Multimegabase-scale DNA hypermethylation and condensed chromatin correlate with chromosome-specific rRNA gene silencing in Arabidopsis

Gireesha Mohannath[1,*], Anastasia McKinlay[2,3,†], Navinchandra Venkata Puppala[1,,†], Ramya Enganti[2,†], Gargi Prasad Saradadevi[1,†], Todd Blevins[4] and Craig S. Pikaard[2,3,]

[1]Department of Biological Sciences, Birla Institute of Technology and Science-Pilani, Hyderabad campus, Hyderabad, Telangana, 500078, India.

[2]Department of Biology and Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, IN, 47405, United States.

[3]Howard Hughes Medical Institute, Indiana University, Bloomington, IN, 47405, United States.

[4]Institut de Biologie Moléculaire des Plantes, CNRS, Université de Strasbourg, F-67084 Strasbourg, France.

* To whom correspondence should be addressed.

Tel: +91-40-66303697

+1-812-855-1874

Fax: +91-40-66303998

Email: gireesha@hyderabad.bits-pilani.ac.in

† Joint second authors

ABSTRACT

In eukaryotes, hundreds of ribosomal RNA (rRNA) genes are clustered at chromosomal loci called nucleolus organizer regions (NORs). *Arabidopsis thaliana* has two NORs, one on chromosome 2 (NOR2) and the other on chromosome 4 (NOR4). Each NOR consists of ~ 400 rRNA gene copies. We recently showed that rRNA gene subtypes that map to the NOR2 are silenced during development, whereas those that map to NOR4 are active. In several DNA methylation mutants of Arabidopsis, we show release of the NOR2 silencing to varying degrees. Significantly, the highest release of NOR2 silencing correlates with a maximum loss of cytosine methylation in the CHH context followed by CG, but independent of RNA-directed DNA methylation (RdDM). Next, we show that NOR2 is relatively hypermethylated and NOR4 is hypomethylated using methylation analysis of genomic DNA carried out with different types of methylation-sensitive restriction enzymes. We demonstrate similar differential methylation status between NOR2 and NOR4 in genotypes carrying NORs from multiple ecotypes. Lastly, using Tn5 transposon-mediated transposition assay, we show that NOR2 chromatin is in more condensed state than NOR4.

INTRODUCTION

In eukaryotes, hundreds to thousands of 45S ribosomal RNA (rRNA) genes are arranged as tandem repeats at chromosomal loci known as nucleolus organizer regions (NORs) (Fig. 1A). RNA polymerase I (Pol I) transcribes rRNA genes to produce 45S pre-rRNA transcripts that are then processed into 18S, 5.8S, and 25-28S rRNAs that comprise the catalytic core of ribosomes, which synthesize proteins in cells. (1-3). Although eukaryotic genomes have many rRNA genes, not all of them are transcribed all the time. Rather, the number of genes transcribed varies during development, dictated by the developmental stage-specific need for ribosomes and protein synthesis (4-6). The rRNA gene regulation in eukaryotes has been found to be a form of 'dosage control' that involves selective gene silencing, not selective gene activation (7). Our recent findings in *Arabidopsis thaliana* (ecotype Col-0) revealed that such selective silencing is determined by chromosomal location and not by the rRNA gene sequence variation (8,9).

Ribosomal RNA genes in *A. thaliana* are organized into two NORs; one NOR on chromosome 2 (NOR2) and the other NOR on chromosome 4 (NOR4). Each NOR consists of ~ 400 rRNA gene copies. The 18S, 5.8S, and 25S rRNA sequences are identical among 45S rRNA genes, but some variation exists in the internal transcribed spacer (ITS) and the external transcribed spacer (ETS) regions. Four distinct variants had been defined based on the variation in the 3'ETS region named VAR1 through 4 (Fig. 1B-C) (10). Of these, VAR1, VAR2, and a subset of VAR3 (VAR3a) lack a HindIII site in the 3'ETS region, but the remaining variants (VAR3b,c and VAR4) have a HindIII site (8,9) (Fig.1C).

Earlier, we had shown that the rRNA gene variants that map to NOR2 (VAR1, VAR3a) are selectively silenced as a consequence of dosage control during development. In contrast, the variants that map to NOR4 (VAR2, VAR3b,c, and VAR4) remain active (Fig. 1D)(8). Silencing of many genes in eukaryotes involves cytosine methylation and chromatin condensation. First, we sought to know the specific role of cytosine methylation in all three contexts (CG, CHG and CHH) in selectively silencing NOR2. To address this question, we analyzed various DNA methylation mutants of *A. thaliana*, including mutants of DECREASE IN DNA METHYLATION 1 (*ddm1-2*), METHYLTRANSFERASE 1 (*met1-7*), DOMAINS REARRANGED METHYLASE 2 (*drm2*), CHROMOMETHYLASE (*cmt2-3*) and CMT3 (*cmt3*). DDM1 is a SWI2/SNF2 family chromatin remodeler which acts as a master regulator of DNA methylation (11). It promotes DNA methyltransferases' access to the H1 histone-containing heterochromatin to enable silencing of transposable elements in cooperation with RNA-directed DNA Methylation (RdDM) pathway (12). Loss of DDM1 function resulted in genome-wide demethylation of 70% of all cytosines (13,14). In plants, MET1 is involved in cytosine

methylation in CG context, Chromomethylases (CMT) in CHG context and DRM in CHH context (15). However, another study on CMT2 and DDM1 revealed that most of the cytosine methylation in CHH context is facilitated by DDM1 and mediated by CMT2, but independent of RNA-directed DNA methylation (RdDM). Using RT-PCR analysis, we show that the highest release of NOR2 silencing correlates with a maximum loss of cytosine methylation in the CHH context, followed by CG, but independent of RNA-directed DNA methylation (RdDM).

Further, to determine if DNA methylation and chromatin condensation are involved in the selective silencing of NOR2, we analyzed the cytosine methylation and chromatin status of NOR2 and NOR4 regions using multiple molecular approaches. First, using methylation-sensitive restriction digestion analyses, we show hypermethylation of silent rRNA genes at NOR2 and hypomethylation of the active rRNA genes at NOR4. We demonstrate similar results in different Arabidopsis genotypes containing NORs from multiple ecotypes. Lastly, using Tn5 transposase-mediated transposition assay of native chromatin (16), we show that active NOR4 exists as open chromatin while the silenced NOR2 exists as condensed chromatin.

MATERIAL AND METHODS

Plant material and growth

The *A. thaliana* ecotype Col-0 (stock no. CS22681), and mutants *met1-7* (stock no. SALK_076522), *cmt2-3* (stock no. SALK_012874C), *cmt3*, and *hda6-7* (stock no. CS66154) were obtained from the Arabidopsis Biological Resource Center (ABRC; http://abrc.osu.edu). *drm2-2b* (stock no. SAIL_70_E12) and $Col^{Sf\text{-}NOR4}$ were provided by Andrejez Wierzbicki and Scott Michaels, respectively, and *ddm1-2* and *met1-1* were provided by Eric Richards. All plants were grown under standard growth conditions for Arabidopsis in a plant growth chamber.

RNA isolation

Four-week-old leaf tissue of individual plants was frozen in liquid nitrogen and ground into a fine powder. RNeasy plant minikit (Qiagen) was used to extract total RNA from ground powder. Total RNA was then treated with Turbo DNA-free kit (Invitrogen) to eliminate contaminating DNA.

132 ***PCR and RT-PCR assays***

133 For PCR of genomic DNA to amplify all rRNA variants, PCR amplification was conducted
134 using 22-30 cycles (depending on the source of DNA) of 30 sec at 95°C, 30 sec at 55°C–
135 60°C, and 60 sec at 72°C. For RT–PCR, the reverse transcription reaction was performed
136 using 1 µg of total RNA using SuperScript III reverse transcriptase (Invitrogen). One
137 microliter (~100 ng) of reverse transcription product was then used for PCR amplification of
138 rRNA variable regions or an actin control using 30 cycles of 30 sec at 94°C, 30 sec at 55°C,
139 and 60 sec at 72°C. In all the PCR or RT-PCR reactions, primers 5'allvar and 3'allvar were
140 used to amplify rRNA variable regions (all the 4 variant types). PCR and RT–PCR products
141 were resolved by electrophoresis on 2%–2.5% agarose gels in TBE buffer. Sequences of
142 primers used in this study are given in Figure S1.

143 ***Generation of methylation heat maps***

144 Whole genome bisulfite sequencing data (GSE39901) published by Stroud et al. (2013)(17),
145 was used to assess the methylation state of rDNA in various methylation mutants. The
146 sequencing data were analyzed using Methylpy (18). The reads were mapped to a
147 converted 45S rDNA reference sequence allowing a single mismatch. The methylation state
148 of all cytosines in CG, CHH and CHG contexts was then analyzed and represented as a
149 fraction of methylated cytosines over the total cytosines in each context. Replicates for a
150 given genotype were merged into a single file.

151 ***DNA isolation from leaf tissue and agarose gel pieces (gel elution).***

152 Four-week-old individual plant leaf tissue was incubated for 15 min at 99°C in extraction
153 buffer (200 mM Tris-HCl at pH 8.0, 250 mM NaCl, 25 mM EDTA, 0.5% SDS). Genomic DNA
154 from the supernatant was recovered by ethanol precipitation and resuspended in 60 µL of
155 Tris-HCl (pH 8.0). In some cases, DNA was isolated from leaf tissue of 4-wk-old plants by
156 using Illustra Nucleon Phytopure extraction kit reagents (GE Healthcare). DNA was isolated
157 from agarose gel pieces using QIAquick gel extraction kit (Qiagen).
158

159 ***Nuclei isolation***

160 Nuclei were isolated from 3-week-old Col-0 plant seedlings as described previously (22) with
161 a few modifications. Briefly, 3-week-old plants were harvested and washed in ice-cold water.
162 Then plants were submerged in cold diethyl ether for 5 min and washed twice in ice-cold
163 water. Homogenization of plant material was performed in 100 mL of nuclei isolation buffer

164 (1M sucrose, 50 mM Tris-HCl pH 7.2, 5 mM $MgCl_2$, 5 mM KCl, 10 mM 2-mercaptoethnol, 0.2
165 mM PMSF) with a motorized homogenizer (Fisher Scientific Powergen 700). Homogenate
166 was filtered through 1 layer of Miracloth and 1 layer of nylon mesh (pore size of 50 μm).
167 After centrifugation of the filtered homogenate at 14,000 g for 15 min at 4°C the pellet was
168 resuspended in 10 ml of nuclei isolation buffer and prepared for loading into discontinuous
169 Percoll gradient (15%, 30%, 45%, and 60%) by adding one volume of "100% Percoll
170 solution" (34.23 g sucrose, 5 mL 1 M Tris-HCl pH 7.2, 0.5 mL 1 M $MgCl_2$, 0.5 mL 1 M KCl,
171 34 μL 2-mercaptoethnol, 100 μL 2 mM PMSF and Percoll to 100 mL) to 19 volumes of nuclei.
172 Gradients were centrifuged at 2,000 rpm for 10 min, then 8,000 rpm for 20 min. Nuclei were
173 harvested from 30%/45% and 45%/60% boundaries and pooled. Nuclei were diluted with 5
174 volumes of nuclei isolation buffer, mixed by inversion and collected by centrifugation at 1,500
175 g for 10 min at 4°C. A second wash of nuclei was performed with 25 volumes of nuclei
176 isolation buffer. Final wash was done with 50 mL of Nuclease digestion buffer (50 mM Tris-
177 HCl, pH 8.0, 0.3 M sucrose, 5 mM MgCl2, 1.5 mM NaCl, 0.1 mM CaCl2, 5 mM beta-
178 mercaptoethanol). The resulting pellet was resuspended in 2 mL of Nuclease digestion
179 buffer.

180 ***Transposase-accessible chromatin assay***

181 The assay was carried out based on the methodology described by Buenrostro et al. (16).
182 Oligos (5'-Ez-Tn5 and 3'-Ez-Tn5) were annealed by heating a mixture of equal volumes of
183 100 mM oligos at 95 °C for 5 min, followed by letting the mixture cool down gradually by
184 turning the heating off followed by another incubation at RT for 10 min. Transposition
185 reactions were carried out using 2 μL of purified nuclei, 20 μL of EZ-Tn5 transposase with
186 varying amount of the oligos 5'-Ez-Tn5 and 3'-Ez-Tn5, according to the manufacturer
187 instructions (Epicentre). Incubation was performed at 37°C for 90 min. DNA was isolated
188 using cetyl trimethylammonium bromide (CTAB) method. PCR1 and PCR2 were carried out
189 as described above. For the sequences of the primers and oligos used in this study, refer to
190 the table (Fig. S1).

191

192 **RESULTS**

193 **Selective NOR2 silencing is released in DNA methylation mutants**

194 To determine the relative impact of loss of cytosine methylation in three different contexts
195 (CG, CHG and CHH) on the release of selective NOR2 silencing, we analysed rRNA variant

196 expression in several DNA methylation mutants. We performed RT-PCR analysis using RNA
197 isolated from 4-week-old leaf tissues of various mutants. VAR1 type of rRNA genes
198 constitute ~50% of total 45S rRNA genes, and all of them map to NOR2, which undergoes
199 selective silencing during development (8,10). Therefore, VAR1 expression levels are used
200 as a measure of the release of NOR2 silencing. Among the DNA methylation mutants, VAR1
201 expression levels was found to be the highest among *cmt2-3* and *ddm1-2* mutants followed
202 by *met1-7* (Fig. 2 and Fig. S2). We confirmed high VAR1 expression in these mutants by
203 analysing multiple independent mutant plants using RT-PCR analysis (Fig. S2). However,
204 the *cmt3* and *drm2* (a member of RdDM pathway) mutants displayed low levels of VAR1
205 expression, if any (Fig. 2). The *ddm 1-2* mutation is known to reduce cytosine methylation in
206 all three contexts (CG, CHG and CHH) while *cmt2-3* is known to reduce most of cytosine
207 methylation in the CHH context, independent of RdDM (12). Using published whole-genome
208 bisulfite sequencing data (17), we generated heatmaps of cytosine methylation for a 45S
209 rRNA reference gene sequence (Fig. S3) using Methylpy (18) for multiple DNA methylation
210 mutants and wild-type Col-0 (Fig. 3). The cytosine methylation heatmaps confirm loss of
211 most of the cytosine methylation in CHH context in *cmt2-3* and *ddm1-2* mutants (Fig. 3), in
212 which release of NOR2 silencing was observed to a higher degree (Fig. 2 and Fig. S2). The
213 *hda6-7* mutant was included as a positive control which has already been known to release
214 NOR2 silencing to a very high degree (8,23). Similarly, heatmaps confirm the loss of most of
215 the cytosine methylation in CG context in *ddm1-2* and *met1-7* mutants (Fig. 3). Overall, the
216 patterns of cytosine methylation loss for the rDNA loci align with what has been known of
217 these mutants regarding their effect on genome-wide methylation status. Nevertheless,
218 generating these heatmaps for the rDNA loci was important as most of the earlier genome-
219 wide studies excluded rRNA genes due to their repetitive nature.

220 **NOR2 is hypermethylated while NOR4 is hypomethylated in Col-0**

221 First, we carried out methylation analysis of rRNA variants using HpaII-digested genomic
222 DNA (DNA isolated by grinding leaf tissue) (Fig. 4). Because rRNA variants-defining 3' ETS
223 region lacks HpaII site, HpaII digestion will not have any impact on amplification of rRNA
224 variants. Therefore, we chose to separate HpaII-digested and HpaII-undigested fractions of
225 genomic DNA by resolving them on an agarose gel. The fragments that remained at the top
226 of the gel (>20 Kb) constituted the HpaII-undigested fractions while those resolved at 0.6-1.0
227 Kb range constituted the HpaII-digested fractions (Fig. 4A, left panel). Both the digested and
228 the undigested fractions were gel-eluted, followed by PCR amplification of rRNA variants.
229 Through this analysis, we were able to identify the rRNA variants that were enriched in
230 HpaII-digested or undigested fraction. The PCR data again revealed VAR2 genes from

231 NOR4 to be the most predominant variant among the HpaII-digested products, indicating
232 relative hypomethylation of NOR4 compared to NOR2 (Fig. 4A, right panel).

233 In another experiment, we subjected gDNA to digestion by McrBC, an enzyme which
234 preferentially digests methylated DNA (cytosine methylation). Because rRNA variants-
235 defining 3' ETS region has multiple McrBc sites ($R^mC(N_{40-3000})R^mC$), McrBC digestion will
236 impact PCR amplification of rRNA variants. Therefore gel-elution is not required in this case.
237 McrBC digestion of gDNA followed by PCR amplification of rRNA variants revealed that
238 VAR1 was digested the most and VAR2 the least (Fig. 4B), again indicating
239 hypermethylation of NOR2 and hypomethylation of NOR4. HpaII digestion indicates cytosine
240 methylation levels in the CG context only, but McrBc digestion indicates cytosine methylation
241 levels in all three contexts (CG, CHG, and CHH). Methylation maps of centromere-proximal
242 regions of NORs show that ~ 75 Kb of the NOR2-adjoining region is heavily methylated
243 compared to the NOR4-adjoining region (Fig. S4), which has a protein-coding gene just 3 Kb
244 away (24). It is possible that the heavily methylated NOR2-adjoining region might act as a
245 source for spreading methylation-mediated silencing into the NOR2 region, but this
246 hypothesis remains to be tested.

247 **NOR2 is hypermethylated while NOR4 is hypomethylated in $Col^{Sf\text{-}NOR4}$**

248 An introgression line named $Col^{Sf\text{-}NOR4}$ consists of NOR2 from Col-0 (VAR1 and VAR3 genes)
249 and NOR4 from another ecotype called Sf-2 (VAR1 genes) (8) (Fig. 5A). Sf-2 consists of
250 only VAR1 type of genes, but they can be differentiated from VAR1 genes of Col-0 by HindIII
251 digestion. Sf-2 VAR1 genes have a HindIII site in their 3'ETS region, while Col-0 VAR1
252 genes lack it (8). We also showed that in $Col^{Sf\text{-}NOR4}$, Sf-2 NOR4 is active and Col-0 NOR2 is
253 selectively silenced (8). The methylation status of rRNA variants was determined using
254 similar experiments as described above. In one experiment, genomic DNA (DNA isolated by
255 grinding leaf tissue) of $Col^{Sf\text{-}NOR4}$ was subjected to HpaII digestion, followed by gel elutions of
256 the undigested and the digested fractions. The gel-eluted DNA samples were then subjected
257 to PCR amplification of rRNA variants followed by HindIII digestion. In control samples (no
258 HpaII), HindIII digestion of PCR-amplified rRNA variants resulted in the expected products,
259 which include digested Sf2-VAR1 genes, undigested Col-0 VAR1 genes, and partially
260 digested Col-0 VAR3 genes (Fig. 5B, lanes 1 and 2, top right gel image) (8). However,
261 HindIII digestion of the PCR products obtained from the gel-purified HpaII-digested fraction
262 as template, revealed that Col-0 VAR1 genes predominantly constitute the HpaII-undigested
263 fraction (Fig. 5B, lanes 3 and 4, top right gel image). We also observed a significant increase
264 in the Sf-2 VAR1 genes in the HpaII-digested fraction (Fig. 5B, lanes 3 and 4, bottom right

265 gel image) compared to the undigested fraction (Fig. 5B, lanes 3 and 4, top right gel image).
266 These data indicate hypermethylation of Col-0 NOR2 and hypomethylation of Sf-2 NOR4,
267 despite the presence of the VAR1 genes on both the NORs in the $Col^{Sf\text{-}NOR4}$ introgression
268 line.

269 In another experiment, gDNA (DNA isolated by grinding leaf tissue) of $Col^{Sf\text{-}NOR4}$ was
270 first digested with McrBc, followed by PCR amplification of rRNA variants. These PCR
271 products were then digested with HindIII and resolved using regular gel electrophoresis.
272 HindIII-digested control PCR products (no McrBc digestion) behaved as expected (Fig. 5C,
273 lanes 1 and 2). However, the McrBC digestion patterns of the rRNA variants (Fig. 5C) turned
274 out to be exactly the opposite of the experimental results involving HpaII digestion (compare
275 to Fig. 5B). Predominantly, the Sf-2 VAR1 genes constituted the McrBc-undigested fraction,
276 while the Col-0 VAR1 genes constituted the McrBc-digested fraction (Fig. 5C, lanes 3 and 4).
277 We also observed complete digestion of the Col-0 VAR3 genes by McrBC (Fig. 5C, lane 3).
278 These data again indicate that NOR4 is hypomethylated and NOR2 is hypermethylated in
279 the $Col^{Sf\text{-}NOR4}$ introgression line despite the presence of VAR1 type of genes on both the
280 NORs.

281 **NOR2 exists as condensed chromatin while NOR4 exists as open chromatin**

282 Next, to determine the chromatin condensation status of NOR2 and NOR4, we carried out
283 Tn5 transposase-mediated transposition using nuclei isolated from 3-week-old Col-0 plants.
284 Tn5 is a prokaryotic hyperactive transposase that has been developed as a tool to insert
285 short DNA sequences into open chromatin regions (16,25,26) where in 'outside ends' (OEs)
286 sequences are sufficient for Tn5 recognition for transposition and gDNA nicking (27). We
287 designed short DNA sequences consisting of OEs for Tn5 recognition and used them in a
288 transposition assay. Purified Col-0 nuclei were incubated with either no or 5 pmol oligos
289 containing Tn5 recognition sites for transposition, followed by DNA extraction from the nuclei
290 and PCR analysis of rRNA variants. The PCR data did not reveal differences between NOR2
291 and NOR4 (data not shown). Then, we tried an indirect approach to focus on the chromatin
292 accessibility state of the promoter regions. In this approach, the DNA samples isolated from
293 nuclei subjected to Tn5 transposition were subjected to two rounds of PCR amplification.
294 The first round of PCR (PCR1) amplifies ~2.8 Kb region that spans 3'ETS region, IGS and
295 the promoter of the following gene repeat (Fig. 6B). Due to lack of variant-specific sequence
296 variation in the promoter region, we exploited the proximity of the rRNA variant-defining
297 3'ETS region of one rRNA gene to the promoter of the following rRNA gene. The occurrence
298 of rRNA genes as tandem repeats (Fig. 1A) at NORs permitted us to explore this approach.

299 The idea is that, if promoters of VAR1 genes (NOR2) are more condensed than the
300 promoters of VAR2 genes, then the 2.8 Kb PCR products should be enriched for VAR1
301 genes. Because the 2.8 Kb PCR products obtained during PCR1 (using primers 5'allvar and
302 p34) can't discriminate between NOR2 and NOR4, they were resolved on a 1% agarose gel
303 followed gel-purification of 2.8 Kb amplicons. These amplicons were then subjected to the
304 second round of PCR amplification (PCR2) with primers that either amplify only VAR1 and
305 VAR2 genes (Fig. 6C) or primers that amplify all rRNA variants (Fig. 6D). The results
306 showed severe depletion of NOR4-specific VAR2 genes in the reactions with 5 pmol of Tn5
307 oligos (Fig. 6C&D), indicating depletion of NOR4-specific fragments in the PCR1 amplicon
308 pool. This depletion could be presumably due to higher transposition rates in the promoter
309 regions of NOR4-specific rRNA variants compared to the NOR2 variants. Collectively, these
310 data indicate that NOR4 is predominantly in a more accessible open chromatin state and
311 NOR2 in a more condensed state.

312

313 **DISCUSSION**

314 Our earlier findings revealed that during development, NOR2 is silenced and NOR4 remains
315 transcriptionally active (8,9). In eukaryotes, often gene silencing involves DNA methylation
316 and chromatin condensation. First, we analysed several DNA methylation mutants to know
317 their effect on the release of NOR2 silencing (Fig. 2 and Fig. S2). In plants, cytosine
318 methylation occurs in all three contexts (CG, CHG and CHH). MET1 is primarily involved in
319 cytosine methylation in CG context whereas CHROMOMETHYLASES (CMT) are mainly
320 involved in cytosine methylation in CHG and CHH context, and DRM in cytosine methylation
321 in CHH context (12,15). DDM1, a SWI2/SNF2 family chromatin remodeler, is required for
322 cytosine methylation in all three contexts (CG, CHG, and CHH).

323 Although the release of NOR2 silencing was observed in most of the cytosine
324 methylation mutants studied, it was found to be the highest among *cmt2-3* and *ddm1-2,*
325 followed *met1-7* mutants (Fig. 2). The levels of release of NOR2 silencing in *ddm1-2* can be
326 comparable to that of *cmt2-3* if you look at multiple *cmt2-3* plants (Fig. S2). Notably, *ddm1-2*
327 and *cmt2-3* mutants also exhibit the highest loss of CHH methylation (Fig. 3). Interestingly,
328 an earlier study in Arabidopsis has shown that most asymmetric cytosine methylation is
329 mediated by CMT2 and DDM1, independent of the RNA-directed DNA methylation (RdDM)
330 pathway (12). This could also explain why the release of NOR2 silencing is the least in *drm2*
331 mutants compared to the other mutants (Fig. 2). Note that DRM2 is a methyltransferase
332 primarily involved in RdDM. Although *ddm1-2* loses most of cytosine methylation in all

333 contexts (CG, CHG, CHH) (13,14) (Fig. 3), why the levels of release of NOR2 silencing is
334 not greater than *cmt2-3,* which significantly affects cytosine methylation only in CHH context?
335 One possible reason is that loss of CG and/or CHG methylation could dysregulate the
336 expression of some of the positive regulators of NOR2 silencing, limiting the release of NOR
337 silencing to some extent. Therefore, the effect of loss of cytosine methylation in all cytosine
338 contexts in *ddm1-2* could be redundant than cumulative. Overall, the loss of cytosine
339 methylation in the CHH context appears to have the maximum effect on the release of NOR2
340 silencing.

341 Next, to determine if DNA methylation and chromatin condensation are involved in
342 selectively silencing NOR2, we thoroughly analysed DNA methylation and chromatin
343 condensation status of NOR2 and NOR4 using various approaches in multiple genotypes
344 carrying NORs from different ecotypes. To analyze DNA methylation status of NOR2 and
345 NOR4 regions, we carried out rRNA variant analysis upon digestion of genomic DNA by
346 different types of methylation-sensitive restriction enzymes (HpaII and McrBC) (Fig. 4). The
347 HpaII does not digest DNA if its recognition site (CCGG) has cytosine methylation in the CG
348 context. In contrast, McrBc preferentially digests DNA if its recognition site ($R^mC(N_{40\text{-}3000})R^mC$)
349 consists of methylated cytosines. From these experiments involving two methylation-
350 sensitive restriction enzymes with contrasting mechanisms of detecting cytosine methylation,
351 we again showed that NOR2 is hypermethylated and NOR4 is hypomethylated (Fig. 4).
352 HpaII digestion reveals the methylation status of cytosines only in the CG context, but McrBc
353 digestion reveals the methylation status of cytosines in all three contexts (CG, CHG, CHH).
354 Therefore, hypermethylation of cytosine on NOR2 could likely have occurred in all three
355 contexts.

356 We conducted similar experiments using an introgression line named Col-$^{\text{Sf-NOR4}}$
357 which consists of NOR2 from Col-0 (VAR1 and VAR3 genes) and NOR4 from another
358 ecotype called Sf-2 (8) (Fig. 5). Interestingly, Sf-2 consists of only VAR1 type of genes, but
359 they can be differentiated from VAR1 genes of Col-0 by HindIII digestion. We have earlier
360 shown that in Col$^{\text{Sf-NOR4}}$, Sf-2 NOR4 is active and Col-0 NOR2 is selectively silenced (8).
361 Results from these experiments involving digestions with methylation-sensitive HpaII and
362 McrBc, again showed hypermethylation of NOR2 and hypomethylation of NOR4 (Fig. 5).
363 Based on these data, we conclude that NOR2 is hypermethylated compared to NOR4, and
364 DNA methylation is involved in selectively silencing NOR2.

365 Further, we examined the chromatin condensation status of NOR2 and NOR4 using
366 a Tn5 transposase-mediated transposition assay. Tn5 is a prokaryotic hyperactive

367 transposase that can insert short DNA sequences into open chromatin regions (16,25,26)
368 where in 'outside ends' (OEs) sequences are sufficient for Tn5 recognition for transposition
369 (27). The underlying principle is that the higher the frequency of insertion, the more open the
370 chromatin will be. PCR analysis of DNA obtained from Tn5-treated nuclei did not reveal any
371 difference in chromatin condensation states of NOR2 and NOR4. However, an indirect
372 approach involving two rounds of PCR reactions, as described above, revealed that NOR2 is
373 more condensed than NOR4 (Fig. 6). Well, one possibility for the observed result is that
374 there is not much difference between chromatin states at/near 3'ETS of NOR2 and NOR4,
375 but most of the differences could relatively be more at the promoter regions.

376 In conclusion, hypermethylation and chromatin condensation of NOR2 appears to be
377 primary contributing factor for selective NOR2 silencing. However, despite striking
378 similarities in rRNA gene sequences of NOR2 and NOR4, why does NOR2 undergo
379 selective hypermethylation, condensation and silencing? Interestingly, the immediate
380 centromere-proximal region of NOR2 is composed of transposable elements and transposon
381 remnants that stretch for ~ 60 kb (Fig. S4). The corresponding centromere-proximal regions
382 of NOR4 do not contain any stretch of transposons or their remnants (Fig. S4). In plants,
383 transposons are silenced through methylation-mediated mechanisms. Could such silencing
384 of the NOR2-adjacent transposon-rich region spread to NOR2, resulting in its selective
385 silencing? The question remains to be answered.

386

387 **DATA AVAILABILITY**

388 Whole genome bisulfite sequencing data published by Stroud et al. (2013)(17) with a GEO
389 accession number GSE39901 can be found in NCBI database
390 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39901).

391

392 **ACKNOWLEDGEMENT**

393 We thank the members of the Pikaard lab and the Mohannath lab for their critique of the
394 manuscript.

395 **FUNDING**

396 G.M. is thankful to the Birla Institute of Science and Technology (BITS) Pilani, Hyderabad
397 campus for the research grant (BITS/GAU/ACRG/2019/H0576) to G.M. and an Institute

398 Fellowship to G.P.S., and to the Science and Engineering Research Board (SERB),
399 Government of India, for the Ramanujan Fellowship Research Grant (SB-S2-RJN-062-2017)
400 and CRG grant (CRG/2020/002855). C.S.P. is thankful to research funding from Howard
401 Hughes Medical Institute, Indiana University (USA), National Institute of Health, and Gordon
402 and Betty Moore Foundation.
403

404 **CONFLICT OF INTEREST**

405 The authors declare no conflict of interest.

406 **REFERENCES**

407 1. Saradadevi, G.P., Priyadarshini, N., Bera, A. and Mohannath, G. (2020) Together we are on
408 together we are off -a conserved rule for ribosomal RNA (rRNA) gene regulation? *Journal of*
409 *Plant Biochemistry and Biotechnology*, **29**, 743-753.
410 2. Turowski, T.W. and Tollervey, D. (2015) Cotranscriptional events in eukaryotic ribosome
411 synthesis. *Wiley interdisciplinary reviews. RNA*, **6**, 129-139.
412 3. Viktorovskaya, O.V. and Schneider, D.A. (2015) Functional divergence of eukaryotic RNA
413 polymerases: unique properties of RNA polymerase I suit its cellular role. *Gene*, **556**, 19-26.
414 4. Grummt, I. (2013) The nucleolus—guardian of cellular homeostasis and genome integrity.
415 *Chromosoma*, **122**, 487-497.
416 5. Hannan, K.M., Sanij, E., Rothblum, L.I., Hannan, R.D. and Pearson, R.B. (2013) Dysregulation
417 of RNA polymerase I transcription during disease. *Biochimica et biophysica acta*, **1829**, 342-
418 360.
419 6. Moss, T., Langlois, F., Gagnon-Kugler, T. and Stefanovsky, V. (2007) A housekeeper with
420 power of attorney: the rRNA genes in ribosome biogenesis. *Cell Mol Life Sci*, **64**, 29-49.
421 7. Chen, Z.J. and Pikaard, C.S. (1997) Epigenetic silencing of RNA polymerase I transcription: a
422 role for DNA methylation and histone modification in nucleolar dominance. *Genes Dev*, **11**,
423 2124-2136.
424 8. Chandrasekhara, C., Mohannath, G., Blevins, T., Pontvianne, F. and Pikaard, C.S. (2016)
425 Chromosome-specific NOR inactivation explains selective rRNA gene silencing and dosage
426 control in Arabidopsis. *Genes Dev*, **30**, 177-190.
427 9. Mohannath, G., Pontvianne, F. and Pikaard, C.S. (2016) Selective nucleolus organizer
428 inactivation in Arabidopsis is a chromosome position-effect phenomenon. *Proc Natl Acad Sci*
429 *U S A*, **113**, 13426-13431.
430 10. Pontvianne, F., Abou-Ellail, M., Douet, J., Comella, P., Matia, I., Chandrasekhara, C., Debures,
431 A., Blevins, T., Cooke, R., Medina, F.J. *et al.* (2010) Nucleolin is required for DNA methylation
432 state and the expression of rRNA gene variants in Arabidopsis thaliana. *PLoS genetics*, **6**,
433 e1001225.
434 11. Brzeski, J. and Jerzmanowski, A. (2003) Deficient in DNA methylation 1 (DDM1) defines a
435 novel family of chromatin-remodeling factors. *The Journal of biological chemistry*, **278**, 823-
436 828.
437 12. Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer,
438 S.L. and Zilberman, D. (2013) The Arabidopsis nucleosome remodeler DDM1 allows DNA
439 methyltransferases to access H1-containing heterochromatin. *Cell*, **153**, 193-205.
440 13. Jeddeloh, J.A., Stokes, T.L. and Richards, E.J. (1999) Maintenance of genomic methylation
441 requires a SWI2/SNF2-like protein. *Nature genetics*, **22**, 94-97.

14. Vongs, A., Kakutani, T., Martienssen, R.A. and Richards, E.J. (1993) Arabidopsis thaliana DNA methylation mutants. *Science*, **260**, 1926-1928.
15. Law, J.A. and Jacobsen, S.E. (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature reviews. Genetics*, **11**, 204-220.
16. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. and Greenleaf, W.J. (2013) Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods*, **10**, 1213-1218.
17. Stroud, H., Greenberg, M.V., Feng, S., Bernatavichute, Y.V. and Jacobsen, S.E. (2013) Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell*, **152**, 352-364.
18. Schultz, M.D., He, Y., Whitaker, J.W., Hariharan, M., Mukamel, E.A., Leung, D., Rajagopal, N., Nery, J.R., Urich, M.A., Chen, H. *et al.* (2015) Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature*, **523**, 212-216.
19. Mohannath, G. and Pikaard, C.S. (2016) Analysis of rRNA Gene Methylation in Arabidopsis thaliana by CHEF-Conventional 2D Gel Electrophoresis. *Methods in molecular biology (Clifton, N.J.)*, **1455**, 183-202.
20. Chu, G., Vollrath, D. and Davis, R.W. (1986) Separation of large DNA molecules by contour-clamped homogeneous electric fields. *Science*, **234**, 1582-1585.
21. Copenhaver, G.P., Doelling, J.H., Gens, S. and Pikaard, C.S. (1995) Use of RFLPs larger than 100 kbp to map the position and internal organization of the nucleolus organizer region on chromosome 2 in Arabidopsis thaliana. *Plant J*, **7**, 273-286.
22. Gaudino, R.J. and Pikaard, C.S. (1997) Cytokinin induction of RNA polymerase I transcription in Arabidopsis thaliana. *The Journal of biological chemistry*, **272**, 6799-6804.
23. Blevins, T., Pontvianne, F., Cocklin, R., Podicheti, R., Chandrasekhara, C., Yerneni, S., Braun, C., Lee, B., Rusch, D., Mockaitis, K. *et al.* (2014) A two-step process for epigenetic inheritance in Arabidopsis. *Molecular cell*, **54**, 30-42.
24. Lewis, M.S. and Pikaard, C.S. (2001) Restricted chromosomal silencing in nucleolar dominance. *Proc Natl Acad Sci U S A*, **98**, 14536-14540.
25. Adey, A., Morrison, H.G., Asan, Xun, X., Kitzman, J.O., Turner, E.H., Stackhouse, B., MacKenzie, A.P., Caruccio, N.C., Zhang, X. *et al.* (2010) Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome biology*, **11**, R119.
26. Goryshin, I.Y. and Reznikoff, W.S. (1998) Tn5 in vitro transposition. *The Journal of biological chemistry*, **273**, 7367-7374.
27. Ahmed, A. (2009) Alternative mechanisms for tn5 transposition. *PLoS genetics*, **5**, e1000619.

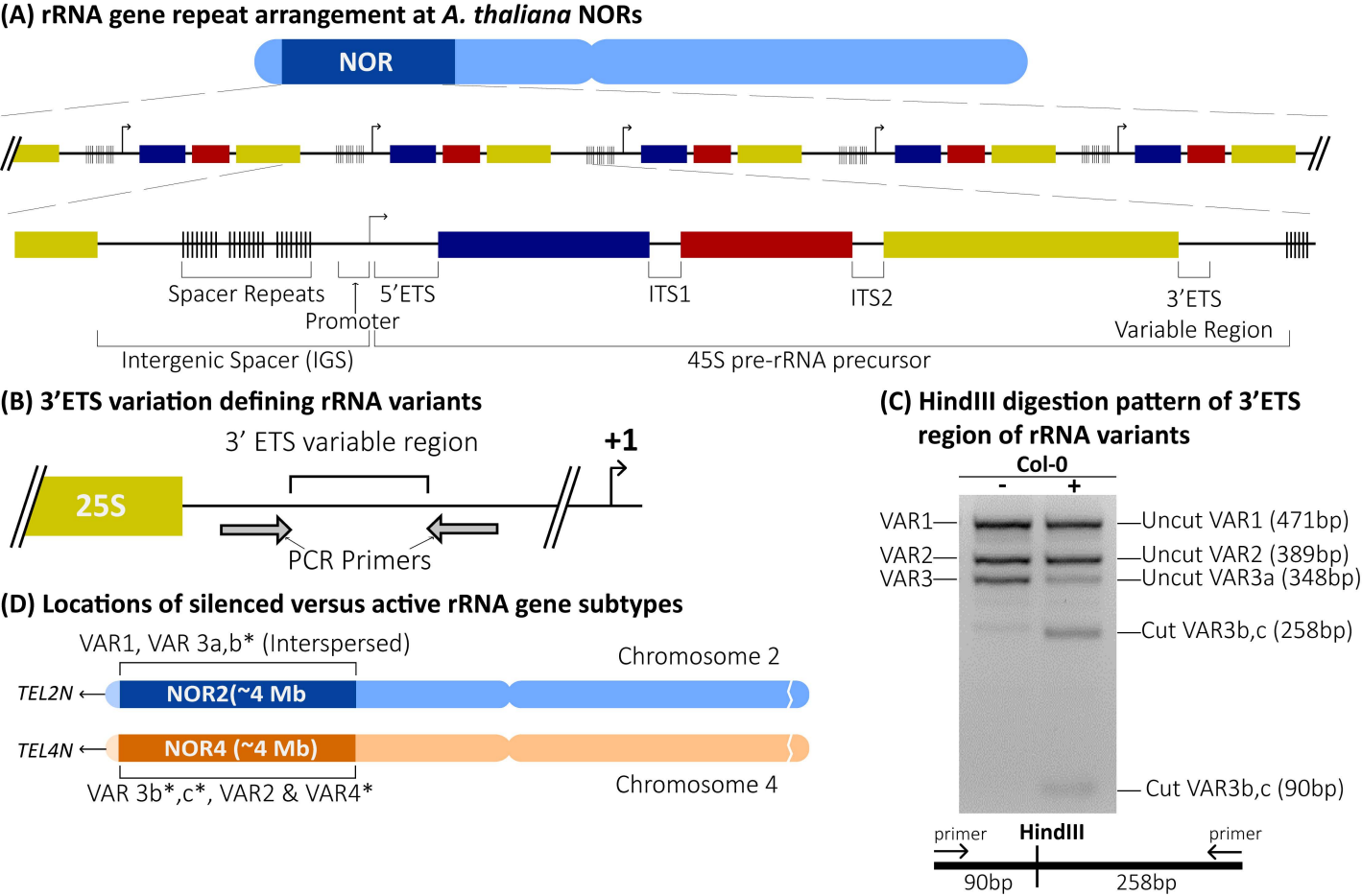

Figure 1. Ribosomal RNA (rRNA) gene organization, subtype variation, localization and their relative arrangement

rRNA subtype expression in DNA methylation mutants

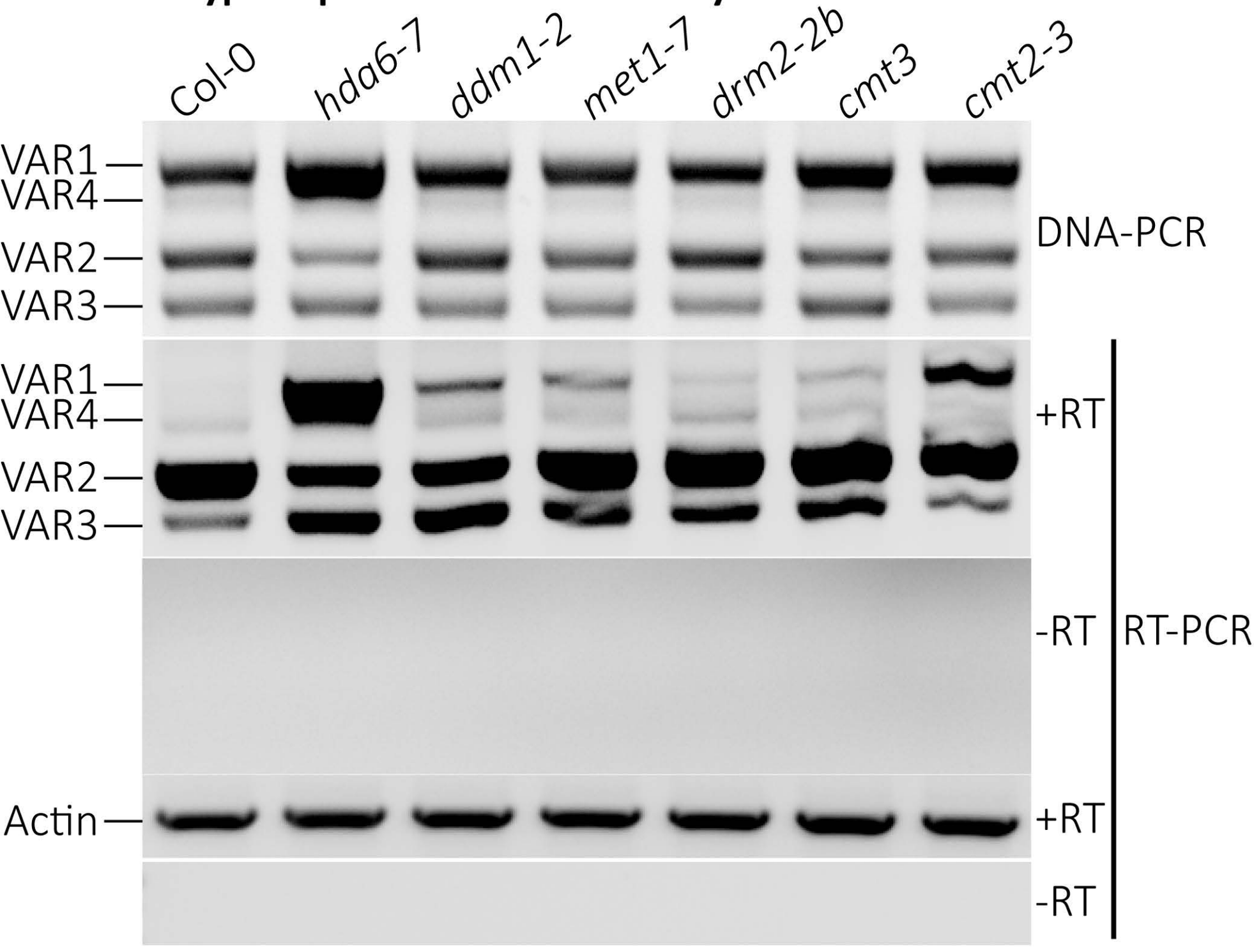

Figure 2. Release of NOR2 silencing to varying degrees in DNA methylation mutants.

The gel images show DNA-PCR data for rRNA gene subtypes and RT-PCR expression data for rRNA gene subtypes and actin.

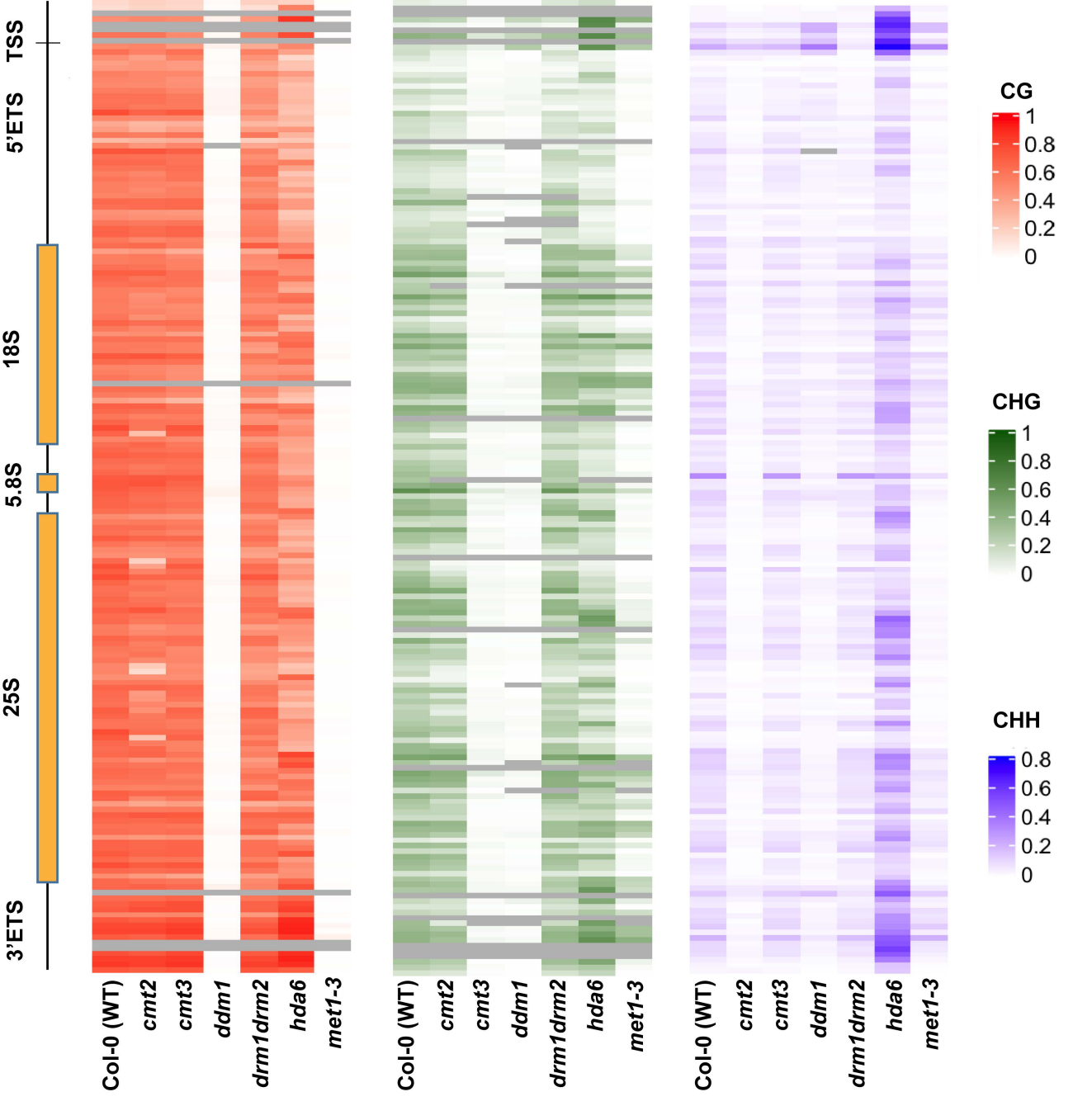

Figure 3: Methylation state of rDNA in CG, CHG and CHH contexts. Methylation state of rDNA in CG, CHG and CHH contexts. Whole genome bisulfite sequencing data from various mutants was analyzed using Methylpy and the overall cytosine methylation state of the rDNA is shown as heatmaps in all three contexts. The Y-axis is a single rDNA unit starting at the promoter and ending in the 3'ETS. Each line on the heatmap represents the methylation state of cytosines in a given methylation context in 50 nucleotide bins. CG is shown in red, CHG in green and CHH in blue. Grey indicates that no methylation sites were present in that bin for a given context.

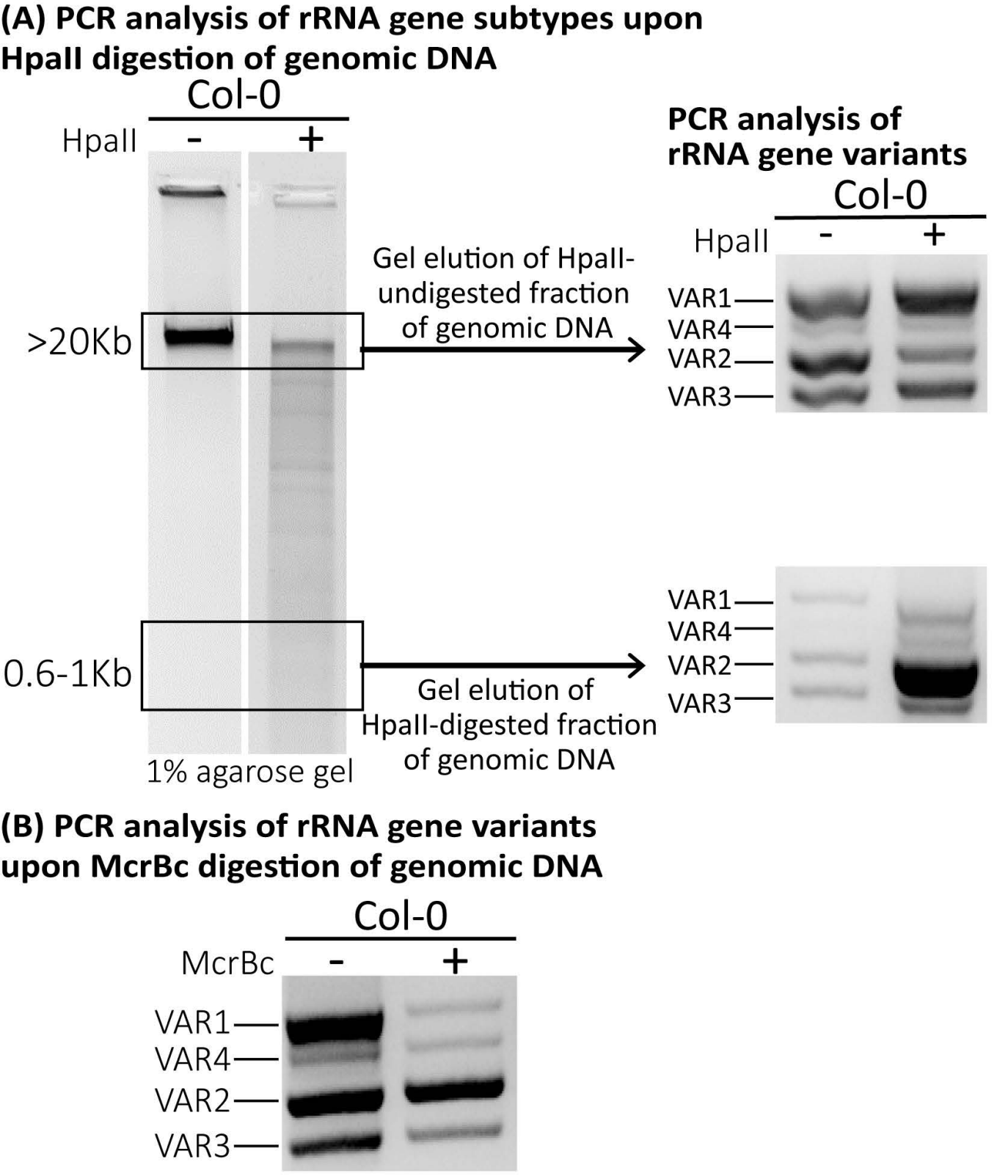

Figure 4. NOR2 is hypermethylated and NOR4 is hypomethylated

(A) Gel images on the left show resolution of HpaII- or buffer-treated Col-0 DNA samples and the gel images on the right show PCR amplification of rRNA gene subtypes. For these PCR reactions, gel-eluted DNA samples representing the HpaII-undigested (≥ 20 kb) and the HpaII-digested (0.6-1.0 kb) fractions were used. (B) The gel image shows rRNA gene subtype PCR products obtained from McrBc-digested or no enzyme-treated Col-0 DNA.

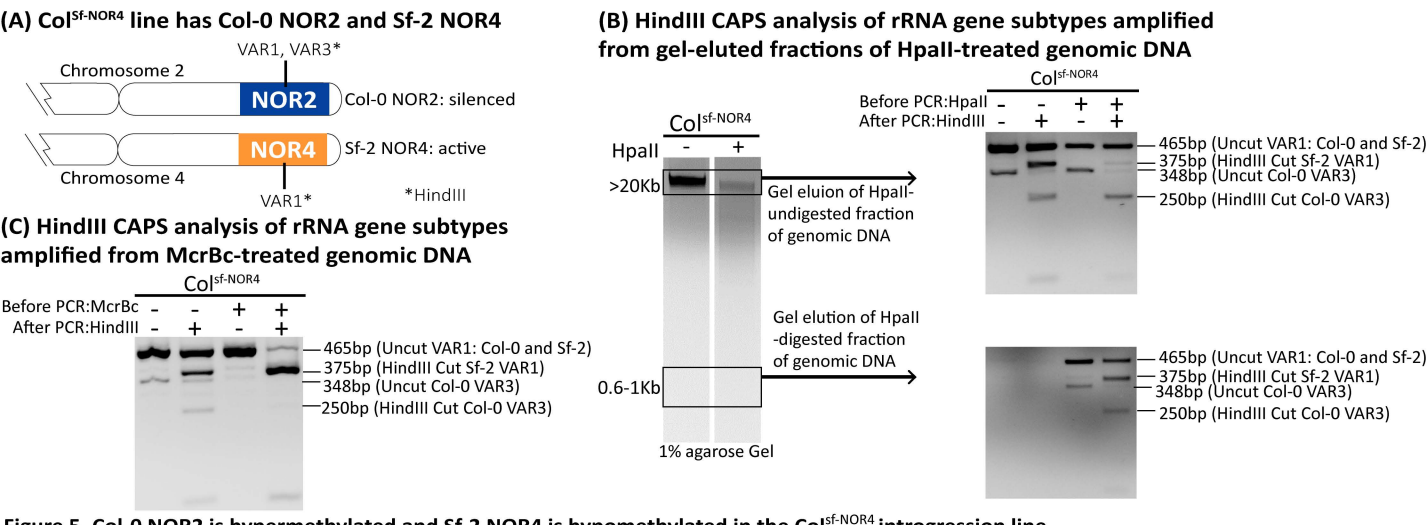

Figure 5. Col-0 NOR2 is hypermethylated and Sf-2 NOR4 is hypomethylated in the Col$^{sf\text{-}NOR4}$ introgression line.
(A) Diagram showing rRNA subtype composition of the Arabidopsis introgression line called Col$^{Sf\text{-}NOR4}$ in which Sf-2 NOR4 and adjacent sequences were introgressed into the Col-0 genetic background. (B) The agarose gel images on the left show resolution of HpaII- or buffer-treated Col$^{Sf\text{-}NOR4}$ gDNA samples, and the gel images on the right show PCR amplification of rRNA variants with or without HindIII digestion. For these PCR reactions, gel-eluted DNA samples representing the HpaII-undigested (≥ 20 kb) and the HpaII-digested (0.6-1.0 kb) fractions were used. (C) The Gel image shows PCR amplification of rRNA variants obtained from McrBc- or buffer-treated Col$^{Sf\text{-}NOR4}$ gDNA with or without HindIII digestion.

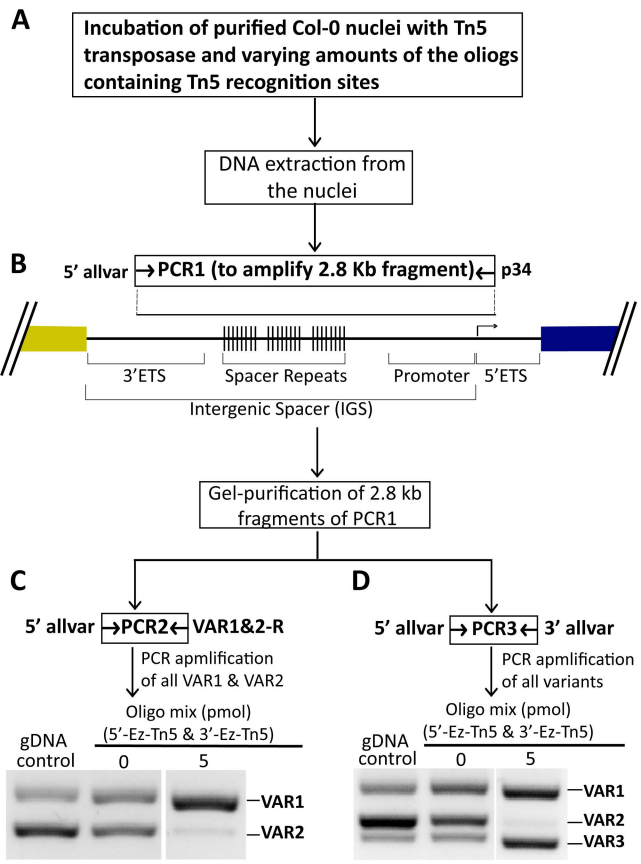

Figure 6. NOR2 exists as condensed chromatin while NOR4 exists as open chromatin.

(A) Tn5 Transposase-accessible chromatin assay (ATAC) was performed on nuclei isolated from wild-type Col-0 plants. (B) DNA isolated from the transposase-treated nuclei was used in PCR1 to amplify ~2.8 Kb of rRNA gene promoter-proximal upstream regions involving the variable 3'ETS of the preceding gene and IGS. (C) The PCR1-amplified 2.8 Kb fragments were gel-eluted followed by PCR2, products of which are visualized on a 2.8% agarose gel stained with ethidium bromide. Amplification of VAR1 and VAR2 genes from 3'ETS variable region from genomic DNA (gDNA) as a control is shown in the first lane. (D) The PCR1-amplified 2.8 Kb fragments were gel-eluted followed by PCR3, products of which are visualized on a 2.8% agarose gel stained with ethidium bromide. Here, PCR3 involves amplification of all rRNA variants. Amplification of all rRNA variants from 3'ETS variable region from genomic DNA (gDNA) as a control is shown in the first lane.