This is an open access article distributed under the terms of the Creative Commons BY-NC-ND Licence

**Somatic mutations, DNA methylation, and expression of DNA repair genes in *Arabidopsis thaliana* treated with 5-azacytidine**

K.V. KISELEV1,2*, Z.V. OGNEVA1, A.S. DUBROVINA1, N.N. NITYAGOVSKY1,2, and A.R. SUPRUN1,2

Laboratory of Biotechnology, Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, 690022, Russia1

Far Eastern Federal University, Department of Biochemistry, Biotechnology, and Microbiology, Vladivostok, 690090, Russia2

**Abstract**

An inhibitor of DNA methylation 5-azacytidine (5A) is a chemical analog of the nucleoside cytidine. This study investigated the influence of 5A-induced DNA hypomethylation on the accumulation of somatic DNA mutations (nucleotide substitutions, indels) in the *Actin2* 3′ untranslated region, nuclear internal transcribed spacer ITS1-5.8rRNA-ITS2, and the ribulose-1,5-bisphosphate carboxylase/oxygenase gene of *Arabidopsis thaliana* and analyzed concurrent changes in the expression of DNA methyltransferase and DNA repair genes. The 5A treatment (20 mg per 100 g of soil) decreased DNA methylation, and the detected 5A-induced demethylation was associated with the up-regulation of the DNA methyltransferase genes: chromomethylase *AtCMT3*, methyltransferase *AtMET1*, and domains rearranged methyltransferases *AtDRM1* and *AtDRM2*. Cultivation of plants in the presence of 5A led to a considerable increase in the number of single nucleotide substitutions in the analyzed DNA regions of 5A-treated *A. thaliana*. The 5A treatment significantly increased the transcriptions of 7 DNA repair genes (endonuclease *AtARP*, DNA demethylases *AtDME* and *AtDML2*, DNA glycosylase *AtMBD4*, DNA damage-binding protein *AtDDB1*, and photolyases *AtUVR2* and *AtUVR3*) out of the 17 analyzed genes from the base excision repair, nucleotide excision repair, and photoreactivation pathways. However, 5A decreased the transcription of DNA 3′-phosphatase *AtZDP*, DNA repair protein *AtRad23a*, mismatch repair proteins *AtMsh2* and *AtMsh3*. It is possible that the changes in the transcription of the DNA repair genes contributed to the detected increase in the number of single nucleotide substitutions that accumulated in the 5A-treated *A. thaliana*. Taken together, the data indicate that there is an interaction between the processes of DNA methylation and mutation accumulation.

Additional key words: indels, methyltransferase, single nucleotide substitutions.

**Introduction**

The cytosine analog 5-azacytidine (5A) is widely used to study the role of DNA methylation in various biological processes due to its capability of removing methyl groups from DNA. Both plant and animal cells metabolize 5A to 5-aza-2′-deoxycytidine-triphosphate, which then gets incorporated into DNA, and 5A can also be substituted for cytosine. The DNA methyltransferases recognize azacytosine-guanine dinucleotides as substrates and catalyze the methylation reaction by a nucleophilic attack. This results in a covalent bond between the carbon-6 atom of the cytosine ring and the enzyme. The bond is normally removed by β-elimination through the carbon-5 atom, but this latter reaction does not occur with 5A because its carbon-5 is substituted by nitrogen, leaving the enzyme covalently bound to DNA and blocking its DNA methyltransferase function. As a result, the methylation marks are erased during DNA replication (Stresemann and Lyko 2008, Navada et al. 2014). Therefore, 5A is widely used in medicine for anticancer therapy and in different experiments to remove methyl groups from DNA (Estey 2013, Duan et al. 2016).
It has been shown previously that both somatic DNA mutations and changes in DNA methylation accumulate with the age of plant tissues (Boyko et al. 2006, Golubov et al. 2010, Kiselev et al. 2015a, Dubrovina and Kiselev 2016, Ogneva et al. 2016). Furthermore, methylated cytosine (5mc) is known to frequently mutate to thymine in the DNA (Ehrlich et al. 1986). However, little is known on the relationship between DNA methylation and mutation accumulation in plants. Xia et al. (2012) analyzed the mutation rate of methylated cytosines at CpG dinucleotides in human embryonic stem cells and found that mutation rates in methylated CpG sites are higher than in unmethylated CpG sites. To our knowledge, there has been no study correlating a low DNA methylation to mutation rate variations in plants (a single nucleotide substitution, indels). In this study, we analyzed changes in the somatic mutation frequency and transcriptions of DNA repair and DNA methyltransferase genes in response to treatment by 5A, a DNA methyltransferase inhibitor, in Arabidopsis.

Materials and methods

Plants and growth conditions: Seedlings of Arabidopsis thaliana L. ecotype Columbia (7-d-old) were pre-cultured on a 1/2 Murashige and Skoog (MS) medium for 1 week and then planted in pots filled with a commercially available rich soil and grown in a controlled environmental chamber (Sanyo MLR-352, Panasonic, Kadoma, Japan) at a temperature of 22 °C, a 16-h photoperiod, an irradiance of 100 μmol m⁻² s⁻¹, and a relative humidity of 50 %. Aqueous solutions of 5A (TCI, Zwijndrecht, Belgium; 5, 10, 20, 50, and 100 mg of 5A per 100 g of soil) were added to the soil in the pots before planting. The 5A did not considerably affect fresh and dry biomass accumulation of A. thaliana at all doses, but accelerated flowering after three to four weeks of cultivation (data not shown). A moderate concentration of 200 mg of 5A per 1 kg of soil was used for the following experiments.

Extraction of DNA, PCR analysis, and DNA bisulfite sequencing: Total DNA was purified from 20 mg of dried A. thaliana mixed tissue (all types of rosette leaves) using 8-week-old 5A-treated A. thaliana plants as described by Kiselev et al. (2015b). The DNA for control and 5A-treated A. thaliana plants was extracted from three plants growing in different pots. For PCR amplification, partial sequences of the nuclear Actin2 3′UTR (Actin-U1, GenBank NM_112764), a nuclear internal transcribed spacer sequence ITS1 and ITS2 of ribosomal DNA with 5.8S rRNA (ITS, GenBank X52320), and a chloroplast gene encoding the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (rbcL, GenBank U91966) were chosen. The primer sequences and the PCR conditions were given in Kiselev et al. (2015a). For PCR amplification reactions, we used the Pfu polymerase (Silex M, Moscow, Russia). At least eight cloned PCR products derived from individual plants were sequenced for each used DNA region. The 24 clones for each DNA region were obtained from the 3 biological replicates (8 clones per each individual plant).

The total number of small-scale mutations per 1000 nt (single nucleotide substitutions, insertions, and deletions) was determined by DNA sequencing as described earlier (Kiselev et al. 2015a). The cytosine methylation status of the MEA-ISR DNA region of A. thaliana was analyzed using bisulfite sequencing as described in Kiselev et al. 2015b using primers presented in Deleris et al. (2010). Briefly, a total of 1 μg of genomic DNA was subjected to bisulfite modification using an EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA, USA) according to the manufacturer’s instructions. The DNA was converted using the following conditions: 95 °C for 5 min and 64 °C for 2 h. Cloned PCR products of the Actin2 and ITS regions were used as positive controls for the bisulfite chemical reactions. The level of C to T transitions in the converted PCR products was greater than 99 %. After DNA conversion, a 361-bp Actin2 fragment was amplified using primers 5′-AGGAATYGTTYAYAGAAATGTT and 5′-TATACATACATATTTATACACATRCA. Primers 5′-GATGAAGAAAYGTAGYGAATGYGATA and 5′-ACAARCRACCRATAAATRTAAT were used for amplification of a 281-bp product from the ITS DNA region. The primers were designed according to the Zymo Research recommendations. The PCR products were isolated from agarose gels using a Cleanup mini kit (Eurogene, Moscow, Russia) and subcloned as described by Aleyanova et al. (2017). Twenty four clones for each DNA region were obtained from the three biological replicates (eight clones per each individual plant). The BLAST program was used for sequence analysis. Multiple sequence alignments were performed using the Clustal X program.

Extraction of RNA and reverse transcription (RT) real-time quantitative PCR (qPCR): Total RNA was isolated as described Kiselev et al. (2012) from 8-week-old plants growing in the presence of 5A added to the soil (200 mg of 5A per 1 kg of soil) and under control conditions without 5A. The RNA was extracted from three plants growing in different pots. Complementary DNA was synthesized using a cDNA synthesis kit with Oligo(dT)15 primer (Silex, Moscow, Russia) as described by Aleyanova et al. (2016). The real-time qPCRs were performed using a real-time PCR kit (Syntol, Moscow, Russia) and an EvaGreen real-time PCR dye (Biotium, Hayward, USA) as described Dubrovina et al. (2015).
The analysis of qPCR data was performed using the 2−ΔΔCT method (Livak and Schmittgen 2001) with two endogenous controls for cDNA normalization (AtActin2, GenBank NM_112764 and AtGAPDH, GenBank NM_111283, Kiselev et al. 2017). The data shown were obtained from three independent experiments and are averages of eight technical replicates for each independent experiment (four qPCR reactions normalized to Actin and four qPCR reactions normalized to GAPDH expressions for each independent experiment).

The cDNAs of methyltransferases AtCMT3 (GenBank, NM_105645), AtMETI (GenBank, NM_105645), AtDRM1 (GenBank, NM_121542), and AtDRM2 (GenBank, NM_121466) were amplified by real-time qPCR as described in Ogneva et al. 2016. Similarly, real-time qPCR was also used to amplify the cDNAs encoding for endonuclease AtARP (GenBank, NM_129709); DNA damage-binding protein AtDDB1 (GenBank, AY074257); DNA demethylases AtDME (GenBank, NM_120538), AtDML2 (GenBank, NM_111836), AtDML3 (GenBank, NM_119567), and AtROS (GenBank, NM_129207); DNA glycosylases AtMBD4 (GenBank, BT028919) and AtUNG (GenBank, BT029175); mismatch repair proteins AtMSH2 (GenBank, NM_113607), AtMSH3 (GenBank, NM_118686), and AtMSH7 (GenBank, NM_180299); DNA polymerase AtPol (GenBank, DQ446242); DNA repair proteins AtRAD4 (GenBank, BT010359), and AtRAD23 (GenBank, NM_101486); photolyases AtUVR2 (GenBank, NM_101035626); and DNA 3′-phosphatase AtZDP (GenBank, NM_180255).

Results and discussion

As shown by MEA-ISR bisulfite sequencing, the 5A treatment led to a significant decrease in DNA methylation in three different sequence contexts (CG, CHG, and CHH) (Fig. 1A). The analysis of DNA methyl-
transferase transcription reveals that the expressions of the \textit{AtCMT3}, \textit{AtMET1}, \textit{AtDRM1}, and \textit{AtDRM2} genes were up-regulated in the 5A-treated rosette leaves of \textit{A. thaliana} (Fig. 1B). Notably, the 5A treatment also leads to activation of methyltransferase gene expression in grape cell cultures (Tyunin \textit{et al.} 2012). The detected elevation in the transcriptions of DNA methyltransferase genes could serve to recover the DNA methylation and to avoid the negative consequences of total hypomethylation.

![Fig. 3. Expressions of eight genes from the base excision repair (BER) pathway in response to 5-azacytidine (5A) treatment.](image)

We analyzed the effect of the 5A application on the frequencies of nucleotide substitutions, insertions, and deletions in three chosen DNA regions of \textit{A. thaliana}. We sequenced a portion of the nuclear \textit{Actin2} 3′UTR region (\textit{Actin-U1}), the nuclear \textit{ITS1-5.8 rRNA-ITS2} (\textit{ITS}) intergenic region, and the chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase gene (\textit{rbcL}). The \textit{Actin-U1}, \textit{ITS} and \textit{rbcL} regions were selected and used because they showed the highest number of mutations accumulating during the growth and aging of \textit{A. thaliana} in comparison with a mutation number in other seven analyzed DNA regions (Kiselev \textit{et al.} 2015a). In this study, we extracted DNA and RNA from 8-week-old \textit{A. thaliana} exposed to seven weeks of continuous 5A treatment and analyzed the frequency of small-scale mutation events per 1 000 nt of the \textit{Actin-U1}, \textit{ITS}, and \textit{rbcL} nucleotide sequences (Fig. 2).

The total somatic mutation frequency was slightly elevated after the 5A treatment (from 1.17 in control plants to 1.42 nt per 1 000 analyzed nt in 5A-treated \textit{A. thaliana} (Fig. 2A). We observed a significant increase in the number of single nucleotide substitutions (by 1.5-times) after the 5A treatment (Fig. 2B). Also, detected indels decreased by 1.2-times after the 5A
treatment, but this decrease was not statistically significant (Fig. 2B). The mutation accumulation in the protein-coding *rbcL* sequence considerably increased from 0.09 to 0.21 nt per 1 000 nt (2.3-times, Fig. 2C). The mutation accumulation in the *ITS* and *actin-U1* regions also increased by 1.3- to 1.6-times, but this enhancement was not statistically significant (Fig. 2C).

The somatic mutation frequency was higher in the noncoding *ITS* and *actin-U1* regions after the 5A treatment (2.17 and 1.88 nt per 1 000 nt) and in normal conditions (1.38 and 1.51 nt per 1 000 nt) in comparison with the mutation frequency at the *rbcL* region (0.09 - 0.21 nt per 1 000 nt) (Fig. 2C). These results correlate with data that the somatic mutation rate is higher in non-coding DNA sequences than in protein-coding regions in transgenic *Arabidopsis* (Kovalchuk et al. 2000) and with our data reporting that a higher mutation frequency is present in the non-coding DNA regions of *ITS* and *Actin-U1* regions than in coding sequences in *A. thaliana* (Kiselev et al. 2015a).

All detected nucleotide substitutions were unique in the analyzed sequences, and, therefore, they did not result from PCR amplification of one mutated allele as was demonstrated previously (Kiselev et al. 2018). Transitions A:T→G:C were the prevalent type of the detected

### Table 1. Types and percentages of single nucleotide substitutions in the *Actin-U1*, *ITS*, and *rbcL* sequences obtained from the DNA of *Arabidopsis thaliana* shoots growing under the control conditions and in the presence of 5-azacytidine (5A, 200 mg per kg of soil) treatment. Means ± SEs, *n* = 3, significant differences at *P* < 0.05 versus values of controls.

<table>
<thead>
<tr>
<th>Type</th>
<th>Substitution</th>
<th>Control</th>
<th>5A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitions</td>
<td>A→G</td>
<td>27.3 ±2.3</td>
<td>25.2 ±7.9</td>
</tr>
<tr>
<td></td>
<td>G→A</td>
<td>11.2 ±1.5</td>
<td>16.2 ±1.9</td>
</tr>
<tr>
<td></td>
<td>T→C</td>
<td>26.8 ±1.1</td>
<td>19.8 ±4.7</td>
</tr>
<tr>
<td></td>
<td>C→T</td>
<td>11.2 ±1.5</td>
<td>10.8 ±2.6</td>
</tr>
<tr>
<td>Transversions</td>
<td>A→T</td>
<td>3.1 ±0.5</td>
<td>9.0 ±1.4*</td>
</tr>
<tr>
<td></td>
<td>A→C</td>
<td>9.3 ±1.4</td>
<td>13.6 ±1.9*</td>
</tr>
<tr>
<td></td>
<td>T→A</td>
<td>5.0 ±0.4</td>
<td>0*</td>
</tr>
<tr>
<td></td>
<td>T→G</td>
<td>2.0 ±0.4</td>
<td>0.9 ±0.4</td>
</tr>
<tr>
<td></td>
<td>G→T</td>
<td>1.9 ±0.8</td>
<td>0.9 ±0.3</td>
</tr>
<tr>
<td></td>
<td>G→C</td>
<td>1.1 ±0.3</td>
<td>0*</td>
</tr>
<tr>
<td></td>
<td>C→A</td>
<td>0</td>
<td>3.6 ±0.9*</td>
</tr>
<tr>
<td></td>
<td>C→G</td>
<td>1.1 ±0.3</td>
<td>0*</td>
</tr>
</tbody>
</table>

Fig. 4. Expression of nine genes from the nucleotide excision repair (NER), mismatch repair (MMR), and photoreactivation pathways in response to 5-azacytidine (5A). We used whole rosettes (CR, AR) and divided rosettes into leaves (CL, AL) and shoots with inflorescences (CI, AI). 5AL - leaves treated with 5A; 5AI - inflorescences treated by 5A; 5AR - rosettes treated by 5A. Transcription levels of NER genes: *AtDDB1* (A), *AtRad4* (B), and *AtRad23a* (C); MMR genes: *AtMSH2* (D), *AtMSH3* (E), *AtMSH7* (F); photoreactivation genes: *AtUVR2* (G), *AtUVR3* (H); and a DNA polymerase gene involved in all reparation systems (*AtPol*) in two-month-old Arabidopsis plants. Control - the total RNA was extracted from 2-month-old *Arabidopsis* rosettes growing under the standard conditions. 5A: the total RNA was extracted from 2-month-old *Arabidopsis* rosettes growing in the presence of 5A (200 mg per kg of soil). Means ± SEs, *n* = 3; means followed by different letters are significantly different using ANOVA with the Tukey pairwise comparisons, *P* < 0.05.
single nucleotide substitutions accumulating during *A. thaliana* growth under the control conditions and in the presence of 5A (74 - 77 %; Table 1). The 5A exposure significantly increased the number of A→T (up to 9 %), A→C (up to 13.6 %), and C→A (up to 3.6 %) transversions. Also, we detected a significant decrease in the number of T→A, G→C, and C→G transversions (Table 1).

To elucidate molecular mechanisms contributing to 5A-associated somatic mutagenesis, we studied the expressions of 17 DNA repair genes belonging to the base excision repair (BER), nucleotide excision repair (NER), DNA mismatch repair (MMR), and photoreactivation pathways by RT real time qPCR (Figs. 3 and 4; Spampinato 2017). The genes have been selected as well-known DNA repair system representatives belonging to different reparation pathways and with characterized properties and functions (Spampinato 2017). The selected DNA repair genes are known to encode DNA glycosylases (*AtMBD4* and *AtUNG1*), mismatch repair proteins (*AtMsh2*, *AtMsh3*, and *AtMsh7*), DNA damage-binding protein (*AtDDB1*), endonuclease (*AtARP*), DNA polymerase (*AtPol*), DNA demethylases (*AtDML2, AtDML3, AtROS*, and *AtDME*), photolyase (*AtUVR2* and *AtUVR3*), DNA repair proteins (*AtRad4* and *AtRad3α*), and DNA 3'-phosphatase (*AtZDP*).

The 5A treatment significantly increased the transcriptions of six DNA repair genes including BER pathway genes *AtDME, AtDML2*, and *AtMBD4* by 1.2- to 4.4-times (in all analyzed tissues, Fig. 3B,C,E) and *AtARP* by 2.6-times (in leaves, Fig. 3A), a NER pathway gene *AtDDB1* by 1.3- to 3-times (in all analyzed tissues, Fig. 4A), and photoreactivation genes *AtUVR2* and *AtUVR3* (in rosettes, Fig. 4G,H) by 1.7- to 1.8-times. At the same time, the 5A treatment decreased expression of a BER pathway gene *AtZDP* by 1.8-times (in inflorescences, Fig. 3H), a NER pathway gene *AtRad23α* by 1.6-times (in inflorescences), and two MMR genes *AtMsh2* by 1.3- to 1.6-times, and *AtMsh3* by 1.3- to 3.1-times (in all analyzed tissues, Fig. 4D,E). It is possible that MMR genes are mostly responsible for accumulation of base substitutions since MMR genes are known to be involved in the correction of replication errors (Golubov et al. 2010). Thus, the decrease in transcription of MMR genes *AtMsh2* and *AtMsh3* likely resulted in the detected increase of single nucleotide substitutions accumulating in 5A treated plants.

In conclusion, the results show that the 5A-induced hypomethylation in the analyzed DNA regions of *A. thaliana* was accompanied by a higher frequency of single nucleotide substitutions, considerable changes in expression of DNA repair genes, and increased transcriptions of DNA methyltransferase genes. The data suggest that the extent of DNA methylation affected accumulation of single nucleotide substitutions and activity of the DNA repair system in plant DNA.

References


Kiselev, K.V., Shumakova, O.A., Manyakhin, A.Y., Mazeika, A.N.: Influence of calcium influx induced by the calcium
ionophore, A23187, on resveratrol content and the expression of CDPK and STS genes in the cell cultures of Vitis amurensis. - Plant Growth Regul. 68: 371-381, 2012.


