

OPEN

# Mass spectrometry reveals the presence of specific set of epigenetic DNA modifications in the Norway spruce genome

Igor A. Yakovlev <sup>1</sup>, Daniel Gackowski <sup>2</sup>, Abdulkadir Abakir <sup>3</sup>, Marcos Viejo <sup>1</sup>, Alexey Ruzov <sup>3\*</sup>, Ryszard Olinski <sup>2</sup>, Marta Starczak <sup>2</sup>, Carl Gunnar Fossdal <sup>1</sup> & Konstantin V. Krutovsky <sup>4,5,6,7\*</sup>

5-Methylcytosine (5mC) is an epigenetic modification involved in regulation of gene expression in metazoans and plants. Iron-(II)/ $\alpha$ -ketoglutarate-dependent dioxygenases can oxidize 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Although these oxidized forms of 5mC may serve as demethylation intermediates or contribute to transcriptional regulation in animals and fungi, experimental evidence for their presence in plant genomes is ambiguous. Here, employing reversed-phase HPLC coupled with sensitive mass spectrometry, we demonstrated that, unlike 5caC, both 5hmC and 5fC are detectable in non-negligible quantities in the DNA of a conifer, Norway spruce. Remarkably, whereas 5hmC content of spruce DNA is approximately 100-fold lower relative to human colorectal carcinoma cells, the levels of both - 5fC and a thymine base modification, 5-hydroxymethyluracil, are comparable in these systems. We confirmed the presence of modified DNA bases by immunohistochemistry in Norway spruce buds based on peroxidase-conjugated antibodies and tyramide signal amplification. Our results reveal the presence of specific range of noncanonical DNA bases in conifer genomes implying potential roles for these modifications in plant development and homeostasis.

Epigenetic mechanisms are instrumental in plant development and adaptation to environmental stress<sup>1,2</sup>. DNA methylation (5-methylcytosine, 5mC) is a conserved epigenetic mark that is associated with transcriptional regulation in both plants and animals<sup>3-5</sup>. The patterns of DNA methylation dynamically change during cellular differentiation contributing to tissue- and developmental stage-specific gene expression in these organisms<sup>3</sup>. In metazoans, fungi and algae, 5mC can be oxidized to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) by the members of the TET/JBP (ten-eleven translocation/J-binding proteins) family of iron-(II)/ $\alpha$ -ketoglutarate-dependent dioxygenases<sup>6-10</sup>. Since both 5fC and 5caC can be recognized and removed from DNA by thymine-DNA glycosylase (TDG) followed by subsequent regeneration of the abasic site with unmodified cytosine via the base excision repair (BER) pathway, these modifications may serve as intermediates in active DNA demethylation in mammals<sup>11</sup>. Moreover, as the maintenance DNA methyltransferase, DNMT1, prefers hemi-methylated (5mC/C) substrates over hemi-hydroxymethylated (5hmC/C) ones, 5hmC may be instrumental in passive replication-dependent demethylation of DNA in vertebrates<sup>12</sup>. In contrast with

<sup>1</sup>Norwegian Institute for Bioeconomy Research, Pb 115, NO-1431, Ås, Norway. <sup>2</sup>Department of Clinical Biochemistry, Collegium Medicum, Nicolaus Copernicus University, ul. Karłowicza 24, PO-85-092, Bydgoszcz, Poland. <sup>3</sup>Wolfson Centre for Stem Cells, Tissue Engineering and Modelling (STEM), Division of Cancer and Stem Cells, School of Medicine, Centre for Biomolecular Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK. <sup>4</sup>Department of Forest Genetics and Forest Tree Breeding, Georg-August University of Göttingen, Büsgenweg 2, Göttingen, 37077, Germany. <sup>5</sup>Laboratory of Population Genetics, N. I. Vavilov Institute of General Genetics, Russian Academy of Sciences, Moscow, 119991, Gubkina Str. 3, Russian Federation. <sup>6</sup>Laboratory of Forest Genomics, Genome Research and Education Center, Institute of Fundamental Biology and Biotechnology, Siberian Federal University, Akademgorodok 50a/2, Krasnoyarsk, 660036, Russian Federation. <sup>7</sup>Department of Ecosystem Science and Management, Texas A&M University, College Station, Texas, 77843-2138, USA. \*email: [Alexey.Ruzov@nottingham.ac.uk](mailto:Alexey.Ruzov@nottingham.ac.uk); [Konstantin.Krutovsky@forst.uni-goettingen.de](mailto:Konstantin.Krutovsky@forst.uni-goettingen.de)

animal systems, in plants, 5mC can be directly recognized and excised from DNA by specific DNA glycosylases, such as ROS1 and DME, without the preceding enzymatic oxidation<sup>3,13</sup>. In line with these differences in the mechanisms of DNA demethylation, TET/JBP dioxygenases have not been identified in plants to date<sup>14</sup>, and experimental evidence supporting the presence of the oxidized forms of 5mC (referred together as oxi-mCs) in the plant genomes is very limited and ambiguous<sup>15–18</sup>. Thus, although an early study based on chemical derivatization coupled with liquid chromatography/tandem mass spectrometry reported the presence of all three oxi-mCs in a number of plant species<sup>17</sup>, and, according to another study, 5hmC was found to be localized in transcriptionally silent transposable elements in rice<sup>18</sup>, there are conflicting reports on the prevalence of this DNA modification in *Arabidopsis*<sup>15–17</sup>. In the present study, we aimed to systematically examine the levels of non-canonical DNA bases in the buds of Norway spruce (*Picea abies*), a coniferous tree that is extensively studied in long-term breeding programs and is often regarded as one of the “model” species for plant epigenetics<sup>19,20</sup>. In addition, we wanted to compare the abundance of these DNA modifications in spruce with those of human colorectal carcinoma (HCT 116) and human embryonic stem cells (hESCs), which are well-studied regarding DNA modifications<sup>21</sup>.

## Results and Discussion

To determine the prevalence of 5mC, its oxidized derivatives as well as a thymine base modification, 5-hydroxymethyluracil (5hmU)<sup>22</sup> together with the products of DNA base damage (deoxyuridine and 8-oxo-2'-deoxyguanosine)<sup>23–25</sup> in our samples, we employed a sensitive two dimensional ultra-performance liquid chromatography tandem mass-spectrometry (2D-UPLC-MS/MS) that we previously successfully used for quantification of these DNA modifications in cancer cell lines and tumor tissue<sup>26,27</sup> (Supplementary Fig. S1 and Table S1). In addition, we confirmed the presence of modified DNA bases by immunohistochemistry in Norway spruce buds based on peroxidase-conjugated antibodies and tyramide signal amplification.

We did not find any significant difference in general modification levels between epitypes. In agreement with previous studies reporting elevated levels of cytosine methylation in plants<sup>28</sup>, we found that global 5mC content of the spruce DNA was considerably (3.5–4.5 fold) higher compared with human cells (Fig. 1A). Interestingly, previously published genome-wide bisulfite sequencing profiling of Norway spruce also revealed very high levels of 5mC in CG and CHG sequence contexts across the whole genome (74.7% and 69.1% of all the cytosine residues in the corresponding contexts, respectively)<sup>29</sup>, as well as relatively high levels of DNA methylation in the coding regions (21% in CG, 11% in CHG and 1.3% in CHH contexts)<sup>30</sup>. Remarkably, our analyses demonstrated that, 5hmC, albeit detectable in the spruce DNA, was present in the plant samples at approximately 100-fold lower level relative to human colorectal carcinoma cells that are, present in most of the cancers, characterized by global loss of this modification compared to normal mammalian tissue<sup>31</sup> (Fig. 1B). Accordingly, we detected fluorescence signal for both 5mC and 5hmC in the bud tissues (Fig. 2). Unlike 5hmC, the 5fdC content of Norway spruce buds was comparable with that of HCT 116 cells (Fig. 1C), whereas 5caC was undetectable in the all examined conifer samples (Fig. 1D). Correspondingly, both 5fC and 5caC were immunodetected in the bud tissues (Fig. 2), contradicting the mass spectrometry results regarding 5caC and suggesting that the content of these variants might be development-dependent since there was one-month difference between collection dates for both techniques. Moreover, the levels of 5hmU in the DNA derived from Norway spruce and from the human cell lines were also similar (Fig. 1E), and it was also possible to immunodetect it in the tissues (Fig. 2). A recent observation demonstrated that 5hmU and 5fC are recognized by specific proteins<sup>32</sup> and that 5fC is rich in active enhancers involved in tissue development/differentiation<sup>33</sup>.

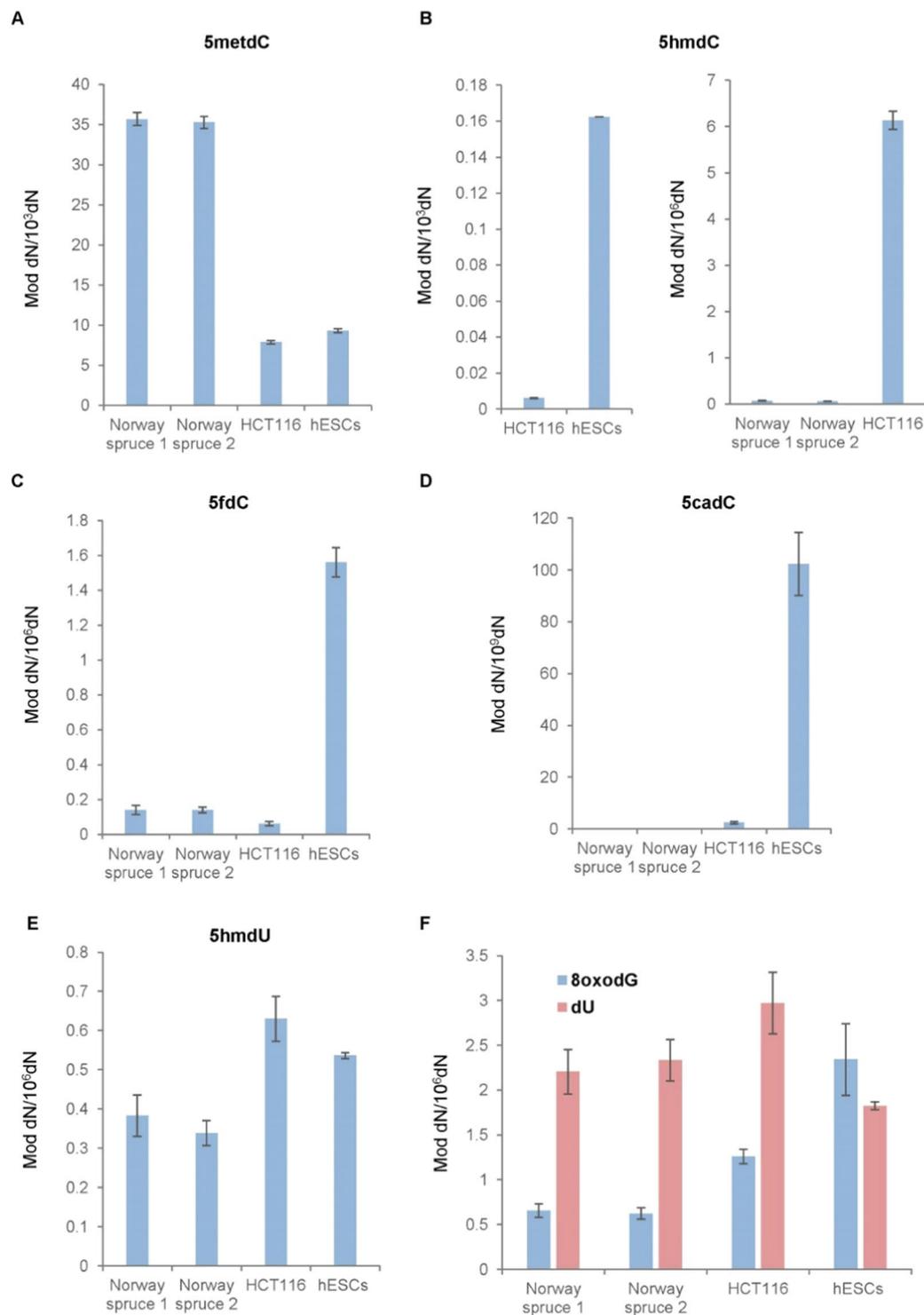
Importantly, analogous to the previously published results<sup>34,35</sup>, none of the oxidized derivatives of 5mC correlated with the levels of the products of DNA damage (deoxyuridine and 8-oxo-2'-deoxyguanosine) in the spruce samples (Fig. 1F) implying that spontaneous free radical-dependent generation of 5hmC and 5fC in plant DNA is extremely unlikely. Although we were unable to identify TET/JBP homologues in Norway spruce using currently available gene models<sup>36</sup>, the specific pattern of oxidized forms of 5mC that we observed in the spruce DNA (near absent 5hmC paralleled by relatively high 5fC and undetectable 5caC) also suggests probable enzymatic origin of these modifications in conifers. In this context, it is worth noting that fungal TET homologue from *Coprinopsis cinerea* exhibited a preference to producing 5fC over other oxi-mCs<sup>9</sup>. Furthermore, a recent report on TET-mediated epimutagenesis of *Arabidopsis* methylome implies the existence of efficient enzymatic machinery allowing removal of 5hmC from DNA and, thus, effectively 5hmC-dependent demethylation, in plants<sup>37</sup>. Correspondingly, as 5hmU is produced via TET/JBP-mediated oxidation of thymine in both kinetoplastids<sup>38</sup> and also, likely, in mammalian cells<sup>39</sup>, our results may indicate both enzymatic origin and potential biological function of this DNA modification in Norway spruce.

## Conclusions

Collectively, our data reveal and confirm the presence of specific set of modified DNA bases in the spruce genome implying their probable non-spontaneous generation in conifers. It is also possible that these epigenetic modifications may play some role to sense environmental changes and cope with the harsh conditions the spruce trees have to pass through. Therefore, further studies are warranted to understand potential roles of these modifications in plant development and homeostasis.

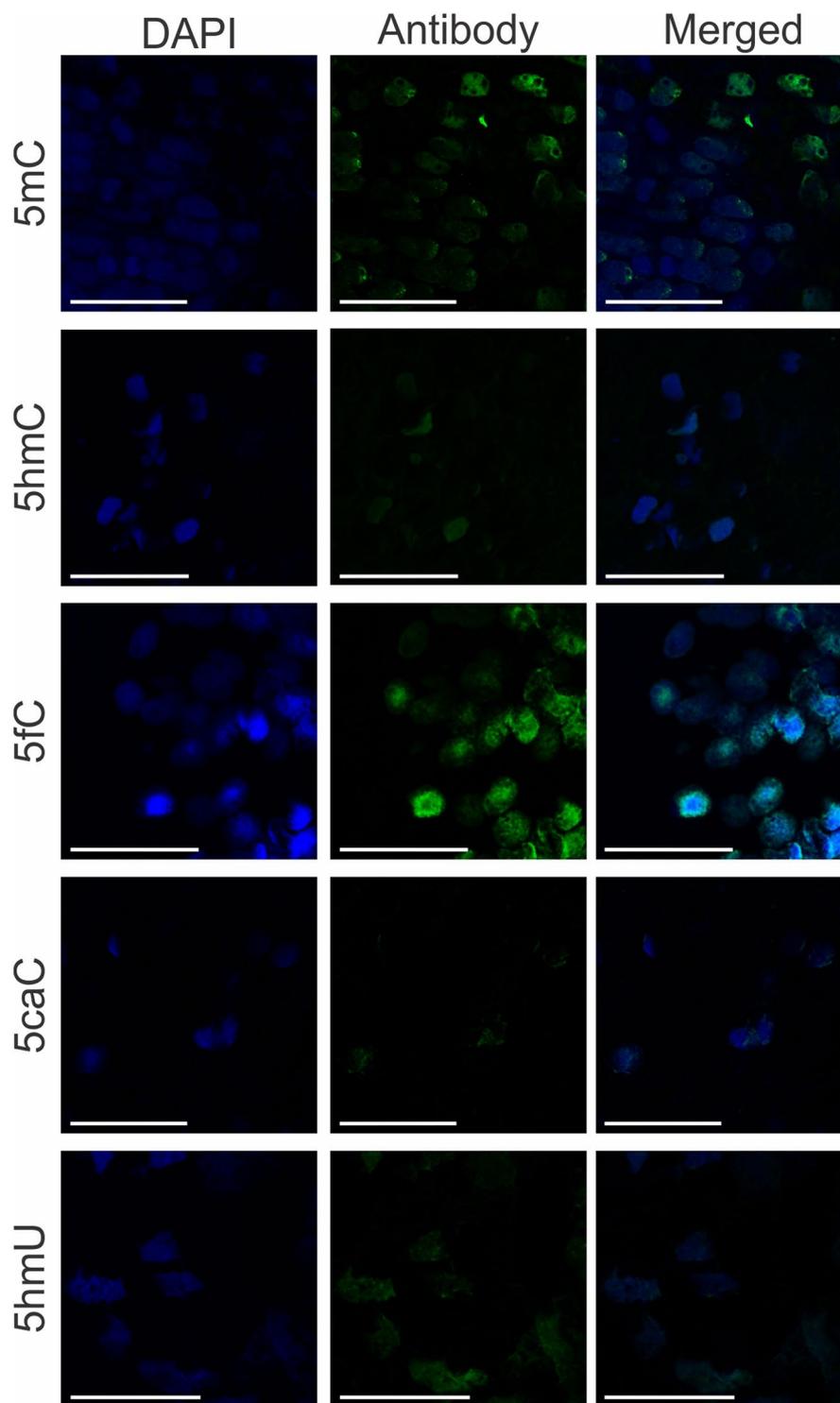
## Materials and Methods

**Plant material and DNA extraction.** DNA samples were collected from the two different epitypes of Norway spruce at the experimental plot in Hoxmark (Norway) in late June after growth cessation and bud formation. Buds were collected from 13-year-old Norway spruce trees produced *in vitro* at two culturing conditions (18°C – cold epitype (1) and 28°C – warm epitype (2) from somatic embryos obtained from a single seed originated in a controlled cross of defined parents (♀#2650 × ♂#2707) of Norway spruce performed in outdoor



**Figure 1.** Mass-spectrometry (MS) quantification of noncanonical bases in the DNA isolated from the Norway spruce bud tissues 1 and 2 and from human colorectal carcinoma (HCT 116) and embryonic stem cells (hESCs): (A) 5-methyl-2'-deoxycytidine, (B) 5-(hydroxymethyl)-2'-deoxycytidine, (C) 5-formyl-2'-deoxycytidine, (D) 5-carboxy-2'-deoxycytidine, (E) 5-(hydroxymethyl)-2'-deoxyuridine, (F) 2'-deoxyuridine and 8-oxo-2'-deoxyguanosine. Vertical line across the top of the bars represents  $\pm$  standard deviation.

conditions, as previously described<sup>40</sup>. Genomic DNA was isolated from terminal and lateral buds of individual trees of Norway spruce using DNeasy Plant Mini Kit (#69104, Qiagen, UK) according to the manufacturer's instructions in several rounds in order to obtain around 50  $\mu$ g of total DNA. The samples were pooled, precipitated with ethanol and dissolved in deionized water.



**Figure 2.** Immunodetection of DNA modified bases in the Norway spruce bud tissue. “DAPI” demonstrates nuclei position, “Antibody” - the modified base fluorescence, and “Merged” - the overlapping of “DAPI” and “Antibody” signals. Antibodies signal coincides with the nucleus as expected. 5mC and 5fC display the highest fluorescence, whereas 5hmU and 5hmC are much lower, and 5caC is the lowest. Microscope magnification is  $63\times$ . Bars are  $50\mu\text{m}$ . Negative controls for DNA modified bases immunodetection shown in Fig. S3.

**Cell culture.** HCT 116 cells were maintained on DMEM (GIBCO) supplemented with 10% bovine serum. HUES7 hESCs were cultured in Essential 8™ (E8) medium with supplement (#A1517001) on Matrigel™-coated tissue culture flasks at  $37^{\circ}\text{C}$  with 5%  $\text{CO}_2$ . Cells were passaged every 3–4 d using TrypLE™ Select Enzyme (#12563029). Genomic DNA from cell cultures was isolated according to standard procedures.

**Mass spectrometry.** DNA samples were incubated with 1 U of nuclease P1 (Sigma-Aldrich) and tetrahydrodouridine (Calbiochem) (cytidine deaminase inhibitor, 10 µg per sample) for 1 h at 37 °C followed by addition of 12 µl of 5% (v/v) NH<sub>4</sub>OH (JT Baker) and 1.3 U of alkaline phosphatase (Sigma-Aldrich) and additional 1 h incubation at 37 °C. The DNA hydrolysates were acidified with CH<sub>3</sub>COOH (Sigma-Aldrich) to final v/v concentration of 2% and ultrafiltered prior to injection. The 2D-UPLC-MS/MS analyses were performed according to the method described earlier by Gackowski *et al.*<sup>27</sup> with some modifications.

Briefly, DNA hydrolysates were spiked with a mixture of internal standards in volumetric ratio 4:1, to concentration of 50 fmol/µL of [D<sub>3</sub>]-5-(hydroxymethyl)-2'-deoxycytidine, [<sup>13</sup>C<sub>10</sub>, <sup>15</sup>N<sub>2</sub>]-5-formyl-2'-deoxycytidine, [<sup>13</sup>C<sub>10</sub>, <sup>15</sup>N<sub>2</sub>]-5-carboxy-2'-deoxycytidine, and [<sup>13</sup>C<sub>10</sub>, <sup>15</sup>N<sub>2</sub>]-5-(hydroxymethyl)-2'-deoxycytidine, [<sup>13</sup>C,<sup>15</sup>N<sub>2</sub>]-deoxyuridine and [<sup>15</sup>N<sub>5</sub>]-8-oxo-2'-deoxyguanosine. Chromatographic separation was performed with a Waters Acquity 2D-UPLC system with photo-diode array detector for the first dimension chromatography (used for quantification of unmodified deoxynucleosides and 5-methyl-2'-deoxycytosine) and Xevo TQ-S tandem quadrupole mass spectrometer (used for second dimension chromatography and compounds analyzed in positive mode after first dimension: 5-(hydroxymethyl)-2'-deoxycytidine and 8-oxo-2'-deoxyguanosine, to assure better ionization at higher acetic acid concentration). At-column-dilution technique was used between first and second dimension for improving retention at trap/transfer column. The columns used were: a Waters Cortecs T3 column (150 mm × 3 mm, 1.6 µm) with precolumn at the first dimension, a Waters X-select C18 CSH (100 mm × 2.1 mm, 1.7 µm) at the second dimension and Waters X-select C18 CSH (20 mm × 3 mm, 3.5 µm) as trap/transfer column. Chromatographic system operated in heart-cutting mode, indicating that selected parts of effluent from the first dimension were directed to trap/transfer column *via* 6-port valve switching, which served as "injector" for the second dimension chromatography system. The flow rate at the first dimension was 0.5 mL/min and the injection volume was 2 µL. The separation was performed with a gradient elution for 10 min using a mobile phase 0.05% acetate (A) and acetonitrile (B) (0.7-5% B for 5 min, column washing with 30% acetonitrile and re-equilibration with 99% A for 3.6 min). Flow rate at the second dimension was 0.3 mL/min. The separation was performed with a gradient elution for 10 min using a mobile phase 0.01% acetate (A) and methanol (B) (1-50% B for 4 min, isocratic flow of 50% B for 1.5 min, and re-equilibration with 99% A up to next injection). All samples were analyzed in three to five technical replicates of which technical mean was used for further calculation. Mass spectrometric detection was performed using the Waters Xevo TQ-S tandem quadrupole mass spectrometer, equipped with an electrospray ionization source. Collision-induced dissociation was obtained using argon 6.0 at 3 × 10<sup>-6</sup> bar pressure as the collision gas. Transition patterns for all the analyzed compounds, as well as specific detector settings were determined using the MassLynx 4.1 Intelli-Start feature in quantitative mode to assure best signal-to noise ratio and resolution of 1 at MS1 and 0.75 at MS2 (Table S2). Calibration curves for MS-detected compounds, recovery, limits of detection and quantitation are presented in Fig. S2. Data were processed in Excel. Means, SDs and RSDs were calculated from 3-5 technical replicates.

**Immunodetection of DNA modifications.** The presence of 5mC, 5hmdC, 5fC, 5caC and 5hmU in Norway spruce was assessed by immunohistochemical analysis in dormant apical buds collected in mid-July. We used a Tyramide-based amplification method<sup>41</sup> to be able to detect the low-abundant DNA modifications in the tissues (5hmC, 5caC, 5fC and 5hmU) while the detection of 5mC did not need of signal amplification, following the below procedure.

Buds were immediately fixed after collection in paraformaldehyde 4% overnight at 4 °C in vacuum. Then, buds were sectioned at 12 µm in a cryotome and placed on glass slides at -20 °C until use. The immunodetection was performed at room temperature unless indicated. Ultrapure water was used to prepare all the solutions. Solutions were removed after each incubation step.

Tissue sections were first covered with phosphate buffer saline (PBS; NaCl 0.137 M, KCl 2.7 mM, Na<sub>2</sub>HPO<sub>4</sub> 0.01 M, KH<sub>2</sub>PO<sub>4</sub> 1.8 mM, pH 7.4) for 5 min and then permeabilized with increasing and decreasing ethanol solutions (25%, 50%, 75%, 100% in PBS) for 5 min each. After washing in PBS for 5 min, cell walls were permeabilized for 30 min at 45 °C with 2% cellulase Onozuka R-10 in PBS (w/v), pH 4.5 in a humid chamber. Sections were washed for 5 min in PBS prior to cell membrane permeabilization with PBS containing 0.05% Tween 20 (v/v) for 30 min. Then, they were washed twice in PBS for 10 min each. The antigen retrieval was performed with a proteinase K treatment for 2 min at 55 °C (20 mg/ml proteinase K in 40% glycerol and 10 mM Tris-HCl, 1 mM calcium acetate and pH 7.5). After washing in PBS for 5 min, the DNA was denaturated for 30 min in a 4 N HCl solution. The sections were washed in PBS for 5 min and then incubated in BLOXALL™ for 20 min to inactivate endogenous peroxidases and washed again in PBS for 5 min. Then, they were incubated in a blocking solution to avoid unspecific antibody binding (5% BSA in PBS; w/v) for 30 min and incubated in 50 µl of a 1:200 dilution of the primary antibody (mouse anti-5mC clone 33D3 [cat. no. MABE146; Sigma-Aldrich], goat anti-5hmC [cat. no. 39092; Active motif], rabbit anti 5fC [cat. no. 61223; Active motif], rabbit anti 5caC [cat. no. 61225; Active motif], mouse anti 5hmdU [cat. no. #ab19735; Abcam, UK] in blocking solution (1% BSA in PBS; w/v) for 1 hour. The excess of antibody was washed twice with PBS containing Tween (as described above) for 10 min each and then were incubated in 50 µl of the corresponding secondary antibody (anti-mouse conjugated with Alexa Fluor 633 (cat. no. A21050, Thermo Fisher) for 5mC, anti-goat horse radish peroxidase (HRP) conjugated (cat. no. A5420, Sigma-Aldrich), anti-rabbit HRP conjugated (cat. no. A0545) in blocking solution (10% BSA in PBS, w/v) for 45 min. Another double wash with Tween 20 was performed (as described above) and sections containing 5hmC, 5caC, 5fC, and 5hmU primary antibodies were incubated in 50 µl of a freshly prepared 1:400 dilution of Tyramide in TSA amplification buffer for 2 min [Tyramide Signal Amplification (TSA) Plus Cyanine 5, cat. no. 415001KT, PerkinElmer]. Immediately after, the sections were washed twice in PBS plus Tween 20 (as previously described) before incubation in a DAPI solution for counteracting staining [1 µg/ml DAPI, 0.1% Tween 20 in water (v/v)] for 15 min. After washing with water 3 times for 5 min each, the samples were let to dry and the mounting medium (cat. no. S3023, Dako) and cover glass were applied.

The samples were visualized in a Leica TCS SP5 running LAS AF software using the appropriate wavelengths for exciting each of the fluorophores. Z-stack function was used to take images in the Z axis. Images were processed with LAS X software from Leica. Negative controls for the secondary antibodies were performed (Fig. S3).

## Data availability

The datasets generated and/or evaluated during the current study are available from the corresponding author on reasonable request.

Received: 27 December 2018; Accepted: 29 November 2019;

Published online: 17 December 2019

## References

- Richards, C. L. *et al.* Ecological plant epigenetics: Evidence from model and non-model species, and the way forward. *Ecol. Lett.* **20**, 1576–1590, <https://doi.org/10.1111/ele.12858> (2017).
- Moler, E. R. V. *et al.* In *Population Genomics: Concepts, Approaches and Applications* (ed Rajora Om) 1–82 (Springer, Cham, 2018).
- Bartels, A. *et al.* Dynamic DNA Methylation in Plant Growth and Development. *International Journal of Molecular Sciences* **19**, 2144 (2018).
- Feng, S., Jacobsen, S. E. & Reik, W. Epigenetic Reprogramming in Plant and Animal Development. *Science* **330**, 622–627, <https://doi.org/10.1126/science.1190614> (2010).
- Finnegan, E. J., Peacock, W. J. & Dennis, E. S. DNA methylation, a key regulator of plant development and other processes. *Curr. Opin. Genet. Dev.* **10**, 217–223 (2000).
- Chan, S. W. L., Henderson, I. R. & Jacobsen, S. E. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nat. Rev. Genet.* **6**, 351–360 (2005).
- Klungland, A. & Robertson, A. B. Oxidized C5-methyl cytosine bases in DNA: 5-Hydroxymethylcytosine; 5-formylcytosine; and 5-carboxycytosine. *Free Radical Biol. Med.* **107**, 62–68, <https://doi.org/10.1016/j.freeradbiomed.2016.11.038> (2017).
- Nabel, C. S., Manning, S. A. & Kohli, R. M. The Curious Chemical Biology of Cytosine: Deamination, Methylation, and Oxidation as Modulators of Genomic Potential. *ACS Chem. Biol.* **7**, 20–30, <https://doi.org/10.1021/cb2002895> (2012).
- Zhang, L. *et al.* A TET Homologue Protein from *Coprinopsis cinerea* (CtTET) That Biochemically Converts 5-Methylcytosine to 5-Hydroxymethylcytosine, 5-Formylcytosine, and 5-Carboxycytosine. *J. Am. Chem. Soc.* **136**, 4801–4804, <https://doi.org/10.1021/ja500979k> (2014).
- Iyer, L. M. *et al.* Lineage-specific expansions of TET/JBP genes and a new class of DNA transposons shape fungal genomic and epigenetic landscapes. *PNAS* **111**, 1676–1683, <https://doi.org/10.1073/pnas.1321818111> (2014).
- He, Y.-F. *et al.* Tet-Mediated Formation of 5-Carboxycytosine and Its Excision by TDG in Mammalian DNA. *Science* **333**, 1303–1307, <https://doi.org/10.1126/science.1210944> (2011).
- Hashimoto, H. *et al.* Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* **40**, 4841–4849, <https://doi.org/10.1093/nar/gks155> (2012).
- Jang, H., Shin, H., Eichman, B. F. & Huh, J. H. Excision of 5-hydroxymethylcytosine by DEMETER family DNA glycosylases. *Biochem. Biophys. Res. Commun.* **446**, 1067–1072, <https://doi.org/10.1016/j.bbrc.2014.03.060> (2014).
- Iyer, L. M., Tahiliani, M., Rao, A. & Aravind, L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* **8**, 1698–1710, <https://doi.org/10.4161/cc.8.11.8580> (2009).
- Liu, S. *et al.* Detection of Oxidation Products of 5-Methyl-2'-Deoxycytidine in *Arabidopsis* DNA. *PLoS ONE* **8**, e84620, <https://doi.org/10.1371/journal.pone.0084620> (2013).
- Erdmann, R. M., Souza, A. L., Clish, C. B. & Gehring, M. 5-Hydroxymethylcytosine is Not Present in Appreciable Quantities in *Arabidopsis* DNA. *G3: Genes|Genomes|Genetics* **5**, 1–8, <https://doi.org/10.1534/g3.114.014670> (2015).
- Tang, Y. *et al.* Determination of oxidation products of 5-methylcytosine in plants by chemical derivatization coupled with liquid chromatography/tandem mass spectrometry analysis. *Analytical chemistry* **86**, 7764–7772, <https://doi.org/10.1021/ac5016886> (2014).
- Wang, X.-I. *et al.* Genome-wide mapping of 5-hydroxymethylcytosine in three rice cultivars reveals its preferential localization in transcriptionally silent transposable element genes. *J. Exp. Bot.* **66**, 6651–6663, <https://doi.org/10.1093/jxb/erv372> (2015).
- Carneros, E., Yakovlev, I., Viejo, M., Olsen, J. E. & Fossdal, C. G. The epigenetic memory of temperature during embryogenesis modifies the expression of bud burst-related genes in Norway spruce epitypes. *Planta* **246**, 553–566, <https://doi.org/10.1007/s00425-017-2713-9> (2017).
- Yakovlev, I. A. *et al.* An adaptive epigenetic memory in conifers with important implications for seed production. *Seed Sci. Res.* **22**, 63–76, <https://doi.org/10.1017/S0960258511000535> (2012).
- Stöger, R. & Ruzov, A. Editorial: Beyond CpG Methylation: New Modifications in Eukaryotic DNA. *Frontiers in cell and developmental biology* **6**, 87–87, <https://doi.org/10.3389/fcell.2018.00087> (2018).
- Gommers-Ampt, J. H., Teixeira, A. J., van de Werken, G., van Dijk, W. J. & Borst, P. The identification of hydroxymethyluracil in DNA of *Trypanosoma brucei*. *Nucleic Acids Res.* **21**, 2039–2043 (1993).
- Krokan, H. E., Drablos, F. & Slupphaug, G. Uracil in DNA – occurrence, consequences and repair. *Oncogene* **21**, 8935, <https://doi.org/10.1038/sj.onc.1205996> (2002).
- de Souza-Pinto, N. C. *et al.* Repair of 8-Oxodeoxyguanosine Lesions in Mitochondrial DNA Depends on the Oxoguanine DNA Glycosylase (OGG1) Gene and 8-Oxoguanine Accumulates in the Mitochondrial DNA of OGG1-defective Mice. *Cancer Res.* **61**, 5378–5381 (2001).
- Cadet, J. & Wagner, J. R. DNA Base Damage by Reactive Oxygen Species, Oxidizing Agents, and UV Radiation. *Cold Spring Harbor Perspectives in Biology* **5**, a012559, <https://doi.org/10.1101/cshperspect.a012559> (2013).
- Fokinski, M. *et al.* Profiles of a broad spectrum of epigenetic DNA modifications in normal and malignant human cell lines: Proliferation rate is not the major factor responsible for the 5-hydroxymethyl-2'-deoxycytidine level in cultured cancerous cell lines. *PLOS ONE* **12**, e0188856, <https://doi.org/10.1371/journal.pone.0188856> (2017).
- Gackowski, D. *et al.* Accurate, Direct, and High-Throughput Analyses of a Broad Spectrum of Endogenously Generated DNA Base Modifications with Isotope-Dilution Two-Dimensional Ultraprecision Liquid Chromatography with Tandem Mass Spectrometry: Possible Clinical Implication. *Analytical Chemistry* **88**, 12128–12136, <https://doi.org/10.1021/acs.analchem.6b02900> (2016).
- Rangwala, S. H. & Richards, E. J. The value-added genome: building and maintaining genomic cytosine methylation landscapes. *Curr. Opin. Genet. Dev.* **14**, 686–691, <https://doi.org/10.1016/j.gde.2004.09.009> (2004).
- Ausin, I. *et al.* DNA methylome of the 20-gigabase Norway spruce genome. *PNAS*. <https://doi.org/10.1073/pnas.1618019113> (2016).
- Heer, K. *et al.* Detection of somatic epigenetic variation in Norway spruce via targeted bisulfite sequencing. *Ecology and Evolution* **0**, 1–11, <https://doi.org/10.1002/ecc3.4374> (2018).
- Ficz, G. & Gribben, J. G. Loss of 5-hydroxymethylcytosine in cancer: Cause or consequence? *Genomics* **104**, 352–357, <https://doi.org/10.1016/j.ygeno.2014.08.017> (2014).

32. Spruijt, C. G. *et al.* Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. *Cell* **152**, 1146–1159, <https://doi.org/10.1016/j.cell.2013.02.004> (2013).
33. Iurlaro, M. *et al.* *In vivo* genome-wide profiling reveals a tissue-specific role for 5-formylcytosine. *Genome Biology* **17**, 141, <https://doi.org/10.1186/s13059-016-1001-5> (2016).
34. Modrzejewska, M. *et al.* Vitamin C enhances substantially formation of 5-hydroxymethyluracil in cellular DNA. *Free Radical Biol. Med.* **101**, 378–383, <https://doi.org/10.1016/j.freeradbiomed.2016.10.535> (2016).
35. Starczak, M. *et al.* *In vivo* evidence of ascorbate involvement in the generation of epigenetic DNA modifications in leukocytes from patients with colorectal carcinoma, benign adenoma and inflammatory bowel disease. *Journal of Translational Medicine* **16**, 204, <https://doi.org/10.1186/s12967-018-1581-9> (2018).
36. Nystedt, B. *et al.* The Norway spruce genome sequence and conifer genome evolution. *Nature* **497**, 579–584, <https://doi.org/10.1038/nature12211> (2013).
37. Ji, L. *et al.* TET-mediated epimutagenesis of the Arabidopsis thaliana methylome. *Nature. Communications* **9**, 895, <https://doi.org/10.1038/s41467-018-03289-7> (2018).
38. Luonen, V. *et al.* Glucosylated Hydroxymethyluracil, DNA Base J, Prevents Transcriptional Readthrough in Leishmania. *Cell* **150**, 909–921, <https://doi.org/10.1016/j.cell.2012.07.030> (2012).
39. Pfaffeneder, T. *et al.* Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat. Chem. Biol.* **10**, 574, <https://doi.org/10.1038/nchembio.1532> (2014).
40. Yakovlev, I. A., Carneros, E., Lee, Y., Olsen, J. E. & Fossdal, C. G. Transcriptional profiling of epigenetic regulators in somatic embryos during temperature induced formation of an epigenetic memory in Norway spruce. *Planta* **243**, 1237–1249, <https://doi.org/10.1007/s00425-016-2484-8> (2016).
41. Abakir, A., Wheldon, L. M. & Ruzov, A. In *Epigenetic Methods in Neuroscience Research* (ed Nina Karpova) 125–137 (Springer New York, 2016).

## Acknowledgements

We thank Yupeng Zhang (Norwegian Institute for Bioeconomy research) for the help during DNA extraction. The study was supported by the Research Council of Norway (FRIBIO Grant#191455/V40 and FRIMEDBIO Grant#240766/F20). A.R.'s lab is supported by Biotechnology and Biological Sciences Research Council [grant number BB/N005759/1] to A.R. A.A. is supported by Medical Research Council IMPACT DTP PhD Studentship [grant number MR/N013913/1] to A.A. We also acknowledge support by the German Research Foundation (DFG) and the Open Access Publication Funds of the University of Göttingen.

## Author contributions

I.A.Y., M.V., A.A. and C.G.F. designed the study, performed the experiments and wrote the draft manuscript together with A.R. and R.O., D.G., M.S. and R.O. performed M.S. K.V.K. and A.R. conceived the study and contributed to writing the manuscript. All authors have reviewed the final version of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-019-55826-z>.

**Correspondence** and requests for materials should be addressed to A.R. or K.V.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019