline→P-Cho) or CTP: P-Cho cytidylyltransferase (P-Cho→CDP:choline) (Summers and Weretilnyk, 1993). The lower activities of these latter enzymes would insure that a large pool of free choline remains available for conversion to GB.

In plants that do not accumulate GB such as *Arabidopsis* and tobacco, the apparent absence of the enzyme CMO limits their ability to accumulate GB (Nuccio *et al.*, 1998). These plants do have a functional BADH and can convert betaine aldehyde to GB but, without CMO activity, no betaine aldehyde is produced (Nuccio *et al.*, 1998). Attempts have been made to engineer a transgenic plant capable of accumulating GB (McNeil *et al.*, 2001; Sakamoto *et al.*, 2001). Transgenic tobacco plants engineered to constitutively express spinach CMO can produce GB but they are still unable to accumulate significant amounts of this organic solute. This lack of accumulation was determined to be due to

insufficient amounts of choline (Nuccio *et al.*, 1998). When EA was supplied to leaf discs prepared from the transgenic tobacco there was still little GB accumulation. This finding suggests that ethanolamine kinase or P-EAMeT may be the limiting factor(s) to the transgenic plants preventing them from producing enough free choline for GB accumulation (Nuccio *et al.*, 1998). A refinement using the transgenic tobacco involved over-expression of spinach P-EAMeT in order to augment the endogenous tobacco P-EAMeT activity and increase the pool of free choline available to be converted to GB (McNeil, *et al.*, 2001). These transgenic tobacco plants accumulated up to 30-fold more free choline than control plants and greater amounts of GB were accumulated in these plants. However, these plants did not accumulate GB to levels approaching GB-accumulating species such as spinach. McNeil *et al.*, (2001) concluded that feedback inhibition by P-Cho and or the transport of choline to the chloroplast could be limiting the accumulation of GB in these transgenic plants.

*Arabidopsis* plants have also been transformed with a bacterial choline oxidase enzyme and although only small amounts of GB were accumulated, improved stress tolerance was demonstrated (Sakamoto *et al.*, 2001). These engineering attempts were successful at producing plants capable of synthesizing GB but not accumulating GB to the levels seen in GB-accumulators such as spinach. Thus a greater understanding of factors regulating choline synthesis in non-accumulating species is required in order to engineer the trait of GB accumulation in plant species that do not normally accumulate this osmoprotectant.

**Phosphatidylcholine synthesis in plants**

Ptd-Cho is the predominant non-plastid phospholipid in plants and comprises 40-60% of the total phospholipid pool (Moore, 1976). In plants that do not accumulate GB it has been demonstrated that Ptd-Cho content can be increased in response to abiotic stress (Pical *et al.*, 1999). Freezing tolerance has been shown to be associated with the amount and degree of polyunsaturation of Ptd-Cho in the membrane (Uemura *et al.*, 1995). In *Arabidopsis* plants undergoing acclimation to cold temperatures, the proportion of polyunsaturated Ptd-Cho in the membrane increased resulting in a more fluid membrane at low temperatures (Uemura *et al.*, 1995). Also, Tasseva *et al.* (2004) showed that the Ptd-Cho is produced in greater amounts in response to salt stress. Ptd-Cho is also an important reservoir for lipid second messengers such as phosphatidic acid (PA) (Cruz-Ramirez, *et al.*, 2004). The enzyme phospholipase D uses Ptd-Cho as a substrate to produce free choline and PA (Qin and Wang, 2002).

The pool of Ptd-Cho in plants is believed to be tightly regulated and critical to maintaining cell structure and function during periods of stress (Tasseva *et al.*, 2004). Ptd-Cho is produced via two different routes in plants. In the CDP-choline or Kennedy pathway choline is phosphorylated by choline kinase and then phosphocholine is converted to Ptd-Cho using cytidine-5'-diphosphate choline (CDP-choline) as an intermediate (Tasseva *et al.*, 2004). The alternative pathway involves sequential methylations of phosphatidylbase intermediates to yield Ptd-Cho (Datko and Mudd, 1988a). The phosphatidylbase methylation pathway is responsible for the majority of Ptd-choline synthesis in soybean. Mudd and Datko (1986) showed that at isotopic

equilibrium, 34% of labelled ethanolamine was incorporated into Ptd-Cho while only 6% was found in choline. They also detected no radioactivity in CDP-choline which would suggest that the CDP-choline pathway contributes little to the total pool of Ptd-choline or that flux through this step is very rapid (Mudd and Datko, 1986). This pathway predominates in some plants, yeast and animals.

The phosphatidylbase methylation pathway is reported to only account for about 10% of the Ptd-Cho pool in *Arabidopsis* (Tasseva *et al.*, 2004). The CDP-choline (Kennedy) pathway is responsible for the majority of Ptd-Cho production. Choline kinase (CK) and CDP-choline cytidyltransferase (CTT) appear to exert control over the pathway. The transcripts encoding CK and CTT are maintained at a constant level in *Arabidopsis* (Tasseva *et al.*, 2004). Ptd-Cho production through the CDP-choline pathway is increased in response to cold and salinity stress due to post-translational up regulation of CTT genes and a combination of transcriptional and post-translational upregulation of CK isozymes (Inatsugi *et al.*, 2002). In contrast, the expression of enzymes involved in the methylation of phosphobases and phosphatidylbases are downregulated (Tasseva *et al.*, 2004). Cruz-Ramirez *et al.*, (2004) showed that knocking out the gene *XPL1* encoding *Arabidopsis* P-EAMeT reduced the total Ptd-Cho pool by 23%. This reduction suggests *Arabidopsis* P-EAMeT plays a significant role in the synthesis of P-Cho that is later converted to Ptd-Cho.

In *Arabidopsis* P-EA can undergo direct methylation to yield P-Cho. These three reactions can be catalyzed by *Arabidopsis* NMT1 (Bolognese *et al.*, 2000). Alternatively, a second pathway involves the conversion of P-EA to P-MEA with the P-MEA

subsequently converted to Ptd-MEA via CDP-MEA intermediate (Keogh *et al.*, 2009).

Ptd-MEA produced from CDP-MEA then undergoes two methylation events, first to Ptd-DEA and finally Ptd-Cho (Keogh *et al.*, 2009). The methylation of Ptd-EA to Ptd-MEA does not occur in Arabidopsis.

In Arabidopsis the methylations at the phosphatidyl level are catalyzed by the phospholipid methyltransferase enzyme AtPLMT (Keogh *et al.*, 2009). Unlike the soluble P-EAMeT family proteins, AtPLMT is an integral membrane protein. When the gene encoding AtPLMT was knocked out with a dissociator transposon insertion, the abundance of the phosphatidyl intermediates Ptd-MEA and Ptd-DEA increased 9-fold and 3.5-fold, when quantified using electron spray ionization/mass spectrometry/mass spectrometry (ESI/MS/MS) (Keogh *et al.*, 2009). AtPLMT knock-out plants showed no visible phenotype and there was no change in the amount of Ptd-Cho produced. It appears that flux through the phosphobase methylation and CDP-choline pathway can compensate for the loss of AtPLMT in this plant (Keogh *et al.*, 2009).

In many plants that accumulate GB it appears that the methylation of the phosphatidyl intermediates Ptd-MEA and Ptd-DEA contributes little to the total pool of Ptd-Cho. Labelling experiments using sugar beet show that  $^{14}\text{C}$ -labelled EA is incorporated steadily into Ptd-Cho (Hanson and Rhodes, 1983). Interestingly, in Gramineae species such as wheat and barley it has been shown that Ptd-Cho actually acts as a precursor to free choline that is later used for the synthesis of GB (Hanson and Rhodes, 1978). This could be a result of P-Cho phosphatase having low activity in these species (Tanaka *et al.*, 1966). The GB-accumulating species of the Chenopodiceae have a

very active P-Cho phosphatase that produces choline for the synthesis of GB.

### **Phosphatidylcholine synthesis in yeast**

The inability to convert Ptd-EA to Ptd-MEA is not a property of all organisms. Many eukaryotic organisms including yeast can synthesize Ptd-Cho via three sequential methylations of Ptd-EA (Ptd-EA→Ptd-MEA→Ptd-DEA→Ptd-Cho).

In yeast Ptd-Cho synthesis occurs mainly via the methylation pathway for which the phosphatidylbase route is favoured. Yeast defective in the production of Ptd-Cho have been shown to arise with the loss of function of one or both the genes *Cho2* or *Opi3* that encode phosphatidyl *N*-methyltransferase enzymes (Summers *et al.*, 1988, McGraw and Henry, 1989). Both yeast enzymes are membrane bound, in contrast to the P-EAMeTs of castor bean and spinach that have been shown to be cytosolic (Prud'homme and Moore, 1992; Weretilnyk *et al.*, 1995). The enzyme encoded by *Cho2* catalyzes the methylation of Ptd-EA→Ptd-MEA (Summers *et al.*, 1988). The enzyme encoded by *Opi3* catalyzes the methylation of Ptd-MEA→Ptd-DEA and Ptd-DEA→Ptd-Cho. The *opi3* mutant was shown to incorporate only 2-3% of <sup>14</sup>C- labelled ethanolamine into Ptd-Cho while over the same time course 40-50% was incorporated into Ptd-Cho in WT cells (Greenberg *et al.*, 1983). The *opi3* mutant was also shown to be heat sensitive, likely due to the accumulation of Ptd-MEA and possibly Ptd-DEA in the membrane. These intermediates are believed to be less stable than Ptd-EA or Ptd-Cho at higher temperatures resulting in a temperature-sensitive phenotype as a result of compromised membrane function (McGraw and Henry, 1989). Yeast also produces Ptd-Cho via the

CDP-choline pathway but the methylation pathway is favoured (Greenberg *et al.*, (1983).

### **The P-EAMeT family**

The proteins of the plant P-EAMeT family are approximately 500 amino acids in length and in spinach leaves these enzymes have been localized to the cytosol (Weretilnyk *et al.*, 1995). The larger size of plant P-EAMeTs relative to many other methyltransferases is due to the bipartite nature of this family of proteins (Nuccio *et al.*, 2000). Each protein consists of two catalytic domains of approximately 110 amino acids in length. These domains are annotated as SAM binding regions (interpro accession: IPR013216) and are part of the methyltransf\_11 protein superfamily (Protein Data Bank - PF08241). These domains share the alpha/beta Rothman fold that is characteristic of most nucleotide binding domains (Zhang and Cheng, 2003). A single stranded beta-sheet encompasses each of the SAM binding domains (SBDs; Bolognese and McGraw, 2000). These SBDs consist of three motifs (I, post I, II and III) which have been proven to be highly conserved for both plant and non-plant P-EAMeT proteins (Charron *et al.*, 2002). Each SAM-binding motif is separated by a number of conserved amino acids (Van Gemen and Van Knippenberg, 1990). Despite the widespread conservation of amino acids comprising SBDs, little is known about which amino acids are critical to SAM binding (Zhang and Cheng *et al.*, 2003).

### **Spinach P-EAMeTs**

Spinach P-EAMeT can catalyze all three methylations from P-EA→P-Cho. When

the N-terminal half of spinach P-EAMeT was assayed it was found that this domain alone could methylate P-EA→P-MEA (Nuccio *et al.*, 2000). This finding suggests that SBD1 of spinach P-EAMeT is capable of methylating P-EA→P-MEA while SBD2 must methylate the subsequent reactions of the pathway (P-MEA→P-DEA→P-Cho). During time-course experiments it was shown that P-MEA and P-DEA accumulate first and to a greater extent than P-Cho. This finding supports the model of domain function whereby one reaction, namely the conversion of P-EA→P-MEA is performed more efficiently by SBD1 than the two subsequent reactions from P-MEA→P-DEA→P-Cho performed by SBD2 (Nuccio *et al.*, 2000). The product P-Cho is a strong inhibitor of spinach P-EAMeT at 10 mM, levels found in spinach leaves (Smith *et al.*, 2000). However, removal of the C-terminal domain of spinach P-EAMeT alleviates some of the inhibitory effects by P-Cho whereby the N-terminal half remains twice as active in the presence of 10 mM P-Cho (Nuccio *et al.*, 2000). This observation suggests that SBD2 may regulate the activity of SBD1. Importantly, the initial methylation of P-EA is reported to be the rate limiting step in the synthesis of P-Cho so factors regulating the operation of SBD1 conceivably regulate the rate of P-Cho biosynthesis. It has been demonstrated that exposure to saline conditions led to a 10-fold increase in the abundance of RNA transcripts encoding spinach P-EAMeT (Nuccio, *et al.*, 2000) so regulation of the activity of spinach P-EAMeT also takes place at the level of transcript production and/or turnover.

The existence of more than one P-EAMeT enzyme has been shown for spinach (Weretilnyk *et al.*, 1995). P-EA methylation activity was highest in extracts prepared from leaves and not detected in extracts from roots. However, the ability to methylate P-



MEA and P-DEA was found in leaves and roots suggesting the presence of another enzyme capable of methylating phosphobases (Weretilnyk *et al.*, 1995).

### **Wheat P-EAMeTs**

Wheat P-EAMeT (TaPEAMT1) has been cloned and characterized biochemically (Charron *et al.*, 2002). In wheat it was shown that the N-terminal half of TaPEAMT1 is capable of all three methylation reactions ( $\text{P-EA} \rightarrow \text{P-Cho}$ ) while the C-terminal half can catalyze only two methylations starting with the substrate P-MEA ( $\text{P-MEA} \rightarrow \text{P-Cho}$ ) (Charron *et al.*, 2002). As such, the function for the C-terminal half of TaPEAMT1 matches the predicted activity of the C-terminal half of spinach P-EAMeT (Nuccio *et al.*, 2000). Interestingly when motif I of SBD1 was mutated and assayed for phosphobase activity, all methyltransferase activity was abolished. In contrast, when motif I of SBD2 was mutated no change in activity was seen. The conclusion drawn by the authors was that SBD1 was essential for phosphobase catalysis while SBD2 likely plays more of a regulatory role in the plant (Charron *et al.*, 2002). The authors also suggested that SBD2 may be involved in making methyl groups available to SBD1 or that SBD2 is required during times of stress when more choline production is required (Charron *et al.*, 2002). With respect to regulation, *TaPEAMT1* transcripts were detected in leaf and crown tissues only, not the roots.

A wheat P-EAMeT isoform (TaPEAMT2) has been recently cloned and characterized (Jost *et al.*, 2009). TaPEAMT2 has 77% identity to TaPEAMT1 at the amino acid level. Like TaPEAMT1, TaPEAMT2 can methylate all three reactions from  $\text{P-EA} \rightarrow \text{P-Cho}$ . The SBDs of TaPEAMT2 were not assayed individually so it is unknown

if the functions associated with the domains are similar to that of spinach P-EAMeT or TaPEAMT1. TaPEAMT2 is a more efficient enzyme with a specific activity four times greater for P-EA methylation activity than TaPEAMT1 when assayed *in vitro*.

TaPEAMT2 is less susceptible to inhibition by SAH but is more strongly inhibited by P-Cho relative to TaPEAMT1. TaPEAMT1 has a lower affinity for P-EA than TaPEAMT2 but a greater affinity for the substrate SAM. Due to this property, Jost *et al.* (2009) propose that TaPEAMT1 functions when SAM is abundant and permits the accumulation of P-Cho to higher levels because it is less sensitive to inhibition by this product than TaPEAMT2. When SAM is less available and SAH accumulates, TaPEAMT2 will be the more active enzyme as it is not as susceptible to inhibition by SAH. Both TaPEAMT proteins are directly inhibited by PA *in vitro*. Lipid binding assays show that both proteins bind PA via an uncharacterized lipid binding motif. TaPEAMT1 was more sensitive to inhibition by PA and accordingly had a higher binding affinity for PA.

Jost *et al.* (2009) have shown that the levels of TaPEAMT1 and TaPEAMT2 are regulated at both the protein and transcript level in response to cold exposure. Increases in the amount and activity of the TaPEAMT proteins were detected after only 2 d of exposure to cold. This early response appears to be mediated by a post-transcriptional or post-translational mechanism as translation of both proteins increased in response to cold without a corresponding increase in transcript of either gene. After 6 d of cold exposure, TaPEAMT activity had increased 4-fold while the amount of total TaPEAMT protein had increased 6-fold. Analysis of *TaPEAMT* expression by real-time PCR indicated that *TaPEAMT2* expression was induced after the 6 d cold exposure while *TaPEAMT1*

expression was not significantly altered. The authors concluded that the increase in TaPEAMT protein expression and activity after 6 d of cold exposure is likely due to the increased expression of the *TaPEAMT2* gene.

### **Non-plant P-EAMeTs**

*C. elegans* has two P-EAMeT like proteins, together they are able to produce P-Cho from P-EA (Brendza *et al.*, 2007). Phosphoethanolamine N-methyltransferase 1 (PMT-1) contains SBD1 and catalyses the methylation of P-EA→P-MEA. PMT-1 was shown to have a 6 to 10 fold higher affinity for P-EA compared to spinach P-EAMeT. PMT-2 contains SBD2 and catalyzes the remaining two methylation steps: P-MEA→P-DEA→P-Cho (Brendza *et al.*, 2007). Although each protein has only one SBD they are both nearly 500 amino acids long. In contrast, *P. falciparum* has a phosphoethanolamine methyltransferase protein (PfPMT) with only one SBD that contains similarity to both SBDs from plants. This enzyme performs all three methylation steps despite being half the molecular mass of plant P-EAMeTs (Pessi *et al.*, 2004).

### **Arabidopsis P-EAMeTs**

There are three genes annotated as encoding putative P-EAMeT/P-EAMeT-like enzymes in Arabidopsis. The enzyme orthologous to the P-EAMeT enzymes is encoded at locus At3g18000. The ability of P-EAMeT to synthesize P-Cho from P-EA has been demonstrated (Bolognese *et al.*, 2000) but the domain-associated activities of SBD1 or SBD2 are unknown.

The loci encoding two paralogous enzymes are found on chromosome 1 at At1g48600 and At1g73600. Arabidopsis P-EAMeT has been dubbed NMT1 (Bolognese *et al.*, 2000) and following this nomenclature, the enzyme encoded by At1g48600 will be referred to as NMT2 and the enzyme encoded by At1g73600 as NMT3. Alignment of the three proteins with ClustalW (<http://align.genome.jp/>) indicates that Arabidopsis NMT1 shares 87% and 82% amino acid identity with NMT2 and NMT3, respectively, and NMT2 shares a 78% amino acid identity with NMT3. BeGora *et al.* (2010) have shown that Arabidopsis NMT2 is a phosphobase methyltransferase capable of methylating P-MEA and P-DEA but not P-EA. Arabidopsis NMT3 has not been confirmed yet as a phosphobase methyltransferase.

Suppressed expression of Arabidopsis *NMT1* (*XPL1*) produces a stunted root phenotype and reduced levels of P-Cho in epidermal cells in roots (Cruz-Ramirez *et al.*, 2004). These *xpl1* plants also tend to flower late, have a reduced root growth rate, an increased number of lateral roots, and fewer root hairs. RNA *in situ* hybridization shows that *NMT1* is expressed primarily in the apical meristem of roots and shoots and that this gene is only expressed to a lesser extent in leaves, inflorescences, and siliques which supports the predominantly root-specific phenotype exhibited by *xpl1* plants (Cruz-Ramirez *et al.*, 2004). Roots and leaves of *xpl1* showed a 23% and 18% decrease, respectively, in Ptd-Cho content relative to WT. Overall *xpl1* mutant plants showed a 5% and 7% decrease in P-Cho and choline, respectively, when compared to WT plants (Cruz-Ramirez *et al.*, 2004). The mutant root phenotype can be rescued by the addition of choline or P-Cho indicating there is a functional choline kinase enzyme in Arabidopsis.

The mutant phenotype was also complemented by the addition of PA indicating that Arabidopsis NMT1 is required for the synthesis of Ptd-Cho in roots of Arabidopsis consistent with PA being produced from Ptd-Cho (Cruz-Ramirez *et al.*, 2004).

The NMT proteins are critical for choline synthesis in Arabidopsis. Mou *et al.*, (2002) created an RNAi knockdown line that reduced total choline content by 64% relative to WT levels. These knockdown plants likely had reduced amounts of all three *NMT* transcripts due to the high degree of sequence identity shared between them. Knockdown plants experienced early senescence and a temperature-sensitive male sterility phenotype (Mou *et al.*, 2002). The yeast NMT knockout *opi2* also has a temperature sensitive phenotype which is associated with the accumulation of Ptd-MEA and Ptd-DEA in the membranes (McGraw and Henry, 1989). McGraw and Henry (1989) proposed that the altered membranes may be detrimental at high temperatures.

An upstream open reading frame (uORF) has been shown to regulate the post-transcriptional activity of Arabidopsis NMT1. Tabuchi *et al.*, (2006) posit that post-transcriptional regulation is exerted by the uORF in a feedback inhibition manner. This uORF is also found in other plant P-EAMeT's including spinach and corn (Tabuchi *et al.*, 2006). The Arabidopsis *NMT2* and *NMT3* genes do not have this conserved uORF. The uORF does appear to be involved in feedback inhibition of NMT1 by choline (Tabuchi *et al.*, 2006). Transgenic Arabidopsis calli were generated that expressed the uORF sequence fused to the reporter gene encoding  $\beta$ -glucuronidase. When calli are exposed to 10 mM choline both the transcription and translation of the chimeric gene were reduced. Moreover, when spinach plants were grown in the presence of choline, translation of

*P-EAMeT* transcripts was reduced in the roots. These finding indicate that choline plays a role in the regulation of *P-EAMeT* expression (Tabuchi *et al.*, 2006).

## MATERIALS AND METHODS

Chemicals and enzymes were purchased from Sigma (Sigma-Aldrich Canada Ltd., Oakville, ON) unless otherwise stated. All solutions were prepared with de-ionized water purified by a Barnstead NANOpure II water purification system (SYBRON / Barnstead, Dubuque, IA). All primers were synthesized by The Institute for Molecular Biology and Biotechnology (MOBIX) (<http://www.science.mcmaster.ca/mobix/>).

### Plant material and growth

#### *Seeds*

Seeds for the *Arabidopsis thaliana* wild-type CS 60000 and SALK lines were purchased from the Arabidopsis Biological Resource Center (ABRC) at the University of Ohio (<http://www.biosci.ohio-state.edu/~plantbio/Facilities/abrc/abrchome.htm>). The Arabidopsis T-DNA SALK line used in this study was SALK\_062703.

#### *Seed Surface Sterilization*

Approximately 30 seeds were placed in 1.5-mL microfuge tubes for surface sterilization. To each tube 1 mL of 70% v/v ethanol was added. The tubes were inverted periodically for 2 min. The ethanol was removed with a glass pasteur pipette and 1 mL of seed sterilization solution consisting of 30% v/v commercial Javex 5 bleach (Clorox Company of Canada, Brampton, ON) and 0.1 % v/v Tween 20 (Cat. No. 74H101015) was added. Each tube was inverted periodically for 10 min. The sterilization solution was removed, and seeds were rinsed at least five times with 1 mL of sterile H<sub>2</sub>O each time. After the last H<sub>2</sub>O wash was removed, 1 mL of 0.1 % w/v sterile Phytigel™ (Cat. No.

P8169) was added to suspend the seeds then the tubes containing suspended seeds were capped and stored overnight at 4°C.

### *Seed Germination*

Surface-sterilized seeds were transferred to Murashige and Skoog (MS) plates (Murashige and Skoog, 1962) containing (per L volume) 1.0 % w/v sucrose (FisherScientific, Cat. No. 044832, Fair Lawn, NJ), 1.0 mL MS vitamins (Cat. No. 14K2353), and 0.215 g Murashige and Skoog Salt Mixture (Cat. No. 027K23061). The outside of the plates were sealed with Micropore<sup>™</sup> surgical tape (3M Innovation Inc. St. Paul, MN), and placed inside an incubation chamber (Percival 50036, Boone, IA) set for a 24 hour photoperiod at 22°C for 7 to 10 d.

### *Seedling Growth Conditions*

Seedlings from the MS plates were transplanted to sterile Promix BX soil (Premier Horticulture Inc., Quakertown, PA) in black, rectangular 7.7 cm x 4.4 cm plastic pots (The Lerio Corp, Mobile, AL) and placed in a Conviron CMP 3244 controlled environment chamber (Conviron, Winnipeg, MA). Specific light and temperature conditions are described in the results where necessary. The setting used most frequently was 21 °C with a light intensity of  $120 \mu\text{mol m}^{-2} \text{s}^{-1}$  during the 12 h light/12 h dark photoperiod. For testing the effect of heat, the day temperature was raised to 26 °C and the light intensity and day length was  $120 \mu\text{mol m}^{-2} \text{s}^{-1}$  and 12 h, respectively. The light cycle was also adjusted to test the effect of day length on flowering time. Plants were grown at 22°C with a light intensity of  $120 \mu\text{mol m}^{-2} \text{s}^{-1}$  during a light cycle of 8 h light/16 h dark or a 16 h light/8 h dark photoperiod. Plants were watered as required by



adding water to the tray supporting the pots. Plants were fertilized once every two weeks with 1 g/L of 20-20-20 all purpose fertilizer (Plant Products Ltd. Brampton, ON).

### *Seed Collection*

Seeds were collected from plants when the siliques were brown, approximately 8 weeks after transplant to pots. Seeds were stored in 12.7 X 7.6 cm envelopes at room temperature for one week to dry. Envelopes were sealed shut with VWR lab tape (VWR, Mississauga, ON; Cat. No. 36439) for long term storage at room temperature.

### **Salt stress experiment**

Plants to be grown on defined media plates were sterilized and germinated as described above. Seedlings were transplanted from MS plates to defined media plates after three d in 24 hour light and 22°C. Plants were transplanted aseptically to plates containing either 50 mM NaCl or 0 mM NaCl (control). Plants on culture plates were grown in a chamber at 22°C with a light intensity of  $150 \mu\text{mol m}^{-2} \text{s}^{-1}$  and a 12 h light/12 h dark photoperiod. Plates were placed vertically to allow the growth of the roots along the surface of the agar towards the bottom of the plate. Root elongation and shoot phenotype were followed photographically each day until roots reached the bottom edge of the plates. 50 mg of root and shoot tissue was harvested and immediately frozen in liquid nitrogen and stored at -80° C to be used for future analyses.

An experimental trial consisted of 18 plates: 3 control plates containing no NaCl with 8 CS 60000 WT plants on each plate, 3 plates containing 50 mM NaCl with 8 CS 60000 WT plants on each plate, 3 control plates with 8 SALK\_062703 plants

homozygous (Hmz) for a T-DNA insert in At1g73600, 3 plates containing 50mM NaCl with 8 SALK\_062703 Hmz, 3 control plates with 4 CS 60000 WT and 4 SALK\_062703, 3 plates containing 50 mM NaCl with 4 CS 60000 WT and 4 SALK\_062703 plants.

## **Extraction and analysis of nucleic acids**

### *DNA Preparation*

One fully extended rosette leaf was selected for harvest from each plant. The lid of a 1.5 mL microfuge tube was closed on the leaf to excise a leaf disc of approximately 15 mg. The leaf disc was either used immediately for DNA extraction or it was flash-frozen and stored at -80°C prior to extraction. Freezing the tissue was not found to affect the quality of the extraction.

The DNA extraction protocol was adapted from Dellaporta *et al.*, (1983). Tissue was ground by hand with disposable pestles (VWR, Cat No. 70680-122) at room temperature using 400 µL of extraction buffer until the tissue was completely macerated. The extraction buffer contained 200 mM Tris-HCL (BioShop, Burlington, ON; Cat. No. 764661) pH 7.5, 250 mM NaCl (BioShop, Cat. No. 4F1801.), 25 mM EDTA (ethylenedinitrilotetraacetic acid; EMD Chemicals Inc., Darmstadt, Germany; Cat. No. 45082519.), and 0.5% w/v SDS (sodium dodecyl sulfate, BDH Inc., Cat. No. 107318/28914). The macerated tissue was vortexed and the microfuge tubes centrifuged for 5 min at 16,100 g in a 5415 D Centrifuge (Eppendorf, Hamburg, Germany; Cat. No. 0061966.). A 300-µL volume of supernatant was transferred to a new, sterile 1.5 mL microfuge tube and an equal volume of isopropanol (Burdick and Jackson, Muskegon,

MI; Cat. No. COO44) was added to the supernatant. The contents of the tube were mixed by inversion for 5 to 10 sec and then the tube containing precipitated DNA was centrifuged for 5 min at 16,100 g at room temperature. The supernatant was discarded and the pellet was washed once with 300  $\mu$ L of 75% ethanol and the pellet was allowed to air dry. The dried DNA pellet was dissolved in 25  $\mu$ L of autoclaved H<sub>2</sub>O. The solution containing DNA was stored frozen at -20°C prior to use.

### *RNA Preparation*

All steps were carried out at room temperature unless otherwise stated. All centrifugation steps were performed with an Eppendorf 5415 D Centrifuge. RNA was extracted from Arabidopsis leaf tissue using the QIAGEN RNeasy Mini Kit (50) (QIAGEN, Mississauga, ON; Cat. No. 74104,) following the protocol described by the manufacturer (QIAGEN RNeasy Guide). One rosette leaf from an approximately 21 d-old plant (between 25-50 mg Fresh Wt.) was harvested and flash frozen in liquid nitrogen. Leaves were either stored at -80°C or processed immediately. The frozen leaves were ground using a disposable pestle with 450  $\mu$ L of extraction buffer containing guanidinium thiocyanate (RNeasy kit) and 4.5  $\mu$ L of 14.3 M  $\beta$ -mercaptoethanol (Cat. No. 39F-060715). After vortexing, the brei was pipetted into a QIAshredder spin column using a pipette tip with the end cut off to allow for complete transfer of the viscous lysate. The column and a 2-mL collection tube were centrifuged at 16,100 g for 2 min and then the supernatant was transferred to a sterile 1.5-mL microfuge tube. A volume of 225  $\mu$ L of absolute ethanol (Commercial Alcohols Inc, Brampton, ON) was added to the supernatant and the contents mixed thoroughly by repeated pipetting. From this sample, 650  $\mu$ L was

transferred to the membrane of an RNeasy Mini Column and they were centrifuged together for 15 s at 9,300 g. The column flow-through was discarded and 700 µL of “Buffer RW1” containing ethanol and guanidinium thiocyanate was dispensed onto the membrane of the spin column and centrifuged as before at 9,300 g for 15 s. The flow-through was again discarded and 500 µL of “Buffer RPE” containing ethanol was dispensed by pipette onto the membrane. The entire column containing solution was centrifuged for 15 s at 9,300 g. The flow-through fraction was discarded and another 500 µL volume of “Buffer RPE” was added to the membrane and centrifuged at 16,100 g for 1 min. After the flow-through was discarded the RNeasy column was transferred to a sterile, 1.5 mL collection tube and RNA was eluted from the membrane using 30 µL of RNase-free H<sub>2</sub>O. The H<sub>2</sub>O was dispensed directly onto the membrane and the column with contents were centrifuged for 1 min at 9,300 g. The tubes containing the column eluate were closed and RNA samples were stored in a -80°C freezer prior to use.

#### *Screening for T-DNA insertion by Polymerase Chain Reaction (PCR)*

Left and right gene-specific primers and T-DNA-specific left border primers were designed using the T-DNA primer design tool from the Salk Institute Genomic Analysis Laboratory (<http://signal.salk.edu/tdnaprimers.html>; Alonso et al. 2003).

The left and right gene-specific primers were designed to amplify a product of 900 to 1100 bp. The left border primer specific to the T-DNA was selected to amplify a smaller product of about 450 to 800 bp. Identification and basic information on the left and right primers specific to each SALK\_062703 and the left border primers used to screen for T-DNA insertion in the alleles of At1g73600 are reported in Appendix A. The

left and right gene-specific and left border T-DNA primers were synthesized by MOBIX (<http://www.science.mcmaster.ca/mobix/>), and diluted to 10 pmol/ $\mu$ L for use.

Information provided by Oligo calculator was used to determine the optimal melting temperature of each primer for PCR reactions (<http://www.pitt.edu/~rsup/OligoCalc.html>). If multiple products were obtained, the annealing conditions of the primers were empirically tested by increasing the reaction temperature in increments of 2°C until one product was obtained.

The PCR reaction designed to screen for the presence of a T-DNA insertion contained 20 pmol of each primer (left and right gene-specific primers and the left border-specific T-DNA primer), 10 nmol of each dNTP, 25 nmol MgCl<sub>2</sub>, 1 x PCR buffer containing 20 mM Tris-HCl pH 8.8 (from a 10 x buffer stock solution), and 1.25 units of Taq DNA polymerase (Fermentas, Burlington, ON; Cat. No. EP0281) in a total volume of 50  $\mu$ L. An Eppendorf Mastercycler thermocycler (Eppendorf, Cat. No. 5345 003558) was used to carry out the PCR reactions. The initial denaturation step was carried out at 94°C for 2 min and was followed by 30 cycles of denaturation at 94°C for 30 s, annealing at the optimum primer-specific temperature (see Appendix A) for 30 s, and extension at 72°C for 1 min ending with final extension at 72°C for 10 min. DNA products amplified by PCR were stored at -20°C, or separated immediately by agarose gel electrophoresis.

#### *cDNA Synthesis for reverse transcriptase-polymerase chain reaction (RT-PCR)*

RNA prepared using the RNeasy kit was used for cDNA Synthesis using approximately 1  $\mu$ g for the cDNA synthesis reaction. cDNA copies were produced using Superscript<sup>™</sup> II Reverse Transcriptase RNase H (Invitrogen, Cat. No.18064-014). The

first strand synthesis required 10 pmol of forward and reverse primers (see Appendix A), 10 nmol of each dNTP, 1x First-Strand Buffer (Invitrogen; 250 mM Tris-HCL pH 8.3, 375 mM KCL, and 15 mM MgCl<sub>2</sub>), 200 nmol DTT (dithiothreitol; Invitrogen), 40 U of RNaseOUT™ recombinant ribonuclease inhibitor (Invitrogen Cat. No. 10777-019) and 200 U of Superscript™ II reverse transcriptase (Invitrogen, Cat. No.10777-019) in a final volume of 20 µL. The reaction mixture was incubated at 42°C for 50 min and then heated to 70°C for 15 min to deactivate the reverse transcriptase enzyme.

The cDNA control was amplified using left and right primers specific to the Arabidopsis ubiquitin gene on chromosome 4 associated with locus At4g05320 (see Appendix A). For the analysis of RT-PCR results, the product amplified by the ubiquitin primers was used as an internal control to assess the quality of the RNA extraction. The relative intensity of the bands associated with the ubiquitin product was used to interpret any variation in the intensity of the bands associated with the gene-specific primers as a potential difference in transcript abundance.

Gene-specific left and right primers (Appendix A) were designed to anneal to sites flanking introns 5 and 10, respectively, of At1g73600 using information obtained from The Salk Institute Genomic Analysis Laboratory (<http://signal.salk.edu/sect.2.html>). This strategy was used to ensure that cDNA from RNA and not trace contamination by genomic DNA would be amplified in the RT-PCR reaction.

The PCR amplification of first-strand cDNA for RT-PCR involved final concentrations of 20 pmol of each forward and reverse primer, 1X PCR buffer containing 20 mM Tris-HCL pH 8.8, 10 mM dNTP mix, 25 mM MgCl<sub>2</sub>, and 1.25 U recombinant Pfu

DNA polymerase (Fermentas, Cat. No. EP0501). The first denaturation step was carried out at 94°C for 2 min and it was followed by 30 cycles of denaturation for 30 s at 94°C, annealing for 30 s at 62°C, and extension for 45 s at 72°C. The final extension was carried out for 10 min at 72°C. Preparation and dilution of primer stocks to 10 pmol/uL and selection of suitable annealing temperatures were described earlier. RT-PCR products were analyzed by gel electrophoresis on 1.2% w/v agarose gels (see below).

#### *Gel Electrophoresis of PCR and RT-PCR products*

DNA products were separated by electrophoresis using a 1.2% (w/v) agarose (BioShop, Cat. No. 6G1437) gel. A Bio-Rad Mini-Sub<sup>®</sup> Cell GT (BioRad Laboratories, Mississauga, ON; Cat. No. 170-4487EDU) gel apparatus was used for the procedure. The gel was prepared using 800 µL of 50X TAE buffer stock (0.4 M Tris-base pH 8.0, 0.01 M Na<sub>2</sub>EDTA, 0.2 M NaOAc (Caledon Laboratories Ltd., Cat. No. 61362, Georgertown, ON) diluted with 39.2 mL H<sub>2</sub>O to a final volume of 40 mL, to which 0.48 g of agarose (BioShop, Cat. No. AGA001.100) was added. The solution was heated in the microwave for 3 x 16 s, swirling periodically to mix. The agarose solution was then slightly cooled and 2 µL of 10 mg/mL ethidium bromide (Boehringer Mannheim, Darmstadt, Germany; Cat. No. 12499420-19) was added to the molten agarose that was then allowed to solidify at 22°C. A volume of 10 µL of PCR product was mixed with 2 µL of 6X loading dye (10 mM Tris-HCL pH 7.6, 0.03 % w/v bromophenol blue, 0.03 % w/v xylene cyanol FF, 60 mM EDTA, 60 % v/v glycerol) for gel loading. To determine the size of the amplified products a lane containing 0.8 µg of 1 Kb Plus DNA ladder (Invitrogen, Carlsbad, CA; Cat. No. 10787-018) was combined with 2 µL of 6X loading dye and 8 µL of H<sub>2</sub>O and run

on each gel. Electrophoresis was performed at a constant voltage of 70 V for 1 h on a PowerPac 300 V power supply (Biorad, Cat. No. 165-5050). DNA fragments were visualized on an AlphaImager™ 2200 transilluminator with AlphaImager v. 5.5 software (Fisher Scientific, Ottawa, ON), and photos taken using a Mitsubishi CP700D printer/cutter (Mitsubishi Electric Canada Inc., Markham, ON).

*cDNA synthesis of a full length Arabidopsis NMT3 transcript*

RNA prepared with the RNeasy kit was used for cDNA synthesis using approximately 1 µg for the cDNA synthesis reaction. cDNA copies were produced using FirstChoice® RLM-RACE Kit (Ambion/Applied Biosystems Inc., Austin, TX; Cat. No. AM1700). The cDNA synthesis reaction required 10 pmol of forward and reverse gene-specific primers (see Appendix A), 10 nmol of each dNTP, 1x First-Strand Buffer (from a 10 x buffer stock solution, Ambion kit; 500 mM Tris pH 8.3, 750 mM KCl, 50 mM DTT and 30 mM MgCl<sub>2</sub>), 10 U RNase inhibitor (Ambion kit), 4.8 µmol of (D)-trehalose (Cat. No.T0167) and 100 U of Moloney Murine Leukemia Virus Reverse Transcriptase M-MLV RT (Ambion kit) all in a final volume of 20 µL. The reaction mixture was incubated for 1 h at 55°C. Normally the protocol for this reaction recommends incubation at 42°C. However, trehalose is used here to enrich for the synthesis of long (>1 kb) cDNA templates. Trehalose is a compatible solute that stabilizes the reverse transcriptase enzyme at higher temperatures allowing for the reaction to proceed at 55°C. At this temperature, difficult mRNA secondary structures and GC-rich regions are more likely to melt and permit the binding of reverse transcriptase to the template (Spiess and Ivell, 2002). The forward and reverse primers used were designed based upon the 5' and



3' untranslated region (UTR) sequences of At1g73600 in order to amplify a 1684 bp cDNA template. This template was subsequently used in a nested PCR reaction to amplify a PCR product containing the full open reading frame (ORF) of At1g73600.

The PCR amplification of the full length cDNA involved final concentrations of 20 pmol of each forward and reverse primer, 1X Klen-Taq<sup>™</sup> buffer (from a 10 x buffer stock solution, Cat. No.B6178 containing 400 mM Tricine-KOH (pH 9.2 at 25°C), 150 mM KOAc, 35 mM Mg(OAc)<sub>2</sub>, 750 µg/ml bovine serum albumin), 10 nmol dNTP mix, 50 nmol MgCl<sub>2</sub>, and 5 U of Klen-Taq<sup>™</sup> DNA polymerase (Cat. No. D5187 supplied in 50% glycerol, 40 mM Tris-HCl (pH 7.5), 50 mM KCl, 25 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.1mM EDTA, 5.0 mM 2-mercaptoethanol, 0.25% Thesit<sup>®</sup> (polyoxyethylene 9 lauryl ether) in a total reaction volume of 20 µL. The first denaturation step was carried out at 94°C for 1 min and it was followed by 5 cycles of denaturation for 30 s at 94°C, annealing of primers to template at 60°C and extension for 2 min and 30 s at 68°C . The initial temperature for primer annealing was performed at 60 °C instead of 68°C to insure binding of primers to the template and transcription of the proper cDNA template. This was followed by 25 cycles of denaturation for 30 s at 94.0°C and annealing and extension for 3 min at 68°C. The primer annealing temperature was increased for this second step to allow for more specific amplification of the desired template. The preparation and dilution of primer stocks to 10 pmol/µL and selection of suitable annealing temperatures was described earlier. PCR products were analyzed by gel electrophoresis on 1.2% w/v agarose gels in TAE buffer.

This PCR design uses primers that anneal to the template to the inside of the primers used to synthesize the cDNA template. These primers also contain mismatched base pairs used to create restriction sites in the PCR product that can be used for subsequent cloning steps. The forward primer was designed to incorporate an *NcoI* restriction site. Since the restriction sequence of *NcoI* contains an ATG this restriction site overlaps with the translation start codon. The reverse primer was designed to contain a *BamHI* site outside of the translation stop codon. The primers were designed to amplify a 1591 bp PCR product containing an N-terminal *NcoI* site and C-terminal *BamHI* site.

### ***Arabidopsis NMT3* clone preparation**

#### *Restriction enzyme digestion of DNA*

PCR-amplified *NMT3* cDNA or pET-30a(+) vector (EMD Biosciences, Madison, WI; Cat. No. 699099) were dissolved in H<sub>2</sub>O and subjected to restriction enzyme digestion with *NcoI*, *BamHI*, *EcoRI* or *SacI* (Fermentas FastDigest<sup>®</sup> enzymes). The digestion mixture consisted of 2 µL of 10 X FastDigest<sup>®</sup> buffer (Fermentas), 1 µL each of the two enzymes required for that particular reaction and a variable volume of water, PCR amplified insert or vector pET-30a(+) to yield a final volume of 20 µL in a 0.5 mL microfuge tube. Concentration of DNA was approximated through comparison to the band intensity of DNA fragments separated using non-denaturing gel electrophoresis of the Lambda DNA/HindIII 2 DNA marker mix (Fermentas; Cat. No. SM0101). The restriction digest mixture was incubated at 37°C for 5 min. Restriction products were then purified by precipitation or gel extraction. To precipitate the restriction digestion

products, 2.5 x volume absolute ethanol and 0.1 x volume of 3 M sodium acetate were added to the contents of the restriction digestion tube and then tubes were placed at -80°C for 1 h. The precipitated DNA was pelleted by centrifugation at 14,000 g for 5 min. The pellet was then dissolved in a varying amount of H<sub>2</sub>O depending on the desired final concentration of the DNA. Gel extraction was performed using a QIAquick<sup>®</sup> Gel Extraction Kit (QIAGEN, Cat. No. 28704).

*Ligation of purified insert to expression vector*

The cDNA insert encoding NMT3 (or the N/C-terminal half of NMT3) was ligated into the expression vector pET-30a(+). This plasmid carries a gene for kanamycin resistance that can be used to select for successful ligation and transformation. The ligation mixture was composed of 50 ng each of restriction digested and purified insert DNA and pET-30a(+), 2 µL of 10 X ligation buffer (Fermentas, 400 mM Tris-HCl, 100 mM MgCl<sub>2</sub>, 100 mM DTT, 5 mM ATP (pH 7.8 at 25°C)), 1 µL of T4 DNA ligase (provided in 20 mM Tris-HCl (pH 7.5), 1 mM DTT, 50 mM KCl, 0.1 mM EDTA and 50% (v/v) glycerol; 5U/µL, Cat. No. EL0014, Fermentas) and H<sub>2</sub>O to a final volume of 20 µL. The ligation reaction was incubated overnight at 16°C.

*Transformation of E. coli and identification of NMT3-containing colonies*

Ligation products were used to transform the *E. coli* host strain DH5α, a strain that lacks a T7 RNA polymerase gene. 2 mL of liquid LB media (10% (w/v), 5% (w/v) Bacto-tryptone, 5% (w/v) Bacto-yeast extract, 10% (w/v) NaCl) was inoculated with a single colony of DH5α and grown overnight at 37°C with shaking (200 rpm) in an Innova<sup>®</sup> 44 Incubator shaker (New Brunswick Scientific). 0.1 mL of this culture was

used to inoculate 50 mL of LB which was then incubated for 2 h at 37°C with shaking (200 rpm). Cells were harvested by centrifugation at 3,000 g, 4°C for 5 min. The supernatant was decanted and from this point all steps were performed on ice with solutions kept at 4°C. Cells were resuspended in 10 mL of 0.1 M MgCl<sub>2</sub> then centrifuged at 3,000 g for 5 min. The supernatant was discarded and the cells were resuspended in 10 mL of 0.1 mL of CaCl<sub>2</sub> and the suspension was incubated on ice for 20 min. The cells were again pelleted at 3,000 g for 5 min. The supernatant was discarded and the cells were resuspend in 1 mL of 0.1 mL of CaCl<sub>2</sub>. A 200 µL volume of the cell suspension was transferred to a round bottom polypropylene tube (Becton Dickinson Labware, Lincoln Park, NJ ; Cat. No.4-2059-2, Falcon<sup>®</sup>) and 20 µL of the ligation reaction product was added and the mixture was allowed to incubate on ice for 40 min. The cells were heat-shocked at 42°C for 1 min before 1.8 mL of LB was added. This 2 mL mixture was then incubated for 2 h at 37°C with shaking (200 rpm).

Following transformation, DH5α cells were plated to select for those containing the plasmid with the *NMT3* cDNA insert. 100 µL of cells were plated on solid LB media (Liquid LB containing 1.5% (w/v) agar) and 30 µg mL<sup>-1</sup> kanamycin. Plates were incubated overnight at 37°C. Several colonies were picked from the plates with a sterile toothpick and streaked onto a plate divided into grids and the toothpick was then used to inoculate 10 µL of water in a 0.2 mL PCR tube. Up to 10 colonies were pooled for each tube of H<sub>2</sub>O. Tubes were incubated for 10 min at 95°C to lyze the cells, cell debris was pelleted by centrifugation for 1 min at 14,000 g and then 2 µL of the supernatant was used for screening by PCR. The multiple cloning site of the pET-30a(+) vector is flanked by a

T7 promoter and T7 terminator sequence. Primers were used that correspond to each of these regions in order to amplify the cloned DNA between the primer annealing sites.

### Radioassays

The substrate *S*-adenosyl-L-methionine (SAM) was dissolved in 0.01 N H<sub>2</sub>SO<sub>4</sub> : ethanol 9:1 (v/v) and 25 µl aliquots were stored at -20°C. *S*-[methyl-<sup>3</sup>H]adenosyl-L-methionine ([<sup>3</sup>H]SAM) (NEN® Perkin Elmer, Inc., Waltham MA) was purchased as a 10 mM solution in H<sub>2</sub>SO<sub>4</sub> : ethanol 9:1 that had a specific activity of 81.9 Ci nmol<sup>-1</sup>. P-EA substrate was dissolved in 0.1 N HCl to a concentration of 7.5 mM and stored at -20°C. The substrates P-MEA and P-DEA were synthesized according to the method developed by Datko and Mudd (1988b) using 25 mg of dipalmityl-phosphomethylethanolamine (Ptd-MEA, Avanti Polar Lipids, Alabaster, AB; Cat. No.850851) or dipalmityl-phosphodimethylethanolamine (Ptd-DEA, Avanti Polar Lipids, Cat. No.850854). The Ptd-lipids were first dissolved in 1 mL of 5% v/v Triton X-100 by mixing end-over-end. 160 U of phospholipase C (from *B. cereus*, Cat. No. P4014) was mixed with 140 µL of dimethylglutarate buffer pH 7.5 NaOH. The entire mixture was desalted by centrifugation through Sephadex G-25 (equilibrated with dimethylglutarate buffer-NaOH, pH 7.5). Phosphobases were cleaved from the phosphatidylbases by hydrolysis with 360 U of phospholipase C for 40 min at 37°C. The reaction was stopped by the addition of 3.2 mL of methanol and 1.6 mL of chloroform. To assist in phase separation, 2.06 mL of H<sub>2</sub>O and 4.8 mL chloroform were also added. The top phase, containing the Phosphobases, was removed and N-evaporated (Meyer N-EVAP Organomotion Berlin,

MA; Model Bo. 111) at 40°C in a 13 x 100 mm test tube. The residue was dissolved with 150 µL of 0.1 N HCl and stored at -20°C. The concentrations of P-MEA and P-DEA were determined by hydrolyzing the phosphate groups and quantifying the phosphate (Pi) released using the method of Martin and Tolbert (1983).

Dowex 50 W(H<sup>+</sup>) X8-200 (Cat. No. 44504) resin was prepared by resuspending 300 gm of resin in approximately 600 mL of H<sub>2</sub>O. The mixture was stirred and then the beads were allowed to settle before excess water was decanted. In this way the beads were “washed” with H<sub>2</sub>O 50 times and then protonated by successive treatment with 1 N HCl 30 times in a similar manner. To remove any excess acid the beads were again “washed” with water until the pH of the beads was between 6 and 7. The resin was stored in water at 4°C.

#### *Phosphobase N-methyltransferase assays*

The method used to assay for methyltransferase activity was adapted from Datko and Mudd (1988b). Methyltransferase activity was assessed using P-EA, P-MEA and P-DEA substrates. The components of each assay were 100 µL of 150 mM Hepes-KOH buffer (pH 7.8), 1 mM Na<sub>2</sub>-EDTA, 5 µL of 7.5 mM phosphobase (P-EA, P-MEA or P-DEA), 2.5 µL of 12 mM SAM, 1 µL of [<sup>3</sup>H]SAM (0.55 µCi; 1.22x10<sup>6</sup> dpm), 21.5 µL H<sub>2</sub>O, and 25 µL of sample extract to a final volume of 150 µL. For control assays the 25 µL of sample extract was replaced with 25 µL of H<sub>2</sub>O. Assays were prepared in 1.5 mL microfuge tubes and incubated in a water bath at 30°C for 30 min. The reaction was stopped by addition of 1 mL of ice cold H<sub>2</sub>O, the reaction mixture was then vortexed and placed on ice until proceeding to subsequent steps.

Dowex 50 W(H<sup>+</sup>) X8-200 resin columns were prepared in disposable Evergreen columns (Diams, Mississauga, ON; Cat. No.208-3384-060). Disposable columns were first rinsed with H<sub>2</sub>O and then 1 mL of regenerated Dowex resin was added and washed with 5 column volumes of H<sub>2</sub>O before use. To separate the [<sup>3</sup>H] SAM from the [<sup>3</sup>H] phosphobase products, a 1 mL aliquot of the diluted enzyme assay mixture was applied to the Dowex columns. The columns and their contents were then washed three times with 0.5 mL of H<sub>2</sub>O. The columns were transferred to individual 16 x 150 mm disposable borosilicate test tubes and the phosphobases bound to the Dowex resin were eluted by washing the resin with 10 mL of 0.1N HCl. The eluate was vortexed and a 2 mL volume was transferred to a 20 mL disposable scintillation vial (Fisherbrand Cat. No. 333711B) containing 10 mL of Ready Safe<sup>®</sup> fluor (Beckman, Mississauga, ON; Cat. No. 141349). The contents of each vial were vortexed and radioactivity present was quantified with a Tricarb 2900 TR liquid scintillation analyzer (Perkin Elmer, Waltham, MA).

### **Phospholipid profiling**

The protocol reported by Fiehn et al. (2000) to extract metabolites present in the polar and lipid phases was followed but modified. For the analysis plants were grown under 12 h day conditions at 21°C and harvested at 28 d post-germination. Approximately 200 mg of fully expanded rosette leaves from three WT and three SALK\_062703 plants were ground as separate samples, each in a chilled mortar with a pestle and 700 µL of 100% HPLC grade methanol (Caledon Laboratories Ltd, Cat. No. 6701-7). 58 µL of 1M NaCl was added to the mortar and the slurry was transferred to a 15 mL Corex<sup>®</sup> tube.

Another 700  $\mu\text{L}$  of methanol was used to wash the mortar. The remainder of the slurry was added to the Corex<sup>®</sup> tube and placed in a 70°C water bath and shaken for 15 min. The tubes were then centrifuged at 14,000  $g$  for 3 min at 4°C using a Beckman Avanti<sup>™</sup> J-25 centrifuge fitted with a Beckman JA-20 rotor. The supernatant was transferred to a 16 x 100 mm disposable culture tube (Fisherbrand®, catalogue no. 14-961-27) and 750  $\mu\text{L}$  of chloroform was added to the Corex<sup>®</sup> tube to resuspend the pellet. The resuspended pellet was heated at 37°C for 5 min and then centrifuged as above. The second supernatant was pooled with the supernatant from the initial centrifugation. A volume of 1.4 mL de-ionized water was added to the pooled supernatant and this mixture was centrifuged at 3,000  $g$  for 10 min at room temperature using an IEC Clinical centrifuge. The upper, polar phase was removed and dispensed into a Wheaton 1-mL V-vial (Cat. no. 986254) and evaporated to dryness under a stream of  $\text{N}_2(\text{g})$  in an N-evap analytical evaporator. The non-polar (chloroform) portion was removed to a 2 mL screw cap vial (Gerresheimer, Cat. No. 60910, Rochester, NY). The remaining non polar phase and protein interphase was extracted twice more with 750  $\mu\text{L}$  of chloroform as above. The non-polar portions were pooled in the 2 mL vial and evaporated to dryness as above. Samples were stored at -20°C until processing. Lipid profiling was performed by the McMaster Regional Centre for Mass Spectrometry according to the protocol outlined by Basconcillo and McCarry (2008).

### **Construction of phylogenetic relationships among selected NMTs**

The UniProtKB database was searched using the amino acid sequence of Arabidopsis NMT1 (At3g18000) for related protein sequences using the ExPASy/SIB



blast program provided on the ExPASy server (<http://ca.expasy.org/>). All incomplete sequences (ESTs) were excluded. A total of 37 sequences listed in Table 1 were selected for analysis and the list includes both plant and non-plant species. Of the 37 protein sequences 23 were selected for the comparison of the full-length proteins. A sub-group of proteins was selected for which SBD1 and 2 have been assayed to determine their ability to methylate P-EA, P-MEA and P-DEA. Other sequences selected for this analysis include proteins from representative monocot and dicot species including several members of the Chenopodiaceae family. In addition to NMT sequences, several animal proteins that are known to only methylate at the Ptd level were used in the analysis. Only sequences similar to the length of plant P-EAMeTs were used (>450 bp) in this comparison. Sequences were aligned with ClustalX version 2.0 multiple alignment software (Larkin *et al.*, 2007). Tree construction methodology was adapted from Wu and Liu (2005). Phylogenetic relationships were calculated using the PROPARS maximum parsimony program and a tree was constructed with bootstrap confidence values using CONSENS from the PHYLIP 3.67 software package (Felsenstein, 1993). The tree was visualized using TreeView software (Page, 1996).

Trees were also constructed for the N and C-terminal halves of all proteins identified using the method outline above. These halves contain SBD1 and SBD2, respectively. The domains were identified as putative SAM-binding domains through the InterPro database ([www.ebi.ac.uk/interpro/](http://www.ebi.ac.uk/interpro/)). The domains were identified in Arabidopsis NMT1 as spanning amino acids 49 to 158 and 278 to 385, respectively. The sequences representing each half of the full length ORF were created by splitting the protein sequence at a shared valine residue that is amino acid 275 for Arabidopsis NMT1. Nine

Table 1 - Legend of protein sequences included in phylogenetic trees.

Abbreviations include UniProtKB/Swiss-Prot accession number. Sequences that contain only a single SBD are underlined.

---

AnNMT – *Atriplex nummularia* Phosphoethanolamine N-methyltransferase Q5NT83  
 AtNMT1 - *Arabidopsis thaliana* Phosphoethanolamine N-methyltransferase 1 Q9FR44  
 AtNMT2 - *Arabidopsis thaliana* Phosphoethanolamine N-methyltransferase 2 Q944H0  
 AtNMT3 – *Arabidopsis thaliana* Phosphoethanolamine N-methyltransferase 3 Q9C6B9  
 AtrNMT – *Aster tripolium* Phosphoethanolamine N-methyltransferase Q84SA4  
 BnNMT – *Brassica napus* Phosphoethanolamine N-methyltransferase Q7XJJ2  
 BvNMT – *Beta vulgaris* Phosphoethanolamine N-methyltransferase Q4H1G5  
 CbPMT – *Caenorhabditis briggsae* Putative methyltransferase Q61LS8  
 CeNMT1 – *Caenorhabditis elegans* Phosphoethanolamine N-methyltransferase (PMT-1) Q95PW7  
 CeNMT2 – *Caenorhabditis elegans* Phosphoethanolamine N-methyltransferase (PMT-2) Q22993  
 DrPMT – *Danio rerio* putative methyltransferase Q08CI9  
 GhNMT - *Gossypium hirsutum* Phosphoethanolamine N-methyltransferase A9XU50  
 HcNMT - *Hahella chejuensis* SAM-dependent N-methyltransferase Q2SG26  
 NsPMT – *Nodularia spumigena* Putative methyltransferase A0Z9Y0  
 NvPMT – *Nematostella vectensis* Putative methyltransferase A7S0M9  
 OsNMT – *Oryza sativa* Phosphoethanolamine N-methyltransferase Q6QA26  
 PfNMT – *Plasmodium falciparum* Phosphoethanolamine N-methyltransferase Q8IDQ9  
 PvPMT – *Plasmodium vivax* Putative methyltransferase A5K867  
 PtNMT – *Populus trichocarpa* Predicted Phosphoethanolamine N-methyltransferase B9I2F0  
 RcNMT - *Ricinus communis* Phosphoethanolamine N-methyltransferase B9T1H8  
 SbNMT1 - *Sorghum bicolor* Predicted Phosphoethanolamine N-methyltransferase 1 C5XHH3  
 SbNMT2 - *Sorghum bicolor* Predicted Phosphoethanolamine N-methyltransferase 2 C5YUY7  
 SeNMT – *Salicornia europaea* Phosphoethanolamine N-methyltransferase A5X7D6  
 SjNMT – *Sueda japonica* Phosphoethanolamine N-methyltransferase Q852S7  
 SINMT – *Solanum lycopersium* phosphoethanolamine N-methyltransferase Q9AXH3  
 SliNMT – *Suaeda liaotungensis* Phosphoethanolamine N-methyltransferase A0N067  
 SoNMT – *Spinacia oleracea* phosphoethanolamine N-methyltransferase Q9M571  
 TaNMT1 - *Triticum aestivum* Phosphoethanolamine N-methyltransferase Q8VYX1  
 TaNMT2 - *Triticum aestivum* Phosphoethanolamine N-methyltransferase 2 C8YTM5  
 TgUNK – *Toxoplasma gondii* Unknown protein Q1JT77  
 TnPMT – *Tetraodon nigroviridis* Putative methyltransferase Q4RPF1  
 VvNMT1 – *Vitis vinifera* putative methyltransferase 1 A7PH63  
 VvNMT2 - *Vitis vinifera* putative methyltransferase 2 A5ANL8  
 XIPMT – *Xenopus laevis* Putative methyltransferase Q6DCC9  
 XtPMT – *Xenopus tropicalis* Putative methyltransferase A4IHG2  
 ZmNMT1 – *Zea mays* Phosphoethanolamine N-methyltransferase 1 A7XZC6  
 ZmNMT2 – *Zea mays* Phosphoethanolamine N-methyltransferase 2 Q5SDQ0

---

non-plant NMT sequences were selected (underlined in Table 1) that contain only one catalytic domain with a high degree of amino acid sequence similarity to only one of the domains of Arabidopsis NMT1 (either SBD1 or SBD2). Each of the plant NMT sequences and 4 of the 13 non-plant proteins consist of two catalytic domains annotated as SAM binding domains (SBD). Tree construction proceeded with the use of ClustalX and PHYLIP as above.

### **Homology modelling of Arabidopsis NMTs**

Potential templates were identified using the ExPASy/SIB blast program by querying against the PDB (protein data bank) database ([www.rcsb.org](http://www.rcsb.org)). The three Arabidopsis NMT amino acid sequences were searched. The PDB sequences identified through the blast search were downloaded from the PDB and uploaded to Deepview/Swiss-PDB viewer<sup>®</sup> (available for download at [www.expasy.org/spdbv](http://www.expasy.org/spdbv)). The primary amino acid sequence to be modelled was then uploaded to Deepview and threaded through the template using the MagicFit tool. The two sequences, now aligned, were manually adjusted to minimize negative interactions and high threading energy between the two sequences. The adjusted alignment was uploaded to the Swiss model server and a PDB file containing the homology model was created (Arnold *et al.*, 2006).

### **Analysis of Arabidopsis NMT gene expression**

Genevestigator version 3 (Hruz *et al.*, 2008) was used to perform a meta-profile analysis of the expression of all three *NMT* genes. A total of 4070 high quality (22k) microarrays were analyzed for expression of the three genes.

*Predicting secondary structure of RNA*

Full length mRNA sequences were obtained for At3g18000, At1g48600 and At1g73600 from The Arabidopsis Information Resource (TAIR, <http://www.arabidopsis.org/>). Secondary structure prediction was performed using mfold version 3.2 (Zuker, 2003).

*Predicting transcription factor binding sites*

Potential transcription factor binding sites (TFBS) were identified using three separate programs: MATCH, PATCH and ATHAMAP (Kel *et al.*, 2001 ; Steffens *et al.*, 2005). The sequences used to query the programs were the full genomic sequences (including introns) of At3g18000, At1g48600 and At1g73600 and 1000 bp up and downstream of each gene. The match database (Kel *et al.*, 2001) was queried for potential TFBS annotated in plants within the TRANSFAC database. The same sequences were used to search the PATCH database ([www.gene-regulation.com/pub/](http://www.gene-regulation.com/pub/)). The minimum length of detected binding sites was set to 5 while the maximum number of mismatches allowed per site was set to 1. The lowest pattern based score allowed was 87.5. The gene analysis mode was used to deduce potential TFBS with Athamap (Steffens *et al.*, 2005). The scan was performed using a 50% restriction to highly conserved binding sites.

## RESULTS

### Sequence and structure analysis of NMT proteins

#### *Phylogenetic analysis of NMT-like proteins*

In order to develop a phylogenetic tree of NMT-like proteins, only sequences greater than 40% identical to Arabidopsis NMT1 at the amino acid level were chosen. This criterion identified the 37 protein sequences shown in Table 1. For analysis of the full protein sequences, a subset comprised of 23 of the 37 identified proteins was chosen (See Materials and methods). In addition to plant sequences, several animal proteins that are known to methylate only at the Ptd level were used in the analysis.

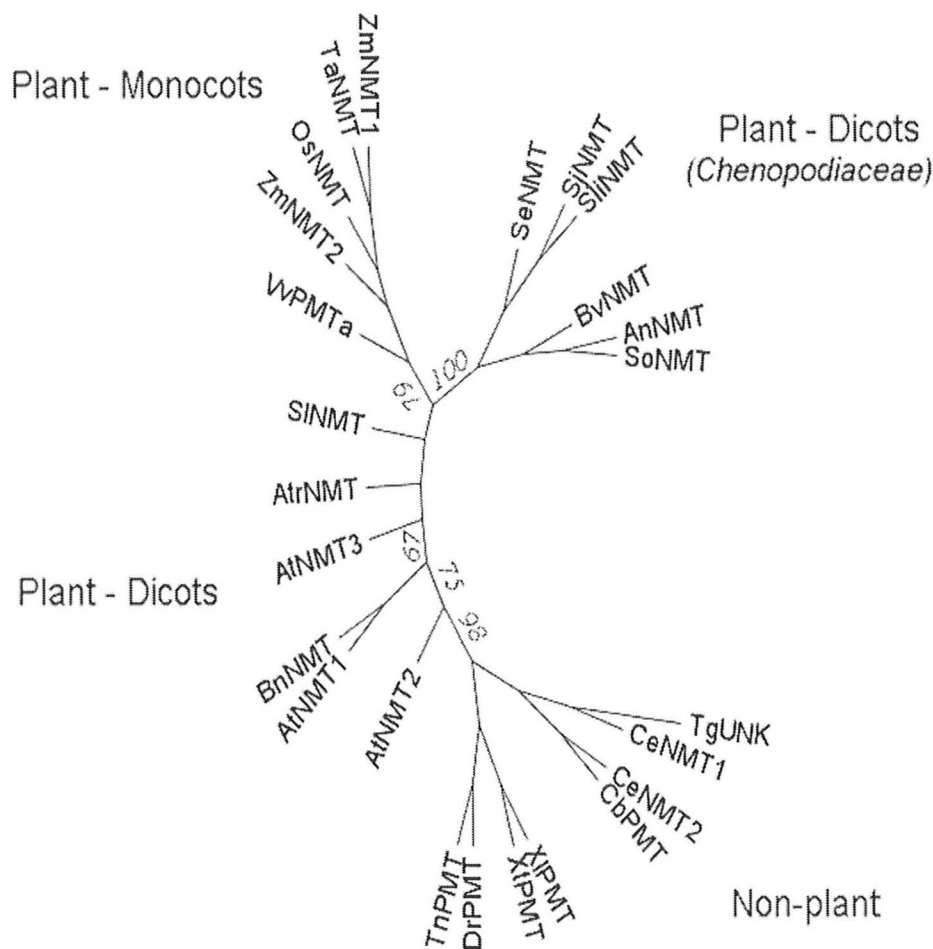
The phylogenetic tree of complete NMT and related protein sequences is shown in Figure 2. The animal NMT sequences XIPMT, XtPMT, DrPMT and TnPMT together form a group found in 98% of bootstrapped trees. Two well supported monophyletic groups are formed that represent monocot and chenopod NMT sequences. Arabidopsis NMT3 is grouped closer to these monophyletic groups than NMT1 or NMT2 in 67% of bootstrapped trees. A notable outcome of this analysis is that 75% of bootstrapped trees contain a monophyletic group containing all plant sequences except Arabidopsis NMT2.

The majority of the 37 sequences chosen for the phylogenetic analysis contain two domains that are predicted to interact with SAM based upon their consensus with amino acid sequences of conserved SAM binding motifs. As such, individual domains were also separated and analyzed as distinct segments. For this analysis, full protein sequences were split into two halves, an N-terminal half and a C-terminal half (Fig. 3 and 4, Panel A). CLUSTAL X alignments of the N and C-terminal halves of the 37 protein sequences

Figure 2 – Consensus phylogenetic tree constructed from full length amino acid sequences of 23 known/putative SAM-dependent methyltransferases.

Numbers and branch lengths represent bootstrap values from 100 replicate analyses.

NMT – annotated phosphoethanolamine *N*-methyltransferase. PMT – putative SAM-dependent *N*-methyltransferase, UNK – Unknown protein function.





were created and used to generate a phylogenetic tree for each functional domain with PHYLIP (Felsenstein, 1993). The nine proteins that contain only one SBD were included in this alignment and were matched with either the N or C-terminal half sequences based on their greatest identity to either half of NMT1. Of the nine proteins, only CeNMT1 has greater amino acid identity to SBD1 of NMT1 while the other eight single domain proteins share greater amino acid identity to SBD2 of NMT1.

Results of the analysis of the N-terminal half amino acid sequences are shown in Figure 3. This tree resembles Figure 2 with non-plant NMT sequences forming a well supported monophyletic group as is the case for monocot and chenopod sequences. Proteins from dicot, non-chenopod species, that includes the Arabidopsis NMTs, do not form a group together and instead form their own clades. Another similarity between phylogenies for the full length and N-terminal half sequences is that >70% of bootstrapped trees have a monophyletic group containing all plant sequences except Arabidopsis NMT2.

The phylogenetic tree based upon comparisons of the C-terminal half sequences is shown in Figure 4 and this tree shares many of the same branches as those given in Figures 2 and 3. The sequences of non-plant NMTs again cluster away from plant NMTs forming two well supported clades with one composed of eight proteins that contain only a single SBD. PfNMT and PvNMT are reported to methylate P-EA, P-MEA and P-DEA and are part of this group of eight proteins (Pessi *et al.*, 2004). The second clade of non-plant NMTs contains the C-terminal halves of non-plant proteins that show the bipartite domain structure common to plant NMT protein sequences.

Figure 3 - Phylogenetic tree constructed from sequences of the N-terminal halves of 30 similar SAM-dependent methyltransferases.

A, SBD organization in plant NMT proteins. Dark grey regions represent region of protein encompassing SBD1/2. The portion corresponding to the N-terminal half of the full-length protein is labelled. B, Numbers and branch lengths represent the bootstrap analysis of 1000 trials. Only bootstrap values referred to in the results are reported here. NMT – annotated phosphoethanolamine *N*-methyltransferase. PMT – putative SAM-dependent *N*-methyltransferase.

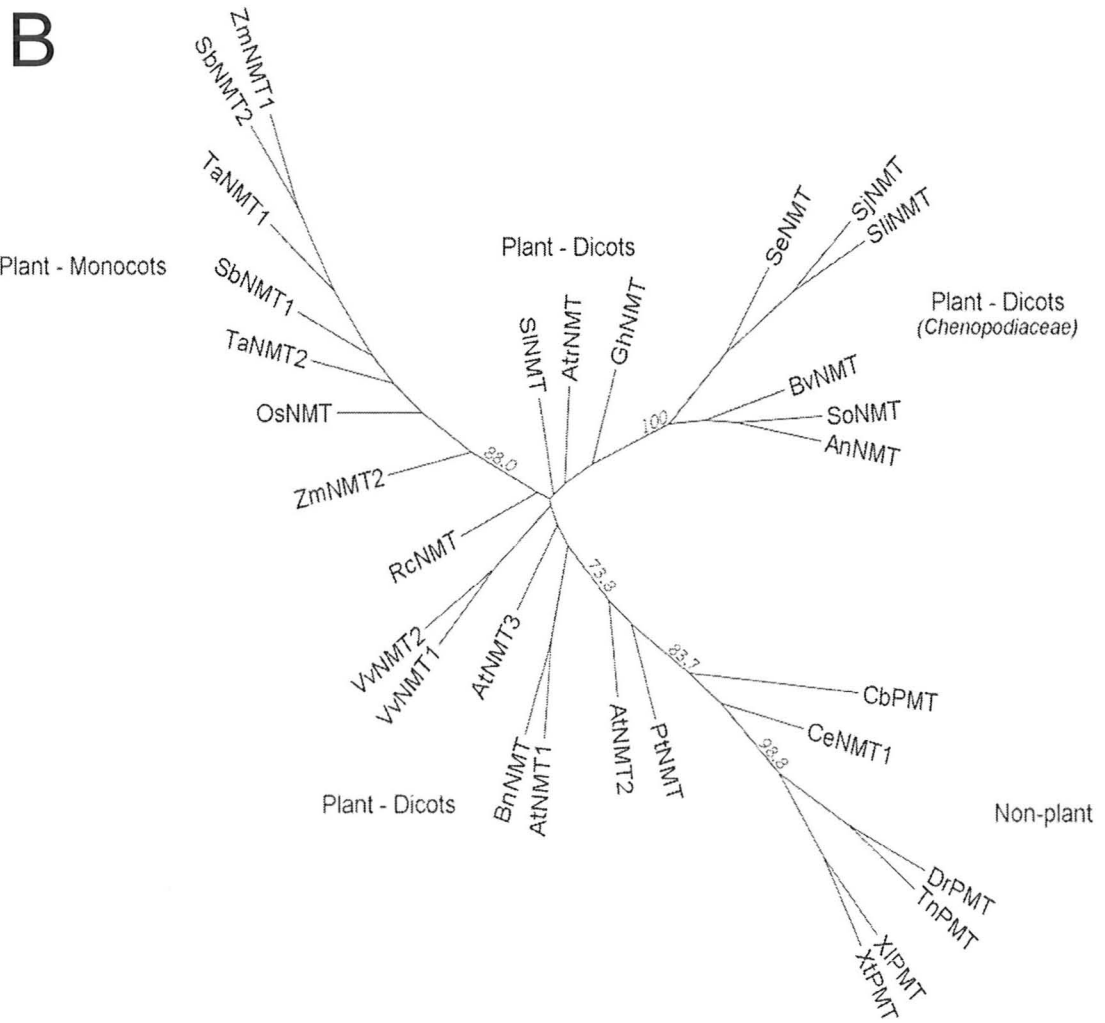
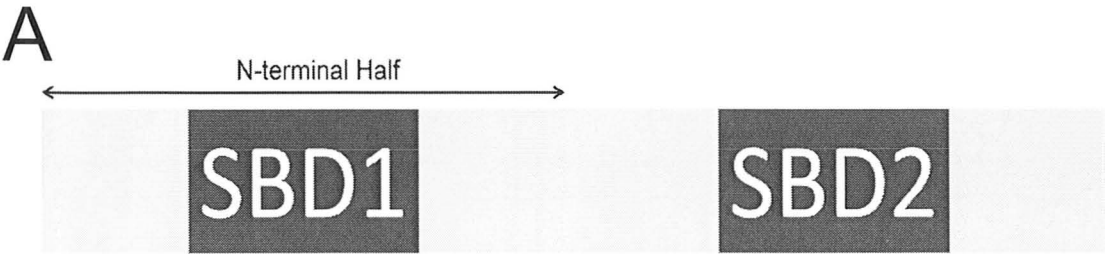


Figure 4 - Phylogenetic tree constructed from sequences of the C-terminal halves of 36 similar SAM-dependent methyltransferases.

A, SBD organization in plant NMT proteins. Dark grey regions represent region of protein encompassing SBD1/2. The portion corresponding to the C-terminal half of the full-length protein used in the phylogenetic analysis is labelled. B, Numbers and branch lengths represent the bootstrap analysis of 1000 trials. Only bootstrap values referred to in the results are reported here. NMT – annotated phosphoethanolamine *N*-methyltransferase. PMT – putative SAM-dependent *N*-methyltransferase. UNK – unknown protein function.

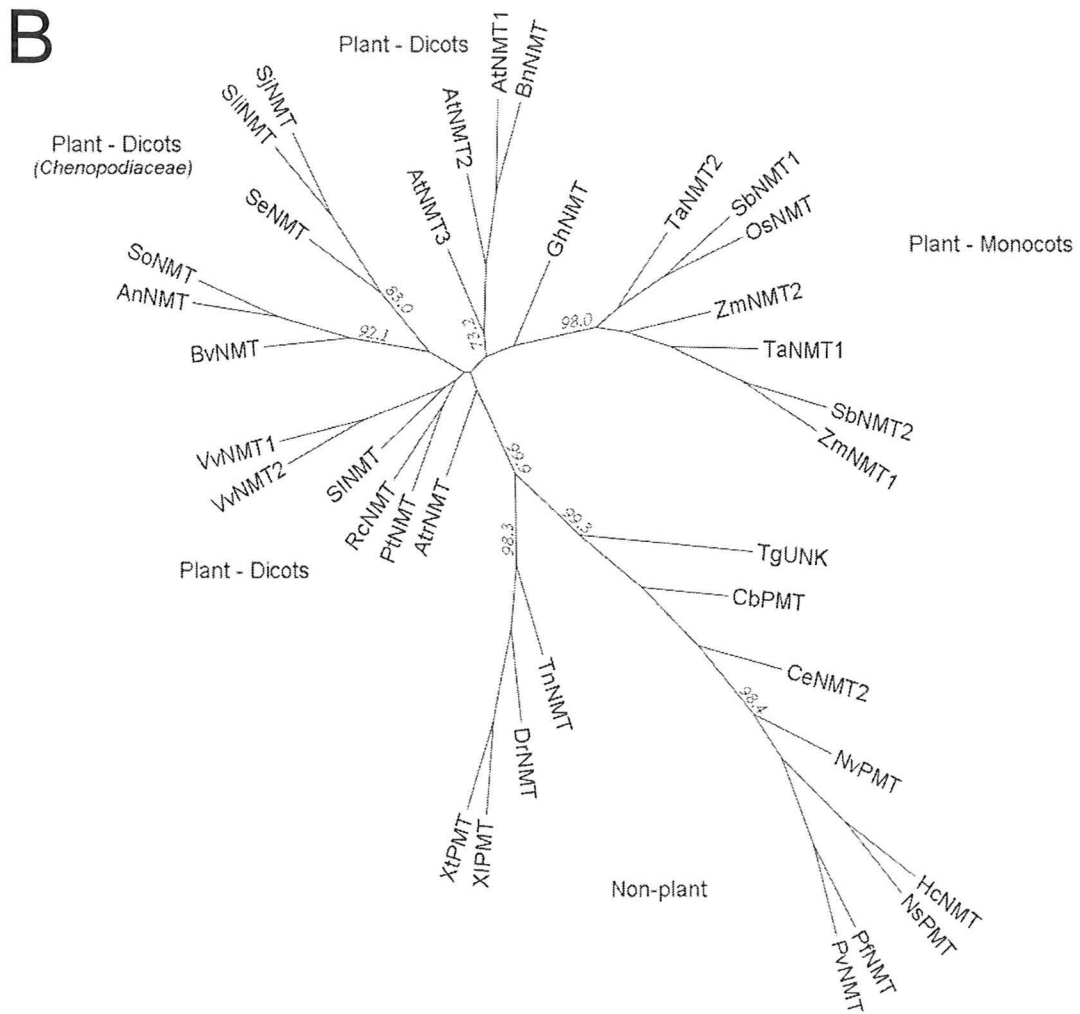


Figure 4 also shows that NMT sequences from monocot plants and dicot chenopod species again form their own distinct clades. Some sequences from dicot, non-chenopod species (including poplar and grape) also form their own branch. In contrast to these similarities, *Brassica* species (including all three *Arabidopsis* NMTs) form a single clade in >70% of bootstrapped trees produced using the C-terminal half NMT sequences (Fig. 4).

#### *Detailed alignment analysis of NMT amino acid sequences*

The amino acid sequence for *Arabidopsis* NMT3 was aligned with the best characterized plant P-EAMeT proteins: *Arabidopsis* proteins NMT1 and NMT2, spinach P-EAMeT (SoNMT) and wheat TaPEAMT1 (TaNMT) (Fig. 5). SoNMT and TaNMT were chosen for this comparison because their capacity to methylate specific phosphobases is reported for both the full length protein and separate domains (Nuccio *et al.*, 2000; Charron *et al.*, 2002). TaNMT and SoNMT should also represent phylogenetically distant sequences as wheat is a monocot and spinach is a dicot member of the family Chenopodiaceae, a group that forms its own clade away from branches formed by other dicot species including *Arabidopsis* (See Fig. 3 and 4). The alignment of the full amino acid sequence of these five NMTs shows that *Arabidopsis* NMT2 lacks approximately 15 amino acids at the N-terminus when compared to the other four proteins (Fig. 5). Also, the N-terminus of TaNMT does not show the same sequence conservation shared by *Arabidopsis* NMTs 1 and 3 and SoNMT. TaNMT is also at least four amino acids longer than all of the other aligned proteins.

Figure 5 – Deduced amino acid sequence alignment of full length Arabidopsis, spinach and wheat NMT proteins.

Conserved amino acids are highlighted in black with semi-conserved residues in gray.

The sequences spanning the SAM-binding motifs of SBD1 and SBD2 are indicated. At – *Arabidopsis thaliana*, So – *Spinacia oleracea*, Ta – *Triticum aestivum*.

At NMT1 - - - - - MAASYEEERDIQKNYWI EHSADLTVEAMMLDSRASDL DKEERP  
 At NMT2 - - - - - MAS YGEEREI QKNYWK EHSVGLSVEAMMLDSKASDL DKEERP  
 At NMT3 - - - - - MAASYGEEREI QKNYWK EHSVGLSVEAMMLDSKASDL DKEERP  
 SoNMT - - - - - MAASAMGVLOEREVFKKYWI EHSVDLTVEAMMLDSQASDL DKEERP  
 Ta NMT MDTI TVVENVFGEVERKVQKS YWE EHSKDLTVESMMLDSRAKDL DKEERP

Motif I Post I

At NMT1 EVLSL LPPYEGKSVLELGAGI GRFTGELAQKAGE LI ALDFI DNVI KKNES  
 At NMT2 EVLSL LPPYEGKSVLELGAGI GRFTGELAQKAGEVI ALDFI ES AI QKNES  
 At NMT3 E L LAF LPP I EGT TVLEF GAGI GRFT T ELAQKAGQVI AVDFI ES VI KKNEN  
 SoNMT EVLSMLPPYEGKSVLELGAGI GRFTGELAEKASQVI ALDFI ES VI KKNES  
 Ta NMT EVL AI LPSYAGKT VLELGAGI GRFTGELAKE AGHVI ALDFI DSVI KKNEE

Motif II

At NMT1 I N- GHYKNVKFMCADVTSPDLKI TDGSLDLI FSNWLLMYLSDK EVEL LAE  
 At NMT2 VN- GHYKN K FMCADVTSPDLKI KDGS I DLI FSNWLLMYLSDK EVEL MAE  
 At NMT3 I N- GHYKNVKF L CADVTSPNMF PNE SMDLI FSNWLLMYLS DQ EVEL LAK  
 SoNMT I N- GHYKNVKFMCADVTSPS L NI SPNSVDI I FSNWLLMYLS DEVEVERLVE  
 Ta NMT I NGDI YKNI T FMCADVTSP ELKI EDNSVDI V FSNWLLMYLNDE EVEKLI G

Motif III

At NMT1 RMV GWI KVGGYI FFRES CFHQSGDSKRKSNPTHYREPRFY SKVFQECQTR  
 At NMT2 RM L GWV KPGGYI FFRES CFHQSGDSKRKSNPTHYREPRFY TKVFQECQTR  
 At NMT3 KML QMT KVGGYI FFRES CFHQSGDNKRKYNP THYREPKFYTK L FKECHMN  
 SoNMT RML KWL KPGGYI FFRES CFHQSGDHKRKSNPTHYREPRFYTK L FKECHMQ  
 Ta NMT RI V KWL KPGGHI F I RES CFHQSGDSKRKV NP THYREPRFYTKVFKECHSY

At NMT1 DAA GNSFELS M I GCKCI GAYVKNKKNNQNI CWL WQKVSS ENDR GFQRF LD  
 At NMT2 DAS GNSFELS M V GCKCI GAYVKNKKNNQNI CWL WQKV S V ENDK DFQRF LD  
 At NMT3 DE D GNSYELSLV S CKCI GAYVRNKKNNQNI CWL WQKVSS D NDR GFQRF LD  
 SoNMT DDS GNSYELSL I GCKCI GAYVKS KKNQNI S WL WQKV D S ED DKGQRF LD  
 Ta NMT DQE GNSFELS LVT S KCI GAYVKS KKNQNI CWL WEKV KCTE DKGQRF LD

Motif I

At NMT1 NVQYKSSGI LRYERVFGQGF VSTGGL ETTKEFVE KMN L KPGQKVL DVGCG  
 At NMT2 NVQYKSSGI LRYERVFGEGYVSTGG F ETTKEFVAKMDL KPGQKVL DVGCG  
 At NMT3 NVQYKSSGI LRYERVFGEG F VSTGGL ETTKEFVDM L KPGQKVL DVGCG  
 SoNMT SS QYKFNS I LRYERVFG P GYVSTGGL ETTKEFVSKL DL KPGQKVL DVGCG  
 Ta NMT NVQYKSTGI LRYERVFGEGYVSTGG F ETTKEFVDKL DL KAGQKVL DVGCG

Post I

At NMT1 I GGGDFYMAEF DVHVVGI DLSVNM I SFALERAI GL S CSVEFEVADCTTK  
 At NMT2 I GGGDFYMAENF DVHVVGI DLSVNM I SFALERAI GLKCSVEFEVADCTTK  
 At NMT3 I GGGDFYMAENF DV DVVGI DLSVNM I SFALE HAI GLKCSVEFEVADCTTK  
 SoNMT I GGGDFYMAENY DV E VVGI DLSI NM I SFALE RSI GLKCAVEFEVADCTTK  
 Ta NMT I GGGDFYMAET Y DVHVLGI DLSI NM V SFAI ERAI GRS CSVEFEVADCTTK

Motif II Motif III

At NMT1 HYPDNSFDVI YSRDTI LHI QDKPALFR T FFKWL KPGGKVL I SDYCRSPKT  
 At NMT2 TYPDNSFDVI YSRDTI LHI QDKPALFR T FFKWL KPGGKVL I T DYCRS AET  
 At NMT3 EYPDN T F DVI YSRDTI LHI QDKPALFR R FFKWL KPGGKVL I T DYCRSPKT  
 SoNMT DYPNSFDVI YSRDTI LHI QDKPALFR S FFKWL KPGGKVL I SDYCKSAGT  
 Ta NMT EYAEN T F DVI YSRDTI LHI QDKPALFR N FFKWL KPGGKVL I SDYCRSPGT

At NMT1 PS AEFS E YI KQRGYDL HDVQAYGQML KDAGF T DVI AE DRTDQF MQVL KRE  
 At NMT2 PS PEF AE YI KQRGYDL HDVQAYGQML KDAGF DDVI AE DRTDQF VOVL RRE  
 At NMT3 PSP D F AI YI KKRGYDL HDVQAYGQML RDAGF E EVI AE DRTDQF MKVL KRE  
 SoNMT PS AEFAAYI RQRGYDL HDV KAYGKML KDAGF V EVI AE NRTDQF I QVL QKE  
 Ta NMT PS E EFAAYI KQRGYDL HDV KTYGKML E DAGF H D V V AE DRTDQF L RVL ERE

At NMT1 L DRVEKEKEKEFI SDFSKE DYDDI VGGWKS KL ERCA S DEQKWGL FI ANKN  
 At NMT2 L E KVEKEKEEF I SDFS E E DYNDI VGGWS AKL ERTAS GEQKWGL FI ADKK  
 At NMT3 L DAVEKEKEEF I SDFSKE DYEDI I GGWKS KL LRSS S GEQKWGL FI AKRN  
 SoNMT L DALEQEKDDFI DDFS E E DYNDI V DGWAKAL VRTTE GEQKWGL FI AKKM  
 Ta NMT L GETEKNKEAFLA DFTQE DYDDI VNGWS AKL KRSS A GEQKWGL FI ATK-



Alignments of the 110 amino acids that are associated with SBD 1 and 2 shows that amino acid identity in the SBD regions of all five proteins is highly conserved (Fig. 6). Only TaNMT has a single aspartic acid residue (D; designated by an arrow in Fig. 6A) in SBD1 that is not found in the other four sequences. A visual comparison of Panels A and B of Figure 6 shows that the amino acid sequences associated with SBD2 are more highly conserved across the five NMTs than SBD1.

*Comparison of tertiary protein structures predicted by homology modelling*

Homology models of the N and C-terminal halves of the Arabidopsis NMTs as well as SoNMT and TaNMT were constructed to determine if small differences in primary amino acid sequence (shown in Fig. 5 and 6) would result in changes in 3D protein structure. The homology models of the N-terminal half and the C-terminal half of all five NMT proteins were modelled based upon templates taken from small, bacterial SAM-dependent methyltransferases from *Leishmania major* and *Pyrococcus horikoshii* (PDBIDs: 1XTP and 1VE3). The selection of these structural templates was necessitated by the lack of X-ray crystal structures of suitable plant methyltransferases. Both X-ray crystal structures are of the enzymes in a conformation where the enzyme is bound to either SAM or SAH, representing a substrate and product, respectively, that compete for the same active site. The bacterial NMTs share 32 and 36% amino acid sequence identity, respectively, to Arabidopsis NMT1. Identity of the bacterial NMT's to Arabidopsis NMT2 and 3, SoNMT and TaNMT all fell within  $\pm 2\%$  of that found for Arabidopsis NMT1.

A composite of the models for the N-terminal halves of the five NMT proteins is shown in Figure 7A. These 3D protein ribbon models show that the predicted 3D structures of the five proteins are very similar. A more detailed view of the N-terminal half of Arabidopsis NMT1 binding SAH is shown in Figure 7B. In this view the structure surrounding the SAH binding cleft is more clearly visible. The  $\alpha$ -carbon backbones of the N-terminal halves of Arabidopsis NMT proteins are shown in Figure 8. NMT2 is clearly divergent from Arabidopsis NMT1 and NMT3 in the structure predicted for the N-terminus (feature designated by an arrow in Fig. 8). Closer inspection of the ribbon models of the N-terminal halves of NMT1 and NMT2 shows that NMT2 is missing an alpha helix (Fig. 9). This divergence in the modelled 3D structure is a result of the missing 15 amino acid sequence identified in the alignment of NMT2 with other NMTs shown in Figure 5. The position of the alpha helix is interesting in that this structure appears to fold down over the SAM binding cleft. The proximity of this alpha helix to the SAM binding cleft suggests a possible interaction between this domain and SAM.

Figure 10A shows the results of modelling the C-terminal halves of Arabidopsis NMTs 1, 2 and 3, SoNMT and TaNMT. A visual comparison of the overall predicted protein structures indicates no major structural differences as was the case for the alpha helix found in the models of the N-terminal halves. For all of the C-terminal halves of NMTs modelled, SAM binds to the N-terminal half portion of a seven-stranded beta sheet (Fig. 10). The modelled structure of the C-terminal half of Arabidopsis NMT1 is shown in greater detail in Figure 10B where SAM appears to bind between the first and second beta strands (indicated by an arrow).

Figure 6 – Alignment of deduced amino acid sequences for *S*-Adenosyl methionine (SAM) binding domains of Arabidopsis, spinach and wheat NMT proteins.

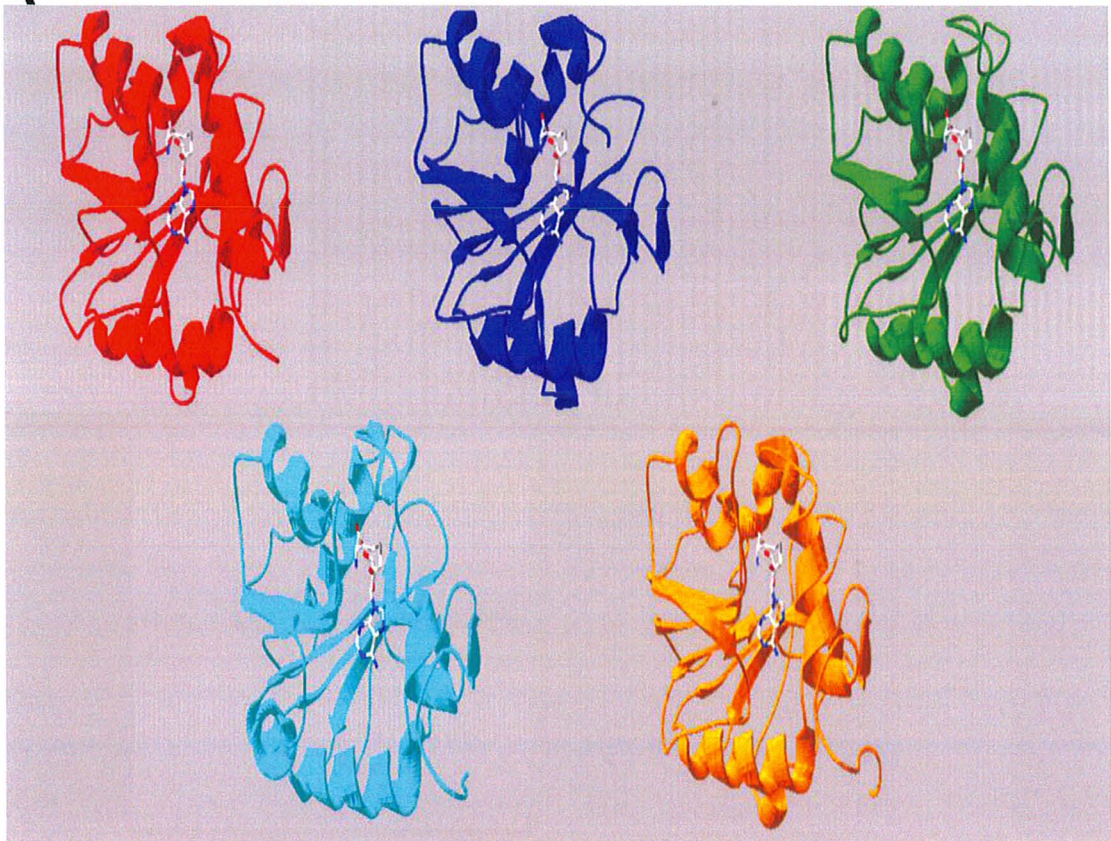
A, alignment for SBD1. B, alignment for SBD2. \* Denotes 100% conservation at a specific site. Arrow indicates aspartic acid residue unique to TaNMT. At – *Arabidopsis thaliana*, So – *Spinacia oleracea*, Ta – *Triticum aestivum*.



Figure 7 – Protein modelling of five plant NMT proteins binding SAH.

A, Protein ribbon diagrams of modelled 3D structures of the N-terminal half of AtNMT1 (red), AtNMT2 (navy), AtNMT3 (green), SoNMT (cyan) and TaNMT (orange) proteins showing the binding cleft conformation incorporating SAH. B, protein ribbon diagram of Arabidopsis NMT1 SAH binding cleft. Diagram is coloured according to secondary structure: helices (red), sheets (violet), coils (grey)

A



B

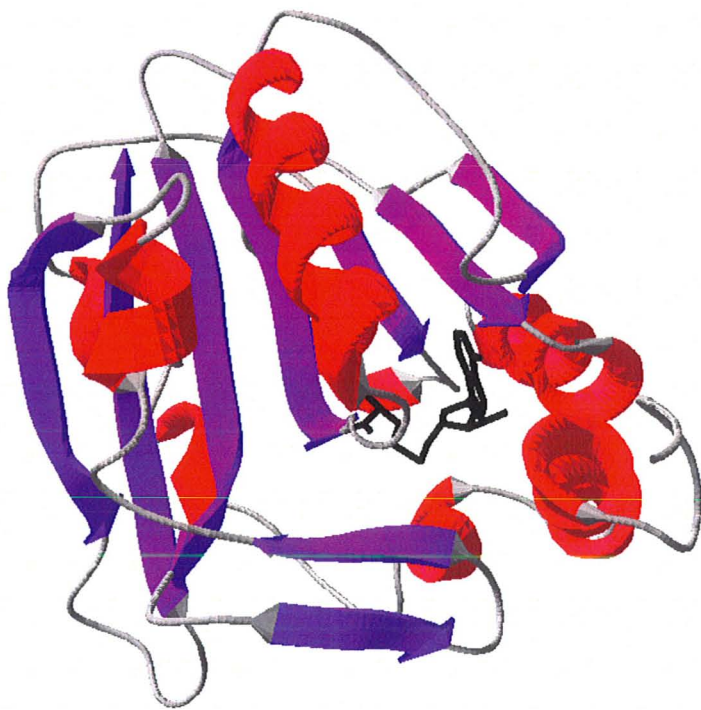


Figure 8 –  $\alpha$ -Carbon backbones of Arabidopsis NMT protein models

Overlap of  $\alpha$ -carbon backbones of N-terminal half homology protein models for AtNMT1 (red), AtNMT2 (blue) and AtNMT3 (green). Arrow indicates position of N-terminal half 15 amino acid extension that is present in AtNMT1 and AtNMT3 but absent from AtNMT2.



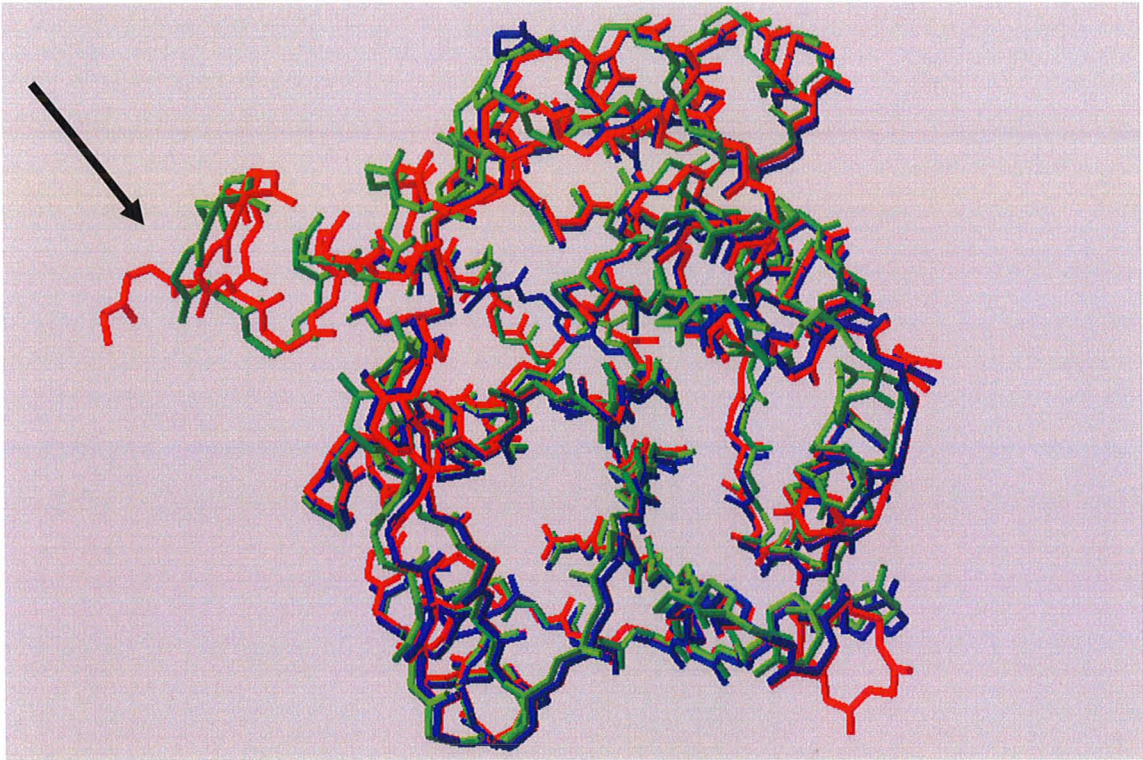
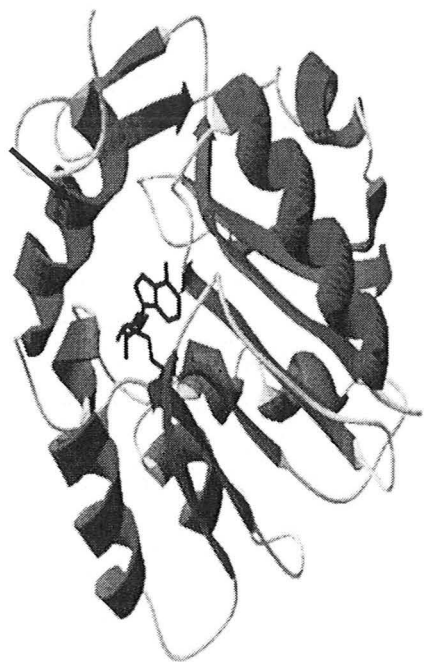




Figure 9 – Comparison of Arabidopsis NMT protein models binding SAH

Model structures of the N-terminal half of AtNMT1 (left) and AtNMT2 (right). The models show the binding cleft formation incorporating SAH. The arrow designates the position of an N-terminal alpha helix that is present in AtNMT1 but absent in AtNMT2.



**AtNMT1**

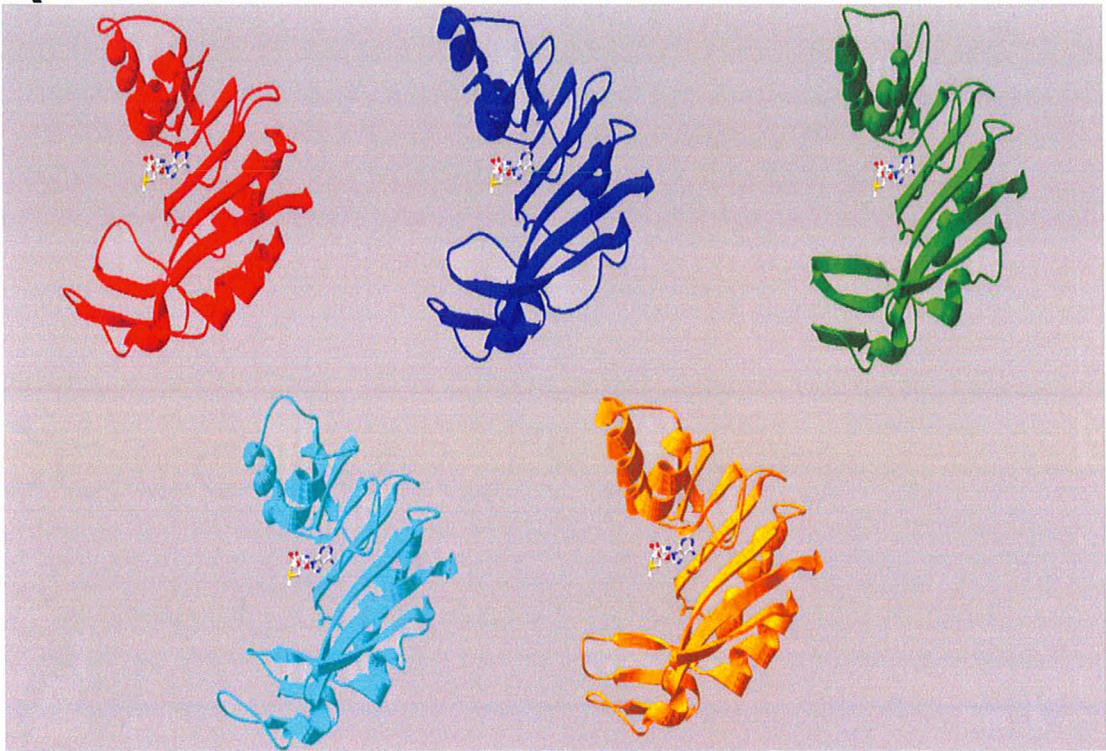


**AtNMT2**

Figure 10 – Protein modelling of five plant NMT proteins binding SAM

A, Protein ribbon diagrams of modelled 3D structures of the C-terminal half of AtNMT1 (red), AtNMT2 (navy), AtNMT3 (green), SoNMT (cyan) and TaNMT (orange) proteins showing the binding cleft conformation incorporating SAM. B, protein ribbon diagram of Arabidopsis NMT1 SAM binding cleft. Diagram is coloured according to secondary structure: helices (red), sheets (violet), coils (grey).

A



B



## Organization and Expression of Genes encoding Arabidopsis NMTs

### *Transcription Factor Binding Site Prediction*

A list of potential transcription factor binding sites (TFBS) within 1000 bp upstream of the translation start codon was compiled (see Materials and Methods) and the findings are summarized in Table 2. The GT1/2 family of transcription factors have several binding sites associated with all three Arabidopsis *NMT* sequences at approximately 200-400 bp upstream of the start site codon. GT1/2 transcription factors have been shown to confer light-inducible gene expression in Arabidopsis (Ni *et al.*, 1996). Similarly, multiple TFBS for the transcription factor AG were found in all *NMT* 5' regions. However, AG TFBS do not map to a single region like the other TFBS associated with these three genes. AG is implicated in the regulation of the flowering and reproduction pathways in Arabidopsis (Huang *et al.*, 1995). TFBS for the transcription factors CBF1/2/3, DREB1a/2a were detected upstream of *NMT1* and *NMT2* and co-localize approximately 500 bp upstream of the first exon of both genes. The CBF and DREB family of transcription factors are known to play a role in regulating response to cold and desiccation, respectively (Wu *et al.*, 2007). TFBS for the factor WRKY18 were detected in the 5' regions upstream of *NMT1* and *NMT3* but not *NMT2*. WRKY18 is reported to play a role in plant hypersensitive response to pathogens and its expression is induced by salicylic acid (Chen *et al.*, 2002). RAV is another pathogen-inducible transcription factor (Sohn *et al.*, 2006) and a RAV TFBS was found only upstream of the gene encoding *NMT1*.

Table 2 – Potential transcription factor binding sites upstream of Arabidopsis *NMTs*.

Analysis included nucleotides found in sequence 1000 bp upstream of transcription start sites of Arabidopsis *NMT1*, *NMT2* and *NMT3*.

Transcription factor	Predicted binding site	Position (strand)	Significance
WRKY18	<i>NMT1,NMT3</i>	Multiple	Role in SA dependent signalling and regulation; response to pathogen invasion (Chen <i>et al.</i> , 2002)
CBF1/2/3	<i>NMT1,NMT2</i>	-400 to -500 (+)	Cold responsive element (Wu <i>et al.</i> , 2007)
DREB1a/2a	<i>NMT1,NMT2</i>	-400 to -500 (+)	Dessiccation responsive element (Wu <i>et al.</i> , 2007)
GT1/2	<i>NMT1,NMT2, NMT3</i>	-200 to -400 (+/-)	Light inducible transcriptional activator (Dehesh <i>et al.</i> , 1996)
AG	<i>NMT1,NMT2, NMT3</i>	Multiple	Regulator of flowering and reproduction (Huang <i>et al.</i> , 1995)
RAV	<i>NMT1</i>	-250 (+/-)	Role in SA dependent signalling and regulation; response to pathogen invasion (Sohn <i>et al.</i> , 2006)

*Expression of Arabidopsis NMT genes*

Multiplex PCR was used to confirm the presence of a T-DNA insertion in exon 7 of the *NMT3* gene (Fig. 11A). Transcript levels for *NMT3* were detected by RT-PCR to determine if the presence of a T-DNA insertion alters the transcription of *NMT3*. A 469 bp cDNA sequence was reverse transcribed from RNA isolated from leaves of WT plants and the SALK\_062703 line containing the T-DNA insertion in locus At1g73600 associated with *NMT3*. This 469 bp cDNA was used as a template for RT-PCR using the same primers to create a 469 bp RT-PCR amplicon. When using the WT cDNA as a template, the 469 bp product was amplified (Fig. 11B, Lane 4) but it was not amplified from cDNA prepared from SALK\_062703 plants found to be homozygous for the presence of a T-DNA insertion in *NMT3* (Fig. 11B, Lane 5). This outcome is evidence that the T-DNA insertion disrupts the transcription or stability of *NMT3* and plants homozygous for the presence of T-DNA have suppressed expression of *NMT3* relative to WT plants. As a result, this T-DNA mutant is referred to in the remainder of this thesis as *nmt3*. As a control for template quality, an Arabidopsis ubiquitin gene was used with the same source of isolated RNA to show that the absence of amplified transcripts associated with *nmt3* plants cannot be explained by RNA degradation during isolation (Fig. 11B, Lanes 2 and 3).

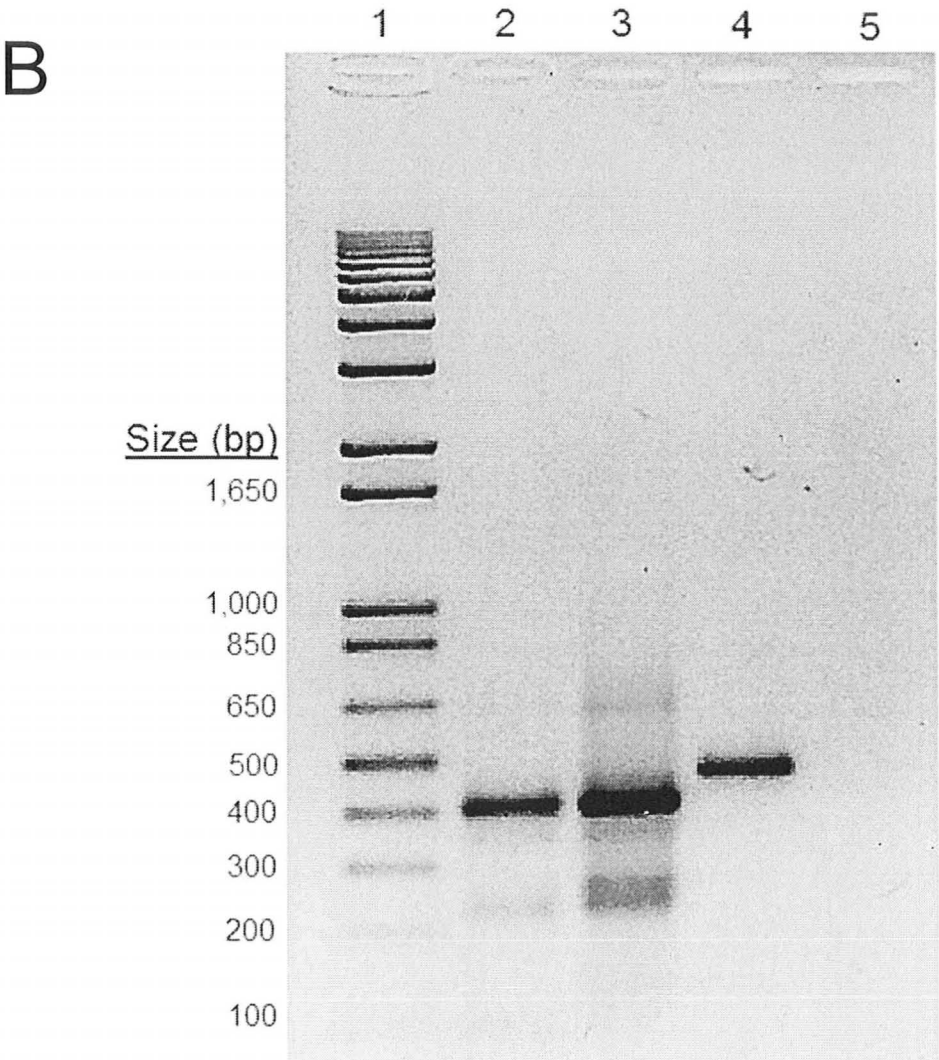
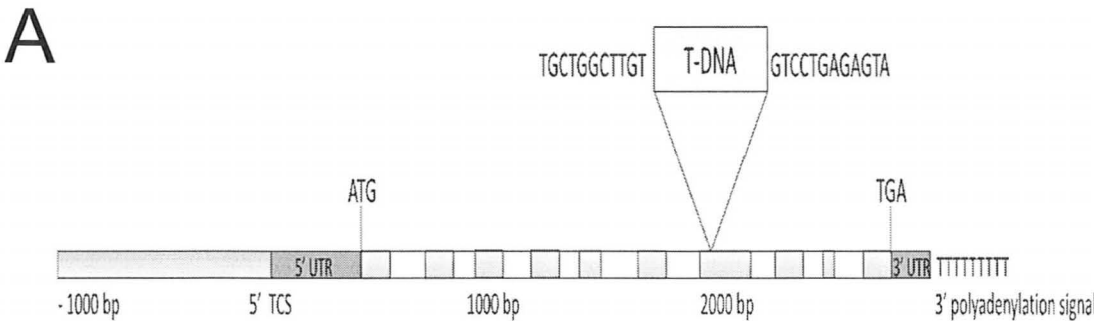
*Differential Expression of NMT Genes in Arabidopsis*

The average expression level of the Arabidopsis *NMT* genes was compiled from over 4070 high-quality microarray datasets (Genevestigator V 3.0 Hruz *et al.*, 2008).



Figure 11 - RT-PCR analysis of *NMT3* transcripts from WT and SALK\_062703 leaf RNA.

A, Map of *NMT3* gene structure showing the position of the T-DNA insert relative to introns (white) and exons (grey). TCS refers to the transcription start site. B, Molecular Size Markers (Lane 1), Steady-state transcript abundance associated with ubiquitin amplified from WT leaf cDNA (Lane 2) and from SALK\_062703 leaf cDNA (Lane 3). *NMT3* gene-specific product amplified from WT leaf cDNA (Lane 4) and its absence from cDNA produced from SALK\_062703 leaves found to be homozygous for T-DNA insertion in *NMT3* (Lane 5). See Materials and methods for details on PCR and electrophoresis conditions.



This meta-analysis shows that the *NMT* genes are differentially expressed in this plant (Fig. 12A). Transcripts associated with *NMT1* are more abundant in root tissues and the shoot apex while *NMT3* transcripts are more abundant in the leaves of the rosette and plants with an inflorescence. While this information suggests that *NMT1* and *NMT3* are differentially expressed in the plant, *NMT2* transcripts appear to be present constitutively in virtually all tissues sampled and, in many cases, at higher relative levels than those associated with *NMT1* and *NMT3* (Fig. 12A).

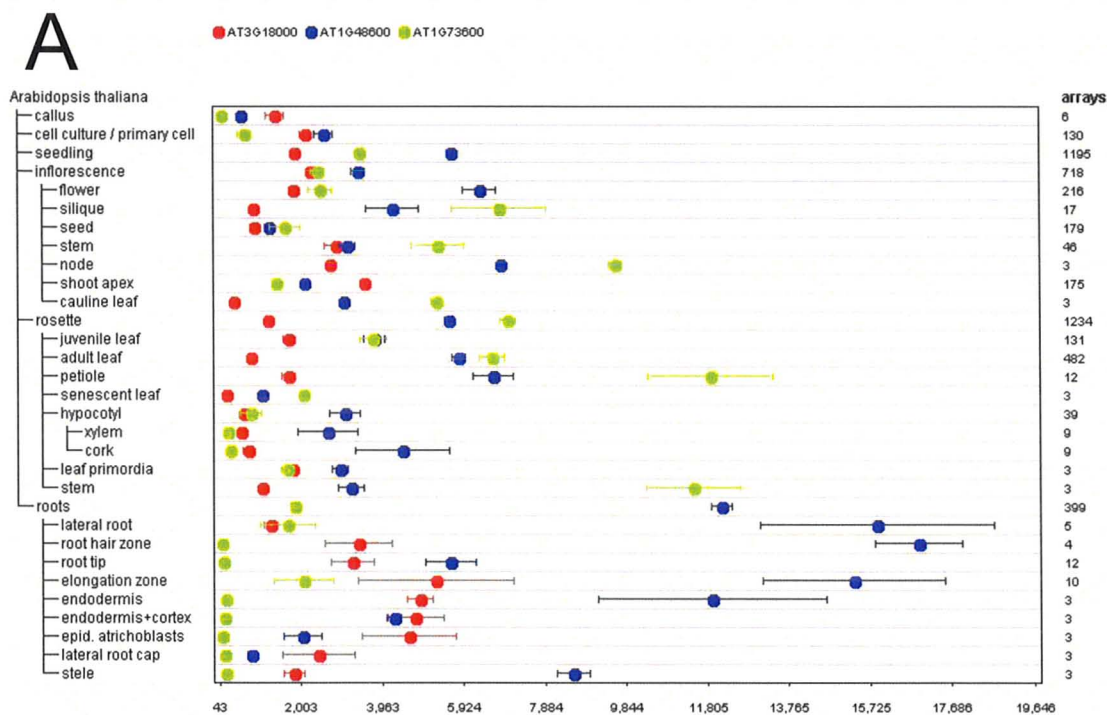
RT-PCR was performed on RNA extracts of leaf and root tissues from WT plants to determine whether transcripts associated with *NMT3* can be detected in these tissues. Both root and leaf extracts have some *NMT3* transcript as indicated by a 469 bp *NMT3* transcript as indicated by a 469 bp *NMT3* gene specific product (Lanes 2 and 3, Fig. 12B). However, the RT-PCR product abundance is comparatively less for the root RNA (Fig. 12B, Lane 2,) than leaf RNA (Lane 3). Both microarray analysis (Fig. 12A) and RT-PCR of leaf and root RNA (Fig. 12B) indicate that *NMT3* is differentially expressed in leaf and root tissues with higher steady state *NMT3* transcript abundance in leaves relative to *NMT1* and the converse is true for roots where *NMT1* transcript abundance is higher.

#### *Salt Stress Experiment*

WT and *nmt3* seedlings were grown on defined media plates to determine if *nmt3* plants are associated with a salt-sensitive phenotype. An example of a plate containing WT and *nmt3* plants on defined media supplemented with 50 mM NaCl is shown in Figure 13. The number of lateral roots at 10 d after germination and the average daily rate of root elongation are not significantly different between WT and *nmt3* plants growing on salt-supplemented media (Table 3). The addition of salt reduces root growth for both

Figure 12 – Expression of *NMT* genes in various developmental tissues of Arabidopsis.

A, *NMT* gene expression shown graphically generated with GENEVESTIGATOR version 3 (Laule *et al.*, 2008). Red dots refer to *NMT1* at locus At3g18000, blue dots refer to *NMT2* at locus At1g48600 and green dots refer to *NMT3* at locus At1g73600. Units on x-axis represent calculated (normal) gene expression values. Datapoints shown  $\pm$  SD with  $n = 4070$  microarrays. B, RT-PCR showing *NMT3* transcript abundance in RNA from root and leaf tissue of WT plants. Molecular Size Markers (Lane 1), *NMT3* (gene-specific) product amplified from WT root RNA (Lane 2) and leaf RNA (Lane 3).



**B**



Figure 13 – Growth of *Arabidopsis* seedlings on defined media plates in response to NaCl.

Three representative 10 day-old WT (left) and *nmt3* (right) seedlings shown growing on a culture plate supplemented with 50 mM NaCl.



Table 3 – Response of WT and *nmt3* Arabidopsis to NaCl exposure on defined media.

Root elongation rate and number of lateral roots in WT and *nmt3* plants grown on culture plates with or without 50 mM NaCl grown for 10 d post germination. Data represents the mean of 6 plants grown on 4 separate plates +/- SE (n = 24).



	WT		<i>nmt3</i>	
	No NaCl	50 mM NaCl	No NaCl	50 mM NaCl
Average root elongation rate (mm)	5.06 +/- 2.25	1.34 +/- 0.904	5.23 +/- 2.01	1.19 +/- 0.942
Number of lateral roots	14.0 +/- 0.505	5.62 +/- 0.992	14.3 +/- 1.96	6.33 +/- 1.44

lines suggesting that at least in this experimental approach, *nmt3* plants are not more salt-sensitive than WT.

#### *Early flowering mutant phenotype*

A visual comparison between WT and *nmt3* plants indicates that an early flowering mutant (EFM) phenotype is associated with the suppressed expression of *NMT3* in *nmt3* plants (Fig. 14). This phenotype is associated with bolting 5 to 7 days before WT plants and it appeared in 100% of *nmt3* plants that were grown together as a first generation (T1) in a single flat (16 plants) along with WT plants present for direct comparison. This phenotype was noted again among single seed descendents of these mutants but the frequency of this phenotype dropped as early as the next generation to only one or two plants out of a flat of 16 *nmt3* plants (Fig. 14E). Rosette leaves of early flowering *nmt3* mutants are stunted relative to those of WT plants (14B). This difference in leaf morphology and size is likely brought on by slowed vegetative growth accompanying the early transition to flowering (Fig. 14A, C and D). Also, EFM plants often develop more than one inflorescence originating from the base of the rosette as opposed to one large inflorescence with many branches originating from a single node as is often seen in flowering WT plants (Fig. 14C and D). The EFM phenotype was originally observed in *nmt3* plants grown under 12 h day-length conditions and this phenotype persisted in plants grown under both 8 and 16 h day-length regimes. WT and *nmt3* plants were also grown to maturity at 26°C and compared to WT and *nmt3* plants grown at 21°C. The EFM phenotype was found in the heat-stressed mutant plants at the

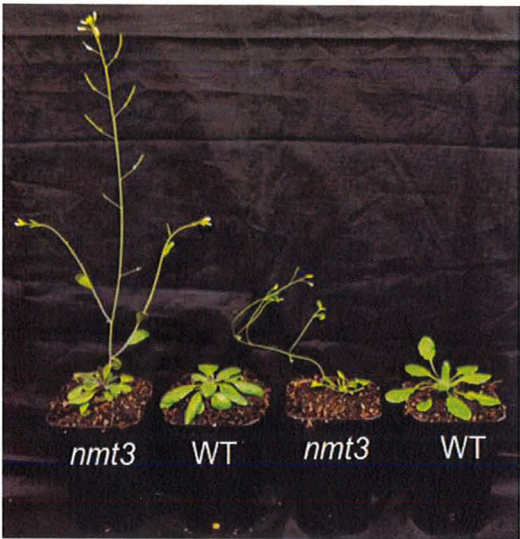
Figure 14 – Developmental phenotype of early flowering *nmt3* plants.

A, *nmt3* and WT plants 14 days post-germination. The *nmt3* plant displays the EFM phenotype seen in some but not all *nmt3* plants. B, WT and *nmt3* plants 28 days post-germination. Two plants (left) were grown at 21° C and two plants (right) were grown at 26° C. C, representative *nmt3* plants showing WT-phenotype (left) and EFM phenotype (right), 21 days post-germination. D, WT and *nmt3* plants 21 days post-germination. E, frequency of EFM phenotype found in *nmt3* single seed decent lines.

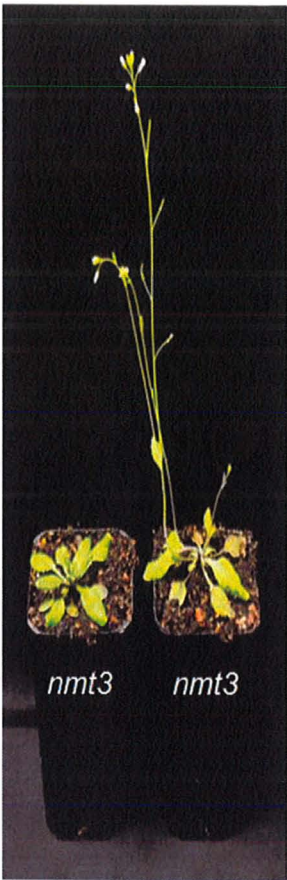
A



B



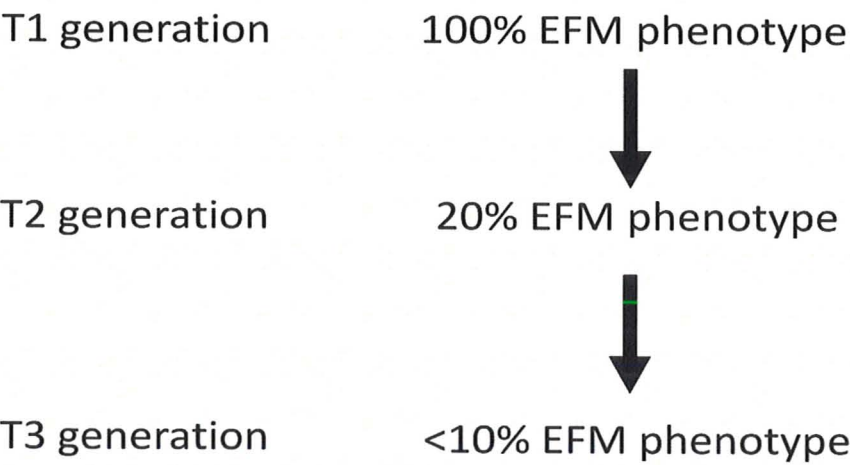
C



D



E



same frequency, only one or two plants out of 16, as observed in *nmt3* plants grown at 21°C (Fig. 14B).

### *Phospholipid profiling*

Shot-gun lipidomic profiling using ESI/MS/MS was performed on non-polar fractions of leaf extracts from WT and *nmt3* plants 28 d post-germination and the results are shown in Figure 15. Neutral loss scans differentiate between polar head groups of phosphatidyl lipids (Basconsillo and McCarry, 2008). Using this approach, both Ptd-EA and Ptd-Cho species were detected in WT and *nmt3* samples. However, no consistent difference among phospholipids could be detected by comparing lipid profiles for Ptd-EA and Ptd-Cho of WT and *nmt3* plants. Neutral loss scans used to detect Ptd-MEA and Ptd-DEA lipid species showed no accumulation of either species in WT or *nmt3* plants.

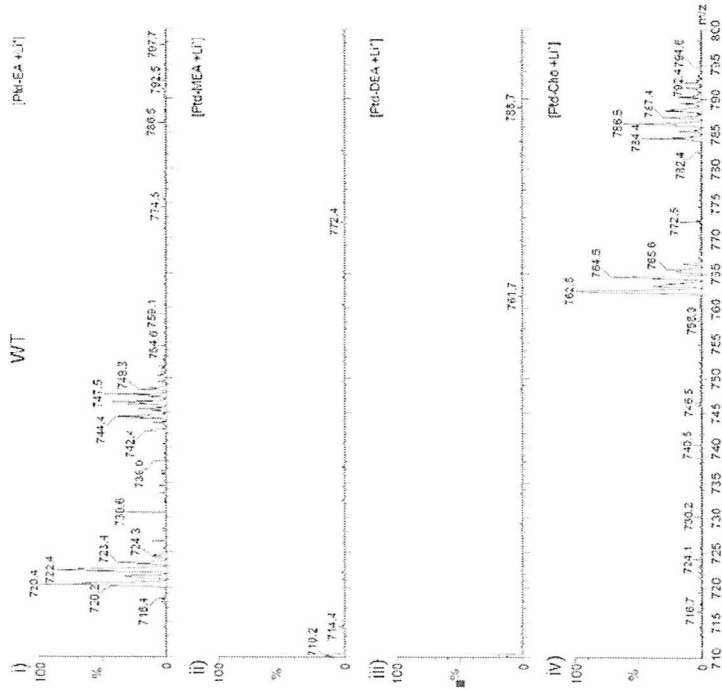
### *Cloning NMT3*

Primers for the synthesis of Arabidopsis *NMT3* cDNA by reverse transcription were designed to create a 1690 bp template that included sequence up and downstream of the translation start and stop codons, respectively (Fig. 16A). In order to verify the production of this 1690 bp cDNA template, PCR was performed using three different primer pairs. The same primers used in the reverse transcription reaction were employed to amplify a 1690 bp PCR product (Fig. 16A; predicted product (left) designated by dark grey arrow). Nested PCR amplification was performed using primers that anneal to sites just outside of the longest ORF to yield a 1608 bp amplicon (Fig. 16A; predicted product (right) designated by light grey arrow). Primers were also designed to amplify a shorter

Figure 15 – Representative phospholipid profiles of WT and *nmt3* leaf extracts.

Positive mode electrospray mass spectra of crude lipid extracts from leaves of (A) WT and (B) *nmt3* plants showing (i) neutral-loss scan of 147 mass units targeting Ptd-EA; (ii) neutral-loss scan of 161 mass units targeting Ptd-MEA; (iii) neutral loss scan of 175 mass units targeting Ptd-DEA; (iv) neutral-loss scan of 189 mass units targeting Ptd-Cho. Lipid extracts were prepared in triplicate from 200 mg of leaf tissue from three individual WT and *nmt3* plants (n = 3).

A



B

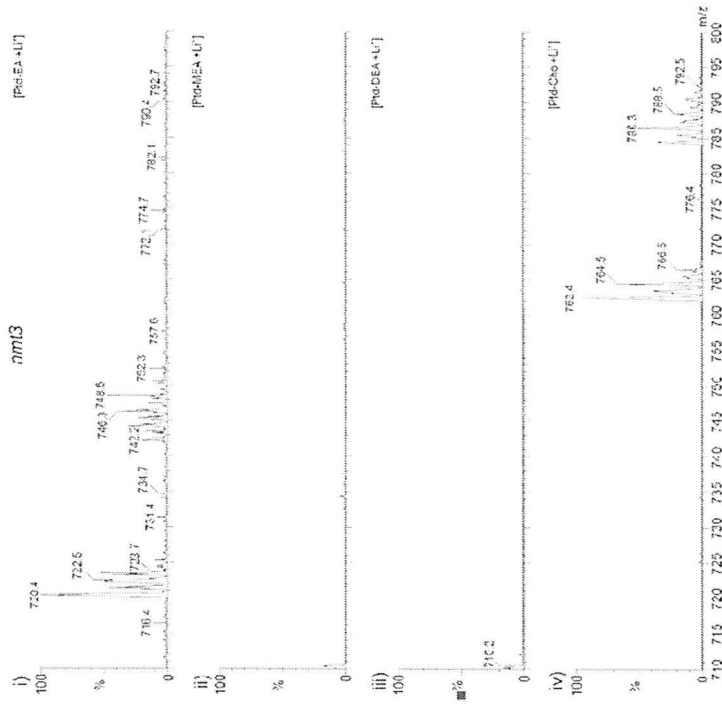
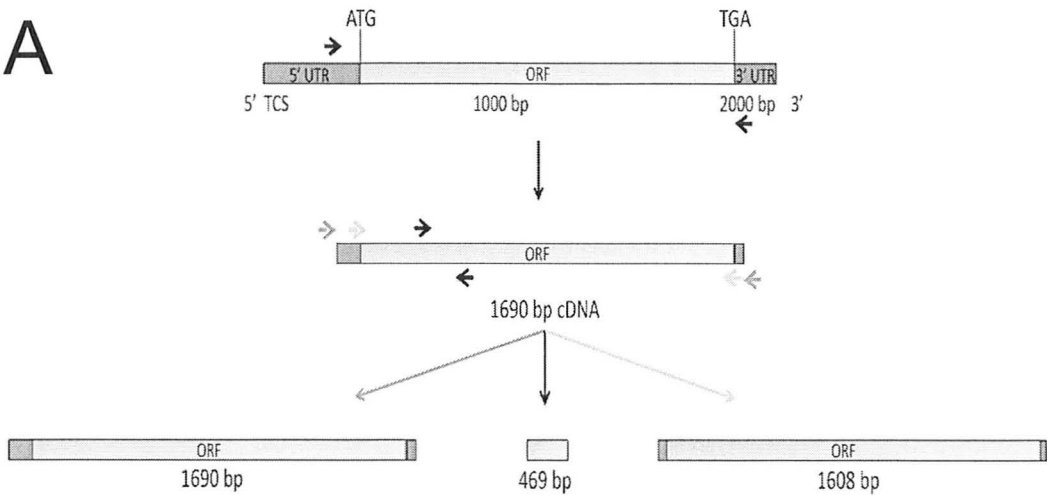


Figure 16 – Isolation of a partial *NMT3* cDNA template

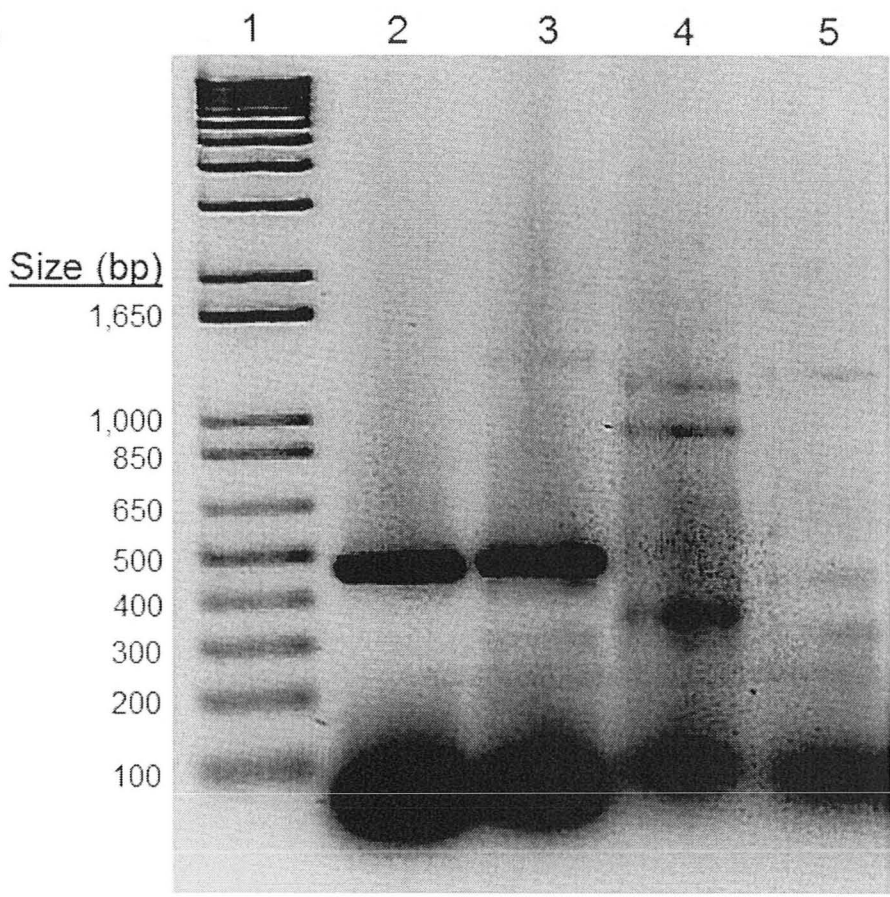
A, Map of *NMT3* gene structure showing the position of primers used to amplify gene-specific products of 1690 bp (left), 469 bp (middle) and 1608 bp (right). B, Molecular Size Markers (Lane 1), amplification of a small *NMT3*-specific product (Lanes 2 and 3). Note the absence of long (>1600 bp) *NMT3*-specific amplicons in Lanes 4 and 5. See Materials and methods for details on PCR and electrophoresis conditions.



A



B



gene-specific *NMT3* product (469 bp) using primers annealing inside the ORF (Fig. 16A; predicted product (middle) designated by black arrow). Lanes 2 and 3 (Fig. 16B) show gene-specific products of 469 bp; an indication that some *NMT3* cDNA was synthesized during the reverse transcription reaction. Lanes 4 and 5 show that no *NMT3* product >1600 bp was made. The absence of longer PCR products indicates that the 1690 bp template may have been incomplete or that PCR conditions were not optimal for the amplification of the predicted >1600 bp products. Annealing temperature and time were altered as well as extension time in an attempt to optimize PCR conditions for the amplification of these long (>1600 bp) *NMT3* products. These strategies did not result in the amplification of either the 1690 bp or 1608 bp *NMT3* amplicon described above. In order to clone the regions of *NMT3* containing SBD1 and SBD2 separately, primers were designed to amplify the N and C-terminal halves of *NMT3* (see primer map, Fig. 17A). These primers were designed to amplify products incorporating translation start and stop codons with the intent that these products could later be subjected to restriction digestion at a shared restriction site followed by ligation to assemble a clone containing the intact *NMT3* ORF from the two halves (Fig. 18). Using the same 1690 bp cDNA template noted above, an N-terminal half product of 681 bp and a C-terminal half product of 1031 bp were amplified (Fig. 17B, Lanes 2 and 3). The identity of these PCR products was confirmed by DNA sequencing to be the N and C-terminal halves of *NMT3* (Figs. 19 and 20). The DNA and predicted amino acid sequences are shown with the translation start/stop and SAM binding motifs highlighted. The N and C-terminal halves of *NMT3* were digested at a shared *EcoRI* restriction site and the halves (with cohesive ends) were

Figure 17 – Isolation of *NMT3* N and C-terminal halves amplified from full-length *NMT3* cDNA template.

A, Map of *NMT3* gene structure showing the position of primers used to amplify gene-specific products of 681 bp (left) and 1081 bp (right). B, Molecular Size Markers (Lane 1), amplification of the C-terminal half (Lane 2) and N-terminal half (Lane 3) of *NMT3* from full-length cDNA. See Materials and methods for details on PCR and electrophoresis conditions.

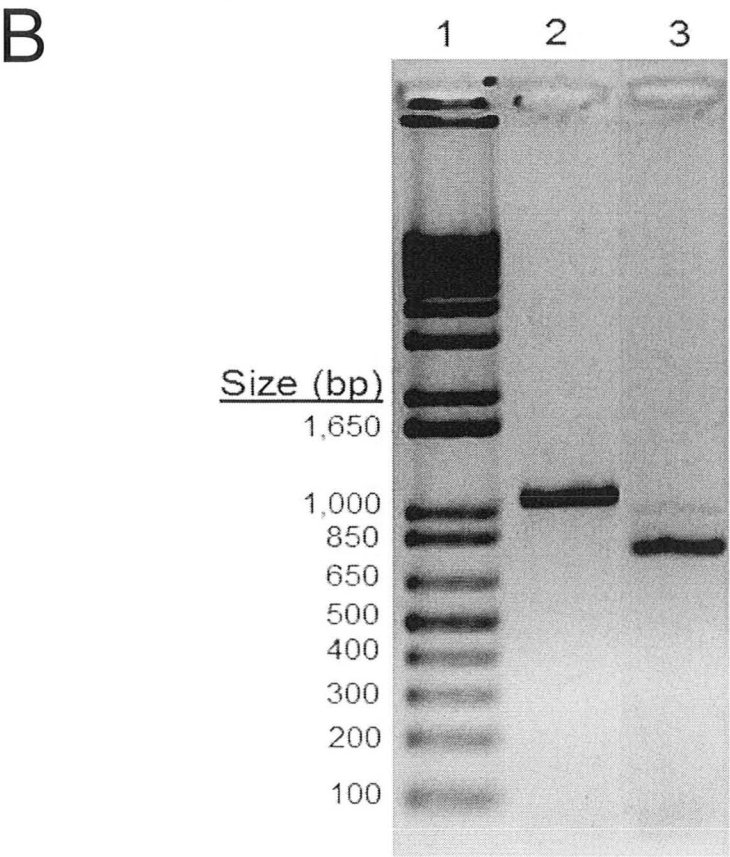
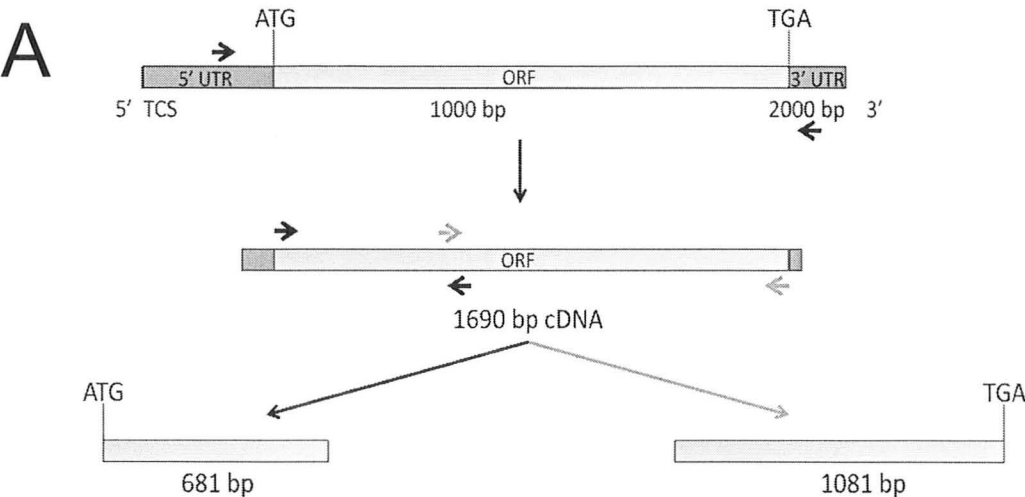


Figure 18 – Cartoon of strategy for cloning *NMT3* from N and C-terminal halves.

Restriction digestion of amplicons representing N and C-terminal halves of *NMT3* using *EcoRI* to create overlapping cohesive ends. A ligation reaction was performed using the N and C-terminal *NMT3* halves to produce an *NMT3* product containing the longest ORF.

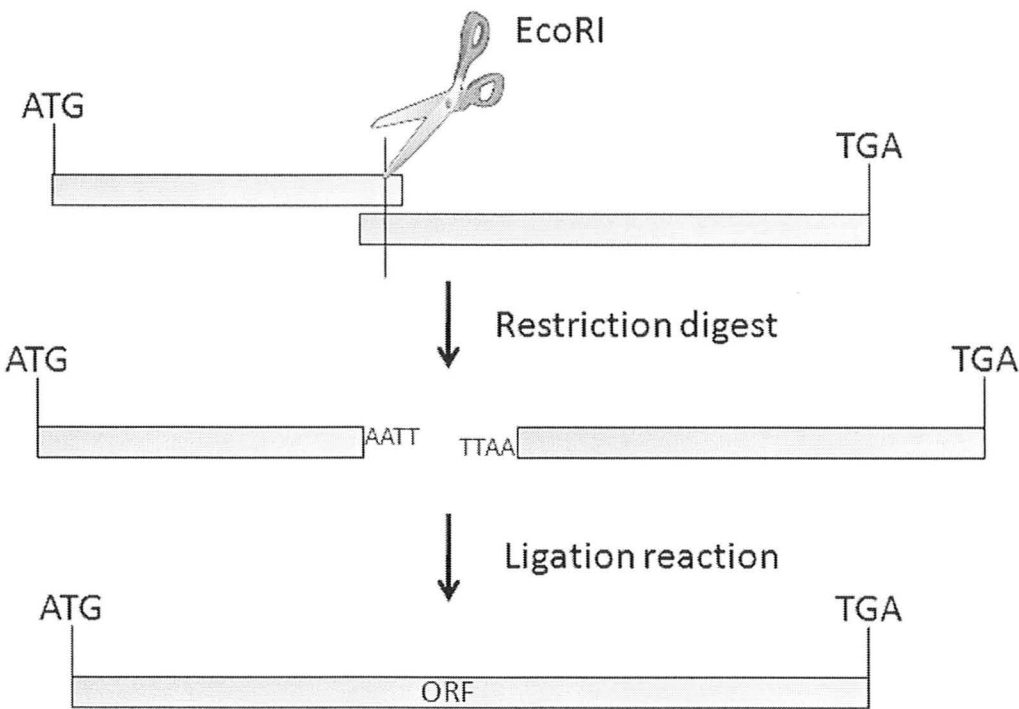


Figure 19 – N-terminal half DNA sequence and translated amino acid sequence.

Translation start as well as SAM binding amino acid motifs of the N-terminal region (SBD1) of NMT3 are highlighted. In order from start of sequence: Translation start, SAM binding motifs I, postI, II and III.

```

1
2  TCCGTCTCGAATTTCTTCTCTATTCATAGATTTTCATTACCCTCGTGAAAAATCGTCTCT

61  M A S Y G E E R E I
62  TTTTGTGTTCTAGTGTGTTCTCAAGGATAATGGCTTCGTATGGCGAGGAGCGTGAAATC

41  Q K N Y W K E H S V G L S V E A M M L D
122  CAAAAGAATTACTGGAAAGAGCATTTCAGTGGGATTGAGTGTGAAGCTATGATGCTTGAT

61  S K A S D L D K E E R P E I L A F L P P
182  TCCAAAGCTTCTGACCTCGACAAAGAAGAAGTCCTGAGATACTTGCGTTTCTTCCACCT

81  I E G T T V L E F G A G I G R F T T E L
242  ATTGAAGGGACAACAGTGCTAGAGTTTGGTGCTGGAATTGGTCGTTTTACTACTGAATTA

101  A Q K A G Q V I A V D F I E S V I K K N
302  GCTCAGAAGGCCGGCCAGGTCATTGCGGTTGACTTCATTGAAAGTGTTATCAAAAAGAAT

121  E N I N G H Y K N V K F L C A D V T S P
362  GAGAACATTAACGGTCACTACAAGAAGTCAAATTTCTGTGCGCTGATGTCACATCACCA

141  N M N F P N E S M D L I F S N W L L M Y
422  AATATGAACTTTCCAAATGAGTCTATGGATCTGATATTCTCCAAGTGGCTGCTAATGTAT

161  L S D Q E V E D L A K K M L Q W T K V G
482  CTCTCTGATCAAGAGGTTGAAGATTTGGCGAAAAAGATGTTACAATGGACAAAGGTTGGC

181  G Y I F F R E S C F H Q S G D N K R K Y
542  GGGTATATTTTCTTTTCGGGAGTCTTGTTTCCATCAGTCTGGTGATAACAAGCGGAAGTAC

201  N P T H Y R E P K F Y T K L F K E C H M
602  AACCCAACACACTACCGTGAACCTAAATTTTACACAAAGCTTTTCAAAGAATGCCATATG

221  N D E D G N S Y E L S L V S C K C I G A
662  AATGACGAAGATGGGAATTCGTATGAACTCTCTTTGGTTAGCTGTAAATGCATTGGAGCT

241  Y V R N K K N Q
722  TATGTGAGAAACAAAAGAACCAG

```



Figure 20 – C-terminal half DNA sequence and translated amino acid sequence.

Translation start and stop as well as SAM binding amino acid motifs of the C-terminal region (SBD2) of NMT3 are highlighted. In order from start of sequence: SAM binding motifs I, postI, II, III and translation stop.

```

1  E X G N S Y E L S L V S C K C I G A Y V
2  GAANATGGGAATTCGTATGAACTCTCTTTGGTTAGCTGTAAATGCATTGGAGCTTATGTG

21  R N K K N X N Q I C W L W Q K V S S D N
62  AGAAACAAAAGAACCCGAAACCAGATATGCTGGCTTTGGCAGAAAGTCAGTTCGGATAAT

41  D R G F Q R F L D N V Q Y K S S G I L R
122  GATAGGGGCTTCCAACGCTTCTTGGACAATGTCCAGTATAAGTCTAGTGGTATCTTACGC

61  Y E R V F G E G F V S T G G L E T T K E
182  TATGAGCGTGTCTTTGGAGAAGGGTTTGTAGCACAGGGGGACTCGAGACAACAAAGGAA

81  F V D M L D L K P G Q K V L D V G C G I
242  TTCGTGGATATGCTGGATCTGAAACCTGGCCAAAAGTTCTAGACGTTGGGTGCGGAATA

101  G G D F Y M A E N F D V D V V G I D L
302  GGAGGAGGGGACTTCTACATGGCTGAGAACTTTGACGTGGATGTTGTGGGCATTGATCTA

121  S V N M I S F A L E H A I G L K C S V E
362  TCTGTAAACATGATCTCTTTTGGCGCTTGAACACGCAATAGGACTCAAATGCTCTGTAGAA

141  F E V A D C T K K E Y P D N T F D V I Y
422  TTCGAAGTAGCTGATTGCACCAAGAAGGAGTATCCTGATAACACCTTTGATGTTATTTAT

161  S R D T I L H I Q D K P A L F R R F Y K
482  AGCAGAGACACCATTCTACATATCCAAGACAAGCCAGCATTGTTTCTGAGAAGATTCTACAAA

181  W L K P G G K V L I T D Y C R S P K T P
542  TGGTTGAAGCCGGGAGGGAAAGTTCTCATCACTGATTACTGCAGAAGCCCCAAAACCCCA

201  S P D F A I Y I K K R G Y D L H D V Q A
602  TCTCCAGACTTTGCAATCTACATCAAGAAACGAGGTTATGATCTTCATGATGTACAAGCA

221  Y G Q M L R D A G F E E V I A E D R T D
662  TACGGTCAGATGCTGAGAGATGCTGGTTTCGAGGAGGTAATCGCGGAGGATAGAACCGAT

241  Q F M K V L X R E L X A V E K E K E E F
722  CAGTTCATGAAAGTCCTGAANC GGGAAGTGGNTGCACTGGAGAAGGAGAAGGAAGAATTC

261  I S D F S K E D Y E D I I G G W E V K A
782  ATCAGTGAAGTCTCGAAAGAGGATTACGAGGATATTATAGGCGGGTGGGAAGTCAAAGCT

281  T
842  ACTTAGGAGCTCAAGTGGTGAGCAGAAGTGGGGTTTGTTCATCGCCAAGAGAAACNGATT

301
902  ANGGCATTCTACTATAAGTNGGAAATAAATTTGGGTATGTATGTANNTCNTCAAACCCAG

321
962  TACCCNGCTNTTATNTNTNNNNCTNNTNNTNTGNGGCNNNNNGATNNTNNTNN

```

ligated to form a full-length *NMT3* cDNA sequence. Two lines of evidence suggest this strategy was unsuccessful: a product was not seen following agarose gel electrophoresis and attempts to use PCR amplification of an *NMT3* gene-specific product spanning the site of ligation was unsuccessful (data not shown).

An alternative strategy using rapid amplification of cDNA ends (RACE) was performed in order to isolate and amplify a full-length, 1473 bp ORF associated with *NMT3*. For this application 3' RACE was used to amplify *NMT3* cDNA. 3' RACE generates cDNA with a poly-T<sub>n</sub> primer that anneals to the poly-A tail found at the 3' end of complete mRNA sequences (Fig. 21A). By targeting only complete (intact) mRNA templates, 3' RACE increases the likelihood of isolating full-length cDNA sequences. Despite using multiple PCR conditions and various gene-specific primers in conjunction with 3' RACE-generated cDNA, I was unable to amplify a full length *NMT3* product (results not shown). Internal *NMT3* primers were then designed to anneal in the middle of *NMT3* and just inside the poly-A tail in order to amplify an approximately 1100 bp PCR product. Using cDNA generated by 3'RACE as a template, I was able to amplify the 1100 bp product (Fig. 21B, Lane 4) which was then shown by sequence analysis to contain 965 bp of *NMT3* protein coding sequence including the translation stop codon. Attempts to assemble a full-length *NMT3* product by ligating the 1100 bp product to the 681 bp N-terminal half shown in Figure 19 were not successful (data not shown).

The mRNA secondary structure for all three Arabidopsis *NMT* genes was predicted using mfold (Fig. 22). *NMT3* RNA contains the largest area of difficult secondary structure (806 bp) residing in the protein coding region with the beginning and

end of this area denoted by an asterisk in the right panel of Figure 22. Given this information and with the failure to amplify a full-length *NMT3* product by conventional PCR and RACE, an alternative strategy was adopted that was suited for producing long templates with difficult secondary structure. This strategy employed longer (>30 bp) primers and an extended range Taq DNA-polymerase engineered to read through areas of difficult secondary structure (see Materials and methods for details). The use of these conditions is reported to enhance the synthesis of full-length cDNA sequences. A similar strategy was used by Sahdev *et al.*, (2007) to synthesize human mRNA that had a G/C content >65% and problematic secondary structure. The same 1690 bp cDNA template described above (Fig. 16A and 17A) was used for the PCR reaction. The longer primers used were designed to anneal just outside of the translation start and stop codons (Fig. 23A) producing a 1554 bp PCR product containing the 1473 bp ORF of *NMT3*. These primers were also designed to incorporate *NcoI* and *BamHI* restriction sites at the N and C-termini of *NMT3*, respectively.

With precautions taken to optimize full length cDNA synthesis as discussed above, a PCR product was amplified that was about 1600 bp long (Fig 23B, Lane 3). The remainder of the PCR reaction product was subjected to restriction by *NcoI* and *BamHI* and the digested product ligated into the pET-30a(+) ® expression vector (Fig. 24). A similar strategy was used to ligate the 681 bp N-terminal half *NMT3* product (Fig. 17B) in the same expression vector. The recombinant pET® vectors containing the full-length and N-terminal half *NMT3* products were sequenced to confirm their identities (Fig. 25 and 26).

Figure 21 – C-terminal half of *NMT3* amplified using an 1800 bp RACE- generated cDNA template.

A, Map of *NMT3* gene structure showing the position of primers used to amplify a gene-specific product of approximately 1100 bp. B, Molecular Size Markers (Lane 1), template free control for DNA contamination (Lane 2), amplification of RACE control product confirming efficacy of RACE reagents (Lane 3). The C-terminal half portion of *NMT3* was amplified from a 3' RACE generated cDNA template (Lane 4). See Materials and methods for details on PCR and electrophoresis conditions.

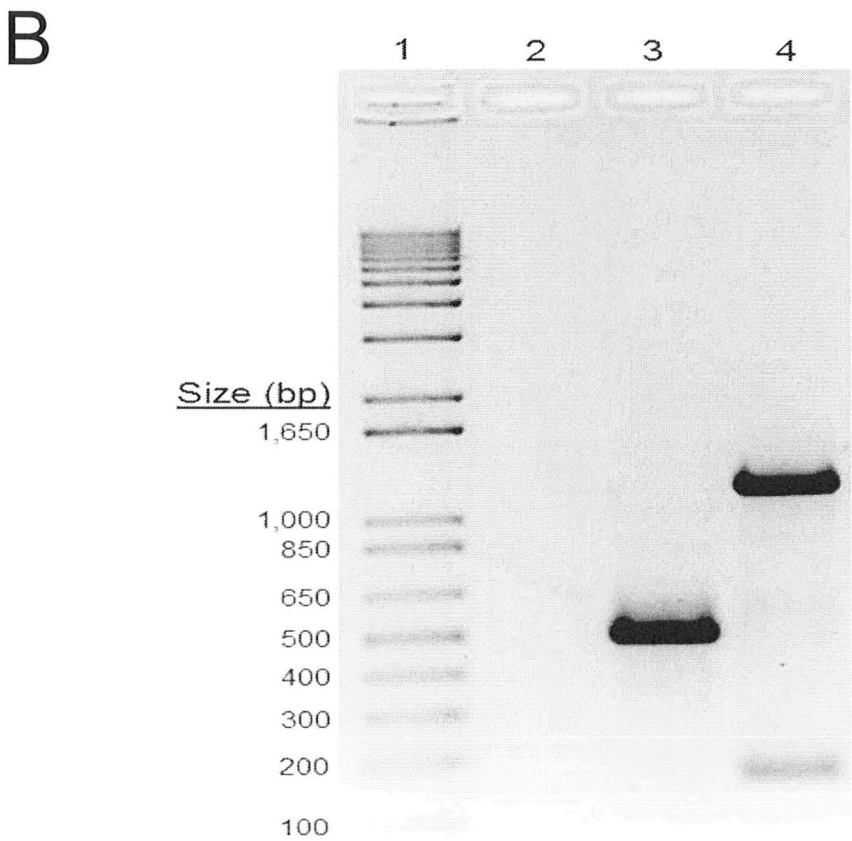
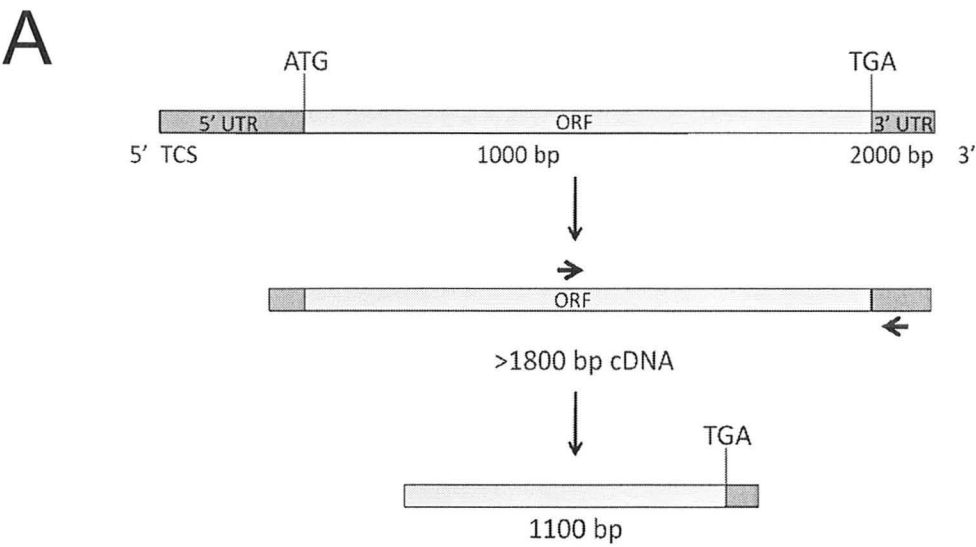


Figure 22 - Predicted RNA secondary structure of the transcripts encoding Arabidopsis NMT proteins.

RNA folding was predicted using the mfold online server (*mfold.bioinfo.rpi.edu*). The prediction was based on the full length mRNA sequence encoding Arabidopsis *NMT1* (locus At3g18000), *NMT2* (locus At1g48600) and *NMT3* (At1g73600). The translation start and stop codons are indicated with an arrow. The beginning and end of a large region of difficult secondary structure in *NMT3* is denoted with a star (★).

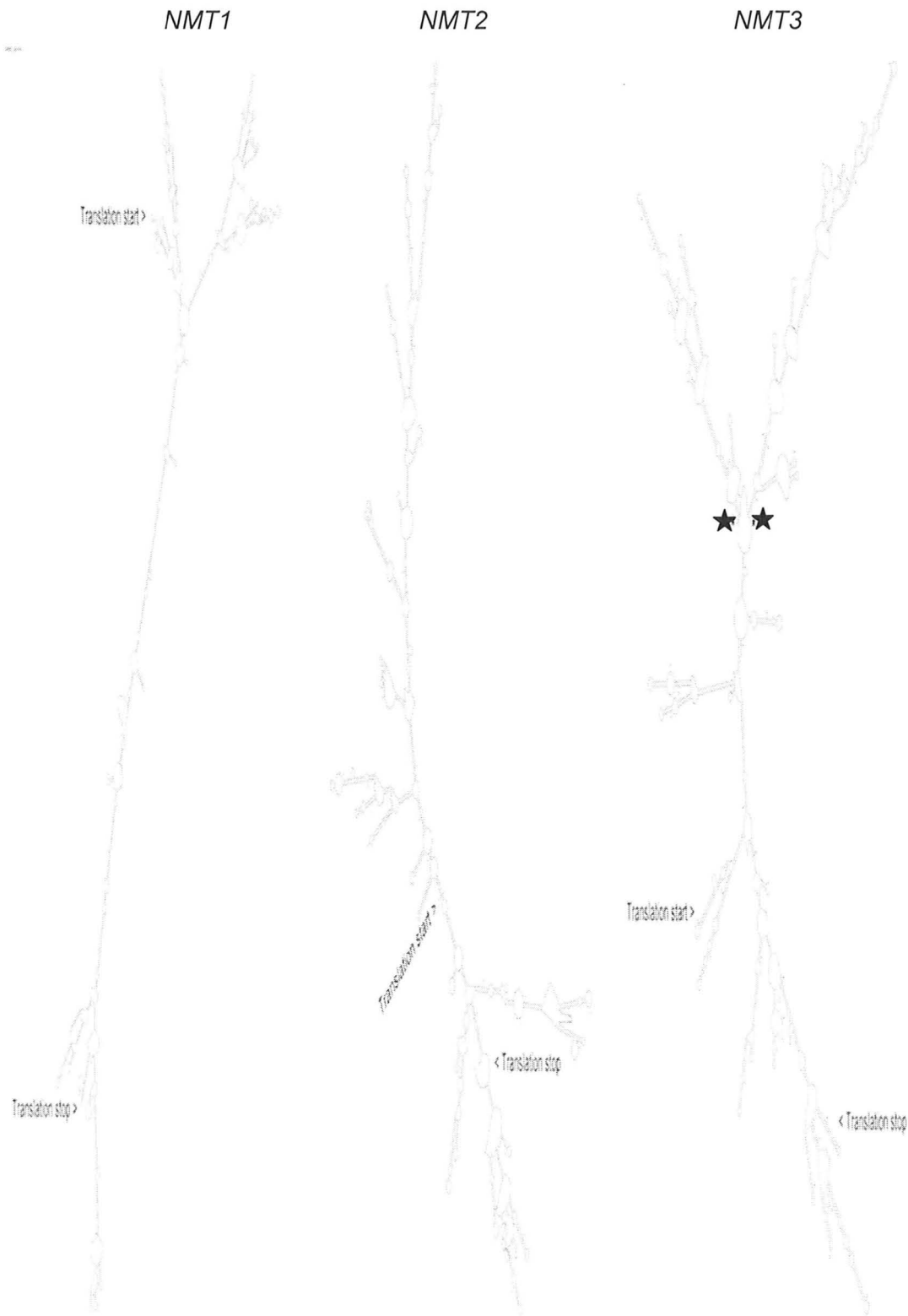




Figure 23 – Amplification of a full length *NMT3* product by RT-PCR.

A, Map of *NMT3* gene structure showing the position of primers used to amplify a gene-specific products of 1554 bp. B, Molecular Size Markers (Lane 1), template free control for the presence of DNA contamination (Lane 2), amplification of a full-length *NMT3* product (Lane 3). See Materials and methods for details on PCR and electrophoresis conditions.

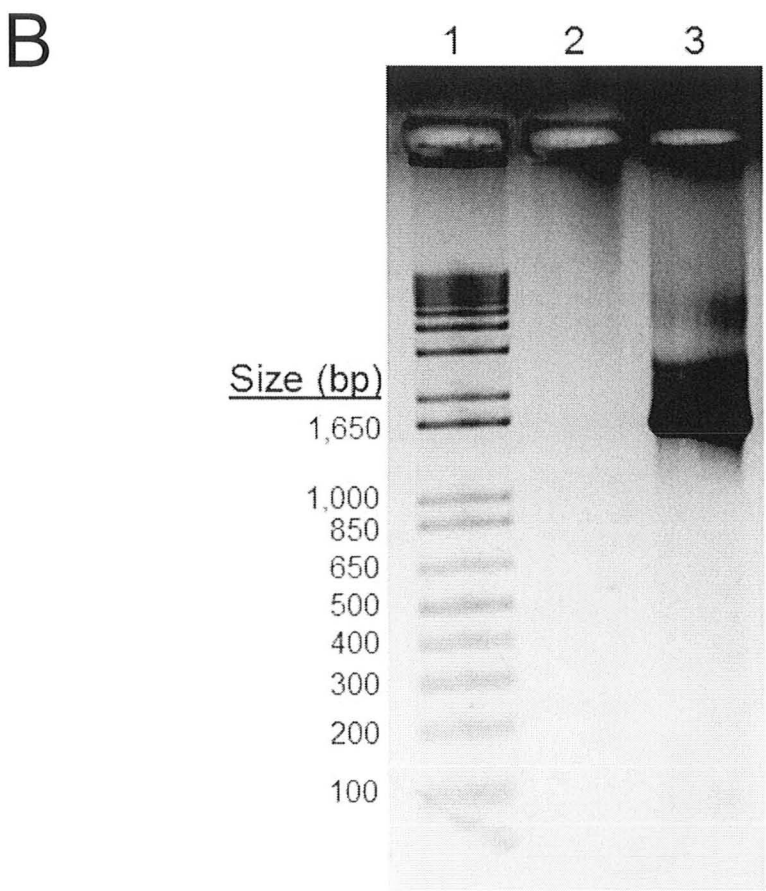
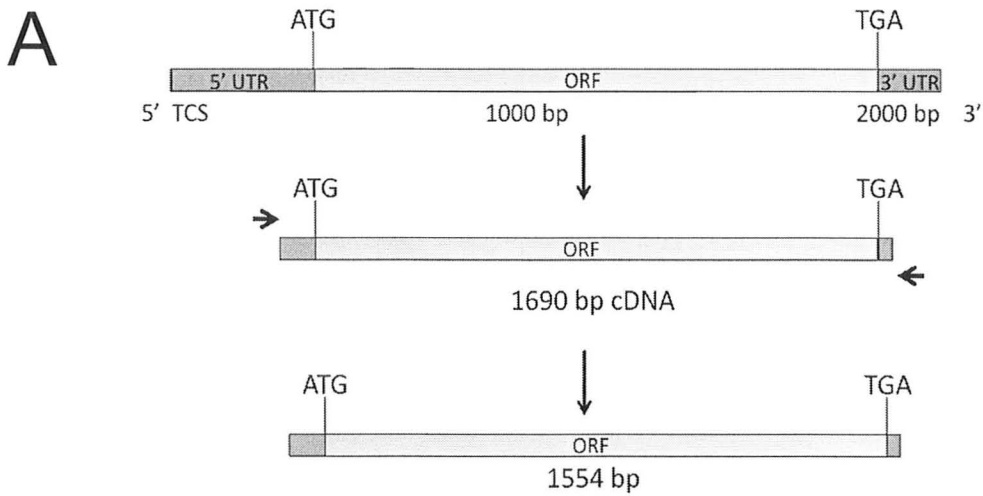


Figure 24 – Approach to cloning *NMT3* and sub-cloning amplicon into pET-30a(+) expression vector.

Full length *NMT3* cDNA sequence cloned using RT-PCR. Primers were designed to incorporate an N-terminus *NcoI* and C-terminus *BamHI* sites. The *NMT3* clone, with incorporated restriction sites, and the pET vector were subjected to double restriction digestion by *NcoI* and *BamHI* and the products used together in a ligation reaction.

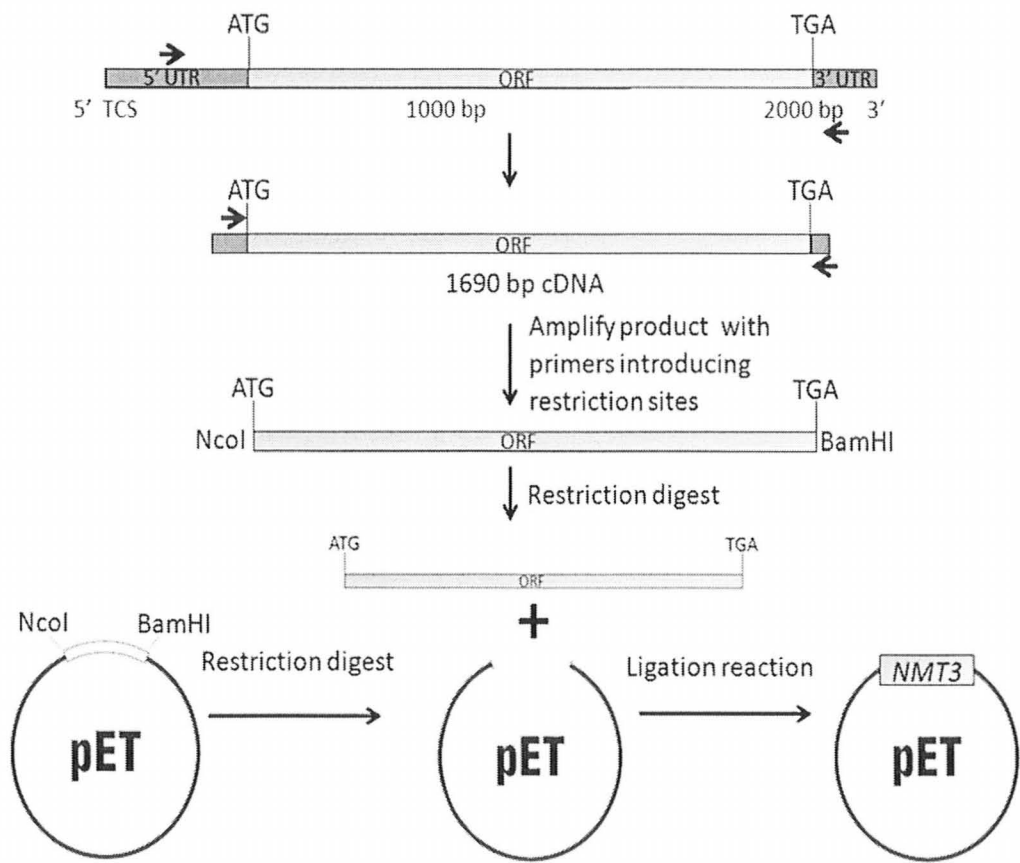


Figure 25 – DNA sequence of full-length *NMT3* ORF inserted into pET-30a(+) vector.

Highlighted in order from start of sequence: the pET-30a(+) His-tag, translation start and SAM binding motifs of SBD1. Each SBD consists of three motifs I, post I, II and III. A premature translation stop codon is indicated with the cartoon stop sign.

1  
3 NNNNNNNNNNNNNNNNNNTNNNNNNNNANNNTTGTTTAACTTTAAGAAGNNNNNATACAT

21 M H H H H H H S S G L V P R G S G M K E  
63 ATGCATCATCATCATCATATTCTTCTGGTCTGGTGCCACGCGTTCTGGTATGAAAGAA

41 T A A A K F E R Q H M D S P D L G T D D  
123 ACCGCTGCTGCTAAATTCGAACGCCAGCACATGGACAGCCCAGATCTGGGTACCGACGAC

61 D D K A M A S Y G E E R E I Q K N Y W K  
183 GACGACAAGGCCATGGCTTCGTATGGCGAGGAGCGTGAAATCCAAAAGAATTACTGGAAA

81 E H S V G L S V E A M M L D S K A S D L  
243 GAGCATTCAGTGGGATTGAGTGTGAAGCTATGATGCTTGATTCCAAAGCTTCTGACCTC

101 D K E E R P E I L A F L P P I E G T T V  
303 GACAAAGAAGAACGTCCTGAGATACTTGCCTTCTTCCACCTATTGAAGGGACAACAGTG

121 L E F G A G I G R F T T E L A Q K A G Q  
363 CTAGAGTTTGGTGCTGGAATTGGTCGTTTTACTACTGAATTAGCTCAGAAGGCCGGCCAG

141 V I A V D F I E S V I K K N E N I N G H  
423 GTCATTGCGGTTGACTTCATTGAAAGTGTTATCAAAAAGAATGAGAACATTAACGGTCAC

161 Y K N V K F L C A D V T S P N M N F P N  
483 TACAAGAACGTCAAATTTCTGTGCGCTGATGTCACATCACCAAATATGAACCTTCCAAAT

181 E S M D L I F S N W L L M Y L S D Q E V  
543 GAGTCTATGGATCTGATATTCTCCAAGTGGCTGCTAATGTATCTCTCTGATCAAGAGGTT

201 E D L V K K M L Q W T K V G G Y I F F R  
603 GAAGATTTGGTGAAAAAGATGTTACAATGGACAAAGGTTGGCGGGTATATTTCTTTTCGG

221 E S C F H Q S G D N K R K Y N P T H Y R  
663 GAGTCTTGTTTCCATCAGTCTGGTGATAACAAGCGGAAGTACAACCCAACACACTACCGT

241 E P K F Y T K L F K E C H M N D E D G N  
723 GAACCTAAATTTTACACAAAGCTTTTCAAAGAATGCCATATGAATGACGAAGATGGGAAT

261 S Y E L S L V S C K C I G A Y V R N K K  
783 TCGTATGAACCTCTTTGGTTAGCTGTAAATGCATTGGAGCTTATGTGAGAAACAAAAAG



281 N Q N Q I C -  
843 AACCAGAACCAGATATGCTGACTTTGGCAGAAAGTCAGTTCGGATAATGATAGGGGCTTC

301  
903 CAACGCTTCTTGGACAATGTCCAGTATAAGTCTAGTGGTATCTTACGCTATGAGCGTGTC

321  
963 TTTGGAGAAGGGTTTGT TAGCACAGGGGACTCGAGACAACAAAGGAATTCGTGGATATG

341  
1023 CTGGATCTGAAACCTGGCCAAAAAGTTCTAGACGTTGGGTGCGGAATAGGAGGAGGGGAC

361  
1083 TTCTACATGGCTGAGAACTTTGACGTGGATGTTGTGGGCATTGATCTATCTGTAAACATG

381  
1143 ATCTCTTTTGC GCTTGAACACGCAATAGGACTCAAATGCTCTGTAGAATTCGAAGTAGCT

401  
1203 GATTGCACCAAGAAGGAGTATCCTGATAACACCTTTGATGTTATTTATAGCAGAGACACC

421  
1263 ATTCTACATATCCAAGACAAGCCAGCATTGTT CAGAAGATTCTACAAATGGTTGAAGCCG

441  
1323 GGAGGGAAAGTTCTCATCACTGATTACTGCAGAAGCCCCAAAACCCCATCTCCAGACTTT

461  
1383 GCAATCTACATCAAGAAACGAGGTTATGATCTTCATGATGTACAAGCATACGGTCAGATG

481  
1443 CTGAGAGATGCTGGTTTCGAGGAGGTAATCGCGGAGGATAGAACCGATCAGTTCATGAAA

501  
1503 GTCCTGAAACGGGAACTGGATGCAGTGGAGAAGGAGAAGGAAGAATTCATCAGTGA CTTC

521  
1563 TCGAAAGAGGATTACGAGGATATTATAGGCGGGTGGAAGTCAAAGCTACTTAGGAGCTCA

541  
1623 AGTGGTGAGCAGAAGTGGGGTTTGTTTCATCGCCAAGAGAAACTGATTATGGCATT TACTA

561  
1683 TAAGTTGGAAATAAATTTGGGTATGTATGTGGANCCGAATTCGAGCTCCGTG CACAAGCT

581  
1743 TCGGGCCGCACNCGAGCACCACCACCACCACCTGANNNCGGCTGCTAACAAAGC NNAN

601  
1803 ANANNNNANGNGCNNNNNNNN

Figure 26 – DNA sequence of N-terminal half of *NMT3* sub-cloned into pET-30a(+) vector.

Highlighted in order from start of sequence: the pET-30a(+) His-tag, translation start, SAM binding motifs of SBD1 found in the N-terminal half region of AtNMT3 as well as the translation stop.



```

1  M H H H H H H S S G L V P R G S G M K E
1  ATGCACCATCATCATCATCATTCTTCTGGTCTGGTGCCACGCGTTCTGGTATGAAAGAA

21  T A A A K F E R Q H M D S P D L G T D D
61  ACCGCTGCTGCTAAATTCGAACGCCAGCACATGGACAGCCCAGATCTGGGTACCGACGAC

41  D D K A M A S Y G E E R E I Q K N Y W K
121  GACGACAAGGCCATGGCTTCGTATGGCGAGGAGCGTGAAATCCAAAAGAATTACTGGAAA

61  E H S V G L S V E A M M L D S K A S D L
181  GAGCATTCAGTGGGATTGAGTGTGAAGCTATGATGCTTGATTCCAAAGCTTCTGACCTC

81  D E E E R P E I L A F L P P I E G T T V
241  GACGAAGAAGAACGTCCTGAGATACTTGCCTTTCTTCCACCTATTGAAGGACAACAGTG

101  L E F G A G I G R F T T E L A Q K A G Q
301  CTAGAGTTTGGTGCTGGAATTGGTCGTTTTACTACTGAATTAGCTCAGAAGGCCGGCCAG

121  V I A V D F I E S V I K K N E N I N G H
361  GTCATTGCGGTGACTTCATTGAAAGTGTTATCAAAAAGAATGAGAACATTAACGGTCAC

141  Y K N V K F L C A D V T S P N M N F P N
421  TACAAGAACGTCAAATTTCTGTGCGCTGATGTCACATCACCAAATATGAACCTTCCAAAT

161  E S M D L I F S N W L L M Y L S D Q E V
481  GAGTCTATGGATCTGATATTCTCCAACGGCTGCTAATGTATCTCTCTGATCAAGAGGTT

181  E D L A K K M L Q W T K V G G Y I F F R
541  GAAGATTTGGCGAAAAAGATGTTACAATGGACAAAGGTTGGCGGGTATATTTTCTTTTCGG

201  E S C F H Q S G D N K R K Y N P T H Y R
601  GAGTCTTGTTTCCATCAGTCTGGTGATAACAAGCGGAAGTACAACCCAACACACTACCGT

221  K P K F Y T K L F K E C H M N D E D G N
661  AAACCTAAATTTTACACAAAGCTTTTCAAAGAATGCCATATGAATGACGAAGATGGGAAT

241  S S S V D K L A A A L E H H H H H H
721  TCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCTGA

```

The full-length *NMT3* construct (Fig. 25) contains the *NMT3* start codon in-frame as well as the SAM binding motifs of SBD1 and SBD2. However, Figure 25 shows that a single nucleotide substitution (T→A) is found 862 nucleotides into the sequence leading to a premature stop codon. The construct encoding the N-terminal half of the protein is in-

#### *Assaying NMT3 activity*

The recombinant plasmids carrying the full-length and N-terminal half constructs for *NMT3* were used to transform competent *E. coli* (BL-21) cells. The transformed cells were screened for the presence of *NMT3* cDNA using *NMT3* gene-specific primers. Four different colonies were shown to be successfully transformed and the plasmids associated with each colony were isolated and sequenced. The cells were cultured and induced by IPTG to over-express the cloned product and cell-free protein extracts were prepared and assayed for phosphobase methyltransferase activity by radioassay.



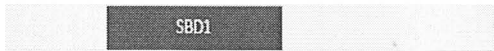


Figure 27 shows a cartoon of the predicted protein products for the four recombinant clones recovered and the enzymatic activity detected in the associated cell-free extracts. Protein A contains a complete SBD1 (all four SAM binding motifs) found in the N-terminal half of *NMT3*. Protein B has approximately 330 amino acids containing an in-frame translation start codon, SBD1, and additional sequence but does not extend to SBD2 (Fig. 27B). Protein C has the same sequence as protein B but two nucleotides inserted upstream of the pET-30a(+) start codon results in a frame shift whereby SBD1 is out-of frame (Fig. 27C). Protein D is a truncated, 266 amino acid long version of *NMT3* with an out-of-frame SBD1 (Fig. 27D). Of all the permutations of *NMT3*, only extracts prepared from clones encoding protein B showed phosphobase methyltransferase activity

and only with P-EA as substrate at a rate of  $0.38 \text{ nmol} \cdot \text{PEA} / \text{mL} \cdot \text{extract} / \text{min}$ . The activity of a full-length NMT3 protein remains unknown as a cDNA encoding the entire protein has yet to be cloned and the product expressed and assayed.

Figure 27 – Cartoon depicting regions of *NMT3* cDNA sub-cloned into the pET-30(a)+ protein expression vector.

Cell-free extracts were prepared from *E. coli* BL-21 cells transformed with plasmids carrying cloned fragments of *NMT3* as shown in panels A to D. Frame-shift mutations are denoted with an asterisk(\*). Desalted extracts were assayed for phosphobase methyltransferase activity (see Materials and Methods for details of assay conditions).

The annotated domain architecture of *NMT3* is given as a reference.

<u>Protein</u>	<u>Domain Architecture</u>	<u>Substrates methylated</u>
NMT3		Unknown
A		None
B		PEA
C		None
D		None

## DISCUSSION

The results presented in this thesis provide several lines of evidence that Arabidopsis NMT3 is a phosphobase methyltransferase. An alignment of the predicted amino acid sequence for Arabidopsis NMT3 with those for other known phosphobase NMTs is given in Figure 5 and shows the percent identity between all of these products to be very high, in the range of 72% to 82%. This alignment also shows that NMT3 has a bipartite structure with conserved SBD1 and SBD2 domains associated with the N-terminal and C-terminal halves, respectively, a feature typical for these enzymes. In terms of biochemical activity, *in vitro* assays show that the N-terminal half of the Arabidopsis NMT3 protein that includes SBD1 can methylate P-EA as a substrate but not P-MEA or P-DEA (Fig. 27). Taken together, these observations provide compelling evidence that NMT3 is a phosphobase methyltransferase like NMT1 and NMT2.

While the N-terminal half was used for *in vitro* enzyme assays, we were unable to clone an in-frame, full-length or C-terminal portion of Arabidopsis NMT3. As a result, it was impossible to directly test the methylation activity of the intact full-length protein or the C-terminal half towards the three phosphobase substrates. However, a comparison with other related NMTs in the literature suggests that Arabidopsis NMT3 can methylate all three phosphobases (P-EA, P-MEA, P-DEA). Support for this proposed property is given by the phylogenetic analyses of these gene products, shown in Figures 2, 3 and 4, as well as the predicted 3D structure of Arabidopsis NMT3 (Fig. 7 and 10).

The C-terminal half of the P-EAMeT protein, that includes SBD2 has been shown to methylate P-MEA and P-DEA for NMT1 and NMT2 from Arabidopsis, SoNMT from

spinach, TaNMT1 and TaNMT2 from wheat, and CeNMT2 from *C. elegans* (Bolognese *et al.*, 2000; BeGora *et al.*, 2010; BeGora, unpublished; Nuccio *et al.*, 2000; Jost *et al.*, 2009; Palavalli *et al.*, 2006). Thus there is precedence for P-MEA and P-DEA methylation activities residing with the SBD2 domain. If Arabidopsis NMT3 was unable to methylate P-MEA and P-DEA it would likely branch away from plant sequences of C-terminal halves of NMTs that have been shown to methylate P-MEA and P-DEA. However, the tree of the C-terminal half of Arabidopsis NMT3 shows that it belongs to a monophyletic group with the two other Arabidopsis NMT proteins that are both capable of methylating P-MEA and P-DEA (Fig. 4).

The predicted tertiary structure of Arabidopsis NMT3 also offers circumstantial evidence that NMT3 is able to methylate all three phosphobase substrates. The modelled structure of the N-terminal half of Arabidopsis NMT3 resembles proteins capable of methylating P-EA. Arabidopsis NMT1 and NMT3, wheat NMT and spinach NMT all have an alpha helix predicted to be positioned near the N-terminus of the protein. This helix appears to interact with or help form the SAM binding cleft (Fig. 10). In contrast, Arabidopsis NMT2 lacks the N-terminal alpha helix and this enzyme is unable to methylate P-EA (BeGora *et al.*, 2010).

The models for all three Arabidopsis NMTs as well as the spinach and wheat NMT proteins show no significant differences with respect to the C-terminal half. The SAM/SAH core binding fold (binding cleft) is reported to be composed of alternating beta strands and alpha helices forming a seven-stranded beta sheet with three helices on either side (Martin and McMillan, 2002). This binding fold is highly conserved and found

in nearly all SAM-dependent methyltransferases (Kagan and Clarke, 1994) and Figure 10 shows this to be the case for NMTs used for modelling in this study increasing confidence in the models presented here. Since no notable differences are seen in the models of the C-terminal half portions of these enzymes (Fig. 10A), their biochemical activities are expected to be the same. It follows, therefore, that Arabidopsis NMT3 would be expected to methylate P-MEA and P-DEA. Indeed, to date all P-EAMeT/NMT family plant proteins have been found to methylate P-MEA and P-DEA (Bolognese *et al.*, 2000; Nuccio *et al.*, 2000; Jost *et al.*, 2009, BeGora *et al.*, 2010). In addition to the conserved structure of the C-terminal half, the primary amino acid sequence encoding the C-terminal half is more conserved than the sequence encoding the N-terminal half sharing >76% compared to >70% identity at the amino acid level, respectively (Fig. 5). The amino acid sequence encoding SBD2 is nearly identical for all three Arabidopsis NMT proteins as well as wheat and spinach NMT proteins (Fig. 6B). The high similarity among the different C-terminal halves of plant NMTs and especially the high sequence identity shared in the SBD2-encoding sequence is consistent with evolution favouring a high degree of conservation for the C-terminal half of the protein.

As stated above, the Arabidopsis NMT protein family members are clearly closely related and have at least some overlapping function with respect to substrate utilization. For example, NMT1 and NMT2 can both methylate P-MEA and P-DEA. A reasonable question to consider is how such highly related paralogs are maintained. Several possibilities have been suggested by other researchers including differential regulation of the enzymes by light, differing sensitivity of the enzymes towards feedback inhibition by



SAH and P-Cho, and/or a differential spatial expression of NMT genes (Cruz-Ramirez *et al.*, 2002; Jost *et al.*, 2009). In spinach the methylation of P-MEA and P-DEA has not been shown to be light-regulated while P-EA methylation activity is absent in assays of extracts prepared from plants at the end of the dark cycle (Weretilnyk *et al.*, 1995). Spinach P-EAMeT and Arabidopsis NMT2 have been shown to differ in their sensitivity to P-Cho and SAH. The activity of NMT2 is 47% of the uninhibited rate with 1 mM P-Cho present in the assay (BeGora *et al.*, 2010). This enzyme is relatively less sensitive to P-Cho than spinach P-EAMeT which shows only 29% of uninhibited activity at the same concentration of product (Smith *et al.*, 2000). Moreover, at a final assay concentration of 0.01 mM SAH, NMT2 has 73% activity remaining as compared to spinach P-EAMeT at 53% (BeGora *et al.*, 2010; Smith *et al.*, 2000). These results are consistent with Arabidopsis NMT2 being less susceptible to inhibition by SAH than P-EAMeT from spinach. If Arabidopsis NMT1 and NMT2 show differential light regulation, different substrate usage and product inhibition properties, then the maintenance of these paralogs can be understood because their roles in the plant would likely be distinct. What is not easily explained is why Arabidopsis *NMT3* is maintained when it is predicted to encode an NMT1 paralog capable of methylating P-EA, P-MEA and P-DEA, a potentially identical biochemical activity as Arabidopsis NMT1.

Differential regulation of Arabidopsis *NMT* genes has been suggested as a possible explanation for the maintenance of three seemingly functionally redundant genes (Cruz-Ramirez *et al.*, 2004). Figure 12A shows the results of a meta-analysis of microarray data that indicates Arabidopsis *NMT1* and *NMT3* are differentially expressed

in this plant with *NMT3* expressed at 5.2 and 1.8-fold higher levels than *NMT1* in the leaves and stem, respectively. Conversely, *NMT1* shows 3.9 and 2.5-fold higher expression levels compared to *NMT3* in the roots and the shoot apex, respectively. The differential expression of *NMT3* in the leaves as opposed to the roots is supported by RT-PCR performed as part of this work. Figure 12B shows that Arabidopsis *NMT3* transcripts are more abundant in RNA extracted from leaves than roots. In agreement with the microarray data, Cruz-Ramirez *et al.*, (2004) found *NMT1* transcripts to be more abundant in roots, siliques and inflorescences than in leaves.

The expression pattern of the Arabidopsis *NMT* genes combined with the known phosphobase methyltransferase activities of the Arabidopsis NMT proteins suggests that *NMT2* may be constitutive and perform a “housekeeping” function (BeGora *et al.*, 2010). Meta-analysis of microarray data shows that *NMT2* is often expressed at high levels (higher than either *NMT1* or *NMT3*) throughout the plant (Fig. 12A). In contrast, *NMT1* and *NMT3* are differentially expressed in various parts of the plant. In a model to explain the roles of these gene products, *NMT1* and *NMT3* would control the methylation of P-EA→P-MEA and hence the entry of carbon into the P-Cho/Ptd-Cho pathway, a step that has been shown to be rate-limiting (Mudd and Datko, 1989; Nuccio *et al.*, 1998). *NMT2*, on the other hand, would function in conjunction with *NMT1* and/or *NMT3* to insure that all the P-MEA is converted to P-Cho, a role proposed by BeGora *et al* (2010). The model of *NMT2* as a “housekeeping” protein for complete conversion of P-MEA is well supported: *NMT2* is expressed throughout the plant, *NMT2* has a greater affinity for P-MEA and P-DEA than *NMT1*, and *NMT2* it is less inhibited by P-Cho than spinach P-

EAMeT allowing this enzyme to convert more substrate to P-Cho before being feed-back inhibited (Begora *et al.*, 2010). Additional evidence for the “housekeeping” role of Arabidopsis NMT2 is offered by the accumulation of Ptd-MEA in *atpmeamt* mutant plants lacking transcripts encoding the NMT2 protein (Begora *et al.*, 2010). This finding suggests that despite the presence of another enzyme capable of methylating P-MEA (NMT1) and possibly two given the predicted NMT3 function, there is still a potential for accumulation of the intermediate P-MEA that could be incorporated into membrane lipids if NMT2 activity is suppressed.

Implicit in this model for NMT function in Arabidopsis is that the efficient production of P-Cho is essential for the plant. Beyond the partially overlapping functions of the Arabidopsis NMT proteins there is further evidence of the importance of this pathway and for NMTs. Recently, the enzyme PLMT was isolated from soybean and Arabidopsis (Keogh *et al.*, 2009). This enzyme has been shown to catalyze the formation of Ptd-Cho via Ptd-MEA or Ptd-DEA intermediates (Ptd-MEA→Ptd-DEA→Ptd-Cho) offering a back-up for Ptd-Cho production by the phosphobase pathway (Fig. 1). However, while PLMT provides redundancy to the conversion of Ptd-MEA to Ptd-Cho, there is no enzyme to methylate Ptd-EA to Ptd-MEA (Fig. 1). Thus despite the back-up offered by PLMT in Ptd-Cho synthesis, the need for an NMT that can convert P-EA to P-MEA in plants is unavoidable (Keogh *et al.*, 2009). This activity would make NMT1 and possibly NMT3 essential enzymes but does raise a question about Arabidopsis NMT2 and whether PLMT alone can meet plant requirements for Ptd-Cho synthesis. Evidence that PLMT is insufficient in this role is again given by *atpmeamt* mutants where PLMT is

assumed to operate in the absence of NMT2. Here, Ptd-MEA accumulation in membranes shows that the presence of a PLMT is unable to compensate for the loss of NMT2 resulting in the accumulation of Ptd-MEA (BeGora *et al*, 2010).

Phospholipid profiles were also analyzed for *nmt3* plants and they did not show accumulation of the Ptd-Cho intermediates Ptd-MEA and Ptd-DEA (Fig. 15). This finding strongly suggests that Ptd-MEA methylation proceeds to completion in leaves of *nmt3* plants. This expectation is consistent with the predicted overlap in the biochemical activity of NMT3 with NMT1 so that NMT1 compensates for the missing NMT3 by producing P-MEA for Ptd-Cho synthesis. In *nmt3* plants P-MEA, once synthesized, can be methylated by NMT1, NMT2 and/or PLMT (via Ptd-MEA) activities. Presumably, a failure to efficiently methylate P-EA would result in an accumulation of Ptd-EA, an already abundant lipid seen in all profiles. A method for the quantification of plant membrane phospholipids would have to be devised in order to determine if membranes of *nmt3* plants were altered with respect to the absolute Ptd-EA or Ptd-Cho content of leaves. However, a comparison between the relative abundance of peaks associated with the phospholipids offers no suggestion that the relative abundance of Ptd-EA and Ptd-Cho is altered between *nmt3* and WT plants (Fig. 15).

Analysis of potential TFBS offers little insight into the differential regulation of Arabidopsis *NMT1* and *NMT3* as summarized in Figure 12A. However, this analysis does show the presence of stress-related and light-inducible TFBS in the promoter regions upstream of all three Arabidopsis *NMT* genes. Included in this list are light-inducible (GT1/2) and flowering-related (AG) TFBS in the predicted upstream regulatory region of

all three *NMT* genes. The presence of P-EAMeT activity in spinach leaf extracts has been correlated to light levels (Weretilnyk *et al.*, 1995) with highest activities found in extracts of plants exposed to light. Thus, given the TFBS prediction, it is probable that *Arabidopsis* *NMT* genes are also light-regulated.

The presence of multiple AG TFBS in all three *Arabidopsis* *NMT* upstream regulatory regions is interesting given the early-flowering phenotype associated with the *nmt3* mutant line (Fig. 14). Indeed, AG TFBS prediction suggests mutant lines deficient in any of the three NMT proteins could be associated with aberrant flowering phenotypes. An *Arabidopsis* line (*xpl*) shown to lack *NMT1* transcripts had an aberrant root phenotype but the aerial part of the plant resembled WT (Cruz-Ramirez *et al.*, 2004). Also, BeGora *et al.* (2010) report that *atpmeamt* mutants deficient in *NMT2* related transcripts resemble wild-type plants and are noteworthy in showing no overt morphological abnormalities, even with respect to flowering and reproduction. It is significant that although *nmt3* plants have been found to have an early-flowering phenotype (Fig. 14A), not all *nmt3* plants are affected to the same degree (Fig. 14C). The apparent reversion of the early-flowering phenotype found at increasing frequency when single-seed decent lines are grown suggests that the genetic perturbation in *nmt3* is epigenetic. A recent report by Yaish *et al.* (2009) associates hyper-methylation of DNA with plants that show early-flowering and multiple lateral branching inflorescences; a phenotype that resembles early-flowering *nmt3* mutants (Fig. 14D). A knock-out of the gene encoding a histone acetylation protein, *AtMBD9*, altered the expression of 333 genes and indirectly led to the hyper-methylation of many regions of DNA (Yaish *et al.*, 2009). According to these authors, hyper-

methylation in the area of the Arabidopsis gene encoding flowering locus C (*FLC*) resulted in the aberrant flowering phenotypes.

How the lack of NMT3 might give rise to the early-flowering phenotype is uncertain. It is possible that the lack of a functional NMT3 protein may result in a build-up of phosphobase intermediates (P-MEA/P-DEA) that could be converted to their phosphatidyl equivalents (Ptd-MEA/Ptd-DEA) and incorporated into the membrane. However, this hypothesis may not explain the ephemeral nature of the early-flowering aspect of the mutant phenotype and no aberrant accumulation of Ptd-MEA or Ptd-DEA was found in phospholipid profiles of the leaves of *nmt3* plants that exhibited the early-flowering phenotype (Fig. 15). Exposing plants to growth at 26°C was used to test for a heat-stress associated phenotype as elevated temperatures are expected to exacerbate the instability of any Ptd-intermediates present in the membrane as is the case for yeast (McGraw and Henry, 1989). However, the same frequency of the early-flowering phenotype was found among the progeny of heat-stressed and unstressed plants (Fig 14B and E). As such, we found no evidence that higher temperature exposure enhanced the expression of a mutant phenotype. We also found no physiological or biochemical reason to connect reduced *NMT3* expression to an early transition to flowering or the apparent loss of this phenotype in successive generations.

A lack of either NMT1 or NMT2 does not result in an early-flowering phenotype (Cruz-Ramirez *et al.*, 2004; BeGora *et al.*, 2010) but there is evidence that when more than one Arabidopsis NMT is likely deficient, an early-flowering phenotype can result. Mou *et al.* (2002) used *RNAi* to examine the effect of reduced expression of *NMT1* in

Arabidopsis. However, this approach could repress the expression of *NMT2* and *NMT3* given the degree of sequence identity shared by the Arabidopsis *NMT* family. In fact, this suggestion was made by Cruz-Ramirez *et al.* (2004) when explaining why their T-DNA insertion mutant for *NMT1* showed few of the aberrant developmental features of the *RNAi* line including early transition to flowering and inflorescences with multiple lateral branches (Mou *et al.*, 2002). This knock-down phenotype closely resembles the early-flowering response found for some *nmt3* plants (Fig. 14). It is tempting to suggest that the methylation status of *nmt3* plants is sufficiently disturbed that DNA methylation is altered giving rise to an epigenetic mutation. How a “reversion” to a WT phenotype with respect to flowering among progeny of successive generations occurs would be a very interesting topic for future studies.

Establishing the biochemical activity of Arabidopsis *NMT3* was an objective of this work that was not met. However, the inability to clone a full-length *NMT3* gene product was likely due to difficulties with the secondary structure of the mRNA that made faithful production of its cDNA template by PCR a challenge even though measures were taken to increase the likelihood of producing this amplicon. Trehalose was added to stabilize reverse transcriptase at higher temperatures thereby allowing for a higher melting temperature to be used (Sahdev *et al.*, 2007). As a consequence, cDNA synthesis was improved as indicated by the ability to amplify the N and C-terminal halves of *NMT3* for the first time (Fig. 17). Similarly, a full length product was synthesized only when a DNA polymerase was used that was engineered to read through areas of difficult secondary structure (Sahdev *et al.*, 2007). From inspection of the RNA secondary

structures of all the *Arabidopsis* *NMT* sequences it is apparent that *NMT3* has the most complex secondary structure (Fig. 22). This may explain why there was no difficulty in cloning *NMT2* (Begora *et al.*, 2010) while the attempted cloning of *NMT3* required many modifications at the procedural level. Even with PCR precautions in place, in only one case was a full-length amplicon successfully sub-cloned into the pET-30a(+) vector and a transformed bacterial colony recovered. Interestingly, when this product was fully sequenced a single base change was found leading to a premature stop codon mutation midway through the gene (Fig. 25). This point mutation created a truncated protein containing only SBD1 out of the two SBD motifs expected to be required for full NMT function (Fig. 27). For all of these attempts, any recombinant clones recovered invariably showed signs of substantive deletions or reversions when they were sequenced. This strongly suggests that expression of this gene in a fully functional form in *E. coli* could be lethal. Future attempts at cloning the gene encoding NMT3 should consider this possibility.

In view of the difficulties encountered in cloning *NMT3*, functional analysis was carried out using two recombinant partial NMT3 proteins designated A and B. Both partial cDNAs encode proteins that contain all three motifs of SBD1 and have an in-frame translation start codon with no nucleotide point mutations suggesting that partial products can be produced following induction by IPTG (Fig. 27). The difference between the two partial proteins is that B is longer in that it contains additional C-terminal amino acid sequence making this protein product around 130 amino acids longer than A. Of these two products only B showed phosphobase methyltransferase activity by in vitro assay



(Fig. 27). This is consistent with the explanation that truncated protein product A is too short and that some or all of the extra 130 amino acids found on B are required to yield an enzyme that can methylate P-EA. This outcome agrees with work by M. BeGora (unpublished) on domain-specific analysis of truncated NMT proteins. When the N-terminal half of a recombinant version of Arabidopsis NMT1 was cloned and expressed, the product was assayed for phosphopase activity and no activity was detected (Begora, unpublished). The length of this clone approximated that for protein product A described above. However, when this half of NMT1 was used to construct a chimeric enzyme by fusion to the C-terminal half of either NMT1 or NMT2, the longer product was shown to have P-EA methylation activity, whether this involved a C-terminal NMT1-N-terminal NMT1 or C-terminal NMT1- N-terminal NMT2 fusion. Charron *et al.* (2002) suggest that the C-terminal half domain (SBD2) may somehow be involved in making the substrate available for the N-terminal half of the protein containing SBD1. My findings support the proposal that at least some contribution from a C-terminal sequence is required for the proper enzymatic activity associated with SBD1.

## CONCLUSIONS

The isolation and analysis of a partial *Arabidopsis* NMT3 protein shows that NMT3 can function as a phosphobase methyltransferase capable of methylating P-EA. The striking amount of conservation in both the predicted structure and function of the C-terminal half of plant NMT proteins strongly suggests that NMT3 also methylates P-MEA and P-DEA. Thus given the similarity of the *Arabidopsis* NMT proteins at the predicted amino acid sequence and structural levels, it is reasonable to propose that all three gene products are involved in choline biosynthesis.

NMT3 would appear to resemble NMT1 with respect to biochemical activity in that both can methylate the phosphobase P-EA. The genes encoding these two proteins are differentially expressed in the plant and this would suggest that their roles are not redundant in *Arabidopsis*. However, *nmt3* plants that show suppressed expression of *NMT3* do not have altered Ptd-Cho composition in leaf membranes where this gene is more highly expressed. It is not clear how these plants are able to compensate for the loss of NMT3 with respect to phospholipid composition. A noteworthy phenotype of the *nmt3* plants is an early transition to flowering relative to WT. This phenotype is less pronounced in progeny of successive generations. Crosses between *atpmeamt* and *nmt3* plants could help determine the extent of overlap in the functional roles of these two NMT enzymes in plant development and metabolism.

The association between *NMT3* expression and flowering would be more easily studied if the exact biochemical role of NMT3 was identified. Future studies should be directed towards cloning a full length version of this methyltransferase that would be

more suitable for assays testing all three phosphobases found in the choline biosynthetic pathway. Once isolated, the kinetic properties of NMT3 can be studied in greater detail and the information used to address the question of whether evolution makes retention of this gene product necessary in plants because of its critical role in choline metabolism or another essential catalytic activity that has yet to be discovered.

## Appendix A – Primer information

Details on primers used for screening SALK\_062703 (*nmt3*) line (A) and cloning and RT-PCR of Arabidopsis *NMT3* (B). The melting temperature ( $T_m$ ) used for each primer during PCR and size of the expected PCR product are shown. The sequence of each primer and if the primer anneals to the forward (LP) or reverse (RP) strand are also indicated.

## A

LP/RP	Sequence (5'-3')	T <sub>m</sub>	Size of product
Left	tcgagctcgagctttctatg	58°C	1105 bp
Right	tatgctggatctgaaacctgg	58°C	1105 bp
Left border	gcgtggaccgcttgctgcaact	52-60°C	450-800 bp depending on LP

## B

Description of primer	LP/RP	Sequence (5'-3')	T <sub>m</sub>	Size of product
<i>NMT3</i> RT-PCR	LP	cgttttactactgaattagct cagaaggcc	62	469
<i>NMT3</i> RT-PCR	RP	agccagcatatctggttctg	62	469
Ubiquitin control RT-PCR	LP	gatctttgccgaaaacaa ttggaggatggt	62	418
Ubiquitin control RT-PCR	RP	cgacttgcattagaaga aagagataacagg	62	418
<i>NMT3</i> full length cDNA (outside ORF)	LP	tccagactcgctatcgaag g	55	1690
<i>NMT3</i> full length cDNA (outside ORF)	RP	cagggtactgggttttgatg	55	1690
<i>NMT3</i> full length cDNA (ORF)	LP	ggttctcaaggaccatggc ttcgatggcga	68	1554
<i>NMT3</i> full length cDNA (ORF)	RP	agggtactgggttttgagg atccacatacatatc	68	1554
<i>NMT3</i> N-terminal half cDNA	LP	gttctcaaggaccatggctt cgtatggcga	65	681
<i>NMT3</i> N-terminal half cDNA	RP	agccagcatatctggttctg	65	681
<i>NMT3</i> C-terminal half cDNA	LP	acaaccaacacactacc gtga	65	1081
<i>NMT3</i> C-terminal half cDNA	RP	agggtactgggttttgagg atccacatacatatc	65	1081

## REFERENCES

- Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science*, 301, 653-657.
- Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, 22, 195-201.
- Bais, A. S., Grossmann, S., & Vingron, M. (2007). Incorporating evolution of transcription factor binding sites into annotated alignments. *Journal of Biosciences*, 32, 841-850.
- Basconcillo, L. S., & McCarry, B. E. (2008). Comparison of three GC/MS methodologies for the analysis of fatty acids in *Sinorhizobium meliloti*: Development of a micro-scale, one-vial method. *Journal of Chromatography B*, 871, 22-31.
- BeGora, M.D., Macleod, M. J., McCarry, B. E. and Weretilnyk, E. A. (2010). Identification of phosphomethylethanolamine *N*-methyltransferase from *Arabidopsis* and its role in choline and phospholipid metabolism. Submitted.
- Block, A., Schmelz, E., O'Donnell, P. J., Jones, J. B., & Klee, H. J. (2005). Systemic acquired tolerance to virulent bacterial pathogens in tomato. *Plant Physiology*, 138, 1481-1490.

- Bolognese, C. P., & McGraw, P. (2000). The isolation and characterization in yeast of a gene for Arabidopsis S-adenosylmethionine:Phospho-ethanolamine N-methyltransferase. *Plant Physiology*, 124, 1800-1813.
- Brendza, K. M., Haakenson, W., Cahoon, R. E., Hicks, L. M., Palavalli, L. H., Chiapelli, B. J., et al. (2007). Phosphoethanolamine N-methyltransferase (PMT-1) catalyses the first reaction of a new pathway for phosphocholine biosynthesis in *Caenorhabditis elegans*. *The Biochemical Journal*, 404, 439-448.
- Charron, J. B., Breton, G., Danyluk, J., Muzac, I., Ibrahim, R. K., & Sarhan, F. (2002). Molecular and biochemical characterization of a cold-regulated phosphoethanolamine N-methyltransferase from wheat. *Plant Physiology*, 129, 363-373.
- Chen, C., & Chen, Z. (2002). Potentiation of developmentally regulated plant defense response by AtWRKY18, a pathogen-induced Arabidopsis transcription factor. *Plant Physiology*, 129, 706-716.
- Coughlan, S. J., & Wyn Jones, R. G. (1982) Glycinebetaine biosynthesis and its control in detached secondary leaves of spinach. *Planta*, 154, 6-17.
- Cruz-Ramirez, A., Lopez-Bucio, J., Ramirez-Pimentel, G., Zurita-Silva, A., Sanchez-Calderon, L., Ramirez-Chavez, E., et al. (2004). The xipotl mutant of Arabidopsis

reveals a critical role for phospholipid metabolism in root system development and epidermal cell integrity. *The Plant Cell*, 16, 2020-2034.

Datko, A. H., & Mudd, S. H. (1986). Uptake of choline and ethanolamine by *Lemna paucicostata* hegel. 6746. *Plant Physiology*, 81, 285-288.

Datko, A. H., & Mudd, S. H. (1988a). Phosphatidylcholine synthesis: Differing patterns in soybean and carrot. *Plant Physiology*, 88, 854-861.

Datko, A. H., & Mudd, S. H. (1988b). Enzymes of phosphatidylcholine synthesis in *Lemna*, soybean, and carrot. *Plant Physiology*, 88, 1338-1348.

Dellaporta, S.L., Wood, J. & Hicks, J.B. (1983). A plant DNA miniprep: Version II. *Plant Molecular Biology Reporter*, 1, 19-21.

Delwiche, C. C. & Bregoff, H. M. (1958) Pathway of betaine and choline synthesis in *Beta vulgaris*. *Journal of Biological Chemistry*, 233, 430-433.

Felsenstein J (1993) Phylogeny Inference Package (PHYLIP). Version 3.5. University of Washington, Seattle.

Galuschka, C., Schindler, M., Bulow, L., & Hehl, R. (2007). AthaMap web tools for the analysis and identification of co-regulated genes. *Nucleic Acids Research*, 35, D857-862.



- Greenberg, M. L., Klig, L. S., Letts, V. A., Loewy, B. S., & Henry, S. A. (1983). Yeast mutant defective in phosphatidylcholine synthesis. *Journal of Bacteriology*, 153, 791-799.
- Hanson, A. D. & Nelsen, C. E. (1978). Betaine accumulation and [14C] formate metabolism in water-stressed barley leaves. *Plant Physiology*, 62, 305-312.
- Hanson, A. D. & Rhodes, D. (1983). 14C tracer evidence for synthesis of choline and betaine via phosphoryl base intermediates in salinized sugarbeet leaves. *Plant Physiology*, 71, 692-700.
- Hitz, W. D., & Hanson, A. D. (1980). Determination of glycine betaine by pyrolysis-gas chromatography in cereals and grasses. *Phytochemistry*, 19, 2371-2374.
- Hitz, W. D., Rhodes, D. & Hanson, A. D. (1981). Radiotracer evidence implicating phosphoryl and phosphatidyl bases as intermediates in betaine synthesis by water-stressed barley leaves. *Plant Physiology*, 68(4), 814-822.
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., & Zimmerman, P. (2008). Genevestigator V3: a reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics* 2008, 420747

- Huang, H., Tudor, M., Weiss, C. A., Hu, Y., & Ma, H. (1995). The Arabidopsis MADS-box gene AGL3 is widely expressed and encodes a sequence-specific DNA-binding protein. *Plant Molecular Biology*, 28, 549-567.
- Inatsugi, R., Nakamura, M., & Nishida, I. (2002). Phosphatidylcholine biosynthesis at low temperature: Differential expression of CTP:Phosphorylcholine cytidyltransferase isogenes in *Arabidopsis thaliana*. *Plant and Cell Physiology*, 43, 1342-1350.
- Jost, R., Berkowitz, O., Shaw, J., & Masle, J. (2009). Biochemical characterization of two wheat phosphoethanolamine N-methyltransferase isoforms with different sensitivities to inhibition by phosphatidic acid. *Journal of Biological Chemistry*, 284, 31962-31971.
- Kagan, R. M., & Clarke, S. (1994). Widespread occurrence of three sequence motifs in diverse S-adenosylmethionine-dependent methyltransferases suggests a common structure for these enzymes. *Archives of Biochemistry and Biophysics*, 310, 417-427.
- Kel, A. E., Gossling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O. V., & Wingender, E. (2003). MATCH: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Research*, 31, 3576-3579.

- Keogh, M. R., Courtney, P. D., Kinney, A. J., & Dewey, R. E. (2009). Functional characterization of phospholipid N-methyltransferases from Arabidopsis and soybean. *Journal of Biological Chemistry*, 284, 15439-15447.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). ClustalW and ClustalX version 2.0. *Bioinformatics (Oxford, England)*, 23, 2947-2948.
- Martin, B. A., & Tolbert, N. E. (1983). Factors which affect the amount of inorganic phosphate, phosphorylcholine, and phosphorylethanolamine in xylem exudate of tomato plants. *Plant Physiology*, 73, 464-470.
- Martin, J. L., & McMillan, F. M. (2002). SAM (dependent) I AM: The S-adenosylmethionine-dependent methyltransferase fold. *Current Opinion in Structural Biology*, 12, 783-793.
- McGraw, P., & Henry, S. A. (1989). Mutations in the *Saccharomyces cerevisiae* Opi3 gene - effects on phospholipid methylation, growth and cross-pathway regulation of inositol synthesis. *Genetics*, 122, 317-330.
- McNeil, S. D., Nuccio, M. L., Ziemak, M. J., & Hanson, A. D. (2001). Enhanced synthesis of choline and glycine betaine in transgenic tobacco plants that overexpress phosphoethanolamine N-methyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 10001-10005.

- Moore, T. S. (1976). Phosphatidylcholine synthesis in castor bean endosperm. *Plant Physiology*, 57, 383-386.
- Mou, Z., Wang, X., Fu, Z., Dai, Y., Han, C., Ouyang, J., et al. (2002). Silencing of phosphoethanolamine N-methyltransferase results in temperature-sensitive male sterility and salt hypersensitivity in Arabidopsis. *The Plant Cell*, 14, 2031-2043.
- Mudd, S. H., & Datko, A. H. (1986). Phosphoethanolamine bases as intermediates in phosphatidylcholine synthesis by *Lemna*. *Plant Physiology*, 82, 126-135.
- Mudd, S. H., & Datko, A. H. (1989). Synthesis of methylated ethanolamine moieties: Regulation by choline in soybean and carrot. *Plant Physiology*, 90, 306-310.
- Murashige, T. & Skoog, F. (1962) A revised medium for rapid growth and bio assays with Tobacco tissue cultures. *Physiologia Plantarum*, 15, 473-497.
- Ni, M., Dehesh, K., Tepperman, J. M., & Quail, P. H. (1996). GT-2: In vivo transcriptional activation activity and definition of novel twin DNA binding domains with reciprocal target sequence selectivity. *The Plant Cell*, 8, 1041-1059.
- Nuccio, M. L., Russell, B. L., Nolte, K. D., Rathinasabapathi, B., Gage, D. A., & Hanson, A. D. (1998). The endogenous choline supply limits glycine betaine synthesis in transgenic tobacco expressing choline monooxygenase. *Plant Journal*, 16, 487-496.

- Nuccio, M. L., Ziemak, M. J., Henry, S. A., Weretilnyk, E. A., & Hanson, A. D. (2000). cDNA cloning of phosphoethanolamine N-methyltransferase from spinach by complementation in *Schizosaccharomyces pombe* and characterization of the recombinant enzyme. *The Journal of Biological Chemistry*, 275, 14095-14101.
- Palavalli, L. H., Brendza, K. M., Haakenson, W., Cahoon, R. E., McLaird, M., Hicks, L. M., et al. (2006). Defining the role of phosphomethylethanolamine N-methyltransferase from *Caenorhabditis elegans* in phosphocholine biosynthesis by biochemical and kinetic analysis. *Biochemistry*, 45, 6056-6065.
- Pessi, G., Kociubinski, G., & Ben Mamoun, C. (2004). A pathway for phosphatidylcholine biosynthesis in *Plasmodium falciparum* involving phosphoethanolamine methylation. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 6206-6211.
- Pical, C., Westergren, T., Dove, S. K., Larsson, C., & Sommarin, M. (1999). Salinity and hyperosmotic stress induce rapid increases in phosphatidylinositol 4,5-bisphosphate, diacylglycerol pyrophosphate, and phosphatidylcholine in *Arabidopsis thaliana* cells. *Journal of Biological Chemistry*, 274, 38232-38240.
- Prud'homme, M., & Moore, T. S., Jr. (1992). Phosphatidylcholine synthesis in castor bean endosperm : Occurrence of an S-adenosyl-L-methionine:Ethanolamine N-methyltransferase. *Plant Physiology*, 100, 1536-1540.

- Qin, C. B., & Wang, X. M. (2002). The Arabidopsis phospholipase D family. Characterization of a calcium-independent and phosphatidylcholine-selective PLD zeta 1 with distinct regulatory domains. *Plant Physiology*, 128, 1057-1068.
- Rontein, D., Nishida, I., Tashiro, G., Yoshioka, K., Wu, W. I., Voelker, D. R., et al. (2001). Plants synthesize ethanolamine by direct decarboxylation of serine using a pyridoxal phosphate enzyme. *Journal of Biological Chemistry*, 276, 35523-35529.
- Rontein, D., Rhodes, D., & Hanson, A. D. (2003). Evidence from engineering that decarboxylation of free serine is the major source of ethanolamine moieties in plants. *Plant and Cell Physiology*, 44, 1185-1191.
- Sahdev, S., Saini, S., Tiwari, P., Saxena, S., & Saini, K. S. (2007). Amplification of GC-rich genes by following a combination strategy of primer design, enhancers and modified PCR cycle conditions. *Molecular and Cellular Probes*, 21, 303-307.
- Sakamoto, A., Valverde, R., Alia, Chen, T. H. H., & Murata, N. (2000). Transformation of Arabidopsis with the codA gene for choline oxidase enhances freezing tolerance of plants. *Plant Journal*, 22, 449-453.
- Smith, D. D., Summers, P. S., & Weretilnyk, E. A. (2000). Phosphocholine synthesis in spinach: Characterization of phosphoethanolamine N-methyltransferase. *Physiologia Plantarum*, 108, 286-294.

- Sohn, K. H., Lee, S. C., Jung, H. W., Hong, J. K., & Hwang, B. K. (2006). Expression and functional roles of the pepper pathogen-induced transcription factor RAV1 in bacterial disease resistance, and drought and salt stress tolerance. *Plant Molecular Biology*, 61, 897-915.
- Spiess, A. N., & Ivell, R. (2002). A highly efficient method for long-chain cDNA synthesis using trehalose and betaine. *Analytical Biochemistry*, 301, 168-174.
- Steffens, N. O., Galuschka, C., Schindler, M., Bülow, L., & Hehl, R. (2005). AthaMap web tools for database-assisted identification of combinatorial cis-regulatory elements and the display of highly conserved transcription factor binding sites in *Arabidopsis thaliana*. *Nucleic Acids Research*, 33, W397-W402.
- Summers, E. F., Letts, V. A., McGraw, P., & Henry, S. A. (1988). *Saccharomyces cerevisiae* Cho2 mutants are deficient in phospholipid methylation and cross-pathway regulation of inositol synthesis. *Genetics*, 120, 909-922.
- Summers, P. S., & Weretilnyk, E. A. (1993). Choline synthesis in spinach in relation to salt stress. *Plant Physiology*, 103, 1269-1276.
- Tabuchi, T., Okada, T., Azuma, T., Nanmori, T., & Yasuda, T. (2006). Posttranscriptional regulation by the upstream open reading frame of the phosphoethanolamine N-methyltransferase gene. *Bioscience, Biotechnology, and Biochemistry*, 70, 2330-2334.

- Tanaka, K., Tolbert, N. E., & Gohlke, A. F. (1966). Choline kinase and phosphorylcholine phosphatase in plants. *Plant Physiology*, 41, 307-312.
- Tasseva, G., Richard, L., & Zachowski, A. (2004). Regulation of phosphatidylcholine biosynthesis under salt stress involves choline kinases in *Arabidopsis thaliana*. *FEBS Letters*, 566, 115-120.
- Uemura, M., Joseph, R. A., & Steponkus, P. L. (1995). Cold acclimation of *Arabidopsis thaliana* (effect on plasma membrane lipid composition and freeze-induced lesions). *Plant Physiology*, 109, 15-30.
- van Gemen, B., van Knippenberg, P.H. (1990). Nucleic acid methylation. In: Clawson, G.A., Weissbach, A., Jones, P.A., editors. Hoffmann-La Roche-UCLA Colloquium on Nucleic Acid Methylation. New York: Wiley-Liss Publishers; pp. 19–36.
- Weigel, P., Weretilnyk, E. A., & Hanson, A. D. (1986). Betaine aldehyde oxidation by spinach chloroplasts. *Plant Physiology*, 82, 753-759.
- Weretilnyk, E. A., & Hanson, A. D. (1990). Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 2745-2749.



- Weretilnyk, E. A., Smith, D. D., Wilch, G. A., & Summers, P. S. (1995). Enzymes of choline synthesis in spinach (response of phospho-base N-methyltransferase activities to light and salinity). *Plant Physiology*, 109, 1085-1091.
- Wu, Y. T., & Liu, J. Y. (2005). Molecular cloning and characterization of a cotton glucuronosyltransferase gene. *Journal of Plant Physiology*, 162, 573-582.
- Wu, S., Yu, Z., Wang, F., Li, W., Ye, C., Li, J., et al. (2007). Cloning, characterization, and transformation of the phosphoethanolamine N-methyltransferase gene (ZmPEAMT1) in maize (*Zea mays* L.). *Molecular Biotechnology*, 36, 102-112.
- Wyn Jones, R. G., & Storey, R. (1978). Salt stress and comparative physiology in the gramineae. Glycinebetaine and proline accumulation in 2 salt-stressed and water-stressed barley cultivars. *Australian Journal of Plant Physiology*, 5, 817-829.
- Yaish, M. W. F., Peng, M., & Rothstein, S. J. (2009). AtMBD9 modulates Arabidopsis development through the dual epigenetic pathways of DNA methylation and histone acetylation. *Plant Journal*, 59, 123-135.
- Zhang, X., & Cheng, X. (2003). Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure*, 11, 509-520.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*, 31, 3406-3415.