Transcriptional response of plants to genotoxic stress
and water deficit

PhD. thesis

Supervisor: RNDr. Helena Štorchová, CSc.

Prague 2007
Acknowledgments

First, I am grateful to my supervisor, Dr. Helena Štorchová and the former director of our institute, Dr. Ivana Macháčková, who jointly enabled me to finish my PhD. studies. Dr. Imre E. Somssich and other staff of MPIZ, as well as Dr. Tomáš Mráček, who substantially facilitated the experimental part of my work. My thanks belong to my parents and grandparents for setting the course of my journey in the direction of natural sciences and to Lenka for supporting me in all aspects of our life. In addition, I owe a lot to many other persons, for making my life easier and more beautiful, also by helping me with my work.
Declaration

I hereby declare that the work presented in this thesis is my own and was carried out entirely with help of literature and aid cited in the manuscript.

Prague, Czech Republic
July 13, 2007
# Table of contents

1 List of abbreviations ........................................................................................................... 6

2 Aims of the work .................................................................................................................. 7

3 Introduction ........................................................................................................................... 8

3.1 Transcriptional response of Arabidopsis thaliana to alkylation stress ................................. 8

3.1.1 DNA damage .................................................................................................................. 8

3.1.2 DNA alkylation .............................................................................................................. 9

3.1.3 Mechanisms of repair of alkylated DNA ....................................................................... 9

3.1.4 Methyl methanesulfonate ............................................................................................. 10

3.1.5 Transcriptional response to MMS .................................................................................. 11

3.1.6 Adaptation to genotoxic stress ...................................................................................... 12

3.1.7 DNA array technology .................................................................................................. 13

3.1.8 Cluster analysis of expression data ................................................................................. 13

3.1.9 Array data validation by real-time PCR ......................................................................... 14

3.1.10 qPCR data normalization ............................................................................................ 14

3.2 Drought stress in tobacco .................................................................................................. 15

3.2.1 Cytokinins ..................................................................................................................... 15

3.2.2 CK metabolism manipulation ...................................................................................... 16

3.2.3 Drought stress and cytokinin signaling related genes ................................................... 17

4 Material and Methods .......................................................................................................... 18

4.1.1 Arabidopsis thaliana suspension culture ........................................................................ 18

4.1.2 MMS treatment .............................................................................................................. 18

4.1.3 RNA isolation ................................................................................................................ 18

4.1.4 Hybridization target labeling ....................................................................................... 19

4.1.5 cDNA array design and hybridization ......................................................................... 19

4.1.6 Expression data acquisition and processing ................................................................. 20

4.1.7 Cloning of cDNA fragments for the cDNA array ........................................................ 20

4.1.8 Reverse transcription ..................................................................................................... 21

4.1.9 Quantitative PCR ......................................................................................................... 21

4.1.10 Total cDNA yield ....................................................................................................... 22

4.1.11 Tobacco plants and drought treatment ....................................................................... 22

5 Results .................................................................................................................................. 26

5.1 Methylation stress in Arabidopsis thaliana cells ................................................................ 26
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA</td>
<td>Abscisic acid</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>BER</td>
<td>Base excision repair</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CK</td>
<td>Cytokinin</td>
</tr>
<tr>
<td>Cp</td>
<td>Crossing point</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>ds</td>
<td>Double-stranded</td>
</tr>
<tr>
<td>GO</td>
<td>Gene ontology</td>
</tr>
<tr>
<td>HR</td>
<td>Homologous recombination</td>
</tr>
<tr>
<td>ICT</td>
<td>Internal control transcript</td>
</tr>
<tr>
<td>LoTrEC</td>
<td>Local Trend based gene Expression data Clustering</td>
</tr>
<tr>
<td>MMR</td>
<td>Mismatch repair</td>
</tr>
<tr>
<td>MMS</td>
<td>methyl-methanesulfonate</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MS</td>
<td>Microsoft</td>
</tr>
<tr>
<td>NER</td>
<td>Nucleotide excision repair</td>
</tr>
<tr>
<td>NHEJ</td>
<td>Non-homologous end joining</td>
</tr>
<tr>
<td>PCD</td>
<td>Programmed cell death</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>qPCR</td>
<td>Quantitative real-time PCR</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RNA</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcription</td>
</tr>
<tr>
<td>RWC</td>
<td>Relative water content</td>
</tr>
<tr>
<td>SCF</td>
<td>Skp1/cullin/F-box protein complex</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>ss</td>
<td>Single-stranded</td>
</tr>
<tr>
<td>STEM</td>
<td>Short Time-series Expression data Miner</td>
</tr>
<tr>
<td>TLS</td>
<td>Translesion synthesis</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type (not transgenic)</td>
</tr>
</tbody>
</table>
2 Aims of the work

1) To study the transcriptional response of thale cress (*Arabidopsis thaliana*) suspension culture to MMS by means of low density cDNA array and quantitative RT-PCR.

   a) To design an algorithm for cluster analysis of the time-series expression data.

   b) To compare the expression profiles characteristic for treatments with high concentration of MMS, low MMS an high MMS with pretreatment.

   c) To confirm selected profiles by qRT-PCR.

2) To explore the transcript abundance changes in response of selected genes in tobacco (*Nicotiana tabacum*) subjected to drought stress.

   a) To compare the effect of drought on gene expression of wild type and *ZOG1*-expressing plants.

   b) To evaluate the effect of leaf position on drought-response.

   c) To compare suitability of different RT-priming strategies for the analysis.
3 Introduction

3.1 Transcriptional response of *Arabidopsis thaliana* to alkylation stress

3.1.1 DNA damage

The vital function of DNA as the principal carrier of genetic information is constantly threatened by various attacks against its integrity. In general, the causative factor can be physical (such as radiation – ultraviolet, ionizing) or chemical. In the aqueous environment inside the cell, hydrolytic damage to various parts of DNA molecule is common, resulting in loss of bases (mainly depurination) and base deamination. Reactive forms of oxygen produced in both mitochondrial and chloroplast (in autotrophic cells) metabolic processes also contribute to DNA alterations. In total, the frequency of spontaneous lesions was estimated to 20,000 per mammalian cell and day (1).

In addition, many chemicals can react with either DNA bases or the backbone, creating diverse types of modification. Whole molecules or just reactive groups of atoms can be added to DNA bases (for example benzo(a)pyrene). If the adduct possesses another reactive part (it is bi-functional), it can react once more with the same or with another molecule, forming thereby a crosslink. Crosslinked DNA strands cannot be properly unwound and therefore transcribed or replicated. Mitomycin C is a popular crosslinking drug employed both experimentally and for cancer therapy. Other compounds can catalyze production of reactive oxygen species - the most abundant representants are hydrogen peroxide, superoxide anion and hydroxyl radical. These rather short-lived agents result in base oxidation or DNA strand breaks, resembling thereby effects of ionizing radiation (e.g. anti-cancer drug bleomycin). DNA damage has two more-or-less separate adverse effects on the cell. First, it is cytotoxic either directly – by preventing replication and transcription of the affected genes, in some cases by fragmenting chromosomes which leads to genome instability - or indirectly, by activation of PCD. Second, the DNA-repair is not always perfect in terms of fidelity and so it causes an increased rate of genetic information change – mutation. In addition, some base modifications directly mispair with one or more canonical counterparts (e.g. 8-oxo-guanine is complementary to any of A, C, G and T). The cytotoxicity of some alkylating, DNA strand breaks or interstrand crosslinks inducing compounds is the basis of function of some anti-cancer chemotherapeutics (temozolomide, bleomycin, mitomycin C). The increased frequency of mutations after application of mutagens has been widely used in basic as well as in applied research and breeding of plants.
3.1.2 DNA alkylation

One mechanism of chemical modification of DNA is its alkylation. The donors of the alkyl groups can be either endogenous (such as S-adenosylmethionine) or from the environment (e.g. naturally occurring organohalogen, tobacco-specific N-nitosamines from tobacco smoke and many artificial chemical pollutants). Reactions of DNA with monofunctional alkylation agents yield mainly bases modified at various positions – either on C or N and O atoms but also alkyl, bis(polynucleotidyl)-phosphotriesters resulting from DNA backbone modification (  ).

3.1.3 Mechanisms of repair of alkylated DNA

Several DNA-repair pathways safeguard the genome integrity, specifically removing distinct classes of the products. They include receptors that recognize the lesion, enzyme machinery to resolve the problem and also members that connect DNA-repair to signaling cascades of the cell, regulating among others cell cycle progression and programmed cell death (PCD). Some of the alkyl-bases destabilize the glycosidic bond between deoxyribose and the base and spontaneously hydrolyze away resulting in an abasic site (e.g. N\(^7\)-methyladenine). Besides of direct reversal of alkylated bases by alkyltransferases and oxidative demethylases (AlkB-homologs; reviewed by ), this type of lesions is predominantly removed by means of base excision repair (BER). The defective base hydrolysis is catalyzed by a family of specific alkyl-base-DNA-glycosylases. Apurinic/apyrimidinic site specific (AP-) endonuclease introduces a single strand break 5’ of the abasic site and DNA polymerase β or λ replaces the abasic nucleotide with a complete one, matching the opposite strand. DNA-ligase then seals the DNA strand break.

Alternatively, nucleotide excision repair (NER) or mismatch repair (MMR) pathways can remove the alkylation damage. NER is accomplished by multisubunit complexes containing XP (Xeroderma pigmentosum) proteins. The damaged DNA strand is incised up- and downstream of the lesion by specific endonucleases and unwinded by a helicase. The resulting gap is then replaced with a stretch of DNA newly synthesized by a repair DNA-polymerase and sealed by a DNA-ligase. The actions performed by MMR are in general similar to NER. The main functional difference is that while NER detects distortion of the double-helix caused by DNA modifications, MMR scans the molecule for mispairing bases. All of BER, NER and MMR produce single strand breaks as an intermediate of their action. If not repaired before the next round of replication, single strand breaks transform to double strand breaks ( ), one of the most toxic types of DNA damage. Faulty repair of double strand
breaks represents the mechanism of inducing chromosomal aberrations by simple alkylating agents.

Homologous recombination (HR) and nonhomologous end joining (NHEJ) are the two possible ways of repairing double-strand DNA breaks. Homologous recombination takes advantage of the other copy of the affected chromosome, using it as a template for break repair. However, any DNA molecule with long enough stretch of sequence homology can be employed. That is why HR is the mechanism utilized for gene-targeting methods. HR results in somatic sister chromatide exchanges but it is generally error-free. Double strand breaks are repaired predominantly by HR in budding yeast.

On the other hand, NHEJ joins the loose ends of DNA molecules at a site of microhomology, which often produces sequence changes (insertions or deletions). However, NHEJ is the preferred pathway in most higher eukaryotes, including plants. The reason for this could be the high proportion of repetitive sequences in most genomes, that could lead to illegitimate HR and thus chromosomal aberrations.

If modified bases or abasic sites persist in the DNA until replication, they block the progression of replicative DNA-polymerase complex. The stalled replication fork can be rescued by translesion synthesis (TLS) DNA-polymerases. These enzymes generally possess low fidelity (and thus produce a considerable number of mutations) but they can replicate a damaged template, thereby retrieving the cell cycle progression. There are several TLS DNA-polymerases, more or less specific to various DNA modifications.

The importance of capability to respond to all possible kinds of DNA damage is illustrated by the severity of syndromes resulting from impaired DNA-repair genes in all organisms. Although the pathways are partially redundant, loss of function of one of them usually causes developmental defects, susceptibility to DNA damage and in animals a high cancer rate. Human syndromes linked with DNA-repair insufficiency include Xeroderma pigmentosum (compromised function of NER or TLS), Cockayne syndrome (impaired transcription-coupled NER), Fanconi anemia (problems in response to DNA crosslinks), Nijmegen breakage syndrome (lack of efficient NHEJ), Hereditary non-polyposis colon cancer (MMR) or Ataxia telangiectasia (mutations in the gene coding for a protein kinase activated after detection of DNA breaks).

3.1.4 Methyl methanesulfonate

Methyl methanesulfonate (MMS, Figure 1) has been used as the most frequent model DNA-damaging (genotoxic) compound for tens of years. It is a methylating agent of SN2 type; that means it attacks predominantly positions N7 of guanine and N3 of adenine. As mutagenicity of
methylation agents correlates with their O-alkylation efficiency (producing O\(^6\)-methylguanine and O\(^4\)-methylthymine), MMS is not as mutagenic as S\(_N\)1 alkylating chemicals (e.g. other common experimental genotoxins, N-methyl-N-nitrosourea and N-methyl-N'-nitro-N-nitrosoguanidine). Ethyl methanesulfonate (EMS), used to mutagenize plants for experimental and breeding purposes is also a more potent mutagen because ethylating agents generally show higher tendency to modify oxygen positions than their methyl- counterparts.

The DNA methylation produced by MMS has been shown to induce DNA strand breaks (as a result of excision repair processing) and subsequent chromosome aberrations. MMS can also methylate RNA and amino-acids cysteine and histidine \textit{in vitro}. Accordingly, various organisms possess mechanisms to repair or eliminate modified mRNA and protein. Mutations in the involved genes produce generally phenotypes sensitive to methylating chemicals.

\[
\text{CH}_3 - \text{S} - \text{O} - \text{CH}_3
\]

\textbf{Figure 1. Structural formula of MMS.}

3.1.5 \textbf{Transcriptional response to MMS}

The global changes in gene expression after treatment with MMS were studied in several organisms. First of them was yeast \textit{(Saccharomyces cerevisiae; \textit{}}. Among the genes induced by 1 hour in 0.1\% (approximately 10 mM) MMS, several were known to be involved in resistance to DNA-damaging agents or directly acting in DNA repair. The treatment led to activation of a group of genes connected with general stress response, proteolysis and also primary metabolism. Among the repressed transcripts, the most prominent group was involved in ribosome biosynthesis.

Transcriptional responses of mouse and human cell cultures to MMS have been also studied. The observed spectra of differential gene activities were significantly different between these organisms, although they were both mammalian leukemia cells. The reason for the fact could be that the mouse culture (L5178Y) possessed a nonfunctional allele of p53, a key modulator of cell-fate after DNA-damage. The human TK6 cells might thus represent a more realistic picture of normal mammalian response to genotoxins. This assumption is consistent
with a much broader transcriptional response of TK6 cells to MMS as well as to bleomycin, encompassing genes known to participate in DNA-repair, cell cycle regulation and apoptosis. The discrepancy could however also relate to the original species. In the mouse L5178Y cells, the set of MMS-induced transcripts did not at all overlap with those upregulated by bleomycin (with a single exception). Also in TK6 culture were most genes possessing differential activity treatment-specific. Anyway, there were 2 genes induced by both treatments after 4 hours and 15 common transcripts with higher abundancy after another 20 hours. Nine of the 15 common activated genes could be annotated as involved in p53, TNF (Tumor Necrosis Factor), ERK (Extracellular-signal-Regulated Kinase) or JNK (c-Jun N-terminal Kinase) pathways, thus connected to cell fate. As many of these demonstrated changes in expression also appeared after application of other genotoxins, it seems to be a general response to DNA damage that takes some time to start. Unfortunately, plants lack some of the key players in the mentioned pathways and so these data cannot be expected to apply directly to them.

Expressions of various plant genes have been shown separately to be influenced by MMS (e.g. ). There is, however, no comprehensive transcriptomic study concerning methylation stress in plants up to date. The only results of macro-/microarray experiments dealing with DNA damaging chemicals, that are either published or accessible in public data repositories, concern Arabidopsis plants treated with a combination of bleomycin and mitomycin C (, AtGenExpress at http://web.uni-frankfurt.de/fb15/botanik/mcb/AFGN/atgenex.htm).

Although both of these compounds are also genotoxic, their main mode of action is very different. Bleomycin induces double-strand breaks while mitomycin C forms predominantly interstrand crosslinks. As the enzymatic machinery necessary to remove these lesions only partially overlaps with that for alkylation damage repair, the spectrum of transcripts induced and repressed by MMS treatment should be presumably also different.

3.1.6 Adaptation to genotoxic stress

The fact that pretreatment with a moderate stress enhances survival of various organisms in conditions of high stress is generally accepted. In the field of DNA damage, studied the effect of preconditioning of Escherichia coli with low concentration of various mutagens on the level of mutagenesis caused by a higher dose. observed a reduction in the frequency of genotoxin-induced mutations and also chromosomal aberrations (clastogenic adaptation) in Chinese hamster cells, when such a preconditioning was applied. The phenomenon can be induced also by a pretreatment with another (but not any) DNA-damaging chemical.
3.1.7 DNA array technology

High throughput methods for assaying simultaneously expressions of many genes have become widely used. They give a global image of the cell processes affected by the particular treatment, mutation, developmental stage transition etc. The most common method to accomplish this task employs hybridization to an array of immobilized probes. Although whole-genome arrays are readily accessible for several species, in some cases it is still effective to use a smaller custom set of selected transcripts spotted on a lower-density array. The main advantages are lower price and no need for highly specialized equipment. Hybridization arrays allow simultaneous assaying of abundances of many transcripts. However, they suffer from a range of artefacts. One of them is hybridization of diverse but similar molecules to a particular probe. If the authentic target is rare, even a limited complementarity of a more abundant transcript can lead to substantial overestimation of its expression level. In addition, this can also mask expression changes (e.g. turning-on from effectively zero to moderate). Therefore it is advisable to test the results at least for a subset of differentially expressed genes using an independent method.

3.1.8 Cluster analysis of expression data

Any array experiment creates a big amount of data. To organize them and make them easier to comprehend, it is useful to find transcripts with similar expression changes in various conditions. Groups of such genes might have a common mode of regulation and/or function. Several possible similarity (or dissimilarity) measures exist that can be used to define the degree of concordance between expression profiles. Most of them treat all the measured expression values as coordinates of a vector in a multidimensional space (with number of dimensions equal to the number of conditions tested). Some of them are based on Euclidean distance, i.e. absolute difference between the compared vectors – they take into account both the changes and absolute intensity of expression. Others use various kinds of correlation – they generally overlook the absolute values, only the shape of the expression profile is considered. Various algorithms then organize the gene expression vectors either to separate clusters (K-means etc.) or to a form in which the groups of similar profiles can be discovered and picked. Examples of the latter strategy are hierarchical clustering (the size and number of the clusters is defined by setting a cut-off value), self-organizing maps and principal component analysis.
3.1.9 **Array data validation by real-time PCR**

Recently, quantitative polymerase chain reaction coupled with reverse transcription (RT-qPCR) is usually the method of choice for the confirmation of DNA-array data (). In contrast to DNA-arrays, this method can be truly quantitative. In comparison with Northern blotting, also sometimes used to confirm array data, RT-qPCR does not depend on hybridization of a long probe and so it can be very specific – even alleles differing by single nucleotide polymorphisms may be distinguished. The key to this specificity is selection of primers and in some experimental setups also a relatively short labeled probe. Another advantage of qPCR-based methods is their outstanding sensitivity (down to several copies of the target) combined with enormous dynamic range of concentrations in which the target stays quantifiable.

However, qRT-PCR reliability depends on, among other factors, appropriate normalization (for a review, see ).

3.1.10 **qPCR data normalization**

Although individual qPCR reactions, if performed properly, generally suffer only from little variability, it is still necessary to compensate (or “normalize”) for differences occurring both in the starting material and qPCR process themselves. Measured data normalization thus allows to directly compare individual samples to each other. The preferred method of quantitative RT-PCR normalization uses housekeeping genes with presumably invariant levels of expression as internal controls. Housekeeping gene-based normalization corrects the measured transcript levels for variable starting RNA amounts and for differences in RT efficiency. However, as there are no universally applicable genes with invariant expression, it is necessary to carefully evaluate the expression of candidate reference genes for every particular experimental system. Normalization with suboptimal house-keeping genes may result in different estimated values and lead to erroneous interpretations (). To avoid a bias caused by the expression fluctuation of a single reference gene, proposed computation of the correction factor from several internal controls. However, this approach may significantly increase the cost and laboriousness of experiments. Another approach derives the correction factor for each sample from the input RNA amount, based either on spectrophotometric ($A_{260}$), or on fluorometric estimation (), meaning that there is no need to select a proper reference gene and verify its expression. However, this method relies upon the reproducibility of the RT reaction, which has been shown to be a major source of quantitative RT-PCR variation (). Several other normalization strategies have been reviewed recently ().
3.2 Drought stress in tobacco

Water represents a major limiting factor for plant growth and development. Accordingly, plant responses to its shortage are complex (). Stress mitigation strategy involves fast changes, e.g. closure of stomata, as well as longer-term modification of plant metabolism and growth. The latter processes involve changes in expression of many genes, as it was shown for Arabidopsis thaliana using whole genome DNA arrays (, Genevestigator at https://www.genevestigator.ethz.ch/; ). Distinct parts of the plant – leaves of different position on the stem – obviously differ in their stress response. To perform multiple analyses from each leaf separately, Arabidopsis appears to be a rather difficult model plant. On the contrary, tobacco plants are significantly bigger and thus allow to quantify various metabolites from a single leaf in addition to RNA isolation for assessing gene expression. This is important for a complex evaluation of metabolism in a precisely specified site of the plant. Tobacco thus represents a more suitable alternative than Arabidopsis. In addition, this model plant has been used for many physiological experiments including stress treatments.

The drought responses are orchestrated by hormone signaling. Phytohormones modulate both fast stress responses and transformation of the stress-related signals into specific changes of gene expression. Most studies on the role of plant hormones in stress physiology have focused on abscisic acid (ABA). Elevation of ABA concentration after initiation of various stresses has been numerous reported (). In response to drought, ABA decreases the transpiration rate by promoting stomatal closure. In addition, expression of many drought- and cold-inducible genes depends on ABA action either via bZIP transcription factors binding to ABRE (ABA response element) present in promoters of the target gene or via MYB and MYC proteins binding to their respective recognition sites MYBRS and MYCRS ().

3.2.1 Cytokinins

In order to achieve precise control in any system, both positive and negative regulators are required. Many physiological processes induced by drought, particularly those mediated by ABA, e.g. acceleration of leaf senescence and closing of stomata, may be counteracted by cytokinins (CKs; for review see ).

CKs are a group of derivatives of adenine with regulatory activities in plant growth and development. They stimulate cell division, reproduction of plastids, chlorophyll synthesis and development of the shoot organs. The latter is connected with the release from apical dominance and inhibition of leaf senescence. CKs occur in several forms. Apart from free CK
bases and ribosides (the active forms), modifications such as phosphates and glucosides exist. In addition, cis and trans isomers differ in biological activity and also metabolic processing. Exogenous CKs were reported to increase the tolerance to mild stress and to speed up the subsequent recovery of the plant (). Applied CKs relieved negative effects of water stress on chlorophyll and carotenoid content and photochemical activity (e.g. ) or improved recovery of stomatal conductance and net photosynthetic rate after rehydration of drought-stressed plants (). Moreover, transcription of many stress-induced genes can be stimulated by CK ()

3.2.2 **CK metabolism manipulation**

It should be noted that the composition and concentration of exogenous phytohormones at the site of action might substantially differ from the applied dose, not only because of the losses during their uptake and translocation, but also due to their metabolism and allocation (e.g. ). These drawbacks may be overcome by altering endogenous CK levels by manipulation of genes coding for CK metabolic enzymes. Plants expressing ectopically the gene *IPT* (coding for isopentenyltransferase, the rate-limiting enzyme of CK synthesis) under the control of either constitutive (Cauliflower mosaic virus 35S promoter) or inducible promoters (responsive to Cu^{2+} or heat-shock) were prepared (for review see ). However, substantially elevated content of active CKs in transformants with increased CK biosynthesis usually causes severe changes in plant morphology, which may complicate evaluation of the potential CK role in response to stress conditions. To minimize this limitation, transgenic tobacco plants containing *IPT* gene driven by a senescence-specific *SAG12* promoter were developed. Activation of *IPT* transcription in these plants is triggered only in senescing leaves and the plants develop normally. *SAG12::IPT* plants were used for evaluation of CK effect on leaf senescence (), photosynthesis and nitrogen partitioning (), flooding tolerance (), seed germination and response to mild drought stress (). An alternative approach is over-expression of the gene coding for *trans*-zeatin O-glucosyltransferase (*ZOG1*). *ZOG1* cDNA was isolated from Lima bean (*Phaseolus lunatus*; ) and transferred to tobacco plants, under the control of either constitutive (35S) or senescence inducible (*SAG12*) promoters (). The levels of physiologically active CKs (bases, ribosides) did not substantially differ between the transgenic and wild type plants, being still under the control of mechanisms regulating CK homeostasis. However, the total CK level (including glucosides) was highly elevated. CK-O-glucosides, responsible for the increase, represent a readily mobilisable CK reserve, capable to provide additional active CKs if a need arises. Conversion of CK O-glucosides to the active CK forms can be accomplished by β-glucosidases (e.g. ). CK O-glucosides are resistant to cytokinin oxidase/dehydrogenase (*CKX*), the main CK degrading enzyme ().
3.2.3 **Drought stress and cytokinin signaling related genes**

*NtERD10B*, one of dehydrins, is a relatively small (169 amino acids), glycine-rich protein. It is a member of the *LEA* (Late Embryogenesis Abundant) family of mainly stress-inducible proteins. It is a group 2 *LEA* protein, as it contains a conservative element EKKGIMDKIKEKLPG at its C-terminus (amino acids 148 – 162) and an additional incomplete copy of the motif (114 - 122). It is presumed to protect the cell components against effects of osmotic stress, possibly as a molecular chaperone. *NtERD10B* was found as one of the *Arabidopsis thaliana* *ERD10* homologues in tobacco by . It was shown to be induced by drought and low temperature treatments and also in tobacco plants over-expressing *Arabidopsis DREB1A* gene. *DREB1A* is a transcriptional activator of stress-induced genes with 41 identified targets in *Arabidopsis* genome (). When ectopically expressed in tobacco, as well as in rice, wheat and *Arabidopsis*, it confers drought tolerance, probably through regulation of expression of the stress mitigating genes (

*cig1*, Cytokinin Induced Gene 1, codes for a putative proline dehydrogenase/oxidase (493 amino acids). Proline dehydrogenases are key enzymes in conversion of proline to glutamate (). As proline is known to serve as an osmoprotectant during various conditions leading to loss of cellular water, its metabolism is directly connected to drought stress. Indeed, observed drought induced decrease in a proline dehydrogenase mRNA concentration in *Arabidopsis*. *cig1* mRNA was shown to accumulate after cytokinin treatment of tobacco suspension culture if the cells were grown in presence of auxin. In medium lacking auxin, *cig1* transcript abundance specifically increased in response to ABA (). *cig1* could therefore represent a functional link between metabolic stress responses and cytokinin signaling.

*CRK1* (Cytokin-in-Regulated Kinase) was predicted to be a transmembrane receptor Ser/Thr-proteinkinase of 794 amino acids. The abundance of its transcript is quickly (within 30 minutes) reduced after addition of cytokinins (benzyladenine) to tobacco suspension culture. However, this reduction seems to be transient. Proteosynthesis and an unidentified phosphatase activity but not a proteinkinase were shown to be necessary for the reaction (). The authors anticipate *CRK1* function in the early steps of cytokinin signaling. Its disappearance could confer a temporary desensitization of the cells towards its ligand. The ligand is however not yet known.
4 Material and Methods

4.1.1 Arabidopsis thaliana suspension culture

The Arabidopsis thaliana suspension culture was obtained from Dr. K. J. Angelis. It was cultivated in a dark orbital shaker at 21°C and 90 rpm in the following medium (per 1 liter):
30 g sucrose, 4.9 g Murashige & Skoog Medium Including Vitamins and MES buffer (Duchefa Biochemie, Haarlem, Netherlands), 1 ml vitamine cocktail and 1 ml of 0.1% (w/v) 2,4-dichlorophenoxyacetic acid in ethanol. After autoclaving, ampicilin was added to the final concentration of 100 mg/l. The vitamine cocktail comprised of 0.5 mg nicotinic acid, 0.5 mg pyridoxine-HCl, 0.1 mg thiamine-HCl and 0.1 mg myoinositol per ml.

4.1.2 MMS treatment

For the MMS treatment (5 mM and 0.5 mM), late logarithmic culture was prepared, 3 days after the last subculture. The cells were sifted out by a stainless steel tea strainer and gently drained away. Approximately 1 g of material in 10 ml of medium in a separate 50 ml Erlenmeyer flask were used for each time point. To reduce the effect of medium exchange, appropriate amount of MMS was added to the original medium, swirled and immediately added to the cells. The samples were sequentially harvested after expiration of 15 min, 30 min, 1 h, 2 h, 3 h, 5 h and 8 h. All the incubations took place in the standard cultivation conditions. The Control sample (supplied with re-used medium only) was harvested after 2 hours (to experience comparable manipulation). Harvested and drained-away cells were aliquoted into 3 microcentrifuge tubes and flash-frozen in liquid nitrogen.

For the combined MMS treatment, 10 g of drained-away cells were resuspended in 100 ml of re-used medium containing 0.5 mM MMS in a 500 ml flask. After 1 hour, the medium was replaced with another 100 ml of medium, lacking MMS. 2 hours later, the culture was treated with 5 mM MMS in the same way as the non-preconditioned samples.

4.1.3 RNA isolation

Total RNA was isolated using RNeasy Plant Mini kit (Qiagen, Carlsbad, USA) with starting amount of about 300 mg of the material per column. RNA quantity was measured as absorbance at 260 nm wavelength by spectrophotometer and its quality checked on an electrophoretic gel. To denature RNA structures, the RNA samples were heated at 75°C for 2 min in presence of 2,5 % SDS (w/v) prior to electrophoresis.
For development and primary demonstration of the RT yield measurement, total RNA was isolated from *Arabidopsis thaliana* suspension culture using acidic phenol-chloroform-isoamylalcohol extraction according to .

In both cases, the RNA was treated with 0.1 units RNase-free DNaseI (Roche Diagnostics, Mannheim, Germany) per μg RNA in 20 mM Tris-HCl (pH = 8.3), 50 mM KCl and 1 mM MgCl₂ for 30 minutes at room temperature. The enzyme was then inactivated by 10 minutes at 65°C after adding EDTA to a final concentration of 1 mM. Alternative treatment with DNA-free (Ambion, Austin, USA) or DNasel on-column digestion (Qiagen) produced similar results.

### 4.1.4 Hybridization target labeling

To produce labeled hybridization target, 20 μg RNA of each sample was reverse transcribed in a 30-μl reaction by 400 u MMLV reverse transcriptase RNaseH Point Mutant (Promega, Madison, USA) in the buffer recommended by manufacturer, supplied with 1 μg anchored oligo dT primer (dT23dV), 0.5 mM dBTP (C, G and T), 0.05 mM dATP and 1 μCi α-³²P-dATP. RNA with the primer were heated to 65°C for 5 min and cooled on ice. After addition of all the remaining components, the reaction mixture was incubated at 42°C for 50 min, then denatured at 96°C for 5 min and again cooled on ice. The probe was ready for hybridization at this stage.

### 4.1.5 cDNA array design and hybridization

The macroarrays carried PCR-amplified inserts from indicated cDNA library clones from the collection at Max Planck Institute for Plant Breeding, Cologne, Germany, spotted in duplicates by the institute Core Facility staff and equipment onto Hybond-N+ Nylon membrane (Amersham-Pharmacia-Biotech, Uppsala, Sweden). In addition, 34 cDNA fragments not present in the cDNA collection were cloned as described below.

Prior to hybridization, each membrane was prehybridized in 20 ml of 6x SSC supplied with 1x Denhardt solution, 0.5% (w/v) sodium laurylsulfate (SDS) and 100 mg/l sonication-sheered salmon sperm DNA (according to ) at 65°C for 2 hours rolling in a hybridization oven.

The hybridization solution was prepared from 5 ml of new prewarmed prehybridization solution and a denatured probe and it immediately replaced the prehybridization solution. After overnight incubation at 65°C with the probe, the membranes were washed twice with prewarmed 2x SSC, 0.1% SDS and once with 0.5x SSC, 0.1% SDS, 10 min at 65°C each.
4.1.6 Expression data acquisition and processing

Still wet washed arrays were heat-sealed into plastic bags and submitted to BAS5000 scanner (Fuji Photo. Film Co., Ltd, Kanagawa, Japan). The densitometry of resulting images was performed using Quantity One software (Bio-Rad, Hercules, CA, USA), subtracting local background surrounding each spot.

Standardization of signal intensities among the arrays was performed by subtracting 2nd percentile (to remove "spot background" corresponding to empty spots) and dividing each value by the median signal of the corresponding array. Median of all standardized values referring to the same transcript-treatment-time point combination (1 to 5 arrays, in most cases 2 to 4; each array possessed 2 spots per transcript) was taken as the mean value for subsequent evaluation.

Visual Basic for Applications implemented in Microsoft Excel versions 2000, XP and 2003 served for programming LoTrEC, the macro presented in this work, aimed for automation of cluster formation from the measured data.

4.1.7 Cloning of cDNA fragments for the cDNA array

cDNA fragments for cloning were PCR amplified by AccuTaq LA (Sigma-Aldrich, St. Louis, USA) from cDNA prepared with SuperScript II reverse transcriptase (Invitrogen, Carlsbad, USA) according to protocols supplied by manufacturers. The list of genes and corresponding primer sequences can be found in Table 1. The PCR products were introduced to pUC57 (Fermentas, Vilnius, Lithuania) plasmid modified for TA cloning as follows. pUC57 was digested by EcoRV, precipitated by ethanol and 2',3'-dideoxy T was added to its 3' ends by 30 u terminal deoxynucleotide-transferase (Promega, Madison, USA) in the buffer recommended by manufacturer, supplied with 100 μM ddTTP and 10 μg linearized plasmid. After 30 min at 37°C, the product was precipitated once more and dissolved in water to a final concentration of 50 ng/μl. 10 μl ligation reaction contained 1 μl of the vector solution, 1 μl 10x ligase buffer, 0.5 μl T4 DNA-ligase (0.5 u; Fermentas, Vilnius, Lithuania) and 7.5 μl of the PCR reaction (containing the product of interest as a single band on a control electrophoresis). The ligation was incubated on ice let to melt at room temperature overnight and stopped by 65°C for 15 min. Electroporation of 1 μl of the reaction mixture usually yielded tens to hundreds of colonies. Clones carrying the desired insert were detected among the white ones on X-Gal/IPTG LB plates (IPTG and X-gal were spread onto the agar medium just before plating the bacteria) by means of PCR. Part of the colony was transferred to 10 μl PCR reaction containing the primes used for amplification of the cloned fragment. The identity of
the cloned fragment was confirmed by sequencing with Autoread Sequencing Kit Auto (Fermentas, Vilnius, Lithuania) and modified M13-universal and/or M13-reverse primers (Table 1) using ALF express II sequencer (Amersham-Pharmacia-Biotech, Uppsala, Sweden).

4.1.8 Reverse transcription

Reverse transcription (RT) reaction for the quantitative polymerase chain reaction (qPCR) of the *Arabidopsis* samples contained 100 u Moloney murine leukemia virus (MoMLV) RNaseH point mutant (Promega, Madison, USA) supplied with its recommended buffer, 0.5 mM each dNTP, 0.5 μg anchored oligo-dT (dT<sub>23</sub>dV) or 2 μg random hexanucleotides (dN<sub>6</sub>) and 1 μg RNA. The procedure followed otherwise the steps described in the hybridization target labelling protocol above with the exception of the final denaturation step. 5 min at 70°C was used instead.

For tobacco RNA, 1 μg of RNA was reverse transcribed by 10 units of Transcriptor Reverse Transcriptase (Roche Diagnostics, Mannheim, Germany) with random hexanucleotides (4 μg) or oligo dT primers (dT<sub>15</sub>; 500 ng). After heating 5 minutes at 65°C, buffer, 0.5 μl of Protector RNase Inhibitor (Roche), 2 μl of 10 mM dNTPs and enzyme were added on ice. The cDNA synthesis proceeded at 55°C for 30 min. If random hexamers were used, a step of 25°C for 10 min was inserted directly after enzyme addition.

4.1.9 Quantitative PCR

The real-time quantitative PCR (qPCR) was performed using LightCycler 1.2, DNA Master Kit PLUS and glass reaction capillaries (Roche Diagnostics, Mannheim, Germany) or polycarbonate capillaries from Genaxxon (Biberach, Germany). contais the list of primers specific for assayed transcripts and their respective annealing temperatures and optimum concentrations. The reaction mixture consisted of 2 μl Master mix (component 1 of the kit), 2 μl of the investigated cDNA (20x or more diluted) and primers. A common mix was prepared for as many samples as possible, avoiding pipetting of volumes lower than 2 μl. The typical PCR program included 10 min at 95°C to activate the enzyme and 50 cycles of 10 s at 95°C, 6 s at primer-specific annealing temperature and 10 s at 72°C. After the PCR, melting curve analysis was performed (95°C, 60°C and heating at 0.1°C/s to 95°C). Annealing temperatures and concentrations of all primers were optimized to get maximum PCR efficiency without nonspecific products. Second derivative maximum (as implemented in LightCycler software v. 3.5) was used to determine crossing points (CP). PCR efficiency (E) was estimated from calibration curve of cDNA serial dilutions. Relative transcript abundance (RA) in sample Y
was calculated from CP in the following way: \( RA_{\text{SampleX}} = E^{(CP_{\text{Calibrator}} - CP_{\text{SampleX}})} \). The result was then standardized by dividing with RA of an internal control transcript (ICT) presumed to have invariable abundance in the experimental conditions. The ICTs used were PDF2 (At1g13320, primers according to , Table 2) for *Arabidopsis* samples and Tac9 (GenBank accession X69885), ) and 18S rRNA (AJ236016), in case of random primed cDNA for tobacco.

### 4.1.10 Total cDNA yield

Total cDNA yield measurement employed Cary Eclipse fluorometer (Varian, Palo Alto, USA) equipped with 40 μl cuvette. MMLV reverse transcriptase RNaseH Point Mutant (Promega, Madison, USA) or Transcriptor (Roche Diagnostics, Mannheim, Germany; similar data were obtained but the yield was generally lower) carried out the RT reaction with random hexanucleotides (dN₆) or anchored oligo-dT primer (dT₂₃dV) essentially according to the manufacturer's instructions.

A 5 μl aliquot from the RT reaction mixture was taken before the reaction started, it was supplied with 15 μl of 4/3x alkaline hydrolysis solution (133 mM NaOH and 1.33 mM EDTA) and incubated at 70°C for 20 min. The mixture was neutralized by adding 6 μl 0.5 M Tris-HCl (pH = 6.4). After the RT had finished, another 5 μl were taken and treated in the same way. 5.2 μl of each hydrolyzed sample (corresponding to 1 μl of the RT reaction) with 0.2 μl RiboGreen solution from RiboGreen RNA Quantitation Kit (Invitrogen, Carlsbad, USA) and 34.6 μl water were submitted to the fluorometer set to 500/520 nm excitation/emission wavelength. The net cDNA-RiboGreen fluorescence (signal from the sample taken after RT minus that of the pre-RT aliquot) divided by the net value of a standard (or calibrator) sample was designated relative cDNA fluorescence and it was proportional to total cDNA yield. The relative cDNA fluorescence was employed as an alternative standardization factor for qRT-PCR.

The techniques not described in detail followed essentially the standard laboratory procedures ()

### 4.1.11 Tobacco plants and drought treatment

*SAG12::ZOG1* and *35S::ZOG1* tobacco lines were a gift from prof. D. Mok, Oregon State University, Corvallis, USA (for detailed information see ). Wild-type tobacco (*Nicotiana tabacum* L. cv. Wisconsin 38) and transgenic plants were grown in soil in a growth chamber (SANYO MLR 350H, Osaka, Japan) at a 16/8 h photoperiod (light/dark) at 130 μmol m⁻² s⁻¹
light intensity, day/night temperature of 25/23°C and relative humidity ca 80%. After 8.5 weeks (wild type and 35S::ZOG1) or 7 weeks (SAG12::ZOG1), half of the plants were transferred into another growth chamber of the same type with the same light and temperature regime, but relative humidity decreased to 35%. Those plants were not watered for 7 days. Leaf samples were collected after 1, 4 and 7 days and always after removal of the main vein and immediately frozen in liquid nitrogen. Relative water content (RWC) in the leaves was calculated as follows: RWC = (fresh weight – dry weight) / (water saturated weight – dry weight). Water saturation was achieved by letting the leaf absorb water from a beaker.

<table>
<thead>
<tr>
<th>AGI code</th>
<th>Gene name</th>
<th>Primer sequence</th>
<th>Genbank</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>UVH1</td>
<td>ATC ggT CTg TTT ggT TAT TTg C</td>
<td>AF089003</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCT Tgg gAg TCT TTT gTT TCT TC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>NTH1</td>
<td>CCT gAA AAC Tgg gTC gAA gTg CTC</td>
<td>AJ272248</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gCA ACC CAT TCT TCC TTT gAA AgC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>BI-1</td>
<td>gAT gCg TTC TCT TCC TTT TTC g</td>
<td>AB025927</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTC CTT TTC TTC TTC TTC TCT TC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>UBC2</td>
<td>gAC TTC AgC gAg gAA gAg</td>
<td>Y13031</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ceg CAg TCC AgC TTT gTT C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>RAD23.</td>
<td>TTT CgT gAA gAC TCT CAg Tgg</td>
<td>BAB09359</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gCg TgT gCT AgA ggA TTg g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>RAD50</td>
<td>gAT CTg AAg AAA ggg gAg Cg</td>
<td>AF168748</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAA AgA TCT CTT ggg CCT Cg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>Rad51</td>
<td>CgT CgT TgA ggA AAg gAA gAg CA</td>
<td>U43528</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCT AAC TAC AAA AAC Cgg Tg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>PR1-like</td>
<td>gTg ggT gCt CTT gTT gTT C</td>
<td>AC005398</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gCA TAg TTC CCC gCA gg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>ARP</td>
<td>TTA CCC ATT ggA CCA CAC C</td>
<td>Z49776</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gTT CCC ATC CCT TCT CTC C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>PARP</td>
<td>gTC AgC ATg ATg gCA Cag C</td>
<td>Z48243</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Agg ATC CTT gAg TgT TgA TgA ATg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>AP-endon.</td>
<td>gAg CTC CgC CAT CAT TgT C</td>
<td>X76912</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ggC AAT gAT CAC TTC CAT Tg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>PolL</td>
<td>gTg TTC CTT CTA CTA gCg</td>
<td>AJ289628</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gAg ATT CCT CTC gTg TgG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>MIM</td>
<td>AAT gAC TCT gTC CCC TCC AAA Tg</td>
<td>AF120932</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CTg gG TgC gTT CgA TCT TgA g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>LIG4</td>
<td>CTC TAg CAg AAg AAA ATg TgC</td>
<td>AF233527</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAT CCC ggT AgC TgA CTT TC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>UBQ3</td>
<td>CAA TCT CTC CCA AAg CTT AAA g</td>
<td>L05363</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TCg ACT CTT TTT gAA TgT TgT Ag</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>MAG</td>
<td>ATg AAA ACg Ccg gCT CgT Cg</td>
<td>X76169</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gTg TTC TTA ggA gCA CTT ATC C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>MPK4</td>
<td>gTC gCg ggA gAg TTg TTT C</td>
<td>AY040031</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>gAT TgA ACT TgA CTT TTT CAC gg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>RAD17</td>
<td>CAA gTg CTC gTC gTC TTC C</td>
<td>AB030250</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAC TCA TgC ACT gTg ACT CC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>SN11</td>
<td>CCA CCg gAg CTA AAg ACA g</td>
<td>AF169596</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>NPR1</td>
<td>gAT CTT gCC gAT gTC AAC C</td>
<td>U76707</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>MRE11</td>
<td>CTC AgA TgT CTg gTT TAg TCC</td>
<td>AJ243822</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>PARP-1</td>
<td>gAg gAT gAA AgT gTgC CAA Cg</td>
<td>AJ131705</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>RAD23</td>
<td>CCC AgT CTg CAC CTg TCC C</td>
<td>AY034912</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>OGG</td>
<td>gAg gat gTT ACg AAg ATg g</td>
<td>AY277400</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>UBC13</td>
<td>CCG CCA TgA ACT CAC AAg Cg</td>
<td>U33758</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>UBC3</td>
<td>TgA TgT ggg ATT TCA AgA gAC</td>
<td>L19352</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>CPR5</td>
<td>CgC TCT ACC ATA gAC ggA Tg</td>
<td>AY033229</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>zf-PARP</td>
<td>gAg gAT Tgg TgT ATT TTC AAg</td>
<td>NPI 188107</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>PMS1</td>
<td>CgA gAg gAT TgT CTT ggA CAA TC</td>
<td>AY047228</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>UVH3</td>
<td>TgA TgT gAT ggg gAT gAC</td>
<td>AF312711</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>ERCC1</td>
<td>gCC AAA CgC ATC AAA CTg</td>
<td>AY050335</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>DRT111</td>
<td>AAT CAg CAC TTT CgC ATC</td>
<td>AY045859</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>TEB</td>
<td>CgT TAT gCT gCT gTg TgA gAA g</td>
<td>CAA18591.1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>LIG1</td>
<td>gTT gTT gCT TCC gAC Agg gAg</td>
<td>X97924</td>
<td></td>
</tr>
</tbody>
</table>

**Sequencing primers**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>M13-forward</td>
<td>gTA AAA CgA Cgg CCA gTg</td>
<td></td>
</tr>
<tr>
<td>M13-reverse</td>
<td>gCT ATg ACC ATg ATT ACg C</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1. List of PCR primers** used for cloning and sequencing cDNA fragments used on the macroarray. If the primer sequences were adopted from a published work, the respective citations are given. Some of the other primers were designed by K. J. Angelis.
<table>
<thead>
<tr>
<th>AGI code</th>
<th>Gene name</th>
<th>Primer sequence</th>
<th>C [µM]</th>
<th>Tₐ [°C]</th>
<th>Genbank accession</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alkylation stress in Arabidopsis</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At1g13320</td>
<td>PDF2</td>
<td>TAA CgT ggC CAA AAT gA TgC</td>
<td>0.2</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>At4g05320</td>
<td>UBQ10</td>
<td>gTT CTC CAC AAC CgC TTg gT</td>
<td>0.5</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>At4g34160</td>
<td>Cyclin</td>
<td>TTT CgT TCg TAG ACC ACA TT</td>
<td>0.2</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>At1g75800</td>
<td>TL1</td>
<td>ACAg AT TTT TgA CTg gTg TT</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At3g28910</td>
<td>Myb30</td>
<td>ggg AgT TCA AgA TCA TAA</td>
<td>0.2</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>At1g02930</td>
<td>GST1</td>
<td>CAA CCA CgA AAC Agg AgT</td>
<td>0.2</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>At3g45640</td>
<td>MPK3</td>
<td>gAg gAT gCg AAA AgA TAC A</td>
<td>0.2</td>
<td>58</td>
<td></td>
</tr>
</tbody>
</table>

**Drought stress in tobacco**

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Primer sequence</th>
<th>C [µM]</th>
<th>Tₐ [°C]</th>
<th>Genbank accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZOG1</td>
<td>ggT CAC CTC AAC CgA CAg T</td>
<td>0.3</td>
<td>58</td>
<td>AF101972</td>
</tr>
<tr>
<td>cig1</td>
<td>CgT gAC CTC CgA TCATTT g</td>
<td>0.3</td>
<td>58</td>
<td>AB046419</td>
</tr>
<tr>
<td>CRK1</td>
<td>TgT ggT Cgg ggAACTAA gAtg</td>
<td>0.3</td>
<td>58</td>
<td>AF302082</td>
</tr>
<tr>
<td>NtERD10B</td>
<td>ATg gAC AgA gcg ggAA gAA g</td>
<td>0.3</td>
<td>58</td>
<td>AB049336</td>
</tr>
<tr>
<td>Tac9</td>
<td>CTA TTC TCC gCT Tg gAC Ttg gCA</td>
<td>0.3</td>
<td>60</td>
<td>X69885</td>
</tr>
<tr>
<td>18S rRNA</td>
<td>CCT CCA ATg gAT CCT CgT TA</td>
<td>0.3</td>
<td>64</td>
<td>AJ236016</td>
</tr>
<tr>
<td>L25</td>
<td>TgC AAT gAA gAT TgA gAA CAA CA</td>
<td>0.3</td>
<td>60</td>
<td>L18908</td>
</tr>
</tbody>
</table>

**Table 2. List of the PCR primers used for quantitative PCR.** The Tac9 and L25 primer sequences were adopted from , those for PDF2 and UBQ10 from . Some primers were designed by Helena Štorchová.
5 Results

5.1 Methylation stress in *Arabidopsis thaliana* cells

To study the transcriptional response of thale cress (*Arabidopsis thaliana*) to methylation stress exerted by MMS, a suspension culture was employed. This system should be more homogeneous than a whole plant consisting of many different cell types. As the plan was to assess the transcript abundances at several time points, it was also advantageous that all the cells were exposed to the compound at the same time. On the other hand, the metabolism and cell regulations might differ between the suspension culture and real tissue cells.

Three different treatment regimes were chosen:

1. High concentration – 5 mM MMS, which was shown to cause a strong DNA damage ( ) and which is finally lethal for *Arabidopsis* cells.
2. Low concentration – 0.5 mM MMS, DNA damage still detectable but much lower; if applied to *Arabidopsis* seedlings their development is normal, the plantlets are just slightly smaller ( ).
3. High concentration with pretreatment (designated as combined treatment) – 0.5 mM MMS for 1 hour, 2 hours of recovery in media without MMS and then 5 mM MMS – a regime based on experiments of that appeared to raise adaptation (and thereby higher resistance to DNA damage) in *Vicia faba* roots.

As the use of whole-genome microarrays for my experiments would be too expensive, a custom macroarray was designed, carrying 376 *Arabidopsis thaliana* cDNAs in duplicates. Most of them originated from the EST collection of Max-Planck-Institute for Plant Breeding, Cologne, Germany, 34 of them were cDNA fragments cloned for this reason. The cDNAs were selected based on their known or assumed functions (by sequence similarity) or on their expression profiles observed in published experiments. The following terms describe principal functional groups of genes that were sought in the collection:

- DNA repair
- Systemic acquired resistance, pathogen resistance
- Apoptosis, hypersensitive response
- Antioxidants
- Detoxification
- Cell cycle
- Signal cascades
- Transcription factors
- Proteolysis
- Secondary metabolism
- Published interesting expression profiles

For a complete list and other details, see the supplementary file SupplementA.doc
5.1.1 Gene expression data processing

Gene expression data sets are often analyzed for the presence of groups of transcripts that possess similar changes in abundance across different cell types, developmental stages, genotypes, treatments etc. Number of methods have been suggested to search for such groups, with many variants. To perform cluster analysis of the data obtained from the experiments with MMS-treated *Arabidopsis* cells, various algorithms were applied that are implemented in web-based tools (such as EPCLUST, http://www.bioinf.ebc.ee/EP/EP/EPCLUST/) as well as in commercial software (for example Vector Expression; Informax, USA). The outcome was however not satisfying. First, the obtained results varied substantially among different algorithms and settings. Even after an extensive search there was no clear indication that any of the procedures would be more appropriate to analyze data that I had acquired than the other ones. Many authors seem to choose the method that gives results supporting best their theories. Second, the results were obviously strongly influenced by random noise – even visually very diverse expression profiles could under certain circumstances cluster together.

At least the latter problem should be partially alleviated by taking into account the time-series nature of the data. Most of the well-established clustering algorithms treat the single values as independent measurements. This means a loss of information that could otherwise be used during the analysis. In addition, the whole expression profile is analyzed globally, making it relatively easy for a single outlier to substantially change the outcome for the gene. Methods for analyzing time-series data have been developed and successfully used in various areas of science and business (*e.g.*). However, most of them tend to be better suited for longer series – optimally in the order of tens of observations.

Therefore, there was a need for an alternative dealing with the above-mentioned issues. The aim was to design an algorithm that would comply with the following objectives:

1. Robust against random variations and outlying values.
2. Yield big clusters of transcripts sharing fundamental trends.

To cover the first point, a novel system for describing the shape of the expression profile has been devised, following the signal changes in time. A simple description using a scale consisting of only few levels is likely to be more stable against random noise than numerical values. In addition, the absolute numbers (or ratios of them) derived from the spot signal densitometry are often misleading due to nonlinearity of hybridization and/or cross-hybridization of similar transcripts (*e.g.*). Time-series character of the data allows using signal processing approach not applicable to sets of independent measurements. To deal with outliers, values most probably resulting from a technical problem that can significantly deteriorate the worth of the data, the concept of local trend has been introduced. The direction of the signal change along the profile is assessed for each segment of the curve separately (*i.e.*
between each pair of consecutive values). The set of directions, or local trends, is used as a part of the profile characteristics. Even if a part of the profile is heavily distorted by an outlier the rest of the curve can remain untouched and still qualify the gene in its original relationships.

To address the second point, the numerical values of parameters of the algorithm used to weight the similarities and dissimilarities between the expression profile descriptions were set in a way preserving the ability to differentiate among distinct shapes of the time-series but not hypersensitive to minor differences. On the top of that, the algorithm prefers bigger clusters.

The whole procedure can be briefly summarized into following steps:

1) Time-series profiles of the signal (expression) are described in terms of:
   a) Relative values compared to Control – in this case a sample not treated with MMS
   b) Local trends of the curve.

2) The descriptions of all the profiles (i.e. for all genes) are compared with each other.

3) Big clusters of similar profile descriptions are formed.

To automate the data processing, an MS Excel macro was written. As it employs a concept of local trends (see below) it was named LoTrEC (standing for Local Trend based gene Expression data Clustering). It can be found on the CD submitted as a part of this thesis (LoTrEC.xls). Results obtained from the experiments comprising methyl-methanesulphonate treatment of *Arabidopsis* suspension culture, processed by LoTrEC, are also present on the CD: SupplementB1_high_MMS.xls, SupplementB2_low_MMS.xls and SupplementB3_combined_MMS.xls.

5.1.1.1 *Data input*

LoTrEC accepts as input a datasheet with one or more columns describing the measured features (e.g. genes) and two or more (up to about 100) columns of experimental values (one time-point each). First of the latter is expected to be the Control sample – beginning of the time series. First row serves as a header containing labels of the time points. Each of further rows of the table is considered to contain an expression profile. Theoretically, there could be over 65,000 profiles. However, gene counts in the order of thousands can produce rather long processing times (depending on the hardware), mainly due to steeply growing number of comparisons to be performed.

If repeated measurements are to be evaluated, which is generally preferable, they must be pre-processed first to produce a single value for each time-point and profile combination. The data presented in this work were standardized as described in Material and Methods chapter.

To summarize the result, no adjustment of data distribution was applied, the distribution is thus left-skewed (approximately log-normal). Median value for each time-point is close to 1
and almost all the values are positive. For optimal performance of LoTrEC, the input data should follow the characteristics mentioned above. However, the macro could be easily adapted to fit other structures of data.

5.1.1.2 Data filtering

To reduce the influence of impulse noise (isolated outlying signals in the time series), low-pass filtering was applied with window size = 3 and double weight for the central value. That means each value was averaged with the two surrounding time points and it was itself counted twice. Arithmetic mean, usually used for low-pass filtering, is rather sensitive to outliers possessing high values that can occur quite often in this type of experiment. Therefore geometric mean was employed that is also not so much disturbed by non-normal distribution of the data. For most of the data, using the two means produced only a very slight difference but in case of sudden isolated peaks (probable impulse noise) the geometric mean behaved more conservatively (Figure 2A). Unfiltered values were only used for the control samples (which integrated highest numbers of arrays - 4 or 5 - to minimize the interference of random noise). The last time points were only averaged (with double weight) with their preceding neighbors. Similarly, the values from 15 minutes of combined treatment were only averaged with the consequent values (30 min), as the control sample in this case cannot be considered to be preceding (see below).

To demonstrate the procedure practically, let there be a set of values corresponding to a time series of signals:

Control – 15min – 30min – 1h – 2h – 3h – 5h – 8h

Low-pass filtering of the data:

\[
\text{Filt}(15\text{min}) = (15\text{min}^2 \times \text{Control} \times 30\text{min})^{(1/4)} \\
\text{Filt}(30\text{min}) = (30\text{min}^2 \times 15\text{min} \times 1\text{h})^{(1/4)} \\
\ldots \\
\text{Filt}(8\text{h}) = (8\text{h}^2 \times 5\text{h})^{(1/3)}
\]

The description is then based on a new series of values:

Control – Filt(15min) - Filt(30min) - Filt(1h) - Filt(2h) - Filt(3h) - Filt(5h) - Filt(8h)

Figure 2B illustrates the effect of filtering on a smooth and a devious (possibly noisy) profiles.
Figure 2. The effect of low-pass filtering applied in the course of describing expression profiles. A – Effect of a single outlying value on the filtered profile. Arithmetic and geometric means are compared. B – Comparison of different shapes of otherwise similar profiles. While relatively smooth profile of the gene *At3g28910* was only slightly affected, a zigzag one of *At5g67300* got straightened to reveal its fundamental trend.

### 5.1.1.3 Profile description

Each of the filtered values is compared with that of Control sample to obtain its relative position. If the higher of the two values being compared does not exceed a certain threshold (by default set to 0.4 i.e. 40% of median of values for all the genes) the result is always “no difference”. In other cases, the outcome of the comparison depends on whether they differ at least by a set constant. The difference can be either absolute (relevant mainly for higher values) or by certain-fold. The possible relative positions are: “higher” (than Control), “lower” or “no difference”.

Each of the filtered values is compared with the value preceding it in the time series to obtain its local trend (“rising”, “falling” or “no trend”; using the procedure described above). The set of thresholds has been devised empirically to describe visually similar profiles in a similar way.

### 5.1.1.4 Comparing profile descriptions

The produced descriptions are compared pair-wise and their similarities rated. The rating depends on the number of segments (time points) at which the relative positions and/or local trends agree (positive rating) or disagree (negative rating) between the two profiles. Illustrates the process of comparing two described profiles. All the mutual similarities among the profile descriptions are ordered to form a matrix of similarities.
Figure 3. Comparison of expression profiles of two genes with attributed values of local trends and relative positions is performed at each segment separately and the results are rated (default parameter settings are used in this example). The sum of all the ratings then forms the final rating. As final rating for this example exceeds the threshold rating (by default set to 0), expression profiles of genes At3g56490 and At3g52560 are similar enough to cluster together (if either of them is the founder of the cluster).

5.1.1.5 Cluster formation

For each profile containing at least one segment with a local trend other than “no trend” a cluster of similar profiles is formed based on the ratings stored in the matrix of similarities. The degree of similarity between profile descriptions necessary to cluster them together depends on threshold rating (see legend of Figure 3). Many of the profiles reside in several clusters at this stage. As bigger clusters are preferable, the cluster with highest number of members is taken and all of its member profiles removed from the initial dataset. In the next round, the biggest cluster of the remaining profiles is taken. This process reiteratively removes clusters of profiles until only single-profile clusters remain. In case of coincident sizes of several clusters, the most compact of them (with highest average rating) is taken. The basic premise explaining why large clusters are preferable is that the expression profiles appear in clusters (or clouds) in the virtual space of similarities (where more similar profiles lie closer to each other). If a fixed diameter part of the space is taken for a cluster then clusters with more members are more likely to contain most of the cloud (or its denser part). In other
words, the profile founding the largest group of similar members is likely to lie near to the center of the cloud.

5.1.1.6 Output

The first sheet of the produced MS Excel workbook serves as an overview of the results. It contains the input data and lists local trends and relative positions for each profile. It is also possible to find there the cluster to which a particular profile belongs. This feature makes it easier to find any profile in the corresponding context. The values of analysis parameters can be found at the bottom of the sheet.

The individual clusters are visualized in separate MS Excel worksheets named by the content of the first column of input table for the founder profile. In a table form, they include all the supplied annotations and the original values for member profiles with color-coded profile descriptions (red or green background for “rising”/”falling” and blue or yellow font for “higher”/”lower”, respectively). For each profile, its similarity rating to the founder profile of the cluster is also given. As a visual impression usually helps to absorb information faster, graphical output in the form of line-charts was also implemented. It illustrates both the non-filtered shapes of the profiles and also the intensities of the signals on the left-side graph. The right-side chart presents ratios of the values to Control. The latter representation is often more transparent but in some cases it can mask absolute expression changes in abundant transcripts.

5.1.1.7 Comparison with a previously published method

There have been several attempts to address the problem of clustering short time-series data (). Recently, published their method based on creating a set of artificial “expression profiles” and distributing the time-series expression data being analyzed among them according to a correlation coefficient. The expected size of every cluster formed just by chance from the analyzed data around each artificial profile is estimated using a Monte Carlo method and compared to the real counts. If calculated probability of the obtained cluster size occurring at random is very low the authors suggest that the enrichment signals a biologically significant process. This approach seems to be more sensitive in discovering nonrandom accumulation of certain expression profiles while reducing false-positive detection of others (that are likely to occur randomly). In addition to publication of the method principle, its Java implementation (Short Time-series Expression data Miner - STEM) has been launched, making it easy to use for anybody. The application adds a statistical analysis of Gene Ontology (GO) categories distribution among the members of particular gene clusters (). STEM was employed as an alternative method of clustering data derived from experiments presented below to compare
the two methods. Unfortunately, the numbers of genes sharing most of the relevant (meaning not very general) annotations are too low to show a statistically significant enrichment.

Table 3. Comparison of clusters of expression profiles after high MMS concentration treatment created by LoTrEC (A) and STEM (B). Color coding represents belonging of the profile (gene) to a particular cluster according to A. Among the most populated clusters in A, expression generally increases in clusters 1, 4, 6 and 9 while decreases in clusters 2, 3, 5, 7 and 8 (all from left; see SupplementB1_high_MMS.xls). For practical reasons, the smallest clusters form second rows. The table serves rather as an overview of the results. For detailed inspection, it is recommended to open it in MS Excel as the supplementary file SupplementC1_high_MMS.xls

To compare clustering results of the presented procedure to those obtained using STEM, genes were color-coded based on the clusters they belong to after LoTrEC analysis (Table 3, Table 4 and Table 5). The presence of groups of genes with the same color in clusters
produced by STEM means that both of the methods rank expression profiles of these genes as similar.

The results are therefore similar, although the two algorithms make use of different similarity measures (correlation coefficient in STEM versus local similarity of trends and relative positions). Some clusters appear split into two or more parts. However, closer look reveals that the sections of a particular color tend to group together with members of certain other clusters (represented by certain different colors) according to the STEM procedure. For an example, see red, light green and yellow sections in Table 4B. The corresponding clusters share the general trends. These are, however, too loosely specified to define a common cluster in any of the programs that each emphasize another set of differences among the profiles. The genes with white background in the Table 3, Table 4 and Table 5 were not detected as differentially expressed by LoTrEC. On the other hand, there are color-coded genes not found in the STEM selection.

The combined MMS treatment is specific in several aspects. There are two controls for it: the common Control not treated with any concentration of MMS and a pretreated control (C-Control) which was subjected to the low dose for 1 hour, let to recover for 2 hours (together with all the combined treatment samples) and harvested after another 2 hours without MMS. Thus it cannot be counted as the beginning of the time-series but it rather illustrates the state of gene-expression in the cells if only the preconditioning were applied. It is de facto a branching of the time-series. Therefore, it is not allowed to form a trend and is only used for relative position assessment (compared to Control). For the analysis of combined MMS treatment results, LoTrEC was modified to handle the data in this way. Using STEM, the full series of samples (Control, C-Control and the samples treated with both the concentrations) was analyzed. However, it is not completely fair as STEM assumes pure time-series data. Therefore, just the time-series (without the controls) were also submitted. As expected, the full series clustering results are more similar to clusters created by LoTrEC which means less cluster fragmentation in Table 5C compared to B. However, both of the STEM options completely missed the third largest cluster (founded by RAD51, At5g20850; blue in Table 5A) of combined treatment expression profiles which is strongly enriched with DNA-repair associated genes (10 out of 22; only 7 of the 10 genes however were entitled with the corresponding GO-term). Most of the genes do not qualify for the clustering and the rest of them (5 profiles) span 2 (full data series) or 3 STEM clusters. The DNA-repair genes are often missed in gene expression analyses because their expression in not-induced cells tends to be rather low and their induction is typically temporally restrained. Another cluster with predominating DNA-repair genes, 11th from the left in combined treatment (founded by DNA polymerase kappa homolog, At1g47980; violet; 3 out of 4 genes) was also completely missed
by STEM with default settings – only a GPX gene (the only one without any known function in DNA-repair) came out through the initial selection of differentially expressed genes. Although under a specific parameter settings some of the missed profiles would be acceptable for STEM, this phenomenon illustrates the global tendency to neglect data of transcripts that are rare in some (or most) samples. Such data were included in the presented analysis although it must be noted that low expression values are more influenced by random fluctuations and various artifacts. However, there are whole groups of proteins that are not necessary in a big amount in the cells, or they even could be harmful (e.g. homologs of DNA repair gene AlkA can catalyze depurination of intact DNA – ). Their transcripts are therefore only accessible for DNA-array analysis if we accept this kind of risk. Anyway, it is still possible to remove these “less reliable” profiles by custom setting of parameters of the procedure, increasing the minimum acceptable signal and/or minimum differences for detection of differential expression. Default parameters were used in both programs.
Table 4. Comparison of clusters of expression profiles after low MMS concentration treatment created by LoTrEC (A) and STEM (B). Color coding represents belonging of the profile to a particular cluster according to A. Among the most populated clusters in A, expression generally increases in clusters 1, 2, 4, 6, 7, 9 and 11 while decreases in clusters 3, 5, 8 and 10 (all from left; see SupplementB2_low_MMS.xls). For practical reasons, the smallest clusters form second and third rows. The table serves rather as an overview of the results. For detailed inspection, it is recommended to open it in MS Excel as the supplementary file SupplementC2_low_MMS.xls.
Table 5. Comparison of clusters of expression profiles after combined MMS treatment
created by LoTrEC (A) and STEM, using only the time-series without controls (B) or the full series (C). Color coding represents belonging of the profile to a particular cluster according to A. Among the most populated clusters in A, expression generally increases in clusters 1, 3, 4, and 7 while decreases in clusters 2, 5, 6 and 8 (all from left; see SupplementB1_c_MMS.xls). For practical reasons, the smallest clusters form second and third rows. In A, the first cluster occupies 2 columns. The table serves rather as an overview of the results. For detailed inspection, it is recommended to open it in MS Excel as the supplementary files SupplementC3_c_MMS1.xls (B) and SupplementC4_c_MMS2.xls (C).
5.1.2 Analysis of cDNA-array results

The algorithm described above selected the differentially expressed genes from the set present on the cDNA array and grouped them according to similarities of the time-course of their transcript abundances. The result comprised 20, 31 and 21 clusters (high concentration of MMS, low and combined doses, respectively). However, 9 – 15 of them (corresponding to 43 to 50% of the cluster numbers) contained less than 3 allocated profiles. For further analysis, it was thus useful to merge several clusters together to form larger groups possessing rather a common fundamental tendency and containing large enough numbers of genes.

The following pages will concentrate on groups of transcripts defined by GO annotations that appear to be more represented among either the up-regulated or the down-regulated lots. Some of these enrichments were statistically significant (based on tests performed by FatiGO+ accessible from Babelomics web-site,) but most of the smaller (more narrowly specified) groups were not due to insufficient number of representatives.

Functional groups of Arabidopsis genes affected by MMS treatment

5.1.2.1 DNA repair

The DNA repair genes induced or repressed by MMS belong to various repair pathways. These subgroups are also mentioned among the GO terms but the number of the transcripts corresponding to each of them is very low. There are genes involved in base excision repair (BER), nucleotide excision repair (NER), mismatch repair, non-homologous end-joining (NHEJ) and homologous recombination. Interestingly, the high MMS concentration only induced a single transcript and 4 were reduced over the period of 8 hours of experiment. A strikingly different regulation takes place in cells treated with low concentration of MMS and with combined dose. They increased abundance of 9 and 13 transcripts, respectively, compared to 1 or 2 decreased ones (Table 6). The expression profiles of majority of the transcripts start at very low values but rise in later time-points.

The only DNA-repair gene that was detected as induced by high MMS dose was one of the Rad23 homologs present in Arabidopsis genome (At3g02540). Rad23 genes generally have a part in NER, the pathway with broadest spectrum of substrates. On the contrary, the other MMS treatments (low and combined doses) resulted in activation of a wide range of DNA-repair genes representing various pathways. While PMS1 (At4g02460) is involved in mismatch-repair, RAD51 (At5g20850) is a key player in homologous recombination. Induction was observed of the transcripts of 3 and 5 genes playing a role in BER (low and combined doses, respectively). BER is the preferred pathway for removing methylated bases.
and it is thus not surprising to find some of BER members to be affected by a DNA-methylating chemical. Three DNA-glycosylases removing modified bases were induced by the treatment – 3-methyladenine-DNA-glycosylase (MAG, At3g12040) by low MMS and endonuclease-three homolog 1 (NTH1, At2g31450) and 8-oxoguanine-DNA-glycosylase (OGG, At1g21710) by the combined dose. The repair DNA-polymerase λ (At1g10520) and nuclease UVH3 (At3g28030) were induced by low and combined MMS treatment. UVH3 is a homolog of RAD2 (Saccharomyces cerevisiae) and mammalian XPG which are endonucleases involved in both BER and NER. Arabidopsis uvh3 mutant is indeed hypersensitive to ultraviolet and ionizing radiations as well as to hydrogen peroxide. It also displays a premature senescence phenotype.

Involvement of NER in repair of DNA-damage inflicted by the low and combined MMS treatments is documented by an increase in abundances of 3 and 4 transcripts. They include in both cases another homolog of Rad23 (At5g38470), ERCC1 (At3g05210, acting as an endonuclease in a complex with UVH1-XPF-At5g41150 which was activated by the combined dose as well) and UVH3.

A TLS gene AtRad30, coding for (At5g44740), a putative homolog of DNA-polymerase η (eta), yeast Rad30 and XP-V, also designated “DNA damage induced protein P” was also induced by the two treatments mentioned above.

The only mRNA of a DNA-repair gene that was consistently repressed by MMS was Rad50 (At2g31970), a subunit of DNA-end processing complex (together with Mre11) employed during NHEJ. Transcript abundance of DRT111 (DNA-damage-Repair-Toleration 111), involved in resolving Holiday structures during homologous recombination, decreased when high and combined doses of MMS were applied.

A homolog of DNA-polymerase θ (theta; PolQ, At4g32700) is one of the transcripts connected with DNA-repair that were cloned (as a fragment of cDNA) for the purpose of the array. It is homologous to Drosophila melanogaster mus308 and mammalian PolQ that were demonstrated to be important for DNA repair and genome stability (chaos1 mouse mutant, an abbreviation of Chromosomal Aberrations Occurring Spontaneously 1) – {Leonhardt 1993; Shima 2004}. Arabidopsis plants mutant in the PolQ-homolog display increased sensitivity to genotoxins mitomycin C and MMS and also an abnormal cell division pattern in meristems. It was designated tebichi (teb; {Inagaki 2006}). The abundance of teb transcript gradually increases from early after the onset of low MMS dose treatment as well as in the later time-points of the treatment with high MMS concentration. However, its non-induced expression is relatively very low and the profile could be missed if a classical approach were used. This is what happened in published Affymetrix data (Genevestigator,
https://www.genevestigator.ethz.ch/at/; ) – the transcript was flagged absent in many experiments and for ratio-based methods it is thus effectively lost. Although teb gene was shown to be connected with DNA-repair in Arabidopsis as well as in other organisms, the corresponding GO-term was not present in the annotation of the gene (as of December 2006). It might be caused by the fact that teb function was revealed first recently and the annotation still needs to be updated.

Table 6. Comparison of sets of mRNAs GO-annotated “DNA repair”, induced (over the black line) and repressed (under the black line) in treatments with high concentration of MMS (left), low concentration of MMS (middle) and the combined treatment (right).

Poly(ADP-ribose)-polymerase (PARP) activity is closely related to DNA-repair, as PARP proteins were shown to serve as sensors of some DNA-damage types, binding to exposed molecule ends. Concurrent poly-ADP-ribosylation of various proteins attracts the repair machinery. PARP interactions were documented for BER () and double strand break repair pathways (). From the 4 PARP genes present on the array, 2 were induced by MMS. APP (At4go2390) by the high concentration and PARP3 (At5g22470) by low dose. Combined treatment elevated abundances of both the transcripts. The activation of Arabidopsis PARP genes after a genotoxic treatment is in agreement with previously published data of , who stressed the plants using ionizing radiation.
5.1.2.2 Programmed cell death

Programmed cell death (PCD) is another process presumably affected by MMS. It is a long-known fact that extensive or persistent unrepaired DNA-damage leads to PCD in various multicellular organisms. Also in plants, including Arabidopsis, genotoxic stress can lead to removal of some affected cells in a regulated and organized way (). Accordingly, MMS was expected to change expression levels of some involved genes. Indeed, several changes were observed (Table 7).

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Symbol</th>
<th>Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>At5g47120</td>
<td>BI-1</td>
<td>High MMS</td>
</tr>
<tr>
<td>At3g28910</td>
<td>Myb30</td>
<td>Low MMS</td>
</tr>
<tr>
<td>At4g19510</td>
<td>protein N-like</td>
<td>Combined treatment</td>
</tr>
<tr>
<td>At2g35520</td>
<td>DAD-2</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Comparison of sets of mRNAs GO-annotated “programmed cell death”, induced (over the black line) and repressed (under the black line) in treatments with high concentration of MMS (left), low concentration of MMS (middle) and the combined treatment (right).

High MMS dose strongly increased abundance of the transcript for BI-1 (Bax-Inhibitor-1; At5g47120) – a known inhibitor of PCD () and decreased that of Myb30 (At3g28910), a transcription factor important for cell death onset (). This set of changes obviously does not support the theory of cell death triggered by a high level of DNA damage. Although another PCD inhibitor, DAD-2 (Defender Against cell Death-2; At2g35520), was repressed, 2 important regulatory transcripts behave in a way clearly contradicting PCD. Using the low concentration, MMS similarly lowered the expression of Myb30, yet also of BI-1. This treatment induced expression of RbohD (Respiratory Burst Oxidase Homolog D; At5g47910) that generates reactive oxygen species, contributing thereby to the cell death program activation. After applying the combined dose of MMS, only RbohD was up-regulated. Neither of the latter two regimes seems to result in PCD as well.

Caspases are the executive proteases accomplishing PCD (apoptosis) in animal cells. The only plant sequences distantly similar to caspases were designated metacaspases (). However, although these proteases are suspected that they might play a role in plant-type PCD (), it has never been proven (). They are thus not annotated as involved in “programmed cell death”. Anyway, transcript abundance of the single metacaspase present on the macroarray (At4g25110) decreased after application of any of the MMS treatments.
Table 8. Comparison of sets of mRNAs GO-annotated “response to wounding”, induced (over the black line) and repressed (under the black line) in treatments with high concentration of MMS (left), low concentration of MMS (middle) and the combined treatment (right).

5.1.2.3 Response to wounding

Induction of genes annotated to be involved in response to wounding documents the fact that general stress response also takes place in MMS treated cells (Table 8).

Among these, several enzymes metabolizing aromatic compounds can be found. These include *phenylalanine-ammonia-lyases 1 (PAL1, At2g37040)* and 2 (PAL2, *At3g53260*) and *anthranilate-synthase alpha 1 (ASA1, At5g05730)*. Interestingly, *ASA1* transcript was detected in an increased amount in the low MMS dose treatment along with *ATR1* (*Altered Tryptophan Regulation 1, At5g60890*) – a Myb transcription factor that regulates expression of *ASA1*.

Wound signaling is known to employ jasmonic acid and ethylene. Indeed, *Ethylene-Response Protein 1 (ETR1, At1g66340)* was induced by all the regimes and *Ethylene Responsive element binding Factor-4 (ERF-4, At3g15210)* by the low and combined treatments. Also an *Allene Oxide-Synthase (AOS, At5g42650)*, a biosynthetic enzyme of jasmonic acid, exhibited an induction in low MMS and combined treatments. Transcription factors of the Myb-family play an important role in stress signaling including wound response. Of these, MMS induced the expression of *Myb91 (At2g37630)*, *Myb95 (At1g74430)*,
and two other Myb-related genes (At1g71030 and At4g37260) in at least two treatments. MYBR1 (MYB-Related protein 1, At5g67300) was repressed in all the samples, Myb30 (At3g28910) and Myb4 (At4g38620) in two of them.

5.1.2.4 Oxidative stress response

Among MMS affected transcripts sorted by enzymatic activity of their protein products (GO-molecular function), groups of glutathione-S-transferases (GSTs) and superoxide-dismutases (SODs) become prominent. There is a large family of GST genes in the Arabidopsis genome with a wide-spread spectrum of roles and activities in the cell. Some GSTs are quite specialized while others respond to a broad spectrum of stress conditions (). However, all members of this family that were detected to change their expression after MMS treatment are annotated as involved in “toxin catabolism”. Each MMS treatment induced 2 to 4 GST transcripts out of the 5 present on the array: increase in transcript levels of GST1 (At1g02930) and GST2 (At4g02520) appeared in all MMS regimes, GST6 (At2g47730) with high and combined doses and yet another GST (At2g29450) in the combined treatment only.

SOD proteins are quenchers of very dangerous superoxide anions formed either by a pathological activity of energy metabolism, or as a part of a defense reaction. Also some chemicals catalyze production of reactive oxygen species including superoxide. SODs contain a metal cation as a cofactor. There are three groups of SODs based on the preferred metal and also on their phylogenetic origin. Arabidopsis cells contain members of all three existing groups and so does the macroarray. There are two cytoplasmic Cu,ZnSODs (At1g08830 and At2g28190), a mitochondrial MnSOD At3g10920 and a plastid FeSOD (At4g25100). All of them are induced by MMS but the FeSOD which is reduced by either low or combined doses of MMS. High MMS concentration resulted only in Cu,ZnSOD (At1g08830) activation.
Table 9. Comparison of sets of mRNAs GO-annotated “response to oxidative stress”, induced (over the black line) and repressed (under the black line) in treatments with high concentration of MMS (left), low concentration of MMS (middle) and the combined treatment (right).

Although SOD proteins play a role in protection of the cells against oxidative stress, only one of them (Cu,ZnSOD At1g08830) reflects this fact in its GO annotation. The GO term “response to oxidative stress” is shared also by catalases 2 (At4g35090, induced by combined MMS treatment) and 3 (At1g20620, activated by high and combined doses) that catalyze decomposition of hydrogen peroxide, PAL1 and PAL2 (up-regulated by high and low MMS) and a group of peroxidases (Table 9). Interestingly, all three peroxidases affected by high MMS concentration (ATP4a - At1g71695, PRXR4 - At5g42180 and chloroplast AtChlGPX - At5g42180) were repressed. On the other hand, each of the other treatments induced 5 (including APX1 - At1g07890, PRXR4 and ATP15a - At5g64120) and repressed 3 “peroxidase” annotated transcripts (e.g. ATP4a). MPK3 (Mitogen-activated Protein kinase 3, At3g45640) that is known to be activated by various stresses () follows an opposite scheme – abundance of its RNA is elevated by high MMS and decreases in the other treatments.

5.1.2.5 Proteolysis

Protein turnover plays a fundamental role in many cell processes (). It is a way of removing inhibitory or unnecessary proteins during cell reprogramming. It also removes damaged and unrepairably denatured protein molecules during stress. For example reports imbalance in DNA repair pathways and consequent increase in killing and mutation rate caused by MMS and UV in yeast (Saccharomyces cerevisiae) with a defect in protein turnover. There are various proteases involved in the system. The most general machinery with broadest spectrum of substrates is the proteasome. It is a large protein complex processing polypeptides “marked” by covalent binding of ubiquitin. The specificity of the system is achieved at the level of ubiquitin conjugation to specific targets. The cascade consists of ubiquitin-activating proteins (E1) ubiquitin conjugating enzymes (E2) and ubiquitin ligases (E3) that are either single polypeptides or make up a part of the anaphase promoting complex or Skp1/cullin/F-box protein complex (APC/SCF). In Arabidopsis, there are 2 different E1s, over 35 E2s and many F-box proteins (). Some of the mentioned genes are represented on the macroarray, as well as several regulatory proteasome subunits and a group of ubiquitin-independent proteases.
While the search for GO-term “proteolysis” among differentially expressed genes after high concentration of MMS yields 5 transcripts induced and 5 repressed, the other treatments caused mainly down-regulation (Table 10). Some of the affected genes can be found in common, anyway. Majority of the affected genes represent known or putative ubiquitin conjugating enzymes (UBC, E2s). All the MMS treatments increased the abundance of \textit{UBC13} (\textit{At3g46460}) transcript and repressed those of \textit{UBC3} (\textit{At5g62540}) and \textit{UBC10} (\textit{At5g53300}). Most of the other UBC mRNA levels decreased in one or two of the MMS regimes, including \textit{UBC2} (\textit{At2g02760}, high and low MMS). \textit{UBC2} is a homolog of yeast \textit{Rad6}, an E2 enzyme involved in DNA-damage resistance and DNA-repair via translesion synthesis (\textit{}). However, its function has not been confirmed in plants and \textit{Arabidopsis} genome also lacks a gene coding for a homolog of \textit{Rad18p} that forms a heterodimer with \textit{Rad6p} and is vital for \textit{Rad6} function in yeast.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textit{At3g46460} & \textit{UBC13} & \textit{At3g46460} \\
\textit{At3g58040} & \textit{?protein} & \textit{At3g46460} \\
\textit{At1g64230} & \textit{UBC-like} & \textit{At3g62980} \\
\textit{At3g62980} & \textit{TIR1} & \textit{At1g12410} \\
\textit{At1g12410} & \textit{nClpP2} & \textit{At5g42990} \\
\hline
\textit{At2g02760} & \textit{UBC2} & \textit{At3g58040} \\
\textit{At5g60360} & \textit{AALP} & \textit{At5g62540} \\
\textit{At5g62540} & \textit{UBC3} & \textit{At5g60360} \\
\textit{At5g53300} & \textit{UBC10} & \textit{At4g25110} \\
\textit{At4g25110} & \textit{Metacaspase} & \textit{At1g12410} \\
\hline
\textit{At2g02760} & \textit{UBC2} (\textit{RAD6}) & \textit{At2g02760} \\
\textit{At5g62540} & \textit{UBC3} & \textit{At5g62540} \\
\textit{At5g41700} & \textit{UBC8} & \textit{At5g42940} \\
\textit{At3g20060} & \textit{UBC} & \textit{At5g42940} \\
\textit{At4g32940} & \textit{VPE gamma} & \textit{At5g42990} \\
\end{tabular}
\caption{Comparison of sets of mRNAs GO-annotated “proteolysis”, induced (over the black line) and repressed (under the black line) in treatments with high concentration of MMS (left), low concentration of MMS (middle) and the combined treatment (right).}
\end{table}

Up-regulated proteolysis-involved genes (other than E2s) were represented \textit{e.g.} by the senescence associated chloroplast protease \textit{nClpP2} (\textit{At1g12410}, by high and low MMS) and \textit{TIR1} (transport inhibitor response 1, \textit{At3g62980}, by high and combined MMS doses). \textit{TIR1} is an F-box containing subunit of SCF-type ubiquitin-ligase complex (E3). SCF\textit{TIR} was shown to be necessary for proper auxin signaling (\textit{}). \textit{Metacaspase} (\textit{At4g25110}) mRNA was down-regulated by all treatments while the transcript abundance of another potential senescence-
associated protease, cysteine protease AALP (At5g60360), was reduced following treatment with high and low concentration of MMS.

5.1.2.6 Cell cycle

Various organisms have been reported to respond to DNA damage by arresting cell cycle progression to prevent propagation of the damage to new cells and also increasing the inflicted injury by the cell division machinery (). The arrest is achieved by a signaling cascade involving proteins able to sense specific DNA lesions, translate the signal and finally members directly regulating cell proliferation. As MMS is a known DNA damaging agent, expression of cell cycle regulators was expected to be influenced by its administration. Cyclins and cyclin dependent kinases are the principal players in progression to the next cell cycle phase. Overexpression of cyclin D3;1 (At4g34160) in Arabidopsis was shown to lead to hyperproliferation caused by induction of both replication and cell division in the affected cells (). Cyclin D3;1 expression is repressed in all the treatments as well as that of cyclin C-like (At5g48630) in low and combined doses (Table 11). On the other hand, combined MMS treatment activates cyclin B2;1 (identical to cyclin 2a, At2g17620) and a pRB (Plant Retinoblastoma protein, At3g12280). RB proteins of animals play a fundamental role in licensing cell cycle progression and thereby act as antioncogenes. They inhibit transition to S-phase until phosphorylated by a cyclin-dependent protein kinase, usually complexed with a D- or E-cyclin. Arabidopsis and maize pRB proteins were demonstrated to interact with D-type cyclins and several components of the system are interchangeable with animal proteins (for a review see ). Strong elevation of pRB transcript level connected with cyclin D3;1 down-regulation is thus likely to signal a cell-cycle arrest inflicted to the cells by MMS treatment.

<table>
<thead>
<tr>
<th>No genes</th>
<th>No genes</th>
<th>At3g12280 RB-related</th>
<th>At2g17620 Cyclin B2;1</th>
</tr>
</thead>
<tbody>
<tr>
<td>At4g34160 Cyclin D3;1</td>
<td>At4g34160 Cyclin D3;1</td>
<td>At4g34160 Cyclin D3;1</td>
<td>At5g48630 Cyclin C-like</td>
</tr>
</tbody>
</table>

Table 11. Comparison of sets of mRNAs GO-annotated “cell cycle”, induced (over the black line) and repressed (under the black line) in treatments with high concentration of MMS (left), low concentration of MMS (middle) and the combined treatment (right).

Cyclin D3;1, rapidly down-regulated by MMS, was shown to be regulated by the cytokinin signaling pathway (). Other transcripts possessing GO annotation “response to cytokinin stimulus” comprise Arabidopsis Response Regulators ARR4 (ATRR1, At1g10470, repressed in
high MMS), **ARR2 (At3g48100)** and **ARR6 (At5g62920)**. The latter two transcripts were also reduced by the high concentration of MMS but induced in both the other treatments.

5.1.3 **Macroarray data evaluation by RT-qPCR**

As discussed in the Introduction, hybridization-based methods suffer from a range of potential problems that can compromise the obtained results. On the other hand, the throughput of the contemporary microarrays allows addressing most of the genes comprising an eukaryotic genome. It is thus advisable to combine the enormous data flow of array hybridization with the precision of another, low-throughput, technique for a subset of transcripts selected on the base of array data. Recently, it is often RT-qPCR, as it is the most quantitative method. RT-qPCR was also used to evaluate the results obtained from the cDNA-array experiment with MMS-treated *Arabidopsis* suspension culture.

![Figure 4. Comparison of three potential internal control transcripts.](image)

Under ideal conditions, an internal control should be represented by a horizontal line.

Three transcripts were tested as internal controls for normalization of the data: **18S rRNA**, protein phosphatase 2A regulatory subunit **PDF2 (At1g13320)** and polyubiquitin 10 **UBQ10** (identical to **SEN3** – senescence-associated protein, *At4g05320*). As the RT reactions contained a constant input of 1µg RNA, a good internal control gene should display an invariable expression level, only disturbed by irregularities in RT. It follows from Figure 4 that **PDF2** had the lowest variability of the three. While **18S rRNA** fluctuated, **UBQ10** transcript was clearly elevated by high concentration of MMS. If pretreatment was used, the **UBQ10** induction appeared even stronger. However, at the low dose, the behavior of **UBQ10** mRNA appeared very similar to that of **PDF2**. Therefore **PDF2** served for normalization obtained for all the other transcripts.

Five transcripts presented in Figure 5 were quantified by RT-qPCR as described in Materials and Methods. The trends of expression profiles are generally similar, although there are also
discrepancies. *Cyclin D3;1* mRNA abundance dropped immediately (before 15 minutes of any treatment) and stayed low for the rest of the experiment, including the whole combined treatment. On the other hand, using array hybridization, the transcript amount decreased gradually for several hours. This might be indicative of another mRNA containing a similar sequence and being subject to slower removal after MMS treatment. However, the trend was retained and the interpretation thereby as well. *TL1 (Thaumatin-Like protein 1, At1g75800)* demonstrated very similar expression profiles between the two compared methods. Progressive lowering of *TL1* mRNA abundance in all treatments was slightly slower in the low MMS dose. The expression of *Myb30* was shown to decrease in high and low MMS by both techniques. However, when the combined treatment was applied, the obtained trends strikingly differ. As results for the other two doses are congruent, it seems that the difference could have arisen from divergent behaviour of the cells subjected to combined MMS treatment. RT-qPCR employing primers specific for *GST1* and *MPK3* cDNAs confirmed induction of these genes by high MMS concentration. However, similar (but weaker) trend in qPCR results was observed for the combined treatment, while no obvious change in expression appeared in the array experiment. Low concentration of MMS additionally elevated mRNA level of *GST1* in the array- but not in the qPCR experiment.
Figure 5. Comparison of expression profiles of 5 genes obtained from macroarray experiment (left) and RT-qPCR (right). The designations are: Cyclin D3;1 - At4g34160, TL1 - At1g75800, Myb30 - At3g28910, GST1 - At1g02930, MPK3 - At3g45640. The curves represent high MMS concentration (♦), low MMS (■) and the combined treatment (▲). Please note that “Control” position of the combined treatment curve corresponds to the pretreated control, C-Control.
5.2 Reverse transcription yield for quantitative PCR normalization

The idea of measuring yield of RT reaction is not completely novel. used PicoGreen for quantifying RNA-DNA hybrid molecules after RT with polyA⁺ RNA added as a template. PicoGreen is a fluorescent dye preferentially binding to double stranded (ds) nucleic acids, however, the fluorescence of RNA-PicoGreen complexes can reach up to 10% of dsDNA-PicoGreen fluorescence at the same concentration (). Their method cannot be thus applied to cDNA reverse transcribed from total RNA. Ribosomal RNA (rRNA), predominant in total RNA, contains many double stranded motifs that bind PicoGreen, and this RNA-borne signal outweighs the cDNA-associated fluorescence. Because total RNA is employed in RT-qPCR assays much more frequently than polyA⁺ RNA, a modification of the method to overcome this limitation would make it more useful.

Removing RNA from the RT reaction mixture is thus a critical step in this procedure. The efficiency of RNA removal was assayed by RiboGreen RNA Quantitation Kit (Invitrogen) as a decrease in fluorescence. Using RNaseA (Qiagen, Carlsbad, USA) resulted in incomplete RNA hydrolysis, even after increasing the reaction temperature to 62°C to reduce secondary structures (according to ). In addition, a component of the reaction partially inhibited PCR intended to prove that cDNA remained intact. As an alternative, alkaline hydrolysis of RNA was employed, using 0.1 M NaOH for 20 minutes at 70°C. The RNA-RiboGreen fluorescence dropped about 100 fold (Figure 6A), and neither higher hydroxide concentration nor increased temperature or extended time lowered the signal further. The cDNA was not affected as assayed by real-time PCR with primers specific for the gene MPK3 (At3g45640), Figure 6B.
Figure 6. RiboGreen fluorescence as a function of first strand cDNA yield. (A) RiboGreen fluorescence of hydrolyzed samples after reverse transcription (RT; ▼), background before RT (●), and the difference between these (■) from reactions at various RNA input. “Mock-hydrolyzed” RT samples in which NaOH was neutralized before adding to the sample (◇) illustrate the efficiency of RNA removal. (B) cDNA survives alkali treatment - relative MPK3 cDNA abundance in the hydrolyzed (●) and mock-hydrolyzed (○) samples from panel A. (C) RiboGreen and PicoGreen staining of identical samples - RiboGreen fluorescence of RT (●) and background (○); PicoGreen fluorescence of RT (▼) and background (▼). (D) RiboGreen fluorescence of reaction mixtures at various Mg²⁺ concentrations - legend is the same as in panel A. In panels A and D, three aliquots of a single RT reaction were independently hydrolyzed, neutralized, and measured; error bars denote the standard deviation.
After the alkaline hydrolysis, obtaining single stranded (ss) cDNA could be expected. As this cannot be measured with PicoGreen, finding another dye was necessary. Although RiboGreen is a part of the RNA Quantitation Kit, it is capable of forming a fluorescent complex with any nucleic acid. Indeed, Figure 7 indicates that while PicoGreen and SYBR GreenI are specific to double-stranded (ds) nucleic acids (in this case DNA), RiboGreen is relatively nonspecific. Its fluorescence is however still about 2-fold higher when it binds to dsDNA than to ssDNA or RNA. For the specificity of the RiboGreen RNA Quantitation Kit is thus critical the use of RNase-free DNase.

Figure 6C documents comparison of the fluorescence of RNA samples hydrolyzed before and after RT with both PicoGreen and RiboGreen. Only RiboGreen fluorescence significantly increased after RT, which confirms single-stranded nature of the hydrolysis-resistant DNA.

**Figure 7. Comparison of PicoGreen, RiboGreen and SYBR GreenI** fluorescence intensity when bound to various types of nucleic acids: dsDNA (▲), ssDNA (■) and total RNA (♦).

The PicoGreen/RiboGreen comparison was performed with a set of RT reactions containing a range of concentrations of free Mg$^{2+}$; since reverse transcriptase requires Mg$^{2+}$ as a cofactor,
as the free Mg$^{2+}$ concentration decreases (and thus RT efficiency declines), the RiboGreen signal is also expectedly reduced (see Figure 6C).

Figure 6A shows RiboGreen signal as measured with a range of different cDNA yields. As might be expected, the amounts of cDNA produced initially increase with the amount of RNA input, then level out as the reaction becomes saturated. In the experiment shown in Figure 6D, RNA input was held constant (1 μg), and the free magnesium ion concentration was decreased by adding various amounts of EDTA. The consequent reduction in RT efficiency was mirrored by lower RiboGreen fluorescence readings.

To demonstrate the utility of the proposed method in quantitative RT-PCR, primers specific for genes MPK3 and UBQ10 were used (see Table 2 in Materials and Methods). For comparison of samples in which RNA input differed, all values were expressed relative to the cDNA abundance measured in the quantitative RT-PCR based on 5 μg input RNA; in the reactions involving different magnesium concentrations, all values were expressed relative to the cDNA abundance measured in the quantitative PCR using the 3 mM Mg$^{2+}$ RT reaction.

The relative percent abundances as measured by quantitative PCR were then compared to the relative fluorescence measured by RiboGreen. For different RNA input amounts in RT reactions primed by the dN$_6$ primer, the relative measures are shown in Figure 8A and the correlation is presented in Figure 8B. Figure 8C shows that the degree of correspondence is equally high in experiments in which RT was performed using the dT$_{23}$dV primer. Similarly, Figure 8D demonstrates a high level of correlation between relative fluorescence and relative abundance in reactions performed using dN$_6$-primed RT with various Mg$^{2+}$ concentrations.

The one-to-one correspondence that was observed in these experiments confirms the suitability of the RiboGreen fluorescence reading as a correction factor: if these samples were normalized according to the RiboGreen readings, the resulting abundance values would vary only slightly. They would thereby confirm that MPK3 and UBQ10 expression levels did not fluctuate under the conditions of the experiment (they only appeared to differ because of the differences in global RNA input or RT efficiency; this is not surprising as we always used the same RNA stock). It should be noted that some data points correspond to conditions in which the RT reaction is saturated. Saturation occurs if the RNA input to enzyme activity ratio is too high, due to either excess RNA (e.g., inputs > 1 μg RNA/50 U MoMLV, dN$_6$ primed, see Figure 8A) or RT inhibition (EDTA chelation of Mg$^{2+}$). Under saturation, normalization to RNA input becomes misleading; however, the fluorescence intensity attributable to cDNA yield correlates with levels of specific cDNA estimated by quantitative PCR whether or not the reaction is saturated.
Figure 8. Correspondence of RiboGreen fluorescence to the specific cDNA abundance.

(A) Relative abundance of MPK3 (black bars) and UBQ10 (green open bars) compared with relative RiboGreen fluorescence (red circles and line) in dN6 primed RT reactions with a range of RNA inputs. The values of 5 μg sample were set to 100. For RiboGreen fluorescence readings, three aliquots of a single RT reaction were independently hydrolyzed, neutralized, and measured; errors bars represent standard deviations. (B) Correlation plot of the data shown in (A); ● = MPK3 and ▽ = UBQ10. (C) Correlation of specific cDNA amount with RiboGreen fluorescence in dT23dV primed RT with a range of RNA inputs. (D) Correlation of specific cDNA amount with RiboGreen fluorescence in dN6 primed RT with a range of free Mg2+ concentration (data from Figure 1D). The values of 3 mM Mg2+ sample were set to 100.
5.3 **Response of tobacco to drought**

To explore the influence of cytokinin storage pool size and/or of the stimulated cytokinin turn-over during drought progression, tobacco lines carrying constructs 35S::ZOG1 and SAG12::ZOG1 were investigated along with wild type plants in our experiment. The two transgenic lines made possible an evaluation of the effect of generalized cytokinin content elevation as well as that of cytokinin overproduction restricted to senescing tissues on tobacco water deficit response.

My part in the project was to estimate the transcript abundances of the ZOG1 transgene and of genes related to drought-response and its potential connection to cytokinin metabolism. Helena Štorchová and Jana Dobrá and Marie Havlová participated on the work as indicated bellow and in Materials and Methods.

5.3.1 **Drought stress application**

Drought stress was achieved by cessation of watering. To illustrate the reduction of water content, decrease of the soil moisture was estimated as a loss of pot weight. A major part of the final reduction of pot weight occurred during the first 3 days; further drought resulted in only marginal differences. This was probably mainly due to a high initial transpiration rate of plants with still fully open stomata. Figure 9A shows the values for SAG12::ZOG1 plants. The reduction of water content in the plant body can be expressed as the relative water content (RWC). Figure 9B documents the RWC trend of SAG12::ZOG1 plants after 1 day, 6 days and 11 days of drought treatment. After 6 days of water deficit, young leaves appeared to retain more water than old ones; however, after prolonged treatment, this difference vanished. In well-watered plants, the leaf number changed during 7 days of the experiment from 10 - 11 to 13-14. The stressed plants did not produce any new leaves; in addition they progressively lost 1-2 oldest leaves.
Figure 9. Time course of the water loss from \textit{SAG12::ZOG1} plants.
(A) Pot weight of control (watered; blue columns) and drought-stressed (magenta columns) plants. (B) Relative water content in control (blue) and drought stressed plants – upper 2 leaves (magenta) and lowest 2 leaves (yellow) are depicted separately. Results from 4 – 9 plants are presented (mean value ± SD are shown). Data for Figure 9 were produced by Marie Havlová.

5.3.2 Quantification of transcript abundances

Quantitative RT-PCR was employed to measure transcript abundances of the genes \textit{Cig1}, \textit{CRK1}, \textit{NiERD10B} and the \textit{ZOG1} transgene in individual leaves of the control (watered) and drought treated plants. The expression levels were normalized to actin (\textit{Tac9}) or to \textit{18S} ribosomal RNA. As the normalization to these two transcripts led to similar results, both of them seem to be useful as reference genes. Data presented below were normalized to \textit{Tac9}. 
The third tested reference gene, coding for ribosomal protein L25, possessed higher variability and therefore was not suitable for data normalization (Figure 10). ANOVA implemented in XLSTAT MS-Excel add-in was employed to test the influence of 4 factors on expression of the examined genes. These factors were: Genotype of the plant, growth conditions (drought or control), time of sampling (i.e. severity of the respective stress) and leaf position on the stem, numbered from the youngest unfolded one downwards. The P-values given below are results of the test, if not stated otherwise. As the experiment performed with SAG12::ZOG1 plants differed in some aspects (see Material and Methods), it would not be fair to directly compare its results to those of the other lines. However, as the basic scheme of the experiments remained the same, the data will be presented simultaneously.

![Comparison of calculated relative abundances of 18S rRNA and L25 transcripts normalized to Tac9. In this projection, a smooth horizontal line would represent a perfect correlation of expression levels of the respective transcript with those of Tac9.](image-url)

**Figure 10**

5.3.2.1 **ZOG1**

The transcript abundance of ZOG1 was assessed in the transgenic lines. 35S promoter is generally considered to be constitutive under various conditions. However, analysis of individual leaves of two plants grown under moderate and severe stress and two control plants revealed that ZOG1 transcript level was significantly higher (2.6-fold) in drought stressed plants than in the watered ones (Students t-test, P < 0.0002).

As SAG12 promoter is switched on by senescence (), quantification of SAG12::ZOG1 expression effectively visualized the onset of senescence in both stressed and control plants. The impact of drought stress on SAG12 promoter activity is shown in Figure 11. Considerable stimulation of ZOG1 expression started in the sixth leaf (from the top) of the drought stressed plants and in the eighth leaf of the control plants. It corresponded
approximately to the same position in an apex-base metabolic gradient because of difference in leaf numbers due to growth cessation under drought conditions. However, the induction of \textit{ZOG1} expression was significantly higher in stressed plants ($P = 0.02$; Students t-test using leaves that reached the plateau of \textit{ZOG1} expression - lowest 5 ones of control and 3 of stressed plants).

**Figure 11. Expression of \textit{ZOG1} in control and drought-stressed plants (35S::ZOG1 - □, SAG12::ZOG1 - ▲).** For \textit{SAG12::ZOG1}, data from 6\textsuperscript{th} day of drought are shown, for 35S::ZOG1 average values from 4\textsuperscript{th} and 7\textsuperscript{th} day of drought are shown. \textit{ZOG1} expression levels were normalized to actin (\textit{Tac9}).

### 5.3.2.2 \textit{NtERD10B}

The abundance of \textit{NtERD10B} transcript was generally induced by drought ($P=0.01$; Figure 12). \textit{NtERD10B} was expressed at a stable level throughout the course of the experiment in the control plants. However, under stress conditions, its induction appeared already after one day and increased even further with the ongoing stress conditions. The mRNA was detected in all the leaves, with a slight ascending tendency towards older leaves. 35S::ZOG1 plants reached higher average \textit{NtERD10B} expression levels than wild type and this comparison held also under stress conditions.
Figure 12. Expression of NtERD10B in control and drought-stressed plants (WT - ▲, 35S::ZOG1 - ▼, SAG12::ZOG1 - ▲▲). For SAG12::ZOG1, data from 6th day of drought are shown, for 35S::ZOG1 average values from 4th and 7th day of drought are shown. The expression levels were normalized to actin (Tac9). Please mention the different scale.

5.3.2.3 cig1

cig1 expression in WT plants followed a profile that rose progressively from low levels in youngest leaves to much higher (up to 20-fold) in the old ones (Figure 13). In stressed plants, the (otherwise low) expression sharply increased only in the 2 lowest leaves. While under control conditions 35S::ZOG1 possessed less cig1 mRNA with the point of abundance increase shifted to older leaves, the expression profile under drought stress was very similar to that of WT. SAG12::ZOG1 possessed increased cig1 mRNA amounts already in relatively young leaves (from leaf number 3) with a peak in middle leaves (5 - 7). Further down along the plant axis, the expression decreased, although it always stayed higher than in the youngest 2 leaves. When drought was applied, cig1 expression throughout the plant decreased to low values, even in the oldest leaves.
Figure 13. Expression of *cig1* in control and drought-stressed plants (WT - •, 35S::ZOG1 - ▼, SAG12::ZOG1 - □). For SAG12::ZOG1, data from 6th day of drought are shown, for 35S::ZOG1 average values from 4th and 7th day of drought are shown. The expression levels were normalized to actin (*Tac9*). Please mention the different scale.

5.3.2.4 CRK1

*CRK1* transcript was generally significantly more abundant in well-watered plants than in the stressed ones under all circumstances (Figure 14). In agreement with previously published results (), it was found throughout the control WT leaves at a relatively low, rather uniform concentration with an indistinct peak around leaf number 5. On average, 35S::ZOG1 reached the highest expression levels of all tested genotypes. The detected mRNA abundances notably varied among individual leaves but also among whole single plants. The control in sampling 2 (pertaining to 4 days of drought) expressed the gene at higher level than any of the other experimental objects, several fold higher than the other 35S::ZOG1 controls. SAG12::ZOG1 plants also appeared to possess a relatively high *CRK1* expression. However, on the contrary to the constitutive *ZOG1* expresser, *CRK1* transcript abundance followed a rather uniform distribution. Under drought conditions, *CRK1* expression in WT dropped approximately 2-fold and in 35S::ZOG1 about 20-fold (still being 2-fold higher than in WT). Also stressed SAG12::ZOG1 plants possessed less *CRK1* mRNA than control. About 10-fold drop was observed in the young leaves, older ones however retained higher transcript levels.
5.3.3 **Comparison of oligo-dT and random hexanucleotide RT priming**

As the RT-qPCR was carried out twice, once with anchored oligo-dT primer (dT23dV) in the reverse transcription and once using random hexamers (dN6), the impact of various priming strategies could be compared on quantification of the transcripts. Figure 15 presents the comparison of obtained crossing points (Cp) from the real-time PCR in a graphical way. While for **CRK1** and **cig1** dN6 priming resulted in lower Cp values (on average by about 1 cycle; Figure 15 A,B), PCR specific for the other transcripts generated Cp higher by 2, 1.9 and 4.5 cycles (**Tac9**, **NtERD10B** and **ZOG1**, respectively; Figure 15 C, D and E) with a random-primed RT. Nevertheless, when observing the trends of gene expression profiles, the two methods provided congruent results.
Figure 15. Comparison of crossing points obtained from real-time PCR of identical RNA samples, reverse transcribed using oligo-dT (●) or random hexamer (■) primers. PCR primers were specific for cDNA of the following genes: *cig1* (A), *CRK1* (B), *Tac9* (C), *NtERD10B* (D) and *ZOG1* (E)
6 Discussion

6.1 Clustering of the expression data

For the purpose of analysis of the data obtained from experiments comprising MMS treatment of Arabidopsis thaliana suspension culture, a novel application named LoTrEC was developed. In contrast to most published procedures, LoTrEC is based on local trends of the expression profile curve combined with low-pass filtering. The low-pass filtering routine is in fact the principle of trend analysis used for various types of noisy or cyclic data. These features improve the resistance of LoTrEC to both random noise and outlying values. The results of LoTrEC processing of the above mentioned data were compared to those produced by STEM, another algorithm dedicated to perform cluster analysis of short time-series gene-expression data.

Both the mentioned clustering algorithms (LoTrEC and STEM) share several characteristics. Either of them deals with time-series data and takes advantage of this fact in the analysis. The expression profiles are transformed to descriptions allowing for only a limited number of levels (states) at each time point. The description specifies two characteristics: the change from the previous value and its relative position compared to other samples (or to Control). Both of the algorithms also seek for large clusters.

However, there are significant differences between the procedures. STEM employs correlation coefficient of the whole time-series as a similarity measure. It was shown that correlation coefficient sometimes fails to recognize an obvious similarity (or dissimilarity) of curve shapes. Outlying values can also severely interfere with the correlation. Indeed, when various similarity measures were tested, variants of correlation coefficient yielded sometimes very doubtful results. It happened mainly in cases where some outlying values occurred. Therefore a novel similarity measure was devised based on comparing corresponding segments of the profile descriptions.

The main difference in the process of cluster formation is that STEM generates theoretical profiles first and then assigns measured data to them. The cluster founders in LoTrEC algorithm are real expression profiles from the analyzed data set. Groups of similar members are formed around each profile that possesses a trend. Clusters representative of largest numbers of profiles are then chosen for further work. As the model profiles in STEM are randomly generated, this adds some variation to the results produced even by repeated analysis of identical data.
STEM offers statistical testing of likelihood that a particular cluster would have so many members by chance. Profile clusters possessing unlikely high counts of members are then suspected to represent biologically relevant pathways or co-regulated genes. This works well for large numbers of genes and for groups of co-regulated transcripts relatively rich in members (in the order of tens). However, for macroarrays possessing just hundreds of genes, only a portion of which are differentially expressed, it is quite rare to obtain a statistically significant enrichment in number of profiles attributed to a single cluster. Similarly, STEM performs testing of randomness of distribution of genes sharing same annotations (based on GO) among clusters. If a single cluster contains e.g. most of the pathogenesis-related genes considered in the study, it is likely to occur for some reason – they are probably co-regulated. Essentially only widely specified (and thus more populated) functional groups could turn out as statistically significantly enriched in some expression profile clusters, if the number of genes under study is not very high. However, such very general GO-terms (“regulation of cellular metabolism” etc.) are often not too informative about the processes occurring in the cell. If this kind of analysis is required for selected clusters (generated by any clustering algorithm) anyway, one can use either STEM (after setting its parameters to effectively skip the clustering routine) or a dedicated tool (such as FatiGO+).

To keep real trends (which are usually plainer than a random zigzag) and to reduce random noise, LoTrEC applies low-pass filtering. This helps to detect subtler expression changes than in standard setups.

Applied cluster analysis differentiates expression profiles that vary not only in the basic trend (up- or down-regulation) but also in the temporal organization of the transcript abundance changes. It means that e.g. a profile possessing a ”falling” trend at 30 min time-point and “rising” at time-points 2 h and 3 h is quite distinct from another one with ”falling” at 2 h (opposite to the previous profile) and “rising” at 5 h and 8 h, although the latter curve might look just shifted to later time points.

The fact that LoTrEC discovered clusters possessing a statistically significant enrichment of a functional group of genes (according to FatiGO+) is indicative of the fact that the obtained clusters have a biological meaning. In addition, one of these groups (cluster number 3 in Table 5A) was not at all detected by the algorithm used for comparison (STEM). As the cluster contained profiles representing a substantial fraction of all the DNA-repair genes considered in the experiment and MMS is known to inflict DNA-damage, the biological relevance of their similar expression profiles appears quite realistic. In addition, there were other examples of GO-term enrichment that were however not significant due to a small size.
of the corresponding clusters (e.g. cluster number 3 of low MMS treatment contained 5 of 10 ubiquitin conjugating enzymes present on the array). LoTrEC allows to adjust the thresholds, sensitivity of trend detection and cluster size or e.g. to increase the influence of relative position on the clustering by setting the parameters used in the cluster analysis.

6.2 Analysis of transcriptomic data of MMS-treated Arabidopsis cells

Arabidopsis suspension culture was incubated in medium supplied with two different concentrations of MMS. To study the phenomenon of adaptation to methylation stress, the combined treatment was also applied, comprising pretreatment with a low concentration followed by the challenge dose. The cells were then harvested at a range of time-points to assess the dynamics of expression changes of selected genes. The transcriptional response differed significantly among the three treatment regimes.

High concentration of MMS that would be lethal for the cells after a prolonged incubation time clearly elicited a stress response. However, the reaction of the cells was not very specific. The predominantly activated functional groups of genes (based on GO annotation) comprised such as “response to wounding”, “response to toxin”, “response to pest, pathogen or parasite”. On the other hand, there was almost no effect (or even a negative one) on the DNA-repair machinery. As MMS is known to damage DNA, the finding is striking. To add more stones to the mosaic depicting state of the cell suffering from high dose of MMS, cell cycle-related transcripts were also investigated. The only observed effect was a fast down-regulation of cyclin D3;1 which plays an important role in the cell cycle progression. The number of differentially expressed proteases and other genes involved in protein catabolism was evenly split to groups of 5 up- and 5 down-regulated members. This contrasts to the preferential reduction of abundance of the transcripts of this type in the other treatments. To summarize it, 5 mM MMS seems to wreck havoc to the cells, as shown by on the level of DNA damage. The only reaction to the treatment that was detected in the presented experiment is a general stress response that does not target the specific lesions produced by MMS. It was previously shown that a massive DNA damage induces programmed cell death. In plants, presented this phenomenon in UV-C irradiated Arabidopsis seedlings. However, the gene expression changes observed with 5 mM MMS obviously contradict such a scenario. 0.5 mM was chosen as the low concentration of MMS because while it still produces a detectable amount of DNA damage, plant cells or even whole plants would survive the treatment. Low MMS produced expression changes in more transcripts than the high dose. In
contrast to the situation described above, 9 genes annotated as involved in DNA repair were activated while only one was repressed. It thus seems that DNA damage inflicted by MMS was properly detected in this case. However, some of the general stress response genes were induced simultaneously. As well as in high MMS concentration, the cell cycle obviously tended to be halted, as an additional cyclin transcript appeared to be down-regulated. This corresponds with the fact that all three affected genes implicated in “cell growth” were repressed. They comprised 2 expansins (At2g39700 and At1g69530) and a peroxidase (At3g49120). Proteolysis was affected by low MMS in a pattern distinct from the high concentration effect. There are more transcripts of this category down-regulated than induced.

The response to low MMS concentration appears to be significantly more specific, aiming, among others, to remove the lesions produced in DNA. However, the pattern of differential expressions does not make PCD likely either in this experimental regime. The combined MMS treatment continues the experiments of . They studied the effect of adaptation to DNA damaging conditions. Assessing the changes in gene expression represents the next step in the elucidation of the adaptation process. In general, the expression pattern obtained from cells subjected to the combined treatment resembles more closely the response to low than to high MMS dose (although the challenge concentration was also 5 mM). Some aspects of the reaction however seem to be even more pronounced in the combined treatment.

For example 7 of the 9 DNA repair genes induced by the low dose are also activated in the combined treatment, along with additional 6 genes. The extent of the general stress response (represented here by transcripts annotated as “response to wounding” and “response to oxidative stress”) remains approximately unchanged. Interestingly, the pattern of transcripts involved in proteolysis was different from the other treatments. There was only a single UBC transcript among the 5 induced proteolysis-related mRNAs (the others being three various proteases and an F-box E3-complex subunit) while 6 UBCs, 2 proteases and a 26S proteasome AAA-ATPase subunit RPT6a were repressed. Metacaspase is a protease that was down-regulated by all the MMS treatments. Although its function in PCD has been never confirmed, its expression profile is consistent with the fact that neither the combined MMS treatment seemed to induce apoptosis. However, reduced sensitivity to induction of PCD might be a trait advantageous for cell survival in a suspension culture.

From the presented data, it appears that preconditioning of the cells with a lower concentration of MMS indeed resulted in an adaptation. The pretreatment rendered the transcriptional response of the cells more similar to the low MMS concentration. The expression profiles of many genes indicated a more specific (and thereby presumably more
efficient) reaction to the stress, contrasting to the high MMS concentration applied alone. As MMS was shown to modify proteins, a possibility arises that molecules involved in a step of signaling could suffer methylation damage that would inactivate them. If the response had been prepared during the preconditioning it could fully develop even in the presence of otherwise prohibitive MMS concentration. This hypothesis would however need to be confirmed.

An alternative theory is that the extent of damage inflicted to an unprepared cell was as big as it disorganized the cell maintenance. The pretreatment induced detoxification and repair pathways in the cells, enabling them to prevent and mend the damage from the very beginning of the high dose application and thereby keep the level of sustained injuries relatively low. The fact that so many transcript abundances were changed following MMS treatments can be attributed to two facts. First, although only a limited number of genes were assayed, their spectrum covered predominantly the ones assumed to have a differential expression. Second, the criteria for detection of the changes in transcript abundances were set relatively sensitive. However, because time-series data and low-pass filtering of the profiles is used, the threshold values for detection are in fact more stringent than they seem to be (e.g. default minimum fold-change of 1.4 would not effectively detect every expression change of 1.4 fold but rather ones significantly higher – depending on the context of the particular profile).

In addition, some genes are up- or down-regulated only in a certain period of time of the treatment. In experiments comprising one or two time-points, many differentially expressed transcripts of this kind could be missed. While not working with ratios (between expression at a particular time-point and in Control, for instance), also such cases can be included in which either of the compared values is very low (effectively zero) as long as the higher of them reaches a set threshold. This is useful to not discard genes that are virtually switched on or off (from or to a very low expression, respectively) and also those possessing a very low expression only eventually rising to reliably measurable levels. DNA-repair genes are a good example of the latter behavior.

6.3 Macroarray data evaluation by RT-qPCR

Biological systems are inherently variable in their forms and also in their reactions to the environment. The experiment addressing the influence of MMS treatment on gene expression of *Arabidopsis thaliana* suspension culture was therefore performed twice. The first experiment provided material for the cDNA-array hybridization. Part of the obtained results were then evaluated by means of RT-qPCR using cells treated with MMS in an independent
experiment following the same procedure. Five genes were selected that possessed clear trends of expression profiles. Both induced and repressed genes were represented. Most of the trends were confirmed, they appeared similar between the two methods. However, there were striking differences as well (Figure 5). These could be explained by either artifacts produced by the employed techniques or by biological variability of the material in conjunction with possible unintended slight differences in the procedure. These inconsistencies comprise $Myb30$ and $MPK3$ (combined treatment) and $GST1$ (low and combined dose). Obviously, the deviation in all three of the transcripts appeared when combined MMS treatment was applied. One possible explanation is that the cells were in a different physiological state and specifically reacted in a distinct way to the combined treatment. However, it might have happened that the deviation resulted from a difference in handling the cells during the experiment. Indeed, the combined treatment comprised higher number of operations than the other two regimes and was therefore more likely to be performed differently. From the results obtained for $GST1$ and $MPK3$, it seems that the pretreatment produced less adaptation to the effects of MMS in the second round of the experiment, rendering the combined treatment more similar to high MMS dose. Anyway, this does not apply to $Myb30$ that possessed a unique expression profile, obviously distinct from those obtained after either high- or low MMS treatments.
6.4 Reverse transcription yield for quantitative PCR normalization

As discussed before, normalization of RT-qPCR results is vital for obtaining meaningful information. Internal control transcripts (ICTs) are the most common means of normalization. However, if not properly evaluated, using an ICT can produce misleading results. For example, one of the ICTs that was thoroughly evaluated in a wide range of conditions and treatments, \textit{UBQ10} appeared to be unreliable for MMS treatment. \textit{UBQ10} mRNA level increased in the later time-points of the high MMS and combined but not low MMS regimes. Additionally, there are no tested ICTs yet for many organisms and establishing a reliable assay can thus be a complicated task.

Therefore, a novel method for RT-qPCR normalization was developed that relies on the total cDNA yield as a factor used for the normalization. The procedure takes advantage of differential susceptibility of DNA and RNA in regard of alkaline hydrolysis. While RNA is damaged to a great extent after 20 min incubation in 100 mM NaOH at 70°C, DNA appears not severely affected. The removal of RNA is a necessary step because there is no fast and easy detection method discriminating between RNA and DNA. Common nucleic acid quantification methods are based either on spectrophotometric (absorbance at 260 nm wavelength) or fluorimetric (using dyes binding to nucleic acids) detection. The latter techniques are generally more sensitive, allowing to detect amounts of cDNA produced in a typical RT reaction. However, RNA usually predominates over the newly synthesized cDNA and most of the measured signal therefore represents the amount of RNA. As the RNA-dependent fluorescence falls to approximately 1 % of the original value after alkaline hydrolysis, even a lower DNA-borne signal can be measured.

Single stranded nature of cDNA resulting from the hydrolysis reaction precludes employing double-stranded DNA specific dyes SYBR Green I and PicoGreen. On the other hand, RiboGreen proved useful for this task. If constant ionic strength and dye concentration are used, RiboGreen fluorescence correlates with RNA input amount and also with RT-efficiency. A range of RT-efficiencies was produced by manipulating concentration of free magnesium ions.

The presented method of direct measurement of cDNA amount makes possible an independent verification of invariant expression of candidate reference genes. Alternatively, it offers an alternative way to normalize mRNA expression levels in the absence of suitable housekeeping genes. As the RiboGreen nucleic acid detection is very sensitive low yield RT reactions may be measured. However, in such cases, random noise influence can increase to unacceptable levels, depending mainly on the sensitivity of applied fluorimeter. As oligo-
dT primed reactions generally produce less cDNA, they tend to suffer from this fact more than random primed RTs. Therefore the gene expression data in this work, when normalized using cDNA yield, possessed substantially more variation than when normalized to our evaluated ICTs.
6.5 Response of tobacco to drought

As cytokinins play a role in drought stress response, it is tempting to speculate that altered cytokinin metabolism could lead to more drought-resistant crop plants. To study the influence of total cytokinin pool on drought response of tobacco, with wild type plants were employed together with two transgenic lines. Both of them express ZOG1, an enzyme converting cytokinins to their inactive storage form, O-glucoside. As the plant compensates for ZOG1 activity by increasing cytokinin production, the total cytokinin pool and turnover is elevated. The transgene was active either in the whole plant (35S::ZOG1) or exclusively in the senescing leaves (SAG12::ZOG1). All the plants were treated with a range of water deficit conditions. Transcript abundances were measured for three tobacco genes connected to drought stress and/or cytokinin signaling. Macroscopically, there was only a marginal difference between the transgenic and wild type plants, however the transcript abundances differed substantially.

To illustrate the expression profiles of ZOG1 transgene, its mRNA level was measured by means of RT-qPCR. The transcript was detected throughout the 35S::ZOG1 plants, although in varying amounts. The fluctuations however did not seem to follow any distinctive pattern. Interestingly, the drought treated plants possessed significantly enhanced ZOG1 expression. However, in a recent experiment performed in our laboratory (Jana Dobrá, personal communication), the ZOG1 expression appeared lower. Therefore, we cannot conclude that 35S promoter driven expression is enhanced by drought but merely that it is influenced by water deficit.

<table>
<thead>
<tr>
<th></th>
<th>Stress</th>
<th>Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 day</td>
<td>6 days</td>
</tr>
<tr>
<td>Leaves</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper</td>
<td>0.0006</td>
<td>0.0036</td>
</tr>
<tr>
<td>Middle</td>
<td>0.0261</td>
<td>0.0283</td>
</tr>
<tr>
<td>Lower</td>
<td>0.9792</td>
<td>1.7981</td>
</tr>
</tbody>
</table>

Table 12. SAG12::ZOG1 relative expression following three intensities of drought stress and subsequent rehydration quantified by RT-qPCR. ZOG1 expression levels were normalized to actin (Tac9). Results from two distinct experiments (each comprising triplicate determination) were combined. “Upper” means the youngest unfolded leaf (number 1), “middle” is leaf number 5 and “lower” is one but lowest of still viable leaves. Data presented in Table 12 were provided by Jana Dobrá.
In SAG12::ZOG1, only older leaves possessed high concentration of the mRNA. This finding confirms the published profile of SAG12 promoter activity. It is also in agreement with the results of Jana Dobrá cited in Table 12 and below.

Comparison of SAG12::ZOG1 expression in upper, middle and lower leaves during the drought stress progression and after 1-day recovery is given in . Apparently, the expression driven by the senescence specific promoter increased under the moderate stress in lower leaves. Indeed, the lowest leaves started to die shortly after the day 6. Under severe stress, at day 11, SAG12 promoter appeared active also in the middle leaves. Interestingly, expression of the transgene was virtually undetectable throughout the whole plant after a single day of rehydration. The leaves that apparently had acquired the fate of senescence were reprogrammed to stay alive.

As expected, water deficit increased the abundance of the mRNA for dehydrin, NtERD10B. The effect of drought was stronger in wild type and SAG12::ZOG1 plants, although 35S::ZOG1 possessed highest average expression of all the lines. This fact is a result of relatively high dehydrin mRNA level in control (well watered) 35S::ZOG1 plants. Detection of higher abundances of NtERD10B transcript in older leaves corresponds well with the fact that old leaves are usually most affected by the stress – they are first to die in conditions of persistent drought.

The observed induction ratio was much lower than some published data on various dehydrins (e.g.). However, NtERD10B transcript was relatively abundant also in the control plants in the presented experiment. Recent data from our laboratory suggest that under modified conditions NtERD10B expression can be induced substantially more.

Transcript abundance of cig1 proline dehydrogenase appears to depend on the leaf age/rank. In standard growth conditions, the gene is active in all but youngest 2-3 leaves. The onset of cig1 expression in 35S::ZOG1 is delayed compared to wild type. Regardless of genotype, cig1 mRNA level is reduced by water deficiency. Only the oldest leaves retain a high cig1 activity, with the exception of SAG12::ZOG1. As the overall cig1 signal is lower in 35S::ZOG1 than in wild type, ZOG1 activity (or high concentration of cytokinin-O-glucosides) appears to negatively influence cig1 expression. In SAG12::ZOG1 plants, possessing high expression of ZOG1 just in the oldest leaves, this could preclude the normal cig1 expression pattern. As other parts of the plant do not express the transgene, a compensatory feedback mechanism of amino acid metabolism might activate proline decomposition in younger leaves than in wild type.
The observed reaction of *cig1* transcript abundance to drought is in agreement with the osmoprotective function of proline (). The oldest leaves are obviously senescing (possessing macroscopic signs of ageing and *SAG12* promoter activation) and they are likely to die if the conditions do not ameliorate. Proline decomposition might then help nutrient recuperation from them by providing glutamic acid, the central point of amino acid metabolism making a connection to sugar metabolism. As proline serves to bind water in the cells, its removal by the dehydrogenase might enable the plant to redistribute water from the dying old leaves to other parts of the body. To support this theory, it would be helpful to know the proline concentration in the leaves. This kind of analysis will be performed soon.

Receptor protein kinases constitute an inevitable component of many signaling pathways in all eukaryotic organisms. Plants possess a large family of receptor-like protein kinases (RLKs). For example about 600 RLK genes were discovered in the genome of *Arabidopsis thaliana* (). Larger genomes can contain even more members of the family – e.g. rice has about 1100 RLKs (). *CRK1* is a RLK, shown to be negatively regulated by addition of an active cytokinin, benzyladenine. However, the reaction was only transient. On the contrary, the expression of *ZOG1* from either of the employed constructs increases the concentration of a putative storage form of cytokinins in a long-term manner. This appeared to strongly induce *CRK1* expression in both the transgenic lines. The induction seems not limited to the area of active *ZOG1* transcription but might be proportional to the total *ZOG1* enzymatic activity or the amount of cytokinin-O-glucoside. Therefore *CRK1* mRNA abundance in *SAG12::ZOG1* background was elevated throughout the whole shoot. Although water deficit significantly reduces *CRK1* expression levels in all lines, the transgenics still maintain notably elevated levels. *CRK1* therefore might play a role in systemic signaling, integrating clues from various sensory pathways. *CRK1* transcript abundance exhibited the highest degree of variation even among control plants, predominantly in *35S::ZOG1* genotype. The fact that the fluctuations appeared among whole plants (and not single samples) and that they were specific for *CRK1* mRNA is our evidence that it was not a technical error of transcript abundance measurement. The discrepancies rather reflected physiological differences of unknown nature present among the plants, although they lived in an apparently homogeneous environment.

The data presented here are a result of a pilot project, finding the optimum conditions for a larger screen being performed recently. That is the main reason why the conditions were changed in the course of the experiment. This made the interpretation of the data more complicated; on the other hand, it provided more experience and a solid fundament for
systematic testing of more genes and transgenic lines. The results of this follow-up study should be published soon.

6.5.1 **Comparison of oligo-dT and random hexanucleotide RT priming**

Reverse transcription output was studied under conditions of two different priming strategies. The results show that there is no best primer generally usable for any transcript, as some mRNAs were more efficiently transcribed with oligo-dT, others with random hexanucleotides. This finding is in agreement with that of . The differences result in different relative expression values among different studied genes. However, as these are relative values, comparable only among measurements of a single mRNA and only if produced with the same reaction setup, the effect on the resulting trends is insignificant. Data derived from oligo-dT primed RT reactions were used for the presented analyses, as this strategy proved superior for the majority of the studied transcripts.
7 Conclusions

1) cDNA array was prepared and used to assay transcript abundances of 376 selected *Arabidopsis* transcripts following various treatments with MMS.

   a) LoTrEC, clustering algorithm based on local trends of expression profiles, was designed and applied to the data. It succeeded to discover functionally relevant clusters of expression profiles.

   b) Transcriptional responses to various MMS treatment regimes were investigated. While high MMS concentration seemed only to induce nonspecific stress reaction, the low and combined MMS resulted in a set of more specific expression changes.

   c) Expression levels of five transcripts were estimated by qRT-PCR. Trends of most of the profiles were confirmed.

2) Expression of 3 genes related to drought stress and/or response to cytokinin were measured by qRT-PCR in wild type and *ZOG1* transgenic plants. Transcript levels of all the genes were altered by water deficit.

   a) Although there are no significant macroscopic differences between wild type and *ZOG1* transgenic plants, the mRNA abundances appeared to be influenced by the genotype.

   b) Leaf position (age) significantly influenced the expression of *cig1* and *ZOG1* driven by *SAG12* promoter.

   c) RT primed with oligo-dT appeared more efficient than random hexanucleotide-primed reaction for 3 out of 5 mRNAs (*ZOG1, Tac9* and *NtERD10B*). The other two genes (*cig1* and *CRK1*) showed marginally more efficiency in random-primed RT.
8 List of cited literature


Kasuga M, Miura S, Shinozaki K, Yamaguchi-Shinozaki K. 2004 Mar. A combination of the Arabidopsis DREB1A gene and stress-inducible rd29A promoter improved drought-


