Doctor's Thesis: 0381040

Change of DNA methylation status in response to environmental stress

Chang-Sun Choi

Research and Education Center for Genetic Information
Graduate School of Biological Sciences
Nara Institute of Science and Technology, JAPAN
(Professor Shinmyo Atsuhiko)

Submitted on 1th June, 2007

ABSTRACT

DNA methylation is a common and rapidly evolving epigenetic mechanism among higher eukaryotic organisms with complex genomes. Cytosine residues in CG and CNG sequences are the main targets for methylation by DNA methyltransferases. Distribution of methylated and unmethylated CGs in genomic DNA is not random, but displays specific patterns in any given type of tissue. The generation of genomic methylation patterns is a dynamic process that requires demethylation and *de novo* methylation. DNA methylation has been suggested to link to various environmental stress responses and specific gene expression. However, only few genes have so far been studied on the actual changes of DNA methylation status.

The first chapter of this study discusses the relationship between gene expression and DNA methylation upon environmental stress in tobacco plants. We previously screened transcriptionally activated genes in hypomethylated transgenic tobacco plants expressing an anti-DNA methyltransferase1 (NtMET1), and found one clone, which encoded an aluminium-stress responsive glycerophosphodiesterase-like protein (NtGPDL). When detached leaves from wild type tobacco plants were treated with aluminium, *NtGPDL* transcripts were induced, and corresponding genomic loci were demethylated at CCGG sites. Direct bisulfite methylation mapping revealed that CG sites in coding regions were selectively demethylated, and that promoter regions were totally unmethylated regardless of the stress. The results suggested a close correlation between methylation and expression of the gene,

with a cause-effect relationship. Since DNA methylation is mutually regulated with histone modification, it was conceivable that demethylation at coding regions might induce alteration of chromatin structure, thereby enhancing transcription.

The chapter II describes isolation and characterization of genes encoding a 5-methylcytocine DNA glycosylase which is similar to ROS1 (repressor of silencing1) from *Arabidopsis*. Four cDNA clones, *NtROS1*, *NtROS2a*, *NtROS2b* and *NtROS3*, were successfully isolated from a cDNA library constructed from whole tobacco plants. NtROSs contains DNA glycosylase domain with significant similarities to *Arabidopsis* ROS1 and DME at the C-terminus. Purified NtROS1 and NtROS2a proteins expressed in Sf9 insect cells clearly exhibited activity of m^5C removal from tobacco genomic DNA *in vitro*. GFP fusion assay showed that NtROS1 and NtROS 2a were localized in nucleus. Expression of *NtROS* genes were induced by abiotic stresses, including aluminium, salt and reactive oxygen species. These observations suggested that NtROS proteins function in demethylating process of genomic DNA during plant stress response, thereby maintaining the balance of gene expression in combination with DNA methyltransferases.

Results of current study indicated a close link between DNA methylation and gene expression upon external stimuli. I propose that environmental response of plants is partly mediated through active alteration of DNA methylation status, which changes chromatin structure and eventually gene expression.

TABLES OF CONTENT

INTRODUCTION **2**

CHAPTER I **Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants**

Introduction **9**

Material and methods **12**

Results **17**

Discussion **24**

CHAPTER II **Identification of tobacco genes encoding proteins possessing removal activity of 5-methylcytosines from intact tobacco DNA**

Introduction **38**

Material and methods **40**

Results **43**

Discussion **46**

CONCLUDING REMARKS **52**

REFERENCS **55**

PUBLICATION LIST **70**

ACKNOWLEDGEMENTS **71**

ABBREVIATIONS

AdoMet	**S-adenosyl-L-methionine**
AGPC	**Guanidium thiocyanate-phenol-chloroform**
CMT	**Chromomethylase**
CTAB	**Cetyl-trimethyl-ammonium bromide**
DDM1	**Decreased DNA methylation 1**
Dnmt	**DNA Methyltransferase (mammals)**
DRM	**Domains Rearranged Methyltransferase**
GFP	**Green Fluorescent protein**
HDAC	**Histone deacetylase**
HPLC	**High pressure liquid chromatography**
MET1	**Methyltransferase1 (plants)**
MBD	**Methyl-CpG binding domain**
m^5C	**5-methylcytosine**
NLS	**Nuclear localization signal**
PTGS	**Post-transcriptional gene silencing**
ROS1	**Repressor of silencing 1**
TDG	**Thymine DNA glycosylase**
TGS	**Transcriptional gene silencing**

INTRODUCTION

DNA methylation is found in the genomes of diverse organisms including both prokaryotes and eukaryotes. In prokaryotes, DNA methylation occurs on both cytosine and adenine bases and encompasses part of the host-specific restriction-modification system of defense against bacteriophage infection (Kuhnlein and Arber 1972). However, in higher organisms, ranging from plants to mammals, DNA methylation is found almost exclusively at cytosine residues. The percentage of methylated cytosines ranges from 0–3% in insects, 5% in mammals and birds, 10% in fish and amphibians to more than 30% in some plants (Gruenbaum et al. 1981, Steward et al. 2002).

Methylation of cytosine residues in DNA is enzymatically catalyzed by methyltransferases, which transfer a methyl-group from S-adenosyl-L-methionine (AdoMet) to the 5-position (Wada 2005). In plants and mammals, symmetric DNA methylation mostly occurs at both strands of CG dinucleotides. CG methylation patterns are heritable in a replication-coupled manner by so-called "maintenance methyltransferases" (Dnmt1 in mammals, MET1 in plants) that have a preference for hemimethylated substrates. In contrast, DNA methylation at CNG (where N is any base) and nonsymmetric CHH trinucleotides (where H is any base except G) is unique to plants. These types of DNA methylation are effected by the plant-specific chromomethylase CMT3, as well as by DRM (homologs of the mammalian de novo

methyltransferases Dnmt3)(Goll and Bestor 2005).

In eukaryotes, methylated cytosines are frequently observed in transposons and other dispersed repeat sequences. These sequences cluster around centromeres but are also found in euchromatin (Singer et al. 2001, Miura et al. 2001, Lippmanet al. 2004). Sequences containing local tandem or inverted repeats are also frequently methylated, and methylated repeats that are dispersed in the genome often contain local repeats. For example, DNA transposons have inverted repeats at their termini, retrotransposons are flanked by tandemly repeated sequences, and some tandem repeats, such as *MEA-ISR* (an intergenic subtelomeric repeat sequence downstream of the *MEDEA* gene (Cao et al. 2002).

Epigenetic regulation plays an important role in mammals and plants development, throughout lifetimes. DNA methylation is associated with numerous functions, depending on the model organism and the experimental context. In mice, DNA methylation has been shown to be essential for proper embryonic development (Li et al. 1992). Similarly, a loss of DNA methylation caused developmental defects in Xenopus embryos (Stancheva and Meehan 2000). Analogous results have been reported in plants, in which reduced levels of DNA methylation were shown to be responsible for a large number of developmental abnormalities (Finnegan et al. 1996, Ronemus et al. 1996, Cao and Jacobsen 2002, Kankel et al. 2003). The triple mutant that lacks *DRM1*, *DRM2*, and *CMT3*, which has greatly reduced non-CG methylation(Cao and Jacobsen, 2002) and the mutant that lacks MET1, in which both CG and,

to a lesser extent, non-CG methylation are decreased (Saze et al. 2003). Superficial similarities in some of the developmental defects have been interpreted to reflect a conserved function of DNA methylation. However, the molecular consequences of genomic DNA methylation have been found to be surprisingly diverse. It has been generally assumed that DNA methylation has a role in regulating global gene activity even if, to date, only a very limited number of genetic loci have been shown to be misexpressed in DNA methyltransferase mutants. These include a subset of imprinted genes (Reik and Walter 2001, Delaval and Feil 2004, Gehring et al. 2006) and the some transposon (Kakutani et al. 1996, Miura et al. 2001). On the cellular level, loss of DNA methylation has been shown to affect apoptosis (Jackson-Grusby et al. 2001, Stancheva et al. 2001), X-chromosome inactivation and chromosomal stability (Gaudet et al. 2003) and the overall chromosome organization (Soppe et al. 2002, Tariq et al. 2003, Tariq and Paszkowski 2004).

Changes in DNA methylation affect chromatin structure and vice versa, suggesting that chromatin structure is intimately linked to the methylation status of DNA. Methylation of DNA has been associated with deacetylation of histone H3 and with methylation of histone H3 at lysine 9 (Ng and Bird 1999, Gender et al. 2002). DNA methyltransferases have been found in protein complexes containing histone deacetylase or histone methyltransferase activity (Rountree et al. 2000, Fuks et al. 2003). Conversely, H3 lysine 9 methylation is essential for methylation of associated DNA under some circumstances (Tamaru and Selker

2001, Jacson et al. 2002). The link between DNA methylation and chromatin structure is reinforced by the finding that mutation of a chromatin remodeling protein, DECREASED DNA METHYLATION 1 (DDM1), results in loss of DNA methylation (Vongs et al. 1993, Jeddeloh et al. 1999, Jerzmanowski 2003). The close connection between these modifications raises the possibility that DNA methylation may play an integral role in the environmental response.

Methylation can repress transcription by blocking transcriptional activators from binding to cognate DNA sequences (Watt et al.1988). It was proposed in 1975 that DNA methylation might be responsible for the stable maintenance of a particular gene expression pattern through mitotic cell division (Riggs 1975, Holliday and Pugh 1975). Although methylation can sometimes affect part of a transgene locus, transcriptional gene silencing (TGS) correlates mainly with methylation of the promoter sequence, whereas post-transcriptional gene silencing (PTGS) correlates with methylation of coding sequence(Fagard and Vaucheret 2000). However, whether methylation is a cause or a consequence of silencing is not known. In eukaryotes, the DNA methylation of promoter regions usually inhibits transcription, but methylation in coding region does not affect gene expression (Park et al. 1996, Jones et al. 1999, Zhang et al. 2006). In genes that are exceptions to this rule, such as *SUPERMAN* (SUP) and *AGAMOUS* (*AG*), DNA methylation in the transcribed portion of the gene probably causes transcriptional shut-down because there are important controlling elements in these

regions (Sieburth and Meyerowitz 1997). *FWA* is an unusual case, because the methylated direct repeats are in the 5' end of the transcribed region, yet DNA methylation inhibits gene expression (Soppe et al. 2000). In this case, the proximity of DNA methylation to the promoter might allow it to inhibit transcription. A resent genome-wide DNA methylation study revealed that about 20% of the *Arabidopsis* genome is methylated. Some regions highly methylated (up to 80%), such as transcriptionally inactive heterochromatic including centromeres, pericentromeres, and the heterochromatic knob on chromosome 4 (Zhang et al. 2006). The situation in *Arabidopsis* contrasts with that in *Neurospora crassa*, where the DNA methylation of coding sequences can attenuate transcriptional elongation (Rountree and Selker 1997). However, in mammalian cells, exons and introns are routinely methylated, but only the methylation status of the promoter seems to affect gene expression (Bird 2002).

Many developmental transitions are triggered by the perception of inductive environmental signals, suggesting that DNA methylation patterns change in response to environmental stimuli. In response to environmental stimuli, the distribution of DNA methylation in many eukaryotic systems can vary. Environmental factors can affect gene expression by altering DNA methylation in animal cells (Minamoto et al. 1999). In maize seedlings, which are exposed to cold stress, methylation decreases by >10% genome-wide (Steward et al. 2002), and cultured maize cells have large variations in DNA methylation levels (Kaeppler and Phillips 1993). The pathogen-resistance gene *Ball* (*BAL*) resides in a

complex gene cluster and is silenced by DNA methylation in *Arabidopsis* (Stokes et al. 2002). Overall, chromatin remodeling activity should change in response to environmental stimuli (Gendall et al. 2001, Stockinger et al. 2001).

DNA demethylation can occur either passively during DNA replication or actively in the absence of DNA replication (Kapoor et al. 2005). In *Arabidopsis*, recent genetic and biochemical studies demonstrated that two bifunctional DNA glycosylase/lyases, ROS1 and Demeter, function as DNA demethylase (Gehring et al. 2006, Agius et al. 2006, Morales-Ruiz et al. 2006). ROS1 can specifically recognize methylated DNA substrate. Its glycosylase activity removes the 5-methylcytosine base, and its lyase activity nicks the DNA backbone at the abasic site by a β, δ elimination mechanism (Agius et al. 2006, Morales-Ruiz et al. 2006). An unmethylated cytosine nucleotide is then, added through the action of other enzymes in the DNA repair pathway. Loss-of-function mutations in *ros1* result in hypermethylation of the specific loci and gene silencing (Zhu et al. 2007). Demeter has largely similar biochemical properties as ROS1, although its role is restricted to the two central cells of the female gametophyte where it is specifically expressed.

In the first chapter of this thesis, I discuss the relationship between gene expression and DNA methylation upon environmental stress in tobacco plants. *NtGPDL* gene was identified as a full-length clone in wild-type tobacco plants and further analyzed for the relationship between expression and DNA methylation. Physiological role of DNA methylation for

regulating gene expression is discussed. In chapter II, I describe isolation and characterization of four ROS1-like proteins from tobacco plants.

CHAPTER I

Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants

Introduction

The typical plant genome contains relatively high levels of 5-methylcytosine (m^5C), ranging from 6 to 25% of total cytosines, depending on the species (Steward et al. 2002). Thus DNA methylation is likely to be a regulatory mechanism. In the plant genome, m^5C is found in three nucleotide-sequence contexts: symmetrical CG, and two categories of non-CG sites, symmetrical CNG and asymmetric CNN sites (where N is A, T or C). Each sequence has different genetic requirements for de novo or maintenance methylation (Chan et al. 2005). The physiological role of DNA methylation is not completely understood, yet it is generally considered to be involved in regulation of gene expression, for example with reference to restriction of foreign DNA species that invade host organisms (Bender 2004), as well as epigenetic inheritance of traits (Zilberman and Henikoff 2005).

DNA methylation has in fact been shown to be involved in gene silencing at both transcriptional and posttranscriptional levels. Transcriptional gene silencing is associated with hypermethylation of promoter sequences, while post-transcriptional gene silencing is linked with hypermethylation of transcribed or coding sequences (Paszkowski and Whitham 2001).

Methylation of both DNA and histone tails is also critical for maintenance or formation of heterochromatin (Peters and Schübeler 2005), and its alteration results in a serious modification of gene expression, and transcriptional and transpositional activation of transposons (Bender 2004). DNA methylation has also been suggested to be involved in epigenetics, defined as mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in the DNA sequence (Bird 2002). This typically occurs during somatic cell differentiation in animal cells (Reik et al. 2001), and in plants and filamentous fungi as well (Kakutani et al. 1996, Martienssen and Colot 2001).

Methylation of plant DNA is catalyzed by three distinct enzymes: DNA methyltransferase 1 (MET1) is primarily responsible for maintenance of CG methylation (Finnegan and Kovac 2000); domains rearranged methyltransferase (DRM) contributes to de novo methylation in all sequence contexts (CG, CNG and CNN) (Wada et al. 2003, Wada 2005); and chromomethylase (CMT) is thought to maintain CNG methylation in heterochromatin and silencing of methylated loci (Bartee et al. 2001, Papa et al. 2001, Lindroth et al. 2001). Genetic analyses using mutants and transgenic plants have revealed close correlations with plant development and morphogenesis (Finnegan et al. 1998; Bender 2004).

In a previous study, we constructed transgenic tobacco plants expressing an antisense RNA for type 1 DNA methyltransferase (NtMET1), and found that growth and organ

development were seriously affected (Nakano et al. 2000). In order to identify genes responsible for the observed abnormalities, differential display was performed and 16 genes were initially isolated (Wada et al. 2004). Sequence analyses indicated that more than half of them were related to stress responses, leading us to speculate that external stress might affect the methylation status of individual genes. Subsequent assays using one representative gene, *NtAlix1*, showed this to indeed be the case. When wild type tobacco was challenged with tobacco mosaic virus (TMV), transcripts of *NtAlix1* temporarily accumulated, and concomitantly its genomic locus showed a temporal change in the DNA methylation pattern (Wada et al. 2004).

In the present study, I selected another so far uncharacterized gene from the above-mentioned 16, designated as hypomethylation associated (HMA) with serial numbers, and further analyzed the relationship between expression and methylation patterns. I screened environmental stress genes (*HMA3, 4, 7, 8*) whose expression is specifically affected by DNA methylation by southern hybridization method, using transgenic tobacco plants in which DNA methylation level were suppressed by expression of anti-sense *NtMET1*. The *HMA3* gene changed expression and methylation patterns in the anti-MET1 tobacco plants but the other genes not changed. I report here that *HMA3* (*NtGPDL*) responds to abiotic stresses, and that its genomic locus is in fact simultaneously demethylated.

Materials and methods

Plant materials and stress treatments

Wild-type tobacco plants (*Nicotiana tabacum* cv Xanthi nc) were grown in a greenhouse at 23°C under a 14 h/10 h light/dark photocycle. About 2 month-old healthy mature leaves were detached and placed in water for 4 h to diminish cutting stress, and then used for subsequent experiments. Samples were then subjected to osmotic stress by transferring them into a solution containing 300 mM NaCl. Similarly, metal stress was applied by transferring leaves into a solution containing 100 μM $AlCl_3$. Low-temperature treatment was performed by exposing plants to 4 ± 1°C under continuous light conditions. Inoculation by pathogenic bacterium, *Pseudomonas syringae pv. glycinea,* was performed by infiltration method as described (Sugimoto et al. 2004). Leaves were harvested at appropriate time points, immediately frozen in liquid nitrogen and stored at -80°C until use.

Gene isolation

Total RNA was isolated from tobacco samples by the acid guanidium thiocyanate–phenol -chloroform (AGPC) method (Chomczynski and Sacchi 1987) with a slight modification, and used for cDNA library construction with the λZapII vector (Stratagene, La Jolla, CA, USA).

In brief, cDNAs were ligated to the vector at the *Eco*RI and *Xho*I sites, and after transformation, the library was screened with ^{32}P-labeled probes. Positive clones were rescued in the pBluescript SK-phagemid vector by *in vivo* excision, and amplified in *Escherichia coli* DH5α. The nucleotide sequences were determined by the dideoxynucleotide chain termination method (PRISM BigDye Terminator, ABI, Sunnyvale, CA, USA), and analyzed with appropriate computer software (GeneWorks, National Center for Biotechnology Information, and PSORT server). Genomic sequence was obtained by PCR using genomic DNA as the template and appropriate primer sets designed after the cDNA sequence. Resulting fragment was 10,243 bp, including a 1918 bp upstream region from the transcription initiation site (accession number AB267678).

DNA and RNA hybridization analysis

For DNA hybridization, genomic DNA was isolated from green leaves by the cetyl-trimethyl-ammonium bromide (CTAB) method (Murray and Thompson 1980). A 25-μg aliquot was digested with an appropriate restriction enzyme at 37°C for 16 h, separated by electrophoresis on a 1% (w/v) agarose gel, and transferred onto a nylon membrane (Hybond N$^+$, Amersham, Buckinghamshire, UK). After cross linking using a UV cross linker (RPN 2501, Amersham), the membrane was subjected to hybridization with appropriate ^{32}P-labeled

probes at 65°C for 16 h. After successive washing with 0.1× standard saline citrate (SSC) and 0.1% (w/v) SDS at 65°C, samples were used to expose either BAS (Fuji Photo Film, Tokyo, Japan) or X-ray film (Eastman-Kodak, Rochester, NY, USA). For RNA hybridization, total RNA was isolated by the aurintricarboxylic acid method (Gonzalez et al. 1980), and 20-μg aliquots per lane were fractionated by formaldehyde/agarose gel electrophoresis and transferred to a nylon membrane (Hybond N^+, Amersham). Hybridization was performed in the same way as for the DNA hybridization described above, except that the hybridizing temperature was 42°C. The hybridization probe was a full length cDNA for *GPDL*, prepared from corresponding plasmids.

Epifluorescence analysis

The entire coding region of *NtGPDL* was amplified by PCR and inserted into the *Sal*I and *Nco*I sites of psGFP(S65T) vector to yield the fusion in frame with GFP. Particle bombardment was performed on a Bio-Rad Biolistic PDS 1000/He system to introduce the fusion construct into onion (*Allium cepa*) epidermal cells. The initial pressure of bombardment was 1100 psi, and the traveling distance of the particles to the plant tissues was 6 cm. Bombarded tissues were placed on the water plates and incubated at 4 °C for 24 h in the light, followed by monitoring the localization of GFP with a confocal microscope (AX70; Olympus, Tokyo, Japan).

Bisulfite methylation mapping

Five microgram of DNA was digested with *Eco*RI and subjected to the bisulfite modification as described (Xiao et al. 2003) with modification. Briefly, 120 μl of 5 M bisulfite was added to 20 μl of denatured DNA solution, incubated under five cycles of 55°C for 3 h and 95°C for 5 min. After samples were desalted with Wizard DNA Clean-up System (Promega, Madison, WI, USA), NaOH was added to a final concentration of 0.3 M and samples were incubated at 37°C for 5 min. DNA was then ethanol-precipitated, and 1-2 μl aliquot containing100 ng DNA was amplified with PCR under a 35-cycle of 95°C 1 min, 60°C 1min, 72°C 2 min. Two regions were selected (see Fig. 3A) for amplification with sets of primers; Primers for region I are: RI-F, 5'-GGATTATGGTTGAATATYYAGGTAT-3' and RI-R, 5'-CTCAATACCATC CTCRCCRARAA-3', where Y stands for C and T, and R for A and G. Primers for region II are: RII-F1, 5'-TGGATGYTYYGATAGTTGAATAG-3', RII-R1, 5'-TTACCRRCTCTTTC TRCCRRCTT-3', RII-F2, 5'-GAAATTAAYYTGATATGYYAAYTAGTGA-3', RII-R2, 5'-A ARCRAACTTARTAATRRCCATTTCCTA-3'. PCR products were then purified using Gel Extraction Kit (Qiagen, Hilden, Germany) and cloned into the pGEMT-Easy vector system (Promega). Individual clone was sequenced using an ABI sequencer (PRISM BigDye Terminator, ABI, Sunnyvale, CA, USA). The conversion efficiency of cytosine into thymine

was over 95% as judged from equal sequences in examined clones (see Fig. 5A).

Reverse transcription-PCR

Total RNA was prepared from aluminium-treated leaves by the acid guanidium thiocyanate-phenol-chloroform method (Chomczynski and Sacchi 1987), and subjected to RT-PCR using an RNA PCR Kit (AMV) ver 2.1 (Takara, Japan) with gene specific primers. Primer #1: 5'-CCAAACACCATCTCTCTCTC-3', Primer #2; 5'-TTACCTCAAGGCCCAAA CAC-3', Primer #3; 5'-AGCCTGAAATTTCCTACAGC-3, Primer RV; 5'-CCGGTCAATGTA AGC CATTTAGAGGT-3. PCR was carried out under the condition of a 30-cycle of denaturation at 94°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 1 min. After fractionation on agarose gel electrophoresis, products were identified by visualization with ethidium bromide staining.

High pressure liquid chromatography analysis

High pressure liquid chromatography (HPLC) was performed as described previously (Steward et al. 2002). A 100-μg aliquot of genomic DNA was incubated with 110 ng RNAaseA (Nacalai Tesque, Kyoto, Japan) at 37°C for 2 h to remove RNA contamination. After ethanol precipitation and denaturation at 100°C for 5 min, DNA was digested with 2

units of nuclease P1 (Sigma, Saint Louis, MO, USA) at 37°C for 20 h, followed by dephosphorylation with 20 units of calf intestine alkaline phosphatase (Takara, Otsu, Japan) at 37°C for 2 h. The sample was fractionated by Ultrafree-MC^{TM} PL-10 microcentrifuge tubes (Millipore, Bedford, MA, USA) and the permeate was injected into a Supelcosil^{TM} LC-18-S column (Supelco, Bellefonte, PA, USA). Separation was performed with a 2.5-20% methanol gradient in the presence of 50 mM KH_2PO_4 (pH 4.3).

Results

Identification of *NtGPDL*

The previously reported *HMA3* was only a fragment of an uncharacterized gene (Wada et al. 2004). In order to further examine its properties, we screened a tobacco cDNA library and isolated a full-length clone of 2439 bp. Homology searches suggested *HMA3* to encode a polypeptide consisted of 752 amino acids with a molecular mass of 82 kDa, with high similarity to glycerophosphodiesterase-like protein (Fig. 1A), and consequently it was designated as NtGPDL (*Nicotiana tabacum* glycerophosphodiesterase-like) (accession numbers, AB120519/AB120520). Phylogenetic analysis indicated that NtGPDL belongs to the glycosylphosphatidylinositol-anchored protein (GAP) family (Fig. 1B), but its physiological function has not yet to be determined. Genomic complexity was analyzed by

genomic DNA hybridization (Fig. 1C), which showed two or at most three signals on digestion with different restriction enzymes. Since *N. tabacum* is an amphidiploid, this rather simple pattern of hybridization suggests that the gene exist as a single copy per haploid genome. Cellular localization was examined by GFP-fusion proteins, which were expressed in onion epidermal cell layers by the bombarding method. Epifluorescence observation suggested that the protein was localized to the nucleus and may also be at the plasma membrane (Fig. 1D).

Stress-induced hypomethylation and expression

As *NtGPDL* (*HMA3*) was initially identified in hypomethylated transgenic tobacco expressing antisense-*NtMET1*, a DNA methyltransferase of maintenance-type 1 (Wada et al. 2004), it was speculated that hypomethylation is correlated with gene expression. Subsequently, the relationship between expression and methylation status was examined in wild type plants. If hypomethylation is linked to gene expression, the question arises as to what are the circumstances under which genes are demethylated. Database surveys indicated that transcripts of *NtGPDL* are induced by aluminium (Wada et al. 2004). This information lead us to speculate that, in wild type plants, such a stress might trigger off demethylation of relevant genes, resulting in transcriptional activation. In order to test this hypothesis, we first examined change in transcript levels for *NtGPDL* upon aluminium stress in wild type tobacco plants.

Transcripts for *NtGPDL* were not detectable by RNA blot hybridization in leaves under non-stressed conditions, but began to accumulate 6 h after aluminium treatment, reaching maximal levels after 12 h with a gradual decline by 24 h (Fig. 2A). In addition, cold and salt stresses also induced transcripts after 3 and 6 h of treatment, respectively (Fig. 2A). Accumulation was temporary, showing gradual decline after 24 h. Hydrogen peroxide generated by paraquat treatment also induced expression within 1 h, indicating oxidative stress to be a powerful inducer. Effects of biotic stress were then examined. Healthy leaves were detached and inoculated with an incompatible pathogen, *Psuedomonas syringae* pv. *glycinea,* which caused severe necrotic lesions on leaves (data not shown). Subsequent RNA blot analysis showed that the level of *NtGPDL* transcripts scarcely changed during the infection up to 24 h (Fig. 2B), suggesting that *NtGPDL* transcription does not respond to biotic-stress. The molecular properties of this gene will be described elsewhere, as the aim of the current study was to determine its response to stressful environments at the methylation and transcription levels.

Methylation status of the genomic locus

The global methylation status of genomic DNA was first examined. Healthy leaves were treated or untreated with aluminium for 12 h, and isolated DNA was hydrolyzed and subjected to HPLC analysis. The amount of m^5C over total cytosine residues was 13.00 ± 0.87% in the

control, while it was 10.95 ± 0.67% in aluminium-treated samples (triplicate measurements). This suggested that global levels of DNA methylation were slightly reduced upon aluminium treatment. The methylation status of the *NtGPDL* locus was then examined during aluminium stress by DNA blot hybridization using a pair of methylation-sensitive restriction enzymes, *Msp*I and *Hpa*II, both recognizing CCGG. *Msp*I is inactive with m^5CCGG and *Hpa*II with both m^5CCGG and Cm^5CGG. Under the non-stressed conditions, digestion of genomic DNA with *Msp*I yielded multiple fragments, ranging between 1.3 and 5.1 kb (Fig 3A, see 0 h of aluminium treatment), indicating the majority of the first cytosine in each CCGG to be unmethylated. In contrast, digestion with *Hpa*II yielded two major fragments of 17 and 8.2 kb and one minor 5 kb fragment (Fig 3B). These patterns clearly indicated that the frequency of Cm^5CGG was extremely high while that of m^5CCGG was low, suggesting that CG methylation is prevalent throughout the locus under non-stressed conditions. Based on this observation, methylation patterns upon aluminium stress were examined at CCGG sites. When leaves were treated with aluminium and periodically isolated DNAs were digested with *Hpa*II, a new fragment of 13 kb was generated within 1 h, in addition to 17 and 8.2 kb fragments which were originally found in non-stressed leaves (Fig 3B). The level of the 13 kb fragment appeared to increase thereafter, with a concomitant decrease of the 17 kb fragment (Fig. 3B, C), suggesting one Cm^5CGG site within the latter to be gradually demethylated, resulting in susceptibility to *Hpa*II. Digestion with *Msp*I equally generated multiple fragments

throughout examined period, indicating m^5CCGG, if any, not to be affected (Fig. 3A). Effects of biotic stress were then examined by inoculating detached leaves with *P. syringae* pv. *glycinea*. Subsequent DNA blot analysis as described above revealed no change of methylation patterns, showing *Hpa*II-fragments with 17 kb and 8.2 kb throughout the infection period (Fig 3D). These results indicated that abiotic stress induced transcripts and demethylation of *NtGPDL*, while biotic stress induced none of them, and therefore suggested a positive correlation between transcript induction and genomic loci demethylation.

Genomic organization

DNA blot analyses suggested methylation level at some CG sites to be decreased upon aluminium stress. However, due to multiple site numbers, identification of particular CCGG sites which were demethylated was difficult. To address this question, direct methylation mapping with the bisulfite method was planned, and the genomic locus covering 10 kb was cloned and the sequence determined. The coding region was found to span over 8 kb with 9 exons and 8 introns (Fig. 4A). The transcription initiation site was determined by PCR with several overlapping forward primers covering positions between –210 and -157, and a fixed reverse primer at +110 from the first ATG codon (Fig. 4B). Results showed that an approximately 280-base fragment was efficiently amplified with the primers #1 and #2, but not at all with the primer #3, indicating transcription to initiate around the position between

#2 and #3 (Fig. 4C). Subsequent computational analysis (http://fruitfly.org/seq_tools/ promote r.pl) suggested that the adenine at position -192 from the ATG codon could be the putative transcription initiation site (Fig. 4B). This position was tentatively assigned as the position +1, from which genome map was numbered (Fig. 4A). The promoter region was scanned for transcriptional cis-elements, but no clear motifs were identified except for one light-responsive GATA box (position -205). However, subsequent RT-PCR assays showed *NtGPDL* not to be responsive to light conditions (data not shown). Based on above-described information, two regions were selected for direct methylation mapping; Region I was at the coding region spanning 389 bases including intron 3 and exon 4 (positions +2263 to +2651), and Region II was selected from the promoter region including exon 1 and upstream (positions -627 to +363) (Fig. 4A).

Direct methylation mapping

Healthy leaves of wild type plants were treated or untreated with aluminium for 12 h, and DNA isolated from each sample was treated with bisulfite, amplified with region-specific primers by PCR, and cloned. For Region I mapping, 25 and 21 clones from untreated and aluminium-treated DNA samples, respectively, were subjected to sequence determination (Fig. 5A). When a 70-base sequence was selected, and aligned, it was evident that the majority of cytosines in CG remained intact in the untreated control, indicating them to be originally m^5C (Fig. 5A, upper panel). In contrast, one third of cytosines in CG were converted into thymine

in aluminium-treated samples, indicating them to be originally cytosines (Fig. 5A, lower panel). Cytosines in CNG and CNN were mostly converted into thymine, indicating them to be originally unmethylated. The distribution and frequency of m^5C were then estimated from the ratio between cytosines and thymines (Fig. 5B). When plotted to the nucleotide position of the region, m^5C was found to be preferentially distributed in CG sites (Fig. 5B). A notable feature is that demethylation appeared to have preferentially occurred at a CG cluster between 309 and 378 within the exon. In this region, there are seven CGs, among which six were found to be demethylated and one appeared to be overmethylated. This site was CCGG, consistent with DNA blot hybridization patterns which showed high methylation levels at many CCGGs (Fig. 3C). To determine whether or not similar demethylation occurs with oxidative stress, leaves were treated with paraquat and DNA was subjected to direct mapping as with aluminium-treated samples. The results were essentially consistent, showing a demethylation pattern at the CG cluster (Fig. 5B). Methylation at the promoter region (Region II) was then examined. Mapping primers were designed to amplify a 990-base fragment which included 627-base promoter region and exon 1 (627 bases upstream and 363 bases down-stream from the transcription initiation site) (Fig. 4A). The numbers of CG, CNG and CNN in this region were 22, 27 and 174, respectively, and CCGG sequences were six (Fig. 4A). Results of direct mapping of 20 clones from untreated and 25 clones from aluminium-treated DNA samples revealed that all cytosine residues were completely free of

methylation in both untreated control and aluminium-treated samples (Table 1).

Differential methylation among sequence contexts

The average frequency of m^5C in Region I was plotted to CG, CNG and CNN contexts (Table 1). The numbers of methylatable CG, CNG and CNN in Region I were 11, 12 and 47, respectively. In untreated samples, the methylatable CG dinucleotides were methylated at over 93% efficiency, while cytosines in CNG and CNN were methylated at 0.7 and 2.8%, respectively. In the same region from aluminium-treated samples, CG methylation was evident for 84%, a reduction of 10%. Methylation frequency at CNG was 0% at CNG and 3.0% at CNN, resembling the control. In paraquat-treated samples, average methylation levels were 82.9% for CG, 0.5% for CNG and 2.9% for CNN, showing again a reduction for CG methylation. Tukey's studentized range test indicated that demethylation frequency is statistically significant at CG site, but not at CNG and CNN sites. Overall, direct mapping experiments indicated that methylation preferentially occurred at CG sites in untreated samples, and that demethylation selectively took place at CG clusters upon aluminium treatment.

Discussion

This paper documents a close correlation between demethylation and expression of a tobacco

gene encoding a glycerophosphodiesterase-like protein upon exposure to abiotic stresses. A recent survey revealed that one of the molecular bases of regulation of gene expression is DNA methylation, the pattern of which spontaneously and reversibly changes during the life time of eukaryotes (Bird 2002) and with disease processes, a notable example being neoplasia (Ehrlich 2000). We also found that pathogen attack simultaneously induced expression and demethylation of one gene, *NtAlix* (Wada et al. 2004). To further examine the relationship between gene expression and demethylation by external stimuli, we here analyzed another gene, *NtGPDL*, for which transcripts were previously found to accumulate upon aluminium treatment (Wada et al. 2004).

GPDL possesses two glycerophosphoesterase-like domains, suggesting them to be involved in phospholipid metabolism, although no functional analyses have so far been carried out. As a protein group, GPDL appears to belong to the glycosylphosphatidylinositol-anchored protein (GAP) family, which are localized at plant cell surfaces and are thought to be involved in extracellular matrix remodeling and cell signaling (Borner et al. 2002, 2003). However, current evidence suggests that they are likely to have a wider range of functions (Borner et al. 2003). This speculation is in accordance with the present finding, showing that NtGPDL is distantly related to other members, and localized not only at cell surface (plasma membrane) but also in nucleus, suggesting a novel, so far unknown biological functions.

In healthy leaves grown under non-stressed conditions, *NtGPDL* was not expressed, and

its coding region was found to be heavily methylated. Upon exposure to abiotic stress, its transcripts were induced and genomic locus was partly demethylated, while upon biotic stress, neither transcripts nor demethylation were induced. This strengthened the close correlation between gene expression and methylation. Direct methylation mapping showed that demethylation predominantly took place at CG doublets, which were clustered in exon but rare in intron. Assuming that examined clones were randomly derived from genomic DNA population, such a demethylation was predicted to have occurred in approximately 10% of total DNA population, or cell population, which might be appropriate for stress response in leaf tissues. This speculation is consistent with current concepts of reverse relationship between expression and genomic methylation (Bird 2002), and involvement of CG methylation in chromatin structure (Fuchs et al. 2006). In addition to methylation of coding regions, gene expression is often correlated with promoter methylation, which sometimes occurs at only one or two cytosine residues (Ko et al. 2005). Methylation mapping at the promoter region of *NtGPDL* surprisingly showed that it was completely free of m^5C regardless of stress treatment. This finding was rather contradictory to general concepts described above, but compatible with recent studies on genome-wide methylation mapping of *Arabidopsis* (Zhang et al. 2006). Heavy methylation was found in heterochromatin, repetitive sequences and small RNA coding regions, consistent with previous studies (reviewed by Wada 2005). However, there were several unexpected findings. For example, over one third of

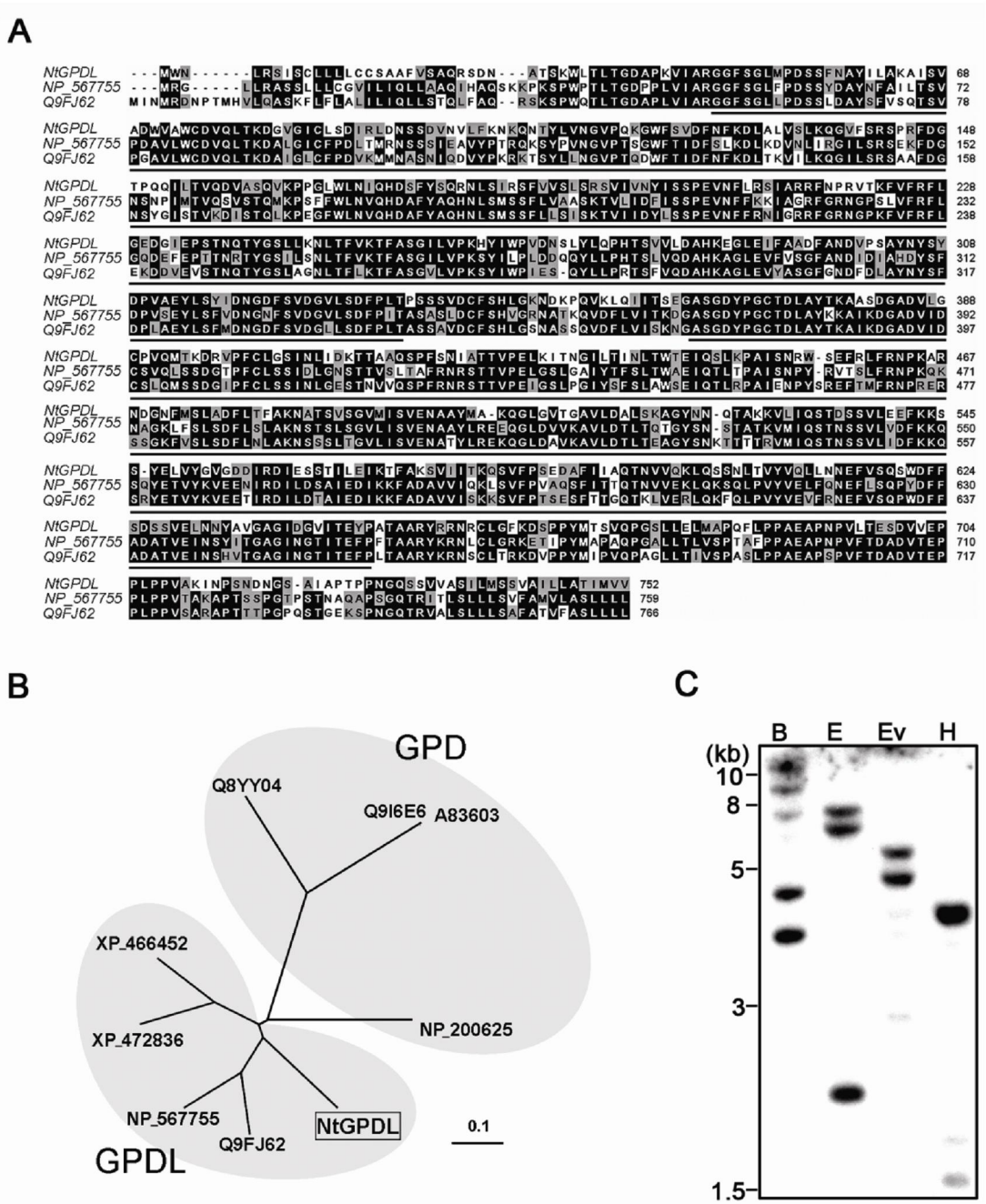
expressed genes were also methylated in coding regions (body-methylated), and 95% genes were free of methylation in their promoter regions (Zhang et al. 2006). This indicates that promoter methylation is not necessarily prerequisite for controlling expression. *NtGPDL* is a heavy body-methylated, but promoter methylation-free gene, and is highly regulated to specifically respond to external stimuli. Although the genome-wide methylation mapping suggested that constitutively active genes were heavily body-methylated (Zhang et al. 2006), cases such as *NtGPDL* were not exemplified. Since cytosine methylation was shown to be closely linked with methylation of histones (Tariq and Paszkowski 2004, Fuchs et al. 2006), demethylation within *NtGPDL* locus might affect chromatin architecture. We speculate that stress-induced demethylation of coding regions may relax chromatin structure, thereby allowing enhanced transcription (Shilatifard 2006).

The question arises as to how demethylation is specifically activated upon stress. Active oxygen radicals, or reactive oxygen species have been proposed as possible triggers (Cerda and Weitzman 1997), formed through the action of environmental factors that cause oxidative stress, such as ozone, intense light, heat and cold, drought, heavy metals and pathogens. They are highly destructive to cellular components, including lipids, nucleic acids and proteins (Bray et al. 2000) and in the case of DNA, ROS attack is known to result in formation of 8-hydroxyguanosine as a major adduct, which is usually repaired by specific mechanisms that correct this oxidative lesion. However, when stress is too heavy, or when repair is inefficient,

methylation of adjacent cytosines is affected, thereby inducing aberrant epigenetic effects, which may lead to carcinogenesis (Cerda and Weitzman 1997). It is tempting to speculate that, in plants, similar events take place, the DNA repair system being activated immediately after injury with consequent alteration of the levels and distribution of m^5C. Indeed, we found similar reduction in methylation and transcriptional activation of *NtGPDL* upon exposure to paraquat, an effective reactive oxygen species generator. Direct methylation mapping showed almost identical demethylation patterns as with aluminium treatment, featuring selective demethylation at CGs in exon regions. Since aluminium is known to generate reactive oxygen species (Boscolo et al. 2003, Apel and Hirt 2004), the present data suggest that the observed demethylation might have been mediated through oxygen radicals. If this were the case, induction of *NtGPDL* transcripts upon drought and cold stresses could also be triggered by similar mechanisms, as both stresses are known to generate reactive oxygen species (Apel and Hirt 2004).

Demethylation of genomic DNA with environmental stress has occasionally been reported in several plant species (Galaud et al. 1993, Steward et al. 2002, Lizal and Relichova 2001, Labra et al. 2002, Alina et al. 2004). Although the underlying process have yet to be clarified in detail, two alternative mechanisms are conceivable: active and passive. The active mechanism involves enzymes that excise m^5Cs from DNA to replace them with cytosines. DNA glycosylases are known to have such activity (Gong et al. 2002), the function being

essentially to maintain DNA repair and to serve as an epigenetic regulator (Gehring et al. 2006). The passive mechanism is considered to be the result of chance 'not-methylation' after DNA replication (Finnegan et al. 1998). In this case, demethylation becomes evident after several cycles of DNA replication, or cell division. Demethylation upon environmental stress must involve the former, since, in stressed tissues, DNA is scarcely replicated, while demethylation takes place rather rapidly (Steward et al. 2000). In this context, the recently identified enzyme, ROS1, is notable, since it shows clear demethylation activity towards m^5C in DNA and ubiquitously expressed in plant tissues (Aguis et al. 2006). Its involvement in stress response has yet to be confirmed, but the present findings indicate that this possibility warrants attention.


A
NtGPDL
NP_567755
Q9FJ62
B
GPD
Q8YY04
Q9I6E6
A83603
XP_466452
XP_472836
NP_200625
NP_567755
Q9FJ62
NtGPDL
0.1
GPDL
C
B
E
Ev
H
(kb)
10
8
5
3
1.5

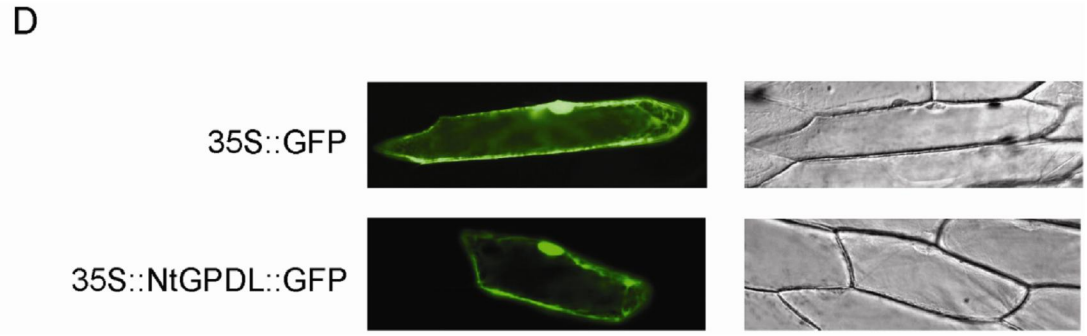


Figure 1. Properties of *NtGPDL*. **A** Amino acid sequence and alignment with related proteins (*Arabidopsis thaliana* NP_567755 and Q9FJ62). *Black and grey boxes* indicate identical and similar amino acids, respectively. Two tandem located conserved GPD motifs are *underlined*. **B** Unrooted phylogenetic tree of *NtGPDL* with related GPDL proteins found in *A. thaliana* (Q9FJ62, NP_567755) and *Oryza sativa* (XP_472836, XP_466452) and GPD proteins from *E. coli* (Q8YY04, Q916E6, A83603) and *A. thaliana* (NP_200625). Each family is indicated by shading. **C** Southern hybridization analysis of *NtGPDL*. A 10-µg aliquot of genomic DNA from tobacco (*N. tabacum*, cv Xanthi) was digested with *Bam*HI (B), *Eco*RI (E), *Eco*RV (Ev) or *Hin*dIII (H), fractionated by agarose gel electrophoresis, transferred to a nylon membrane, and hybridized with radioactively labeled *NtGPDL* as the probe. **D** Intracellular localization. Onion epidermal cell layers were bombarded with gold particles coated with 35S*NtGPDL-GFP* or 35S*GFP* alone, and observed under interference contrast (*right panel*) and by epifluorescence for GFP (*left panel*).

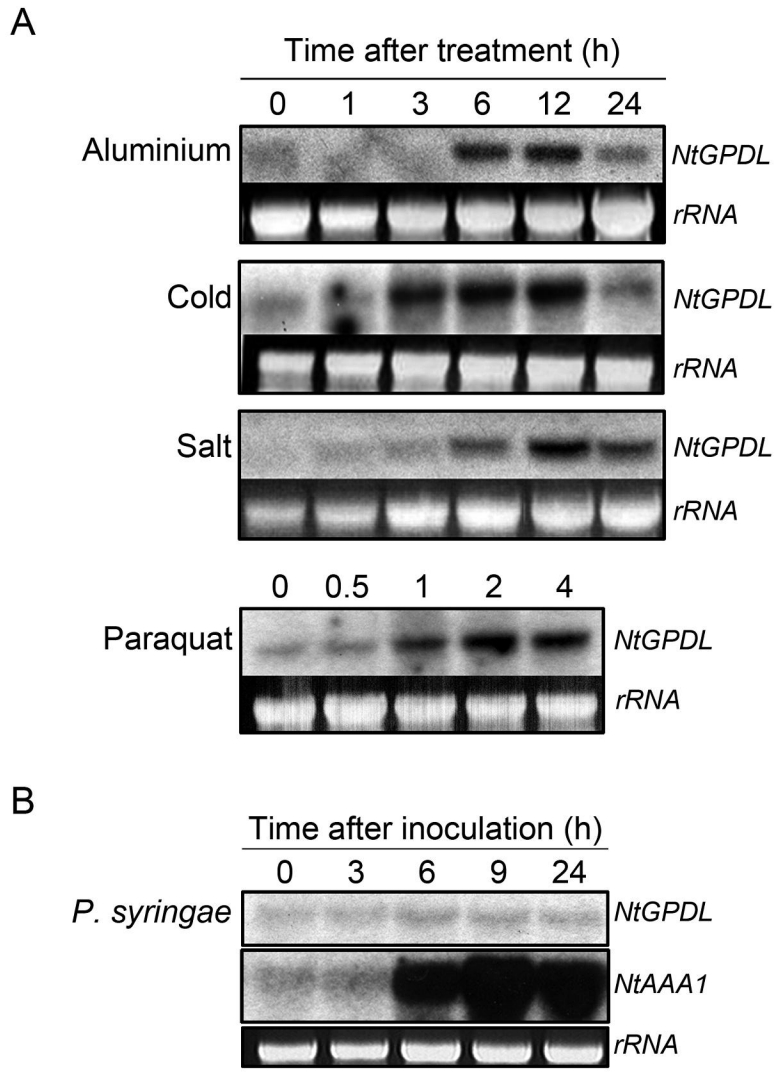


Figure 2. Expression of *NtGPDL* upon stress treatment. A Abiotic stress response. Healthy leaves were detached from wild type mature tobacco plants, and subjected to stress with aluminium chloride (100 μM), cold (4°C), sodium chloride (300 mM) or paraquat (100 μM). Total RNA was isolated at the indicated time points, and subjected to hybridization analyses. A 20-μg aliquot of RNA from leaves after the indicated stress was fractionated on an agarose gel, and subjected to RNA blot hybridization with radioactively labeled full-length *NtGPDL* cDNA as the probe. Equal loading of RNA samples was confirmed by rRNA staining with ethidium bromide. **B** Biotic stress response. Detached healthy leaves were treated with *P. syringae* pv. *glycinea* (2×10^8 cfu/ml) using a syringe without a needle and incubated at 25°C. Total RNA was isolated at the indicated time points after infiltration and subjected to RNA blot hybridization using *NtGPDL* probe (*upper panel*). The successful infection was confirmed by expression of a hypersensitive responsive gene *NtAAA1* (*middle panel*) (Sugimoto et al. 2004). As the loading control, rRNA was used (*lower panel*).

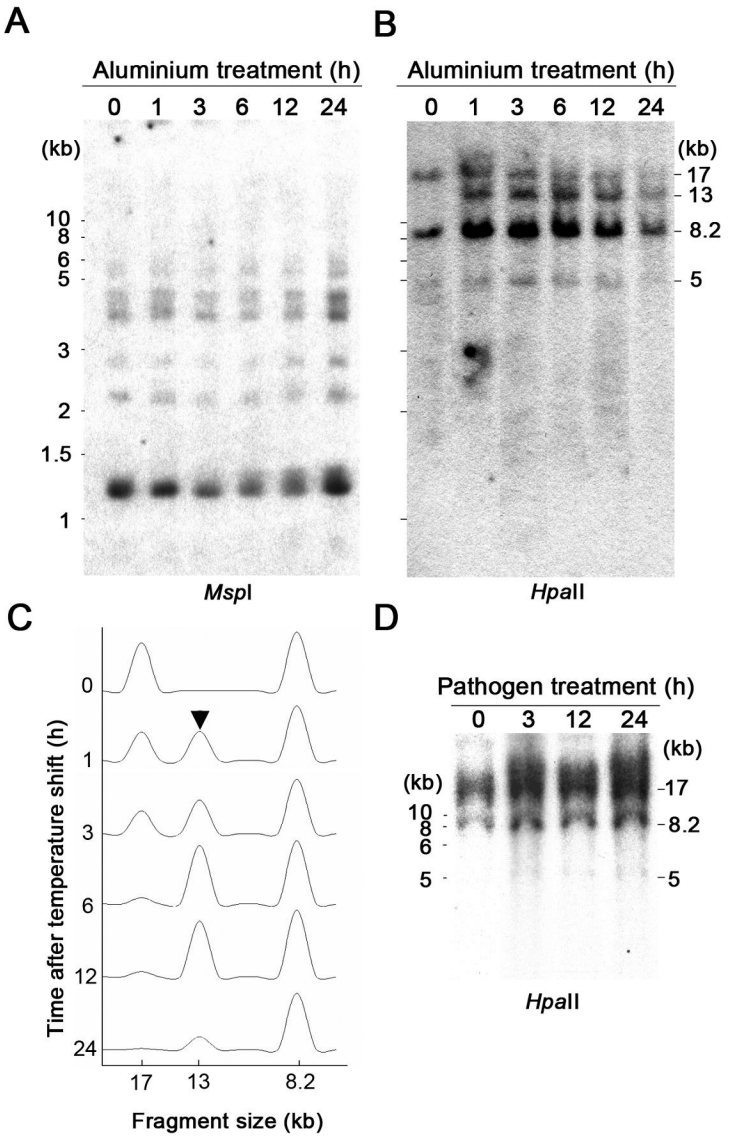


Figure 3. Time course changes in methylation status. A-B Total DNA was isolated from aluminium-treated leaves at the indicated time points, digested with *Msp*I (A) or *Hpa*II (B) for 16 h, and subjected to hybridization with the full length *NtGPDL* probe. Duplicate experiments were performed, and one is shown here. **C** Densitometric quantification of signal intensities observed in (B). The molecular size is shown on the horizontal axis, and the stress treatment time on the vertical axis. Note the progressive increase of the 13 kb fraction as indicated by the *arrowhead*. **D** Methylation status during pathogen infection. Total DNA was isolated from *P. syringae*-treated leaves at the indicated time points, digested with *Hpa*II for 16 h, and subjected to hybridization with the full length *NtGPDL* probe. The position of molecular size marker is indicated at *left*, and that of calculated molecular size of each fragment is indicated at *right* in kb.

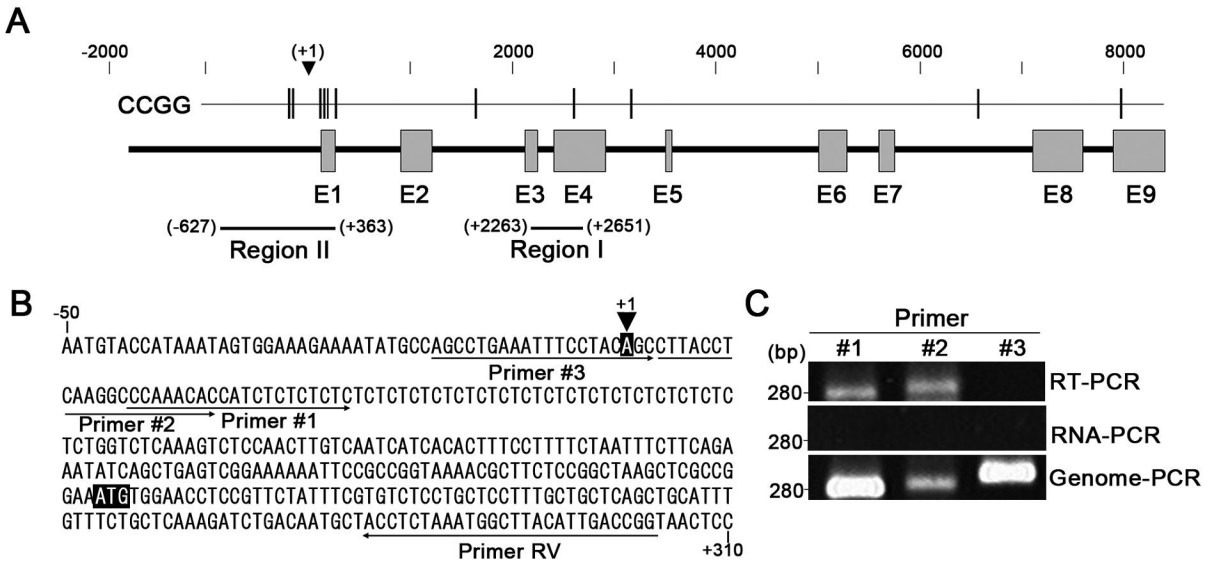
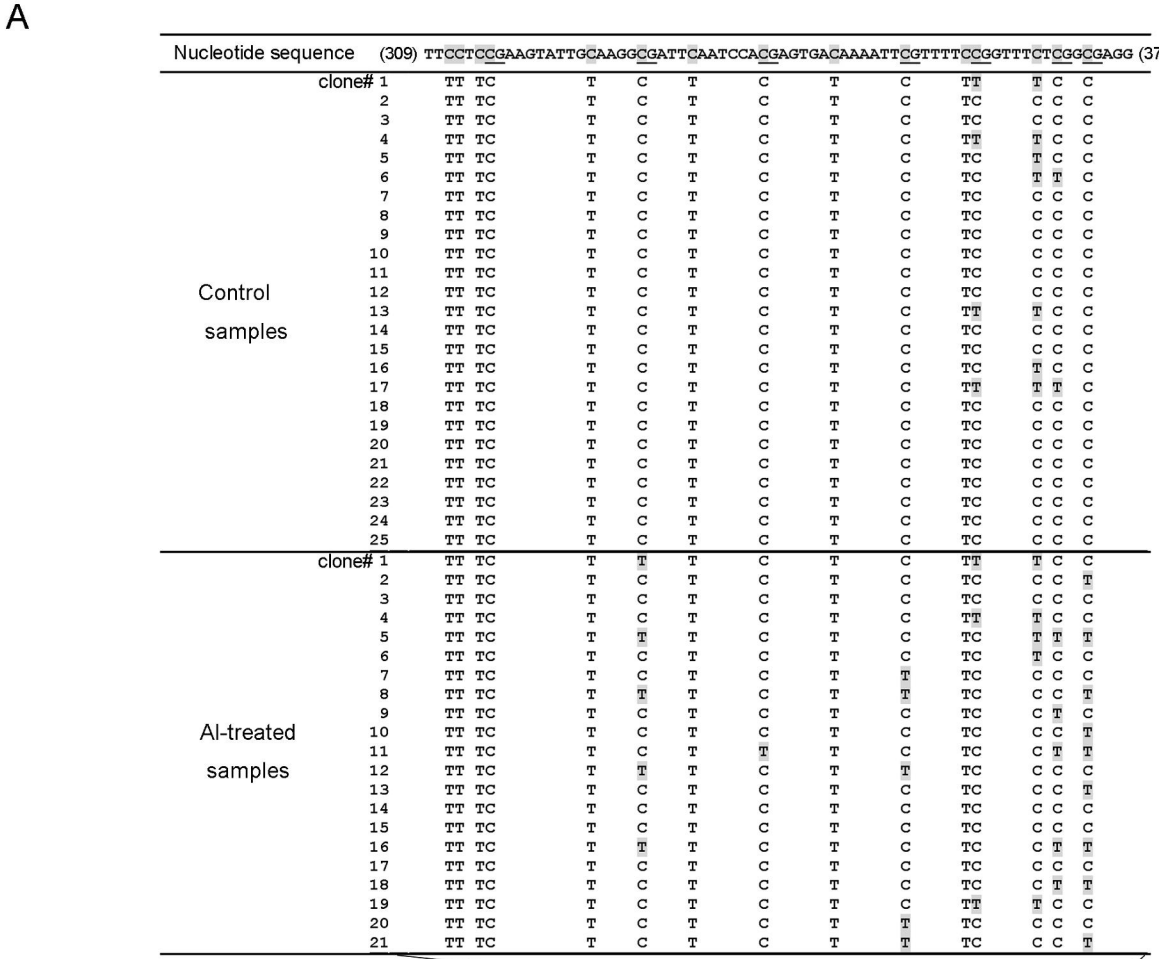


Figure 4. Properties of *NtGPDL* locus. A Genomic organization. A 10 kb genomic locus is illustrated with indication of exons (*box*) and introns (*line*). Exons (*E*) are numbered in Arabic letter starting from the transcription initiation site (+1). CCGG sites are indicated by *vertical bars upper* the sequence and mapping regions are indicated by *horizontal bars*. The regions for direct methylation mapping are located at nucleotide positions 2263-2651 (Region I), and -627 to +363 (Region II). **B** Sequence of the promoter region spanning 360 bases including the ATG site. *Underlines* indicate positions of PCR primers used for determination of transcription initiation site. Three forward (Primer #1 through #3) and one reverse (RV) primers were prepared as shown in Fig. 4c. Nucleotide positions relative to the transcription initiation site are indicated (-50 at the 5'end and +310 at the 3' end). The ATG start codon is shaded, and the putative transcription initiation site is indicated by an *arrowhead*. **C** Total RNA was isolated from aluminium-treated leaves of wild-type plants, subjected to RT-PCR using indicated forward (Primer #1 through #3) and reverse primers (*top panel*, RT-PCR). The possibility of genomic DNA contamination in RNA samples was excluded by no amplification under the same reaction condition without reverse transcription (*middle panel*, RNA-PCR). As the control, genomic DNA was similarly amplified by the same primers (*lower panel*, Genome-PCR). Products were fractionated on gel electrophoresis and visualized with ethidium bromide. Both RT and genomic PCR yielded ca 280 bp fragments (size indicated at the *left side*).

A


Nucleotide sequence (309) TTCCTCCGAAGTATTGCAAGGCGATTCAATCCACGAGTGACAAAATTCGTTTTCCGGTTTCTCGGCGAGG (378)
Control samples
clone# 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
AI-treated samples
clone# 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21

B

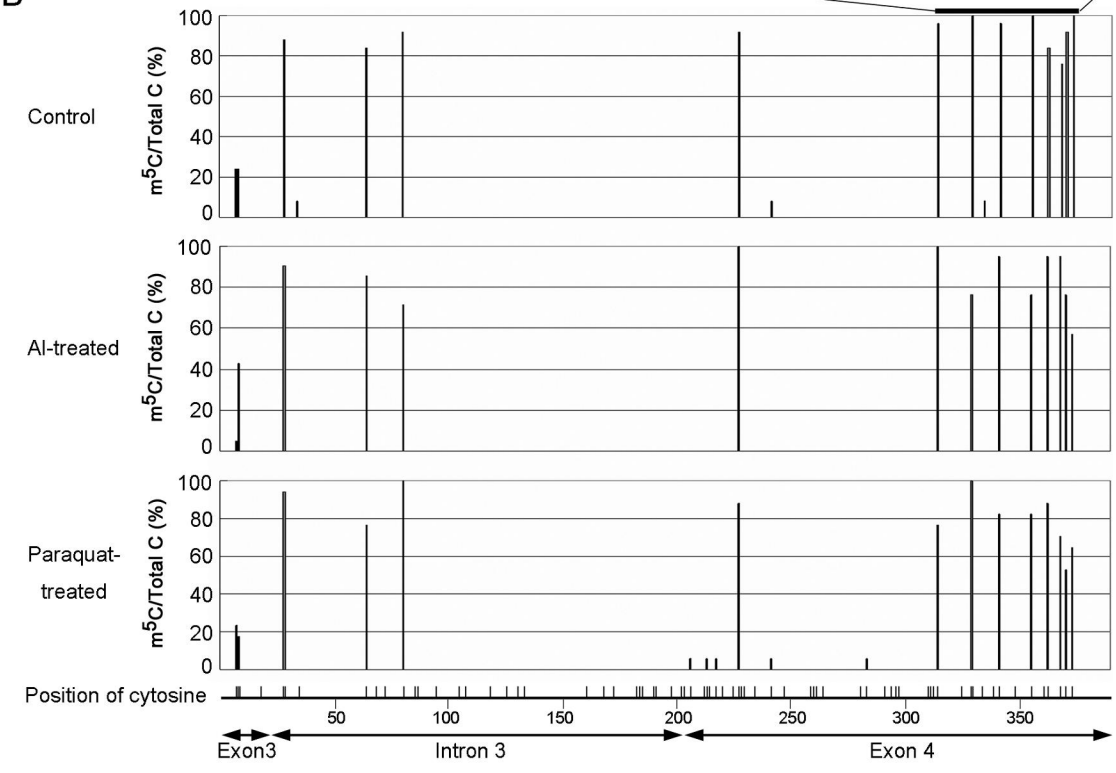

Control
AI-treated
Paraquat-treated
m5C/Total C (%)
0
20
40
60
80
100
Position of cytosine
50
100
150
200
250
300
350
Exon3
Intron 3
Exon 4

Figure 5. Direct methylation mapping. A Identification of m^5C. Healthy leaves were untreated or treated with aluminium for 12 h, and total DNA was extracted and subjected to methylation mapping. After bisulfite treatment, the region I with 389-bp was amplified with PCR, cloned and sequence determined. The nucleotide sequence of 70 bases (positions between 309 and 378) at the 3' end is shown. Cytosines (C) and m^5C in the original sequence were converted into thymine (T) and C, respectively. The original nucleotide sequence of wild-type without bisulfite treatment is shown at the top line with shading of the methylatable cytosines. CG is underlined. Thymines in treated samples which corresponded to cytosines in the control are shaded. Twenty five and 21 clones were analyzed for untreated and aluminium-treated samples, respectively. **B** Distribution and frequency of m^5Cs. Histograms represent the percentages of m^5Cs over the total cytosines in the top strand (*vertical axis*) at the indicated nucleotide positions (*horizontal axis*) in the untreated (*top panel*) and aluminium-treated samples (*middle panel*). The distribution and frequency of m^5Cs in paraquat-treated (4 h) samples are also shown (*bottom panel*). The frame lines on the top indicate the sequence shown in A. All methylatable Cs in are marked by vertical bars over the horizontal axis.

Table 1. Methylation frequency in symmetric and asymmetric sites

Region	Sample	No. of clones	m^5CG/total CG[a]	m^5CNG/total CNG[a]	m^5CNN/totalCNN[a]
I	Control	25	256/275 (93.1)[b,d]	2/300 (0.7)[f]	33/1175 (2.8)[h]
	Al-treated[c]	21	194/231 (84.0)[e]	0/252 (0)[f]	30/987 (3.0)[h]
	PQ-treated[c]	17	155/187 (82.9)[e]	1/204 (0.5)[f]	23/799 (2.9)[h]
II	Control	20	0/440(0)	0/540 (0)	0/3480(0)
	Al-treated[c]	25	0/550(0)	0/675(0)	0/4350(0)

[a] Numbers of CG, CNG and CNN in a single clone were 11, 12 and 47, respectively in Region I, and 22, 27 and 174, respectively in Region II.

[b] The sum of m^5C in the total methylatable Cs in the indicated sequences in a total clones examined are presented as numbers observed and as percentages (%) in parentheses. For example, the 275 CG sites were calculated by multiplying 25 clones by 11 sites of a 389 bases of the Region I.

[c] Al, aluminium; PQ, paraquat.

[d-i] Values with different letters are significantly different at $P<0.05$ by the test of significance.

CHAPTER II

Identification of tobacco genes encoding proteins possessing removal activity of 5-methylcytosines from intact tobacco DNA

Introduction

Eukaryotic DNA contains 5-methylcytosine (m^5C) as a minor base, and it is referred as DNA methylation. DNA methylation has been shown to be involved in gene silencing at both transcriptional and posttranscriptional levels. Transcriptional gene silencing is associated with hypermethylation of promoter sequences, while post-transcriptional gene silencing is linked with hypermethylation of transcribed or coding sequences (Paszkowski and Whitham 2001). DNA methylation is also involved in epigenetics, defined as mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in the DNA sequence (Bird 2002). This typically occurs during somatic cell differentiation in animal cells (Reik et al. 2001), and in plants and filamentous fungi as well (Kakutani et al. 1996, Martienssen and Colot 2001).

Methylation status dynamically changes depending upon cell type, developmental stages and even environmental conditions (Demeulemeester et al. 1999, Steward et al. 2002, Chakrabarty et al. 2003, Wada et al. 2004). Methylation is a post-replication event, catalyzed by DNA methyltransferases which have been relatively well characterized (Bender 2004). In

contrast, demethylation process has largely remained unclear. One of such mechanisms was recently shown to be catalyzed by DNA glycosylases, which excise m^5C from intact DNA and replace it with cytosines (Kapoor et al. 2005). DNA glycosylases recognize and remove a particular subset of base alterations, including demethylation, oxidation, deamination, and alkylation (Lindahl and Wood 1999, McCulloughet al.1999, Sedgwick 2004). A DNA glycosylase specific to m^5C was first identified in chicken embryos (Jost et al. 1995), and an active demethylation pathway initiated by thymine DNA glycosylase (TDG) and/or methyl-CpG binding domain (MBD) protein 4 (MBD4) has been proposed in animal cell (Zhu et al. 2000, Jost et al. 2001, Zhu et al. 2001). Recently, two such DNA glycosylases were identified from *Arabidopsis*, designated as repressor of silencing 1 (ROS1) and DEMETER (DME) (Gong et al. 2002, Choi et al. 2002). Both enzymes remove m^5C at high efficiency through bifunctional DNA glycosylase/lyase mechanism *in vitro* (Gehring et al. 2006, Agius et al. 2006, Morales-Ruiz et al. 2006). They also showed significantly high incision activity in symmetric dinucleotide CpG context, but low activity for G:T mismatch repair. ROS1 was not active against damaged DNA (Agius et al. 2006, Morales-Ruiz et al. 2006). Despite the apparently important function, studies on ROS1/DME from other plants species have so far been limited. In this report, I describe isolation and characterization of four ROS1-like proteins from tobacco plants.

Materials and methods

Plant materials

Wild-type tobacco plants (*Nicotiana tabacum* cv. Xanthi nc) were grown in a greenhouse at 23˚C under a 14/10 h light/dark photo cycle for 1 to2 months. Stress treatments were given as described (Taura et al. 2003). Briefly, detached healthy leaves were detached from wild type mature tobacco plants, and subjected to stress treatments with aluminium chloride (100 μM), sodium chloride (300 mM) or paraquat (100 mM).

Construction and Screening of a cDNA Library

Total RNA was extracted from whole tobacco plants (*Nicotiana tabacum* cv. Xanthi nc) by the AGPC method (Nakano et al. 2000). The mRNA was isolated from total RNA by using a PolyATtract mRNA isolation system III (Promega, Madison, WI, USA) according to the instructions provided. A cDNA library was constructed from five microgram of mRNA using the ZAP cDNA synthesis kit (Stratagene, La Jolla, CA, USA) and Uni-ZAP XR vector, which was packaged in vitro with Gigapack III Gold packaging extract (Stratagene). The amplified library contained approximately 2.9×10^6 recombinant plaques, 5×10^4 from which were plated and transferred to nylon membrane (Hybond N^-, Amersham, Buckinghamshire, UK). Membranes were screened by differential hybridization using ^{32}P-labeled probes.

Tobacco cDNA fragments encoding the m^5C DNA glycosylase domain were obtained by reverse transcriptase (RT)-PCR. PCR primers were designed based on the sequence of plant DNA glycosylase genes in GenBank (accession AAM77215 and AAP37178): DNG1 forward, 5'-CTTACACAGATTGTCTCGGA-3'; DNG1 reverse, 5'-GCTGTTTCACCTGGTG TCCATATA-3'; DNG2 forward, 5'-CTTACACAGATTGTCTCGGA-3'; DNG2 reverse, 5'-GC TGTTTCACCTGGGGTCCATATA-3'. The size of amplified fragments was 2.2 kb, which encoded the m^5C DNA glycosylase domain found in AtROS1 and AtDME. These fragments were used to screen the cDNA library as the probe to obtain the full-length cDNAs.

Prehybridization and hybridization were performed at 60°C in 5× SSC containing 5× Denhardt's solution, 50% formamide, 0.5% SDS, and 0.5% salmon sperm DNA. Filters were washed in 0.1% standard saline citrate (SSC) with 0.1% SDS at 65°C. Positive plaques were isolated and subjected to a second screening under the same conditions. Positive plaques from this second screening were excised in vivo according to the protocols of the manufacturer (Stratagene). The longest plaques of insert, confirmed by PCR-amplified with primers T7 and T3. The nucleotide sequences were determined by the dideoxynucleotide chain termination method (PRISM BigDye Terminator, ABI, Sunnyvale, CA, USA), and analyzed with appropriate computer software.

Expression of *GST:NtROS fusion* proteins and Enzyme assay

Proteins were expressed in vitro using *Spodoptera frugiperda* (Sf9) insect cell line as described previously (Wada et al. 2003). Cells from one culture dish were suspended in 1 ml of lysis buffer (100 mM Hepes, pH 8.0, 240 mM KCl, 1% glycerol, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride), and sonicated twice for 10 sec each. GST-fused NtROS1 and NtROS2a were purified through a glutathione-Sepharose column (Amersham). Approximately 10 μg of genomic DNA were incubated at 37°C for 2 h in a reaction mixture containing 10 mM Bis-Tris propane-HCl (pH 7.0), 10 mM $MgCl_2$, 1 mM dithiothreitol, 0.1 mg mL^{-1} BSA, 1 mM dNTP, 50 pmol of protein in a total volume of 100 μl. Reactions were stopped by adding EDTA to 20 mM, and SDS to 0.6%, and the mixtures were incubated for 10 min at 70°C. DNA was extracted with phenol/chloroform/isoamyl alcohol and ethanol-precipitated. The methylation status was examined by digesting 0.5 μg DNA with *Hpa*II or *Msp*I restriction endonucleases, followed by agarose gel electrophoresis (Gong et al. 2002).

Epifluorescence analysis

GFP fusions were constructed as described (Yap et al. 2005). The entire coding region of *NtROS1* and *NtROS2a* were amplified by PCR. The attB-PCR products were cloned into pGWB6(35S-sGFP-NOS) vector to fuse to green fluorescence protein (GFP) using the

GATEWAY cloning system (Invtrogen, Carlsbad, CA, USA). As the control, a plasmid expressing only GFP was used. Each plasmid was introduced into onion epidermis cell by biolistic bombardment, and observed for localization of GFP with a confocal microscope.

Northern Hybridization

RNA blot analysis was performed using total RNA extracted from samples by the AGPC method (Nakano et al. 2000). Approximately 10-μg of RNA was separated by electrophoresis, transferred to N⁻nylon membrane, and subjected to hybridization with radioactively labeled probes as described (Yoda et al. 2002).

Results

Identification and sequence analysis of *NtROS* genes

Two cDNA fragments were initially isolated tobacco leaves RNA by reverse transcriptase-PCR using a set of degenerate primers which were designed based on the conserved DNA glycosylase domain of known plant 5-methylcytosine DNA glycosylase proteins (http://www.chromdb.org). After intensive screening of the cDNA library, 13 positive plaques were initially obtained, among which seven clones were finally selected and the sequence determined. They were classified into four polymorphic types, consisted of 5940,

5879, 5878 and 5651 bp fragments, and designated as ***N**icotiana **t**abacum* **R**epressor **O**f **S**ilencing 1 (*NtROS1*), *NtROS2a*, *NtROS2b*, and *NtROS3,* respectively (Figure 1A) (accession numbers; AB281587 for *NtROS1*, AB281588 for *NtROS2a*, AB281589 for *NtROS2b*, and AB281590 for *NtROS3*). Each clone encoded a protein of 1796, 1673, 1673, and 1662 amino acids with predicted molecular mass of 201, 186, 186, and 185 kDa, respectively. A pair-wise comparison of their deduced amino acid sequences with AtROS1 showed the similarity of NtROS1 at 38.9%, NtROS2a at 32.2%, NtROS2b at 32.1% and NtROS3 at 31.3%. However, the similarity at the glycosylase domain with AtROS1 was 94.3% for NtROS1 (amino acid positions, 1291-1459), 94.3% for NtROS2a (amino acid positions, 1291-1459), 94.3% for NtROS2b (amino acid positions, 1291-1459) and 90.1% for NtROS3 (amino acid positions, 1305-1433). A pair-wise comparison with NtROS1 indicated the similarity of NtROS2a, NtROS2b and NtROS3 to be 51.1, 51.2 and 49.3%, respectively. Notably, only five amino acids out of total 1673 amino acids (0.4%) differed between NtROS2a and NtROS2b, suggesting them to be isoforms. Conserved lysine and aspartic acid residues located in the helix-hairpin-helix domain in DNA glycosylases (Krokan et al. 1997) and four cysteine residues adjacent to the DNA glycosylase domain were also present in NtROS proteins. The latter cluster is thought to play a role in DNA binding. NtROS proteins also possess nuclear localization signal and histone deacetylase interacting domains (http://smart.embl-heidelberg.de/) (Figure 1A), suggesting them to localize in nucleus and to interact with DNA or chromatin (Gehring et al. 2006).

Methylation-sensitive restriction enzyme assay

Full-length NtROS1 and NtROS2a as glutathione S-transferase (GST) fusion proteins were expressed in baculovirus-mediated expression system in *Spodoptera frugiperda* (Sf9) insect cell line. After 72 h post-infection, cell extracts were sampled, and protein was purified through a glutathione-Sepharose column and analyzed by SDS-PAGE, showing that full-length (ca. 300 kDa) proteins were successfully synthesized (Figure 2A). In vitro demethylation activity was then examined. Tobacco genomic DNA was incubated with or without purified NtROS1 or NtROS2a, purified with phenol extraction and ethanol precipitation, digested with either *Msp*I or *Hpa*II, and fractionated by agarose gel electrophoresis (Figure 2B). The untreated control sample was highly resistant to *Hpa*II but susceptible to *Msp*I, indicating hypermethylation of the second cytosine at CCGG sites. In contrast, samples treated with NtROS1 or NtROS2a were digested by *Hpa*II to the same extent as those by *Msp*I (Figure 2B). Since ROS-treated, but *Hpa*II/*Msp*I-untreated samples were not degraded at all, these results clearly indicated an efficient removal of m^5C in CCGG sites.

Cellular localization

Cellular localization of NtROS1 and NtROS2a was examined using fusion proteins with a green fluorescence protein (GFP) reporter under the control of the cauliflower mosaic virus

(CaMV) 35*S* promoter. Results showed that both proteins were exclusively localized in nucleus (Figure 3). This is consistent with a computational prediction by the Predotar (version 1.03 at http://urgi.infobiogen.fr/predotar/predotar.html), PSORT (Nakai and Kanehisa 1992), and iPSORT (Bannai et al. 2002) programs, which suggested their nuclear localization.

Accumulation of *NtROS* genes upon abiotic stress treatment

Their expression profile was then examined. Transcripts for *NtROS1* and *NtROS2a/2b* in young leaves were scarcely detectable under non-stressed conditions (Figure 4). When subjected to aluminium stress, transcripts for *NtROS1* and *NtROS2a/2b* began to accumulate 24 h and 6 h later, respectively. Salt stress also induced their transcripts after 1 h, reaching maximal levels 3 h later and then gradually declining to the initial level. Similarly, paraquat treatment induced their transcripts within 2 h. These observations indicated that *NtROS1* and *NtROS2a/2b* specifically and temporarily respond to abiotic environmental stresses, and that such response might be mediated through reactive oxygen species. Transcript accumulation profile of *NtROS3* was essentially similar with that of *NtROS1*, except that the amount appeared to be lower (Figure 4).

Discussion

DNA glycosylases are subdivided into four groups according to their substrate specificity:

excision of uracil, uracil-containing mismatches, alkylated bases and oxidized bases (Fortini et al. 2003). NtROS proteins are structurally related to *Arabidopsis* ROS1 and DME, which belong to a large family of uracil-containing mismatch DNA glycosylases. However, apart from the m^5C DNA glycosylase domain and regions at the carboxyl-terminus, their similarity was low, suggesting specific function of NtROS proteins. Since NtROS1 and NtROS2a were shown to possess m^5C removal activity, and to be localized in nucleus, it is conceivable that they are closely associated with DNA and engaged in DNA demethylation. In this context, their induction profile upon abiotic stresses is indicative, as DNA glycosylases are considered to deal with oxidative damages, which are commonly caused by reactive oxygen species (Dizdaroglu 2003, Fortini et al. 2003). To cope with such damages, a set of genes must be expressed to repair and/or strengthen defense reactions. If these genes are locked by methylation under non-stressed condition, demethylation is absolutely necessary to globally switch on them. It is tempting to speculate that NtROS proteins participate in this task, thereby finely adjusting gene expression through methylation/demethylation pathways.

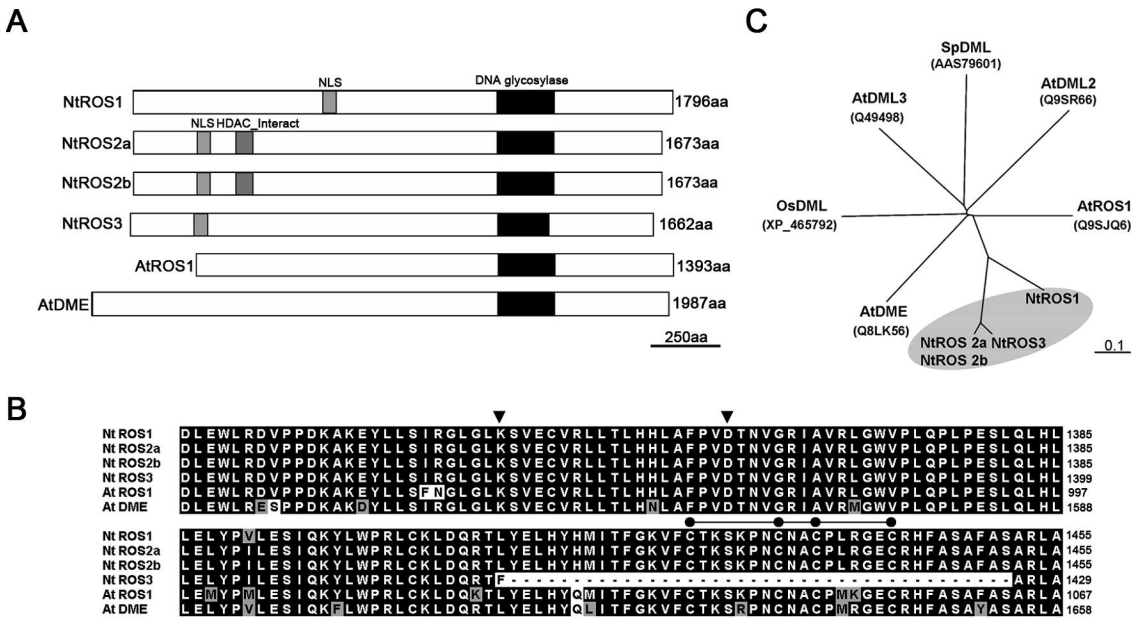


Figure 1. Properties of NtROS proteins. A Structural illustration of m^5C DNA glycosylase proteins. Protein size is shown in number of amino acids (aa), and conserved DNA glycosidase domain is indicated by *closed box*, and specific regions including histone deacetylase interacting (HDAC interact), nuclear localization signal (NLS) are indicated by *shaded boxes*. **B** Alignment of conserved m^5C DNA glycosylase domain of *NtROS1*, *NtROS2a*, *NtROS2b*, *NtROS3*, *AtROS1* and *AtDME*. Common amino acids among five proteins are indicated by *closed box*, and *grey boxes* indicate conservative changes. Conserved lysine and aspartic acid residues in the glycosylase/lyase active site (InterPRO search, IPRO11257) are indicated by *closed arrowheads*, and four cysteine residues predicted to function in DNA binding are indicated by *closed circles*. **C** Unrooted phylogenetic tree. Amino acid sequences were aligned using the ClustalW program (http://www.ddbj. nig. ac.jp/search/clustalw-j.html). Genebank accession numbers are: *NtROS1*, AB281587; *NtROS2a*; AB281588, *NtROS2b*, AB281589; *NtROS3*, AB281590 (N. *tabacum)*; *AtROS1*, Q9SJQ6; *AtDME*, Q8LK56; *AtDML2*, Q9SR66; *AtDML3*, Q49498 (A. *thaliana*); *OsDML*, XP_475792 (O. *sativa*); *SpDML*, AAS79601 (*Ipomea trifida*).

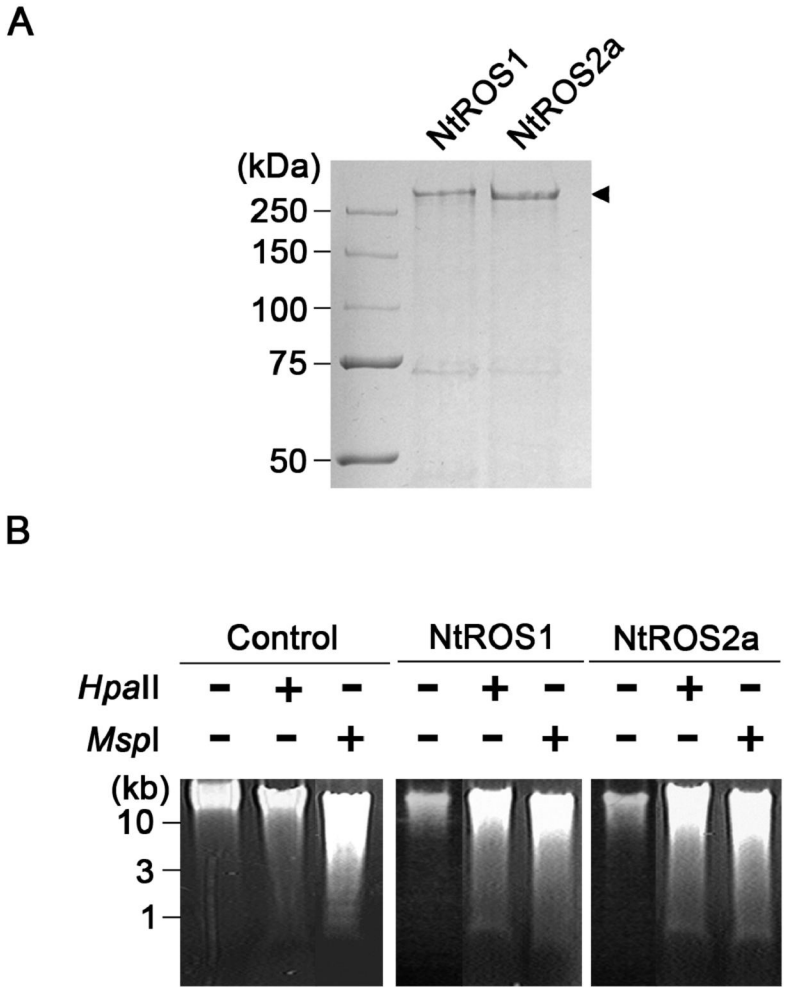


Figure 2. Enzymatic activity. A Expression of NtROS1 and NtROS2a. Corresponding full length cDNA fused to *GST* was expressed in Sf9 cells, and resulting proteins were purified through glutathione-Sepharose column. An appropriated aliquot was fractionated by SDS-PAGE (10%), and stained with Coomassie Brilliant Blue (CBB). The position of marker proteins is indicated at *left*. *Closed arrowhead* indicates expressed proteins with 300 kDa. **B** Removal of m^5C. Tobacco genomic DNA was treated with purified NtROS1 (indicated as NtROS1) or NtROS2a (NtROS2a) at 37°C for 2 h, extracted by phenol/chloroform, subjected to digestion with either *Hpa*II or *Msp*I as indicated by *plus* (+) or *minus* (-), fractionated on 1.2% agarose gel electrophoresis, and stained by ethidium bromide. Untreated DNA was used as the control (Control). The position of molecular marker is indicated at *left*.

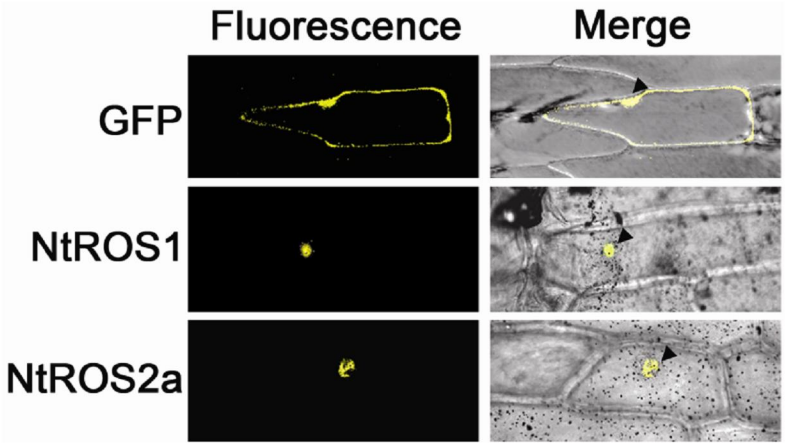


Figure 3. Cellular localization. Onion epidermal cell layers were bombarded with gold particles coated with 35S*NtROS1-GFP*, 35S*NtROS2a-GFP* or 35S*GFP* alone, and observed by epifluorescence for GFP. Bright field mages are merged to show positions of nucleus indicated by *arrowheads*.

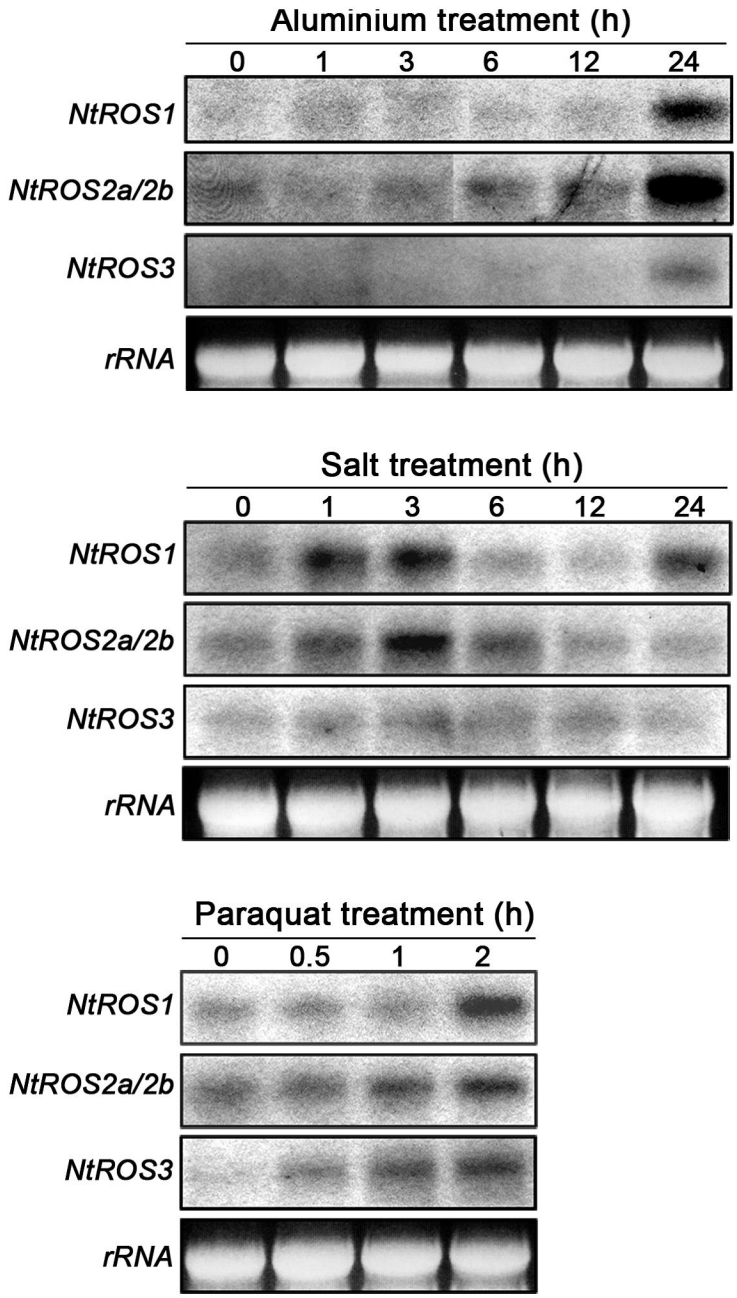


Figure 4. Transcript accumulation of *NtROS* upon stress treatment. Healthy leaves were detached from wild type mature tobacco plants, and petioles were soaked in solution containing aluminium chloride (100 μM), sodium chloride (300 mM) or paraquat (100 μM). Leaf samples were also exposed to cold (4°C). After incubated for indicated time period, total RNA was isolated and a 10-μg aliquot was fractionated on an agarose gel, and subjected to RNA blot hybridization with radioactively labeled probes indicated. Equal loading of RNA samples was confirmed by rRNA staining with ethidium bromide

CONCLUDING REMARKS

Cytosines in eukaryotic DNA are often methylated to yield 5-methylcytosines (m^5C), which play an important role in controlling gene expression. This is referred as DNA methylation, and its status dynamically changes during plant growth by active methylation and demethylation.

In the first part of this thesis, I discuss the relationship between gene expression and DNA methylation upon environmental stress in tobacco. Mutations in the tobacco *MET1* gene cause a global reduction of cytosine methylation throughout the genome and a number of genes expression changes. The antisense *MET1* tobacco plants showed a dwarf phenotype and ectopic expression of genes that are silent in wild type plants without environmental stressing. One such gene was abiotic stresses responding gene (*NtGPDL*). Upon exposure to abiotic stress, transcripts of *NtGPDL* were found to be accumulated and genomic locus was partly demethylated, while upon biotic stress, neither transcripts nor demethylation were induced. Direct methylation mapping showed that demethylation predominantly took place at CG doublets, which were clustered in exon but rare in intron. *NtGPDL* is a heavy body-methylated, but promoter unmethylated, and is highly regulated to specifically respond to external stimuli. In wild-type *Arabidopsis*, many genes are methylated to a higher degree in their transcribed regions than in promoter regardless expression. Of all expressed genes, 95%

are unmethylated in the promoter and 33% are methylated in transcribed region, with a bias toward the 3' portion of the transcribed region. A comparison with accumulated expression data indicates that body-methylated genes tend be expressed at high levels with low tissue specificity, whereas promoter methylated genes are expressed at lower levels in more tissue-specific manner (Zhang et al. 2006). Although the genome-wide methylation mapping suggested that constitutively active genes were heavily body-methylated, cases such as *NtGPDL* were not exemplified. Since cytosine methylation was shown to be closely linked with methylation of histones (Tariq and Paszkowski 2004, Fuchs et al. 2006), demethylation within *NtGPDL* locus might affect chromatin architecture. I speculate that stress-induced demethylation of coding regions may relax chromatin structure, thereby allowing enhanced transcription.

I attempted to identify genes encoding a 5-methylcytocine DNA glycosylase from tobacco plants. Recent studies on *Arabidopsis* suggested DNA glycosylase to be one of the responsible enzymes, with participation in gene silencing by demethylating target genes, such as RD29A gene and other genes (Gong et al. 2002, Agius et al. 2006, Zhu et al. 2007). As described in chapter II, I isolated and characterized four genes encoding ROS1-like proteins from tobacco plants, designated NtROS. NtROS proteins are structurally related to *Arabidopsis* ROS1 and DME, which belong to a large family of uracil-containing mismatch DNA glycosylases. However, apart from the m^5C DNA glycosylase domain and regions at the carboxyl-terminus,

little similarity was observed among them suggesting specific function of NtROS proteins. Since NtROS1 and NtROS2a were shown to possess m^5C removal activity, and to be localized in nucleus, it is conceivable that they are closely associated with DNA and engaged in DNA demethylation. Efforts are also needed to investigate the mechanism of targeting the demethylase to specific loci. It is possible that the demethylase enzyme functions in a complex containing chromatin related proteins that help target the enzyme to specific methylated genes. Transcripts of *NtROS1*, *NtROS2*a, and *NtROS3* were induced by abiotic stresses and reactive oxygen species. These observations suggested that NtROS proteins function in demethylating process of genomic DNA during plant stress response, thereby maintaining the balance of the gene expression in combination with DNA methytransferases. The observed differential methylation at the *NtGPDL* locus may thus reflect crosstalk among these enzymes, whereby fine tuning of the state of de/methylation is achieved. Although my analysis of *NtROS* and *NtGPDL* gene were not complete, I hope that such a study will be of a help for understanding of control of gene expression by DNA de/methylation in plants.

Overall, results of the current study suggested that environmental response of plants is partly mediated through active alteration of DNA methylation status, which changes chromatin structure and eventually gene expression.

REFERENCES

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

Alina, R., Sgorbati, S., Santagostino, A., Labra, M., Ghiani, A., and Citterio, S. (2004) Specific hypomethylation of DNA is induced by heavy metals in white clover and industrial hemp. Physiol. Plant 121, 472-480.

Apel, K., and Hirt, H. (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu. Rev. Plant Biol. 55, 373-399.

Babiushin, B.F. (2006) DNA methylation in plants. Curr. Top. Microbiol. Immunol. 301, 67-122.

Bannai, H., Tamada, Y., Maruyama, O., Nakai, K., and Miyano, S. (2002) Extensive feature detection of N-terminal protein sorting signals. Bioinformatics 18, 298-305.

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

Bender, J. (2004) DNA methylation and epigenetics. Ann. Rev. Plant Biol. 55, 41-68.

Bird, A.P. (2002) DNA methylation patterns and epigenetic memory. Genes Dev. 16, 6-21.

Borner, G.H.H., Sherrier, D.J., Stevens, T.J., Arkin, I.T., and Dupree, P. (2002) Prediction of

glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A genomic analysis. Plant Physiol. 129, 486-499.

Borner, G.H.H., Lilley, K.S., Stevens, T.J., and Dupree, P. (2003) Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. Plant Physiol. 132, 568-577.

Boscolo, P.R.S., Menossi, M., and Jorge, R.A. (2003) Aluminum-induced oxidative stress in maize. Phytochem. 62, 181-189.

Bray, E.A., Bailey-Serres, J., and Wetertilnyk, E. (2000) Responses to abiotic stresses. In: Buchanan, B.B., Gruissem, W., and Jones, R.L. (eds) Biochemistry and molecular Biology of Plants. American Society of Plant Physiologists, Rockville, pp. 1158-1203.

Brzeski, J., and Jerzmanowski, A. (2003) Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. J. Biol. Chem. 278, 823-828.

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

Cerda, S., and Weitzman, A.S. (1997) Influence of oxygen radical injury on DNA methylation. Mut. Res. 386, 141-152.

Chakrabarty, D., Yu, K.W., Paek, K.Y. (2003) Detection of DNA methylation changes during

somatic embryogenesis of Siberian ginseng *(Eleuterococcus senticosus)*. Plant Sci. 165, 61-68.

Chan, SW.L., Henderson, I.R., and Jacobsen, S.E. (2005) Gardening the genome: DNA methylation in *Arabidopsis thaliana*. Nat. Rev. Genet. 6, 351-60.

Choi, Y., Harada, J.J., Goldberg, R.B., and Fischer, R.L. (2004) An invariant aspartic acid in the DNA glycosylase domain of DEMETER is necessary for transcriptional activation of the imprinted MEDEA gene. Proc. Natl. Acad. Sci. USA 101, 7481-7486.

Chomczynski, P., and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem. 162, 156-159.

Delaval, K., and Feil, R. (2004) Epigenetic regulation of mammalian genomic imprinting. Curr. Opin. Genet. Dev. 14, 188-195.

Demeulemeester, M.A.C., Van Stallen, N., and De Proft, M.P. (1999) Degree of DNA methylation in chicory (*Cichorium intybus* L.): influence of plant age and vernalisation. Plant Sci. 142, 101-108.

Ehrlich, M. (2000) DNA hypomethylation and cancer. In: Ehrlich M (ed) DNA Alteration in Cancer. Eaton Publishing, Natick, pp. 273-291.

Finnegan, E.J., Peacock, W.J. and Dennis, E.S. (1996) Reduced DNA methylation in

Arabidopsis thaliana results in abnormal plant development. Proc Natl Acad Sci USA 93, 8449-8454.

Finnegan, E.J., Genger, R.K., Peacock, W.J., and Dennis, E.S. (1998) DNA methylation in plants. Ann. Rev. Plant Physiol. Plant Mol. Biol. 49, 223-247.

Finnegan, E.J., and Kovac, K.A. (2000) Plant DNA methyltransferases. Plant Mol. Biol. 43, 189-201.

Fuchs, J., Demidov, D., Houben, A., and Schubert, I. (2006) Chromosomal histone modification patterns-from conservation to diversity. Trends Plant Sci. 11, 199-208.

Fuks, F., Hurd, P.J., Deplus, R., and Kouzarides, T. (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res. 31, 2305-2312.

Galaud, J.P., Gaspar, T., and Boyer, N. (1993) Inhibition of internode growth due to mechanical stress in *Bryonia dioica*: relationship between changes in DNA methylation and ethylene metabolism. Physiol. Plant 87, 25-30.

Gaudet, F., Hodgson, J., Eden, A., Jackson-Grusby, L., Dausman, J., Gray, J.W., Leonhardt, H., and Jaenisch, R. (2003) Induction of tumors in mice by genomic hypomethylation. Science 300, 489-492.

Gendall, A.R., Levy, Y.Y., Wilson, A., Dean, C. (2001) The *VERNALIZATION 2* gene mediates the epgenetic regulation of vernalization in *Arabidopsis*. Cell 107,

525-535.

Gendrel, A.V., Lippman, Z., Yordan, C., Colot, V., and Martienssen, R.A. (2002) Dependence of heterochromatic histone H3 methylation patterns on the *Arabidopsis* gene *DDM1*. Science 297, 1871-1873.

Gehring, M., Huh, J.H., Hsieh, T.F., Penterman, J., Choi, Y., Harada, J.J., Goldberg, R., and Fischer, R.L. (2006) DEMETER DNA glycosylase establishes *MEDEA* polycomb gene self-imprinting by allele-specific demethylation. Cell 124, 495-506.

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

Goll, M.G., and Bestor, T.H. (2005) Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. 74, 481-514.

Gonzalez, R.G., Haxo, R.S., and Schleich, T. (1980) Mechanism of action of polymer aurintricarboxylic acid, a potent inhibitor of protein-nucleic acid interaction. Biochem. 19, 4299-4303.

Gruenbaum, Y., Naveh-Many, T., Cedar, H., and Razin, A. (1981) Sequence specificity of methylation in higher plant DNA. Nature 292, 860-862.

Holliday, R., and Pugh, J.E. (1975) DNA modification mechanisms and gene activity during development. Science 187, 226-232.

Ito, T., Sakai, H., and Meyerowitz, E.M. (2003) Whorl-specific expression of the

SUPERMAN gene of *Arabidopsis* is mediated by *cis* elements in the transcribed region. Curr. Biol. 13, 1524-1530.

Jackson, J.J., Lindroth, A.M., Cao, X.F., and Jacobsen, S.E. (2002) Control of CpNpG DNA methylation by KRYPTONITE histone H3 methyltransferase. Nature 416, 556–560.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E., and Jaenisch, R. (2001) Loss of genomic methyaltion causes p53-dependent apoptosis and epigenetic deregulation. Nat Genet 27, 31-39.

Jeddeloh, J.A., Stokes, T.L., and Richards, E.J. (1999) Maintenance of genomic methylation requires a SWI2/SNF2-like protein. Nat. Genet. 22, 94-97.

Jones, L., Hamilton, A.J., Voinnet, O., Thomas, C.L., Maule, A.J., and Baulcombe, D.C. (1999) RNA-DNA interactions and DNA methylation in post-transcriptional gene silencing. Plant Cell 11, 2291-2301.

Jost, J.P., Oakeley, E.J., Zhu, B., Benjamin, D., Thiry, S., Siegmann, M., and Jost, Y.C. (2001) 5-Methylcytosine DNA glycosylase participates in the genome-wide loss of DNA methylation occurring during mouse myoblast differentiation. Nucleic Acids Res. 29, 4452-4461.

Kaeppler, S.M., and Phillips, R.L. (1993) Tissue Culture-Induced DNA Methylation Variation in Maize. Proc. Natl. Acad. Sci. USA 90, 8773-8776.

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

Kapoor, A., Agius, F., and Zhu, J.K. (2005) Prevent transcriptional gene silencing by DNA demethylation. FEBS Lett. 579, 5889-5898.

Kuhnlein, U., and Arber, W. (1972) the role of nucleotide methylation in *in vitro* B-specific modification. J. Mol. Biol. 63, 9-19.

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

Ko, Y.G., Nishino, K., Hattori, N., Arai, Y., Tanaka, S., and Shiota, K. (2005) Stage-by-stage change in DNA methylation status of *Dnmt1* locus during mouse early development. J. Biol. Chem. 280, 9627-9634.

Krokan, H.E., Standal, R., and Slupphaug, G. (1997) DNA glycosylases in the base excision repair of DNA. Biochem. J. 325, 1-16.

Labra, M., Ghiani, A., Citterio, S., Sgorbati, S., Sala, F., Vannini, C., Ruffini-Castiglione, M., and Bracale, M. (2002) Analysis of cytosine methylation pattern in response to water deficit in pea root tips. Plant Biol. 4, 694-699.

Li, E., Bestor, T.H. and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase

gene results in embryonic lethality. Cell 69, 915-926.

Lindahl, T., and Wood, R.D. (1999) Quality control by DNA repair. Science 286, 1897-1905.

McCullough, A. K., Dodson, M. L., and Lloyd, R.S. (1999) Initiation of base excision repair: glycosylase mechanisms and structures. Annu. Rev. Biochem. 68, 255-285.

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

Lippman, Z., Gendrel, A-V., Black, M., Vaughn, M.W., Dedhia, N., McCombie, W.R., Lavine, K., Mittal, V., May, B., Kasschau, K.D., Carrington, J.C., Doerge, R.W., Colot, V., and Martienssen, R. (2004) Role of transposable elements in heterochromatin and epigenetic control. Nature 430, 471-476.

Lizal, P., and Relichova, J. (2001) The effect of day length, vernalization and DNA demethylation on the flowering time in *Arabidopsis thaliana*. Physiol. Plant 113, 121-127.

Martienssen, R., and Colot, V. (2001) DNA methylation and epigenetic inheritance in plants and filamentous fungi. Science 293, 1070-1074.

Minamoto, T., Mai, M., and Ronai, Z. (1999) Environmental factors as regulators and effectors of multistep carcinogenesis. Carcinogenesis 20, 519-527.

Miura, A., Yonebayashi, S., Watanabe, K., Toyama, T., Shimada, H., and Kakutani, T. (2001)

Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. Nature 411, 212-214.

Morales-Ruiz, T., Ortega-Galisteo, A.P., Ponferrada-Marin, M.I., Martinez-Macias, M.I., and Ariza, R.R., and Roldan-Arjona, T. (2006) *DEMETER* and *REPRESSOR OF SILENCING 1* encode 5-methylcytosine DNA glycosylases. Proc. Natl. Acad. Sci. USA 103, 6853-6858.

Murray, M.G., and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8, 4321-4325.

Nakano, Y., Steward, N., Koizumi, N., and Sano, H. (2000) A tobacco NtMET1 encoding a DNA methyltransferase: molecular characterization and abnormal phenotypes of antisence transgenic tobacco plants. Plant Cell Physiol. 41, 448-457.

Nakai, K., and Kanehisa, M. (1992) A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics 14, 897-911.

Ng, H.H., and Bird, A. (1999) DNA methylation and chromatin modification. Curr. Opin. Genet. Dev. 9, 158-163.

Papa, C.H., Springer, N.M., Muszynski, M.G., Meeley, R., and Kaeppler, S.M. (2001) Maize chromomethylase Zea methyltransferase2 is required for CpNpG methylation. Plant Cell 13, 1919-1928.

Paszkowski, J., and Whitham, S.A. (2001) Gene silencing and DNA methylation processes.

Curr. Opin. Plant Biol. 4, 123-129.

Peters, A.H., and Schübeler, D. (2005) Methylation of histones: playing memory with DNA. Curr. Opin. Cell Biol. 17, 230-238.

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

Reik, W., Dean, W., and Walter, J. (2001) Epigenetic reprogramming in mammalian development. Science 293, 1089-1093.

Riggs, A.D. (1975) X inactivation, differentiation, and DNA methylation. Cytogenet. Cell Genet. 14, 9-25.

Ronemus, M.J., Galbiati, M., Ticknor, C., Chen, J. and Dellaporta, S.L. (1996) Demethylation-induced developmental pleiotropy in *Arabidopsis*. Science 273, 654-657.

Rountree, M.R., Bachman, K.E., and Baylin, S.B. (2000) DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat. Genet. 25, 269–277.

Sedgwick, B. (2004) Repairing DNA methylation damage. Nat. Rev. Mol. Cell Biol. 5, 148-157.

Shilatifard, A. (2006) Chromatin modifications by methylation and ubiquitination: Implications in the regulation of gene expression. Annu. Rev. Biochem. 75, 243-269.

Sieburth, L.E., and Meyerowitz, E.M. (1997) Molecular dissection of the *AGAMOUS* control region shows that *cis* elements for spatial regulation are located intragenically. Plant Cell 9, 355-365.

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

Soppe, W.J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M.S., Jacobsen, S.E., Schubert, I., and Fransz, P.F. (2002) DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. EMBO J. 21, 6549-6559.

Stancheva, I. and Meehan, R.R. (2000) Transient depletion of xDnmt1 leads to premature gene activation in *Xenopus* embryos. Genes Dev 14, 313-327.

Stancheva, I., Hensey, C., and Meehan, R.R. (2001) Loss of the maintenance methyltransferase, xDnmt1, induces apoptosis in *Xenopus* embryos. EMBO J. 20, 1963-1973.

Steimer, A., Amedeo, P., Scheid, O.M., and Paszkowski, J. (2000) Endogenous targets of transcriptional gene silencing in *Arabidopsis*. Plant Cell 12, 1165-1178.

Steward, N., Kusano, T., and Sano, H. (2000) Expression of ZmMET1, a gene encoding a DNA methyltransferase from maize, is associated not only with DNA replication in

actively proliferating cells, but also with altered DNA methylation status in cold-stresses quiescent cells. Nucleic Acids Res. 28, 3250-3259.

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

Stockinger, E.J., Mao, Y., Regier, M.K., Triezenberg, S.J., Thomashow, M.F. (2001) Transcriptional adaptor and histone acetyltransferase proteins in *Arabidopsis* and their interactions with CBF1, a expression. Nucleic. Acids Res. 29, 1524-1533.

Stokes, T.L., Kunkel, B.N., and Richards, E.J. (2002) Epigenetic variation in Arabidopsis disease resistance. Genes Dev. 16, 171-182.

Sugimoto, M., Yamaguchi, Y., Nakamura, K., Tatsumi, Y., and Sano, H. (2004) A hypersensitive response-induced ATPase associated with various cellular activities (AAA) protein from tobacco plants. Plant Mol. Biol. 56, 973-985.

Tamaru, H., and Selker, E.U. (2001) A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. Nature 414, 277-283.

Tariq, M., Saze, H., Probst, A., Lichota, J., Habu, Y., and Paszkowski, J. (2003) Erasure of CpG methylation in *Arabidopsis* alters patterns of histone H3 methylation in heterochromatin. Proc. Natl. Acad. Sci. USA 100, 8823-8827.

Tariq, M., and Paszkowski, J. (2004) DNA and histone methylation in plants. Trends Plant Sci.

6, 244-251.

Vongs, A., Kakutani, T., Martienssen, R.A., and Richards, E.J. (1993) *Arabidopsis thaliana* DNA methylation deficient mutants. Science 260, 1926-1928.

Wada, Y., Ohya, H., Yamaguchi, Y., Koizumi, N., and Sano, H. (2003) Preferential *de novo* methylation of cytosine residues in non-CpG sequences by a domains rearranged DNA methyltransferase from tobacco plants. J. Biol. Chem. 278, 42386-42393.

Wada, Y., Miyamoto, K., Kusano, T., and Sano, H. (2004) Association between up-regulation of stress-responsive genes and hypomethylation of genomic DNA in to tobacco plants. Mol. Genet. Genomics 271, 658-666.

Wada, Y. (2005) Physiological functions of plant DNA methyltransferases. Plant Biotechnol. 22, 71-80.

Watt, F. and Molloy, P.L (1988) Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. Genes Dev. 2, 1136-1143.

Xiao, W., Gehring, M., Chopi, Y., Margossian, L., Pu, H., Harada, J.J., Goldberg, R.B., Pennell, R.I., and Fischer, R.L. (2003) Imprinting of the *MEA* polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. Dev. Cell 5, 891-901.

Yap, Y.K., Kodama, Y., Waller, F., Chung, K.M., Ueda, H., Nakamura, K., Oldsen, M., Yoda,

H., Yamakuchi, Y., and Sano, H. (2005) Activation of a novel Transcription factor through phosphorylation by WIPK, a wound-induced mitogen-activated protein kinase in tobacco plants. Plant Physiol. 139, 127-137.

Yoda, H., Ogawa, M., Yamaguchi, Y., Koizumi, N., and Sano, H. (2002) Identification of early responsive genes associated with the hypersensitive response to tobacco mosaic virus and characterization of a WRKY-type transcription factor in tobacco plants. Mol. Genet. Genomics 267, 154-161.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, SW.L., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., and Jacobsen, S.E., and Ecker, J.R. (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. Cell 126, 1-13.

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

Zhu, B., Benjamin, D., Zheng, Y., Angliker, H., Thiry, S., Siegmann, M., and Jost, J. P. (2001) Overexpression of 5-methylcytosine DNA glycosylase in human embryonic kidney cells EcR293 demethylates the promoter of a hormone-regulated reporter gene. Proc. Natl. Acad. Sci. USA 98, 5031-5036.

Zhu, J., Kapoor, A., Sridhar, V.V., Agius, F., and Zhu, J.K. (2007) The DNA glycosylase/lyase ROS1 functions in pruning DNA methylation pattern in *Arabidopsis*. Curr. Bio. 17, 54-59.

Zilberman, D., and Henikoff, S. (2005) Epigenetic inheritance in *Arabidopsis*: selective silence. Curr. Opin. Genet. Dev. 15, 557-562.

PUBLICATION LIST

Choi, C.S., and Sano, H. (2007) Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. **Mol. Genet. Genomics. 277, 589-600.**

Choi, C.S., and Sano, H. (2007) Identification of tobacco genes encoding proteins possessing removal activity of 5-methylcytosines from intact tobacco DNA. **Plant Biotechnol. (in press)**

ACKNOWLEDGEMENTS

I would like to express my deep honor and gratitude to Professor Hiroshi Sano of Nara Institute of Science and Technology, who accepted me to study in his laboratory and for his encouragement, guidance, and kind advice during the study. I also wish to thanks Associate professor Nozomu Koizumi, Assistant professor Yuko Wada and Hiroshi Yoda for their helpful advice, and all members of the laboratory of Plant Molecular Breeding, NAST, for their valuable discussion, attention and friendliness during my study and life.

I think Professor Shinmyo Atsuhiko, Professor Takashi Hashimoto, Professor Masaaki Umeda, and Associate professor KazuyaYoshida for critical reading of this thesis and helpful advice. I am greatly in debt to the Japanese government (Monbusho) to provide me a scholarship throughout my study.

I would like to give me thinks to Professor Kyun-Kyu Kang (Hankyong National Univ.), Professor Ill-Sup Nou (Sunchon National Univ.) for continuous guidance and encouragement that enables me to carry out study in Japan. I also wish to thanks Profs Hee-Wan Kang, Kwan-Sik Min(Hankyong National Univ.), Jae-Eul Choi (Chungnam National Univ.) and Yong-Gu Cho(Chungbuk National Univ.) for their helpful advice.

I express think are given to my friends for their encouragement. Lastly, I wash to express my deepest thinks to my family for their love, support, and encouragement throughout the study and life.