



HAL
open science

Construction and characterisation of a knock-down RNA interference line of OsNRPD1 in rice (*Oryza sativa* ssp japonica cv Nipponbare)

Emilie Debladis, Tzuu-Fen Lee, Yan-Jiun Huang, Jui-Hsien Lu, Sandra M. Mathioni, Marie-Christine Carpentier, Christel Llauro, Davy Pierron, Delphine Mieulet, Emmanuel Guiderdoni, et al.

► To cite this version:

Emilie Debladis, Tzuu-Fen Lee, Yan-Jiun Huang, Jui-Hsien Lu, Sandra M. Mathioni, et al.. Construction and characterisation of a knock-down RNA interference line of OsNRPD1 in rice (*Oryza sativa* ssp japonica cv Nipponbare). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2020, 375 (1795), pp.20190338. 10.1098/rstb.2019.0338 . hal-04274817

HAL Id: hal-04274817

<https://univ-perp.hal.science/hal-04274817>

Submitted on 8 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 **Construction and characterisation of a knock-down RNA interference line of *OsNRPD1* in**
2 **rice (*Oryza sativa ssp japonica* cv Nipponbare).**

3

4 Emilie Debladis^{1,2}, Tzuu-Fen Lee⁴, Yan-Jiun Huang⁵, Jui-Hsien Lu⁵, Sandra M. Mathioni⁴, Marie-
5 Christine Carpentier^{1,2}, Christel Llauro^{1,2}, Davy Pierron^{1,2}, Delphine Mieulet⁶, Emmanuel
6 Guiderdoni⁶, Pao-Yang Chen⁵, Blake C. Meyers^{4,7}, Olivier Panaud^{1,2,3}, Eric Lasserre^{1,2,*}

7

8 ¹ Université de Perpignan Via Domitia, Laboratoire Génome et Développement
9 des Plantes, 52, avenue Paul alduy, 66860 Perpignan cedex, France.

10 ² Centre National de la Recherche Scientifique, Laboratoire Génome et Développement
11 des Plantes, 52, avenue Paul alduy, 66860 Perpignan cedex, France.

12 ³ Institut Universitaire de France, Paris, France.

13 ⁴ Donald Danforth Plant Science Center, St Louis, MO, USA

14 ⁵ Institute of Plant and Microbial Biology, Academia Sinica, Taipei 11529, Taiwan

15 ⁶ CIRAD, UMR AGAP, F34398 Montpellier, Cedex 5, France

16 ⁷ University of Missouri-Columbia, Division of Plant Sciences, Columbia, Missouri 65211

17 * Corresponding author, eric.lasserre@univ-perp.fr

18 ABSTRACT

19 In plants, RNA-directed DNA Methylation (RdDM) is a mechanism of silencing that relies on the
20 production of 24-nt siRNAs by RNA POLYMERASE IV (Pol IV) to trigger methylation and thus
21 inactivation of transposable elements (TEs). We present the construction and characterisation of
22 *osnrpd1*, a knock-down RNA interference line of *OsNRPD1* gene that encodes the largest subunit of
23 Pol IV in rice (*Oryza sativa ssp japonica* cv Nipponbare). We show that *osnrpd1* displays a lower
24 accumulation of *OsNRPD1* transcripts, associated to an overall reduction of 24-nt siRNAs and DNA
25 methylation level in all three contexts, CG, CHG and CHH. We uncovered new insertions of known
26 active TEs, the LTR retrotransposons *Tos17* and *Lullaby* and the LINE-type retrotransposon *Karma*.
27 However, we did not observe any clear developmental phenotype, contrary to what was expected
28 for a mutant severely affected in RdDM. In addition, despite the presence of many putatively
29 functional TEs in the rice genome, we found no evidence of *in planta* global reactivation of
30 transposition. This knock-down of *OsNRPD1* likely led to a weakly affected line, with no effect on
31 development and a limited effect on transposition. We discuss the possibility that a knock-out
32 mutation of *OsNRPD1* would cause sterility in rice.

33

34 KEYWORDS

35 Polymerase IV, RNA-dependent DNA Methylation, Transposable elements, RNA interference

36

37 BACKGROUND

38 The DNA of flowering plants as a whole is predominantly made up of transposable elements (TEs),
39 mobile genetic units able to proliferate in their host genomes. First described by B. McClintock
40 more than 60 years ago, they are grouped into two classes based on their transposition mechanisms :
41 Class I elements or retrotransposons transpose via an RNA-mediated copy-and-paste mechanism
42 whereas class II elements or DNA transposons transpose without using an RNA intermediate (1,2).
43 LTR retrotransposons (LTR-RTs) are the most abundant in plants. Non-LTR retrotransposons
44 include long interspersed nuclear elements (LINEs) and short interspersed nuclear elements
45 (SINEs) (3).

46 TEs have been shown to be one of the most important factors driving the structure, function and
47 evolution of eukaryotic genomes (4–6). However, because of their combined mutagenic and
48 replicative properties, they also may threaten the overall structure and function of the genome of
49 their host. A balance between both consequences have been reached : host organisms have evolved
50 strong controls of transposition that allowed the taming of TE proliferation while keeping
51 transposition possible. Young TE copies are thus silenced by epigenetic marks such as cytosine
52 methylation, ensuring a stable repression of TE expression and preventing their proliferation (7).
53 This *de novo* DNA methylation can be initiated via the RNA-directed DNA Methylation (RdDM)
54 mechanism, a plant-specific pathway through which small interfering RNAs (siRNAs) target
55 homologous DNA regions through base-pairing to methylate them. The RdDM pathway is well
56 characterised in *A. thaliana* (8,9). The canonical one is initiated by the RNA POLYMERASE IV
57 (Pol IV) which generates a single strand RNA of the target locus which is a template for RNA-
58 DEPENDENT RNA POLYMERASE 2 to generate a double strand RNA (dsRNA). Then, DICER-
59 LIKE 3 cleaves these dsRNA into 24 nucleotides siRNAs (24nt siRNA) one strand of which are
60 loaded into ARGONAUTE 4 that can target nascent scaffold RNA POLYMERASE V (Pol V)
61 transcripts or genomic DNA by base-pairing. Finally, this targeting leads to the recruitment of the

62 DNA methyltransferase DOMAINS REARRANGED METHYLTRANSFERASE 2 to mediate *de*
63 *nov*o methylation in all three different sequences contexts, CG, CHG and CHH, where H is A, C or
64 T (10). However, *De novo* methylation of cytosines in the CHH context is the specific hallmark of
65 RdDM since methylation at the symmetrical CG and CHG sites can be maintained at each round of
66 replication by METHYLTRANSFERASE 1 and CHROMOMEHTYLASE 3, respectively (11).
67 In addition to the canonical pathway, Pol II expression-dependent forms of RdDM that partly
68 incorporate components typically associated with post-transcriptional gene silencing have recently
69 been identified in *A. thaliana* (9,12).
70 In rice (*Oryza sativa ssp japonica*), TEs account for at least 35 % of the genome (13). Despite
71 evidence of many polymorphic insertions of TEs in different accessions or cultivars of the same
72 species (14,15), showing that *in planta* transposition was ongoing in the field, very few
73 transposition events could be triggered and observed in laboratory conditions. This was the case for
74 three LTR-Retrotransposons, namely *Tos17* (16), *Lullaby* (17) and *Karma* (18) and two related
75 DNA transposons, *mPing* and *Pong* (19,20). However, transposition of these elements did not occur
76 *in planta* but was triggered by an intermediate step of long-term *in vitro* culture of cells or anthers.
77 Only some mutants affected in methylation of histones (21) or DNA (22,23) have been shown to
78 display transposons reactivation *in planta*.
79 Recent advances have shed light on epigenetic regulation and the epigenomic landscape in rice (7).
80 Key actors of DNA methylation have been characterised (23–26). However, only few genes have
81 been fully identified as actors in the RdDM mechanism. OsDCL3a is the rice DICER-LIKE 3
82 homolog involved in 24-nt siRNA processing (26). Reduction of OsDCL3a function reduced the 24-
83 nt siRNAs predominantly from MITEs and elevated expression of nearby genes involved in the
84 homeostasis of the plant hormones gibberellin and brassinosteroid. The *osdcl3a* RNAi lines thus
85 displayed several developmental alterations compared to wild-type. Targeted disruption of
86 *OsDRM2*, coding the rice methyltransferase responsible for *de novo* methylation, lead to a 13.9%
87 decrease in 5-methylcytosine in both CG and non CG contexts and impaired both vegetative and
88 reproductive development (24). However, since none of these studies focused on transposition,
89 nothing is known about the impact of those mutations, and by extension RdDM, on TEs
90 mobilization in rice.
91 Foremost among these factors and central to the mechanism of RdDM, as described above for
92 Arabidopsis, is the RNA POLYMERASE IV because it is needed to produce the siRNA trigger for
93 methylation. Pol IV is a large holoenzyme composed of 12 subunits (27). NRPD1 is the largest one
94 and is derived from the duplication of Pol II subunit NRPB1. It is specific to Pol IV and a
95 component of the catalytic center (27). Two orthologs of NRPD1 have been identified in rice (28)
96 and present the same domain structure than in *A. thaliana*, suggesting a similar molecular function.
97 Interestingly, however, no knockout mutant has been described for any of the primary components
98 of the RdDM pathway, Pol IV and Pol V in rice, suggesting that they could be lethal (29) in contrast
99 to Arabidopsis. This is further supported by the observation that mutations in genes encoding other
100 RdDM factors strongly affected rice development (24,26) when their counterparts in Arabidopsis
101 did not show obvious defects (8).
102 In this paper, we describe the construction and the characterisation of *osnrpd1*, a knock-down RNAi
103 line of *OsNRPD1* gene. Our initial goal was to create a rice line with a broadly relaxed epigenetic
104 control over TEs to study their transcriptional regulation. OsNRPD1, being of central importance in
105 producing the siRNA triggers that maintain TEs silent, was a target of choice. *osnrpd1* did not

106 displayed any obvious growing defects but we show that 24-nt small RNAs are under-accumulated
107 and that the methylation of cytosines residues in all contexts is reduced compared to wild-type
108 plants. However, while we observed new insertions of three known active retrotransposons, namely
109 *Tos17*, *Lullaby* and *Karma*, *OsNRPD1* knock-down was not sufficient to broadly trigger
110 transposition of many putatively functional TEs populating the rice genome.

111

112 METHODS

113 **Plant material**

114 *Oryza sativa* ssp. *japonica* cv. Nipponbare rice plants and the derived *osnrpd1* RNAi line have been
115 obtained from CIRAD, Montpellier, France. They were cultivated in a growth chamber (Percival)
116 under a 12h light-dark cycle (12h-28°C/12h-26°C) and with a relative humidity of 80% during the
117 day and 70% during the night. The light intensity varied gradually in 40 min at the beginning and
118 end of the day.

119 For the construction of *osnrpd1*, a 150 nucleotides-long fragment belonging to a region coding for
120 the RNAP_IV_NRPD1 C-terminal conserved domain of NCBI Gene LOC4336722 (nucleotides
121 3886 to 4036) has been amplified by PCR (primers are described in table S5, additional file 5) and
122 cloned into the hpRNA binary vector pBIOS738 as described in figure S1.

123 To generate the *osnrpd1* RNAi line, an *Agrobacterium*-mediated transformation of mature seed
124 embryo-derived callus was performed as previously described (30). The one T-DNA-containing T0
125 regenerated plants, 6.2, 8.1, 20.1 and 23.1 were selected (Figure 1). Further generations of single-
126 seed descents have been obtained by selfing.

127

128 **DNA and RNA extraction**

129 For DNA-seq and PCR analysis, total DNA was extracted from 50 mg of frozen leaves harvested
130 from one month old plants by a CTAB-based method as previously described (31).

131 For RNA-seq and RT-PCR analysis, total RNA was extracted from 50 mg of leaves using the
132 TRIzol reagent (Invitrogen Life Technologies) and treated with DNase I (RQ1 RNase-Free DNase,
133 Promega). RNAs were checked on 1 % agarose gels and quantified using a Qubit® RNA Assay Kit
134 in a Qubit® 2.0 Fluorometer (Life Technologies).

135

136 **PCR and RT-PCR methods**

137 For PCR analysis of *Tos17* insertions, 20 ng of total DNA were used. 30 cycles of PCR
138 amplifications were performed with a hybridization step at 58 °C.

139 For semi quantitative RT-PCR analysis of *OsNRPD1a* and *ACTIN* transcripts accumulation, cDNAs
140 were synthesized from 800 ng of isolated RNA using an oligonucleotide dT primer and the
141 GoScript reverse transcriptase (Promega). 24 cycles of PCR amplifications were performed with a
142 hybridization step at 58 °C.

143 Analyses by quantitative real-time PCR (qRT-PCR) were established using 7 to 35 ng of cDNA,
144 synthesized as described above. qRT-PCRs were run on a LightCycler 480 (Roche) using Takyon
145 No Rox SYBR MasterMix dTTP Blue Kit (Eurogentec). The qRT-PCR conditions were the
146 following : a first denaturation step at 95°C for 5 min followed by 40 cycles at 95°C for 15s, an
147 annealing and elongation step at 60°C for 60s, and a melting curve analysis at 95°C for 10s, 60°C
148 for 10s, an increase of 0.04°C per second until 95°C and a final step of cooling at 40°C for 30s.
149 Primers were used at a concentration of 2µM. Three biological replicates were analyzed for each

150 plant. *OsNRPD1* expression level relative to *ACTIN* was calculated using the $\Delta\Delta C_t$ method.
151 All primers sequences are listed in table S5, additional file 5.

152

153 **DNA-seq and read mapping**

154 DNA quality and concentration were determined using a high sensitivity DNA Bioanalyzer chip
155 (Agilent Technologies). 350 base pairs DNA libraries have been prepared using Illumina's PCR-free
156 DNA kit and sequenced at 2x102 nt or 2x151 nt on a HiSeq2500 instrument (HudsonAlpha Genome
157 Sequencing Center, Huntsville, USA). Quality control of FASTQ files was evaluated using the
158 FastQC tool (version 0.10.1 www.bioinformatics.babraham.ac.uk/projects/fastqc). To remove any
159 read originating from organelle genomes, reads were mapped against the mitochondria (GenBank
160 NC_0111033) and chloroplast genomes (GenBank X15901) using the program BOWTIE2 version
161 2.2.2 (32) with --sensitive-local mapping. Unmapped reads were then considered for the systematic
162 search of TEs insertions as described in additional file 6.

163

164 **RNA-seq and *OsNRPD1* transcript accumulation**

165 Six stranded cDNA libraries (WT and T1-15, 3 biological replicates each) from poly-A-enriched
166 RNAs have been generated using Illumina's stranded RNAseq kit and sequenced at 2x102 nt on a
167 HiSeq 2500. Quality control of FASTQ files was made as for DNA-seq. Analysis of transcript
168 accumulation of *OsNRPD1a* and *OsNRPD1b* was performed as follows : a BOWTIE2 index has
169 been done for each mRNA sequence. Then, reads from the 6 libraries were mapped on both
170 sequences using BOWTIE2 in the default mode with the --no-unal option to suppress SAM records
171 for reads that failed to align. SAMTOOLS utilities were then used to select and count concordant
172 alignments with no mismatches. Since many transcripts were produced from the inverted repeat of
173 *osnrpd1*, those reads that mapped to the 150 bp region that was used for the RNAi construct were
174 removed in each analysis.

175

176 **sRNA-seq and mapping**

177 Total RNA from the materials described above was isolated using Tri Reagent™ (Molecular
178 Research Center). Small RNA libraries were constructed using the Illumina TruSeq Small RNA
179 Sample Preparation Kit, and sequenced on an Illumina HiSeq2000 instrument. Raw sequencing data
180 were first trimmed of adapter sequences, with trimmed lengths between 18 and 34 nt. The read
181 counts were normalized based on the total abundance of genome-matched reads, excluding
182 structural sRNAs originating from annotated tRNA, rRNA, small nuclear and small nucleolar
183 RNAs. Read counts were normalized to 20M reads per library as well as to 21 nt abundances. The
184 21-nt abundances were used for normalization control as *osnrpd1* knock-down causes a reduction of
185 24-nt siRNA.

186 To assess regions of *OsNRPD1a* targeted by siRNAs, reads were aligned to the variant 6 of
187 *OsNRPD1a* mRNA with BOWTIE. Only reads with perfect match were kept. Coverage at each
188 nucleotide was calculated.

189

190 **Genome-wide methylome profiling by methylC-seq**

191 For methylC-seq library construction, 200 ng of RNA-free genomic DNA (gDNA) were used for
192 library construction following a protocol previously described with some variations (33). Briefly,
193 gDNA in TE buffer was fragmented in a Covaris focused ultrasonicator to generate approximately

194 200 bp fragments. The fragmented gDNA was kit-purified (QIAquick PCR purification kit,
195 Qiagen), end-repaired (End-It DNA end repair kit, Epicentre), and 'A' bases were added to the 3'
196 end of DNA then kit-purified again. Next, the methylated adapters were ligated to DNA fragments
197 using Fast-Link™ DNA ligation kit (Epicentre) then kit-purified. 250-500 bp DNA fragments in the
198 ligated products were extracted from a 2% agarose gel (BioRad) and purified (QIAquick Gel
199 Extraction Kit, Qiagen). The adaptor-ligated gDNA was treated with sodium bisulfite for cytosine
200 conversion with the MethylCode kit (Life Technologies). The bisulfite-converted gDNA was then
201 amplified by PCR using Pfu Turbo Cx Hotstart DNA polymerase and PCR program as described in
202 the above-mentioned paper with 14 cycles of amplification. The amplified library was purified
203 using Agencourt Ampure XP beads (Bechman-Coulter) and sequenced with Illumina technology.
204 Bisulfite converted reads were aligned to the rice reference genome (IRGSP1.0) using BS-Seeker2
205 with default parameters (34). Genome-wide DNA methylation profiles were generated by
206 determining methylation levels for each cytosine in the genome. We only included cytosines that are
207 covered by at least four reads. We estimated the bisulfite conversion rate with respect to rice
208 chloroplast genome : 96.35% for WT, 97.25% for T0, 98.04% for T1-15 and 97.21% for T1-23. BS-
209 seq libraries and mapping data are presented in additional file 3.

210

211 RESULTS

212 **Design of the *osnrpd1* RNAi construct**

213 Two orthologs of *NRPD1* were identified in rice, corresponding to NCBI Gene IDs LOC4336722
214 and LOC4347810, both giving rise to several variants (Figure S1, additional file 1). Hereafter, the
215 corresponding genes or transcripts will be named *OsNRPD1a* and *OsNRPD1b*, respectively, and
216 collectively named *OsNRPD1*. Because we could not identify a viable insertional mutant of
217 *OsNRPD1a* from the available collections (see discussion), we used an RNAi strategy to post-
218 transcriptionally inactivate both *OsNRPD1* genes. We selected a 150 nucleotides-long fragment
219 belonging to a region coding for the RNAP_IV_NRPD1 C-terminal conserved domain of the locus
220 LOC4336722, characteristic of NRPD1 proteins that diverged from NRPB1 of Pol II (35). Both
221 mRNAs are 93% identical in this region (Figure S1, additional file 1). This fragment was cloned
222 into a binary vector in the form of two inverted repeats separated by an intron and inserted into the
223 genome of a *Oryza sativa* cv. Nipponbare plant by transformation with *Agrobacterium tumefaciens*
224 (Figure S1, additional file 1). One T0 regenerant and a progeny of three T1 plants obtained by self-
225 propagation (named T1-15, T1-17, T1-23), homozygous for the T-DNA insertion harboring the
226 RNAi construct, were the objects of this study.

227

228 **The accumulation of *OsNRPD1* transcripts is drastically lowered.**

229 The T0 regenerant described in this study (plant 23.1 on Figure 1A) had been selected among four
230 and self-propagated because it displayed the lowest expression level of *OsNRPD1a* based on semi-
231 quantitative RT-PCR analysis (Figure 1A). Three T1 plants (T1-15, T1-17 and T1-23), displaying a
232 lowest accumulation of the *OsNRPD1a* transcript, have then been selected (Figure 1A). The lower
233 accumulation of the *OsNRPD1a* transcript has been confirmed on T1-15, T1-17, and T1-23 by RT-
234 qPCR (Figure 1B). We also obtained RNA-Seq data from T1-15. These data confirmed that the
235 accumulation of the *OsNRPD1a* transcript was lower (Figure 1C). In addition, they showed that the
236 *OsNRPD1b* transcripts were much less accumulated than the *OsNRPD1a* ones and, more
237 importantly, that the RNAi construct was able to reduce the accumulation of both (Figure 1C). This

238 lower abundance of *OsNRPD1a* transcripts in the RNAi lines is in agreement with the precise
239 targeting of *OsNRPD1a* gene by 21-nt siRNAs at the 150-nt long region used to devise the RNAi
240 construct (Figure S2, additional file 1).

241

242 **The accumulation of 24-nt siRNAs is lowered.**

243 Since *OsNRPD1* is a subunit of the Pol IV enzyme, it was expected that a lower accumulation of its
244 transcripts would impair the accumulation of 24-nt siRNAs. We therefore generated small RNAs
245 libraries (Table S1, additional file 2) from leaves harvested at the same developmental stage on
246 wild-type plantlets, the T0 and three T1 (T1-15, T1-17 and T1-23), sequenced them and compared
247 the abundance of each size class between the libraries. Contrary to 24-nt siRNAs, the 21-nt class
248 were expected to be unimpacted in the RNAi lines. The siRNAs abundances were thus normalized
249 to those of 21-nt ones in the WT library. As expected, we observed that the proportion of 24-nt
250 siRNA abundances were lower in the T0 and all T1 libraries compared with the wild-type one
251 (Figure 2).

252

253 **The overall DNA methylation level is lowered in all three contexts, CG, CHG and CHH.**

254 As a consequence of impairing the accumulation of 24-nt small RNAs, an overall decrease in
255 cytosine methylation was expected. To investigate this aspect, we profiled the whole genome DNA
256 methylation of the T0 regenerant and two T1 individuals of its progeny (T1-15 and T1-23). We then
257 examined the methylation differences between the RNAi lines and the WT on common sites that
258 were available for all samples and found that the RNAi lines were globally hypomethylated in CHH
259 context (Figure 3), the hallmark of RdDM, but also in CG and CHG (Figure S3, additional file 1).
260 Moreover, 10 subclasses of transposable elements were analyzed across all samples and in all
261 contexts and were found to be, on average, hypomethylated (Figure 4).

262

263 **Some new insertions of known active TEs, but no generalized reactivation of transposition in** 264 ***osnrpd1* RNAi lines**

265 We initially hypothesized that *OsNRPD1* was a target of choice to broadly relax epigenetic control
266 over TEs. We then try to systematically assess the transpositional activity in *osnrpd1* and compare it
267 to WT plants. To identify and follow the inheritance of transposition events, we have generated,
268 starting from the T1 plants described above, two more generations of single-seed descents. We thus
269 obtained T2 and T3 generations from T1-15 and T1-23 (Figure 5, top) and until T6 from T1-23. We
270 thus resequenced the genomes of nine individuals, including the T0 as indicated in Figure 5, using
271 the Illumina-based paired-end technology. As a control, we resequenced the genome of the wild-
272 type progenitor of the RNAi lines. All reads were mapped against the reference genome of *Oryza*
273 *sativa* cv. Nipponbare.

274 We then used three different approaches, based on these alignments, to identify TEs mobilization.
275 They are briefly described here, the details can be found in additional file 6. We first used a
276 candidate approach : *Tos17* and *Lullaby* are retrotransposons that have been previously shown to be
277 mobilized both in *in vitro*-cultured cells and *in planta* in mutants impaired in methylation. Based on
278 a simple visual analysis of the alignments in a genome browser (figure S4, additional file 1), we
279 identified 12 new insertions of *Tos17* and one new insertion of *Lullaby* in the T0 plant (Figure 5).
280 Some of them were still present in each of the three T1 plants (Figure 5). All of them were

281 confirmed by PCR (Figure S5, additional file 1) and sequencing. No new insertion was detected in
282 the T1 and further generations except one insertion of *Tos17* (Tos17-13 in Figure 5).
283 Secondly, these 14 insertions were used as a testing set to develop an algorithm to identify
284 insertions of other elements. As depicted in figure S6 (additional file 1), clipped and discordant
285 alignments detected after short reads mapping against the reference genome were used to identify
286 insertion sites and the TE at the origin of each insertion. The parameters of the program were
287 progressively fine-tuned until the whole set of *Tos17* and *Lullaby* insertions described above were
288 identified and all false positives eliminated. In addition, many new insertions of a LINE element,
289 *Karma*, were detected in one T1 and all other lines starting from the T2 and accumulating in the
290 subsequent generations but not in the parental T0 plant (Figure 5). No new insertion from any other
291 TE was detected. Finally, we checked the uncovered neo-insertions by visual inspection of the
292 alignments with a genome browser (IGV) : we confirmed that all of them were present in the RNAi
293 lines and that none of them were present in the wild-type progenitor of *osnrpd1*.
294 Finally, it is well known that short reads mapping is not adapted to the identification of insertions
295 that occur in repetitive regions because this identification rests upon accurate alignments of reads at
296 the insertion site, a prerequisite which is unreachable when the insertion occurs into repeats. To
297 circumvent this problem, we tried to detect TEs that multiply in the genome without identifying
298 their insertion site. We then developed another algorithm based on Depth of Coverage (DOC).
299 Briefly, in a first step, we normalized the read coverage for each analyzed genome by counting the
300 number of reads aligned on a set of 59 unique genes (Table S2, additional file 4) and normalizing to
301 the length of each gene (Table S3, additional file 4). In a second step, the reads aligned on each TE
302 locus were counted and normalized to their length and to the number of reads corresponding to
303 unique regions, giving an estimation of the copy number for each TE (Figure S7, additional file 1).
304 This estimated copy number was obtained for the T0 and the three T1 plants and compared to that
305 of the WT. The method was first validated using the previously known copy number for *Tos17* and
306 *Lullaby*. We then applied the strategy to the whole set of 295 LTR-RTs from our curated database.
307 However, we could not detect a consistent increase in copy number for any of them (Table S4,
308 additional file 4).
309 Therefore, we could not observe any *in planta* generalized reactivation of transposition in *osnrpd1*,
310 except one *Tos17* insertion and the continuous retrotransposition, starting from the T2 generation, of
311 the LINE-type TE called *Karma*. All other neo-insertions (*ie* : detected in *osnrpd1* but not in the
312 WT progenitor) were already present in the T0 regenerant from callus after transformation.

313

314 DISCUSSION

315 In this paper, we have described the effects of a knock-down of *OsNRPD1* gene in rice. The RNAi
316 construct was effective because we could observe the targeting of the *OsNRPD1* gene with 21-nt
317 siRNAs (Figure S2, additional file 1) and the reduction of the accumulation of *OsNRPD1*
318 transcripts (Figure 1). We then observed a reduction of the accumulation of 24-nt siRNAs (Figure
319 2), an overall slight reduction of cytosine methylation in all contexts, CG, CHG and CHH (Figure 3,
320 4, S3) and transposition of three retrotransposons, *Tos17*, *Lullaby* and *Karma*. However, the
321 question arises whether we can attribute these effects to the knock-down of *OsNRPD1* because the
322 construction of this RNAi line included an *in vitro* culture of cells which could explain at least part
323 of them. It has indeed been shown that rice plants that were regenerated after prolonged tissue

324 culture displayed both losses of methylation (36) and new insertions of *Tos17*, *Lullaby* and *Karma*
325 (16–18). Some observations may nevertheless indicate that the knock-down of *OsNRPD1* could
326 enhance these effects.

327 Firstly, we detected 12 new insertions of *Tos17* in the *osnrpd1* regenerant (T0 generation), when the
328 average was 3,37 per line, and no more than 8, in a collection of 384 rice T-DNA mutants which had
329 been obtained at the same facility, with the same protocol of transformation and callus culture (37).
330 In a WT background, where it was established that *Tos17* copy number correlated with the duration
331 of tissue culture, several months were necessary to reach such a level of transposition (16,38) when
332 the transformation and regeneration process to obtain the *osnrpd1* regenerant only took 7 weeks. A
333 similar observation can be made for *Karma* : we have detected as many as 5 new insertions in a
334 single T3 plant (T3-23 in Figure 5) and 20 new other in three further generations (T6-23 in Figure
335 5) when the original publication reported an average number of new insertions per plant lesser than
336 one (18). It should be noted here that transposition events in the papers cited above were evaluated
337 by hybridization after Southern blotting and not genome sequencing as reported here. We however
338 showed in a previous work that both methodologies identified the same number of *Tos17* insertions
339 (39). Therefore, this higher rate of transposition may be a sign of the effect of *OsNRPD1* knock-
340 down.

341 Secondly, we detected a *Tos17* insertion in a T1 plant when it was absent from the analyzed tissues
342 (first leaves of a plantlet) of the parental T0 plant (Tos17-13 in Figures 5 and S5). Such an *in planta*
343 transposition of *Tos17* has never been described in a WT background (16,38) but only in mutants
344 impaired in methylation of DNA (23) or histones (21). Similarly, this event may be interpreted as a
345 transposition that occurred *in planta* in the T0 genome as a consequence of the *OsNRPD1* knock-
346 down, at a late developmental stage, that have been transmitted to some individuals of the progeny.

347 I we assume that, based on these observations, part of the losses in methylation and transpositions
348 events is the consequence of the knock-down of *OsNRPD1*, the effect of the latter is weak. An
349 illustrative aspect is that we were not able to detect any other transposition events than those
350 described above. This is in contrast to what has been reported for *oscm3*, impaired in CHG
351 methylation, where other TEs were found to be mobilized in addition to *Tos17* and *Lullaby* (23). If
352 we can not rule out the possibility that transposition events have occurred that were not detected, it
353 seems that *OsNRPD1* knock-down had a limited impact on transposition. This may be explained by
354 a limited importance of *OsNRPD1* activity in that matter because this selective impact on
355 transposition is reminiscent of what has been observed in *NRPD1* mutants of *Arabidopsis thaliana*
356 (35,40–42). Similarly, the *osdcl3a* mutation, that reduced the expression of *OsDCL3* which encodes
357 the rice DICER-LIKE 3 primarily responsible for 24-nt siRNA processing in canonical RdDM,
358 reduced 24-nt siRNA predominantly from MITEs (26). Although this study did not focus on
359 transposition, this observation also underlined the selective release of the epigenetic control over
360 TEs caused by the impairment of individual components of the RdDM machinery (43,44).

361 Alternatively and more probably, these limited effects are the result of the limited impact of the
362 knock-down itself, raising the possibility, as discussed below, that a null mutant of *OsNRPD1*
363 would be sterile. We indeed gathered several other indications that we only obtained weakly
364 affected *osnrpd1* lines. The most obvious one is that we did not observe any clear developmental
365 phenotype. Pol IV mutants in other species than *Arabidopsis*, whether mono- or dicotyledonous,
366 display severe phenotypes. In maize, the *ZmRPD1* mutant called *rnr6* is pleiotropically affected in
367 development (45,46). Tilling mutants *braA.nrpd1* of *Brassica rapa* notably displayed asynchronous

368 seed abortion (47). Tomato *slnrpd1* null mutants have also been reported to be sterile (48).
369 Moreover, in contrast to *osnrpd1*, pleiotropic developmental abnormalities have been observed in
370 mutants in other components of the RdDM pathway in rice, like *osdrm2* (24) and *osdcl3a* (26). A
371 strong *osnrpd1* mutant would therefore be expected to be affected the same way. In addition, an
372 important reduction of genome-wide methylation was associated with those developmental defects.
373 For instance, a 13.9% decrease in 5-methylcytosine in both CG and non CG contexts was reported
374 for *osdrm2* (24) when *osnrpd1* displayed less than 5% reduction in all contexts (Figure 3, S3) .
375 The fact that we likely obtained an RNAi line only slightly affected in *OsNRPD1* activity indicates,
376 as previously suggested by others (29), that a null mutant would be sterile in rice. It is indeed
377 surprising that a mutant of this central component of the RdDM machinery has never been
378 described in such an important plant, as both a model species and a crop. Data are nevertheless
379 available for other actors of RdDM. The *osdrm2* null mutant, obtained by homologous
380 recombination, is sterile (24). The two described *osdcl3a* lines are knock-down RNAi lines, but
381 they affect phenotypes with a severity correlated with the knockdown level of *OsDCL3a* (26),
382 questioning the possibility to obtain null mutants. Finally, we decided to target *OsNRPD1* by RNAi
383 because we previously failed to find a viable insertional mutant in the available collections
384 (unpublished data). We actually first focus ourselves on a T-DNA insertional mutation into the
385 *OsNRPD1a* gene identified in the Oryza Tag Line library (<http://oryzatagline.cirad.fr/>, line
386 AFVB01) based on the existence of a Flanking Sequence Tag (SAG8G10) corresponding to
387 *OsNRPD1a*. However, we could not confirm the presence of any T-DNA at this position in the T2
388 plants that we could grow. Based on the recent demonstrations, as described above, that *NRPD1*
389 null mutants in tomato and *Brassica* are sterile, it is therefore possible that this insertion gave rise to
390 a sterile phenotype at the homozygous state in rice and that the seeds collected had segregated away
391 this T-DNA. This hypothesis rests upon the speculation that *OsNRPD1b* could not complement the
392 mutation. However, it would explain why no *nrpd1* knock-out mutant is available for rice, and that
393 only weak mutants, like *osnrpd1* we presented in this paper, may survive.

394

395 CONCLUSION

396 We have constructed an knock-down RNAi line of *OsNRPD1*, encoding the largest subunit of
397 POLYMERASE IV, a central component of RdDM in rice. This line displayed reduced 24-nt
398 siRNAs and DNA methylation compared to wild-type. We also detected new insertions of three
399 retrotransposons. However, we could not clearly distinguish between the impact of *in vitro* culture
400 and the knock-down of *OsNRPD1*, even if we have some indications of the effect of the latter. This
401 may be explained by the fact that we only obtained a weakly affected line. We expose reasons that
402 support the hypothesis of the sterility of a null mutant of *OsNRPD1*.

403

404 ACKNOWLEDGEMENTS

405 We thank Moloya Gohain and Pearl Chang for the preliminary analysis of rice methylome, Edouard
406 JOBET for the RT-qPCR analysis, Mayumi Nakano for assistance with data handling and Moaine
407 Elbaidouri for its critical reading of the manuscript.

408

409 FUNDING

410 Emilie DEBLADIS was supported by a grant from the french Ministry of Education. The funding
411 body had no role in the design of the study, analysis and interpretation of data nor in writing the
412 manuscript.

413 REFERENCES

1. Feschotte C, Pritham EJ. DNA Transposons and the Evolution of Eukaryotic Genomes. *Annu Rev Genet.* 2007;41(1):331–68.
2. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhou B, et al. A unified classification system for eukaryotic transposable elements. *Nat Rev Genet.* 2007 Dec;8(12):973.
3. Schulman AH. Retrotransposon replication in plants. *Curr Opin Virol.* 2013 Dec 1;3(6):604–14.
4. Lisch D. How important are transposons for plant evolution? *Nat Rev Genet.* 2013 Jan;14(1):49–61.
5. Bennetzen JL, Wang H. The contributions of transposable elements to the structure, function, and evolution of plant genomes. *Annu Rev Plant Biol.* 2014;65:505–30.
6. Mita P, Boeke JD. How retrotransposons shape genome regulation. *Curr Opin Genet Dev.* 2016 Apr;37:90–100.
7. Deng X, Song X, Wei L, Liu C, Cao X. Epigenetic regulation and epigenomic landscape in rice. *Natl Sci Rev.* 2016 Sep 1;3(3):309–27.
8. Matzke MA, Mosher RA. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet.* 2014 Jun;15(6):394–408.
9. Cuerda-Gil D, Slotkin RK. Non-canonical RNA-directed DNA methylation. *Nat Plants.* 2016 Nov 3;2:16163.
10. Cao X, Aufsatz W, Zilberman D, Mette MF, Huang MS, Matzke M, et al. Role of the DRM and CMT3 Methyltransferases in RNA-Directed DNA Methylation. *Curr Biol.* 2003 Dec 16;13(24):2212–7.
11. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet.* 2010 Mar;11(3):204.
12. Fultz D, Choudury SG, Slotkin RK. Silencing of active transposable elements in plants. *Curr Opin Plant Biol.* 2015 Oct 1;27(Supplement C):67–76.
13. International Rice Genome Sequencing Project. The map-based sequence of the rice genome. *Nature.* 2005 Aug;436(7052):793.

14. Naito K, Cho E, Yang G, Campbell MA, Yano K, Okumoto Y, et al. Dramatic amplification of a rice transposable element during recent domestication. *Proc Natl Acad Sci*. 2006 Nov 21;103(47):17620–5.
15. Petit J, Bourgeois E, Stenger W, Bès M, Droc G, Meynard D, et al. Diversity of the Ty-1 copia retrotransposon Tos17 in rice (*Oryza sativa*) and the AA genome of the *Oryza* genus. *Mol Genet Genomics*. 2009 Dec 1;282(6):633–52.
16. Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M. Retrotransposons of rice involved in mutations induced by tissue culture. *Proc Natl Acad Sci*. 1996 Jul 23;93(15):7783–8.
17. Picault N, Chaparro C, Piegu B, Stenger W, Formey D, Llauro C, et al. Identification of an active LTR retrotransposon in rice. *Plant J*. 2009 Jun;58(5):754–65.
18. Komatsu M, Shimamoto K, Kyojuka J. Two-Step Regulation and Continuous Retrotransposition of the Rice LINE-Type Retrotransposon Karma. *Plant Cell Online*. 2003 Aug 1;15(8):1934–44.
19. Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, et al. An active DNA transposon family in rice. *Nature*. 2003 Jan;421(6919):163.
20. Kikuchi K, Terauchi K, Wada M, Hirano H-Y. The plant MITE mPing is mobilized in anther culture. *Nature*. 2003 Jan;421(6919):167.
21. Ding Y, Wang X, Su L, Zhai J, Cao S, Zhang D, et al. SDG714, a Histone H3K9 Methyltransferase, Is Involved in Tos17 DNA Methylation and Transposition in Rice. *Plant Cell*. 2007 Jan 1;19(1):9–22.
22. La H, Ding B, Mishra GP, Zhou B, Yang H, Bellizzi M del R, et al. A 5-methylcytosine DNA glycosylase/lyase demethylates the retrotransposon Tos17 and promotes its transposition in rice. *Proc Natl Acad Sci*. 2011 Sep 13;108(37):15498–503.
23. Cheng C, Tarutani Y, Miyao A, Ito T, Yamazaki M, Sakai H, et al. Loss of function mutations in the rice chromomethylase OsCMT3a cause a burst of transposition. *Plant J*. 2015 Sep 1;83(6):1069–81.
24. Moritoh S, Eun C-H, Ono A, Asao H, Okano Y, Yamaguchi K, et al. Targeted disruption of an orthologue of DOMAINS REARRANGED METHYLASE 2, OsDRM2, impairs the growth of rice plants by abnormal DNA methylation. *Plant J*. 2012 Jul 1;71(1):85–98.
25. Hu L, Li N, Xu C, Zhong S, Lin X, Yang J, et al. Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proc Natl Acad Sci*. 2014 Jul 22;111(29):10642–7.
26. Wei L, Gu L, Song X, Cui X, Lu Z, Zhou M, et al. Dicer-like 3 produces transposable element-associated 24-nt siRNAs that control agricultural traits in rice. *Proc Natl Acad Sci*. 2014 Nov 3;111(10):3877–82.

27. Zhou M, Law JA. RNA Pol IV and V in gene silencing: Rebel polymerases evolving away from Pol II's rules. *Curr Opin Plant Biol.* 2015 Oct 1;27:154–64.
28. Huang Y, Kendall T, Forsythe ES, Dorantes-Acosta A, Li S, Caballero-Pérez J, et al. Ancient Origin and Recent Innovations of RNA Polymerase IV and V. *Mol Biol Evol.* 2015 Jul 1;32(7):1788–99.
29. Arikkit S, Zhai J, Meyers BC. Biogenesis and function of rice small RNAs from non-coding RNA precursors. *Curr Opin Plant Biol.* 2013 May 1;16(2):170–9.
30. Sallaud C, Meynard D, Boxtel J van, Gay C, Bès M, Brizard JP, et al. Highly efficient production and characterization of T-DNA plants for rice (*Oryza sativa* L.) functional genomics. *Theor Appl Genet.* 2003 May 1;106(8):1396–408.
31. Debladis E, Llauro C, Carpentier M-C, Mirouze M, Panaud O. Detection of active transposable elements in *Arabidopsis thaliana* using Oxford Nanopore Sequencing technology. *BMC Genomics.* 2017 Jul 17;18(1):537.
32. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012 Apr;9(4):357–9.
33. Urich MA, Nery JR, Lister R, Schmitz RJ, Ecker JR. MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. *Nat Protoc.* 2015 Mar;10(3):475–83.
34. Guo W, Fiziev P, Yan W, Cokus S, Sun X, Zhang MQ, et al. BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics.* 2013 Dec;14(1):774.
35. Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi M-A, et al. Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. *Genes Dev.* 2005 Jan 9;19(17):2030–40.
36. Stroud H, Ding B, Simon SA, Feng S, Bellizzi M, Pellegrini M, et al. Plants regenerated from tissue culture contain stable epigenome changes in rice. Baulcombe D, editor. *eLife.* 2013 Mar 19;2:e00354.
37. Piffanelli P, Droc G, Mieulet D, Lanau N, Bès M, Bourgeois E, et al. Large-scale characterization of Tos17 insertion sites in a rice T-DNA mutant library. *Plant Mol Biol.* 2007 Nov 1;65(5):587–601.
38. Hirochika H. Contribution of the Tos17 retrotransposon to rice functional genomics. *Curr Opin Plant Biol.* 2001 Apr 1;4(2):118–22.
39. Sabot F, Picault N, El-Baidouri M, Llauro C, Chaparro C, Piegu B, et al. Transpositional landscape of the rice genome revealed by paired-end mapping of high-throughput re-sequencing data. *Plant J.* 2011 Apr 1;66(2):241–6.

40. Onodera Y, Haag JR, Ream T, Nunes PC, Pontes O, Pikaard CS. Plant Nuclear RNA Polymerase IV Mediates siRNA and DNA Methylation-Dependent Heterochromatin Formation. *Cell*. 2005 Mar 11;120(5):613–22.
41. Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, et al. Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet*. 2005 Jul;37(7):761.
42. Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA Polymerase IV Directs Silencing of Endogenous DNA. *Science*. 2005 Apr 1;308(5718):118–20.
43. Lippman Z, May B, Yordan C, Singer T, Martienssen R. Distinct Mechanisms Determine Transposon Inheritance and Methylation via Small Interfering RNA and Histone Modification. *PLoS Biol*. 2003 Dec 22;1(3):e67.
44. Rigal M, Mathieu O. A “mille-feuille” of silencing: Epigenetic control of transposable elements. *Biochim Biophys Acta BBA - Gene Regul Mech*. 2011 Aug 1;1809(8):452–8.
45. Erhard KF, Stonaker JL, Parkinson SE, Lim JP, Hale CJ, Hollick JB. RNA Polymerase IV Functions in Paramutation in *Zea mays*. *Science*. 2009 Feb 27;323(5918):1201–5.
46. Parkinson SE, Gross SM, Hollick JB. Maize sex determination and abaxial leaf fates are canalized by a factor that maintains repressed epigenetic states. *Dev Biol*. 2007 Aug 15;308(2):462–73.
47. Grover JW, Kendall T, Baten A, Burgess D, Freeling M, King GJ, et al. Maternal components of RNA-directed DNA methylation are required for seed development in *Brassica rapa*. *Plant J*. 2018 May 1;94(4):575–82.
48. Gouil Q, Baulcombe DC. DNA Methylation Signatures of the Plant Chromomethyltransferases. *PLoS Genet*. 2016 Dec 20;12(12).

414

415

416 LEGENDS TO FIGURES

417

418 **Figure 1 : Transcript accumulation of *OsNRPD1* genes is lower in the *osnrpd1* mutants.**

419 A. Top, semi-quantitative RT-PCR analysis of *OsNRPD1a* transcripts accumulation in four T0
 420 regenerants (6.2, 8.1, 20.1 and 23.1) and the wild-type progenitor (WT). Bottom, same analysis in 8
 421 T1 individuals of the T0-23.1 progeny after selfing. The dashed frames indicate the plants that
 422 displayed a strong reduction in transcripts accumulation and that have been selected and further
 423 described in this study.

424 B. RT-qPCR analysis of *OsNRPD1a* transcripts accumulation level relative to *ACTIN* ($\Delta\Delta Ct$
 425 method) in three selected T1 plants (T1-15, T1-17 and T1-23) compared to WT, three biological
 426 replicates each, confirming the results above.

427 C. *OsNRPD1* transcripts accumulation in T1-15 after RNA-Seq analysis showing, based on three
428 biological replicates, that both genes are affected. Bars represent the standard deviation.

429
430 **Figure 2 : Accumulation of 24-nt small RNAs is lowered in the *osnrpd1* mutants.**

431 Small RNA size profiles in wild-type (WT) and *osnrpd1* mutants (T0 and three T1 plants),
432 normalized to the percentage of 21-nt siRNAs abundance in the wild-type library. For each size
433 class, siRNA abundance was calculated as a percentage to the sum of abundances of total genome-
434 matched reads.

435
436 **Figure 3 : The whole genome level of CHH methylation is lower in the *osnrpd1* mutants.**

437 Genome wide view of delta methylation in CHH context, showing that the T0 regenerant and both
438 T1 progenies (T1-15 and T1-23) of the *osnrpd1* mutant have lower methylation than the wild-type
439 (WT). Methylation in CHG and CG contexts is also affected as shown in figure S3, additional file 1.

440
441 **Figure 4 : Methylation in 10 subclasses of transposons is lower in all contexts**

442 To compare the methylation level distribution of differentially methylated transposable elements
443 between each sample (T0, T1-15 and T1-23) and WT, we calculated the methylation differences of
444 each DMTE between each sample and WT for ten known classes of transposable elements and
445 generated boxplots. The analysis has been done for cytosins in CG, CHG and CHH context,
446 respectively, from top to bottom. For each class, the columns 1, 2 and 3 refer to the T0 vs WT
447 comparison, the T1-15 vs WT comparison and the T1-23 vs WT comparison, respectively. The
448 horizontal line indicates the level 0 of delta methylation (no change) all across each boxplot.

449
450 **Figure 5 : List of all TEs insertions identified in *osnrpd1* mutants**

451 The pedigree of all plants tested is indicated at the top. The three columns at the left of the table
452 indicate the chromosome, the coordinates and the name of the new insertion. The coordinates are
453 the start and end positions of the sequence that has been duplicated upon insertion as detected by
454 our program. The true position of the insertion is one of them, depending of the orientation of the
455 TE. For *Tos17* and *Lullaby* insertions, the positions highlighted in bold have been determined by
456 sequencing. A circle (green for *Tos17*, red for *Lullaby*, blue for *Karma*) indicate the presence of the
457 insertion in each plant.

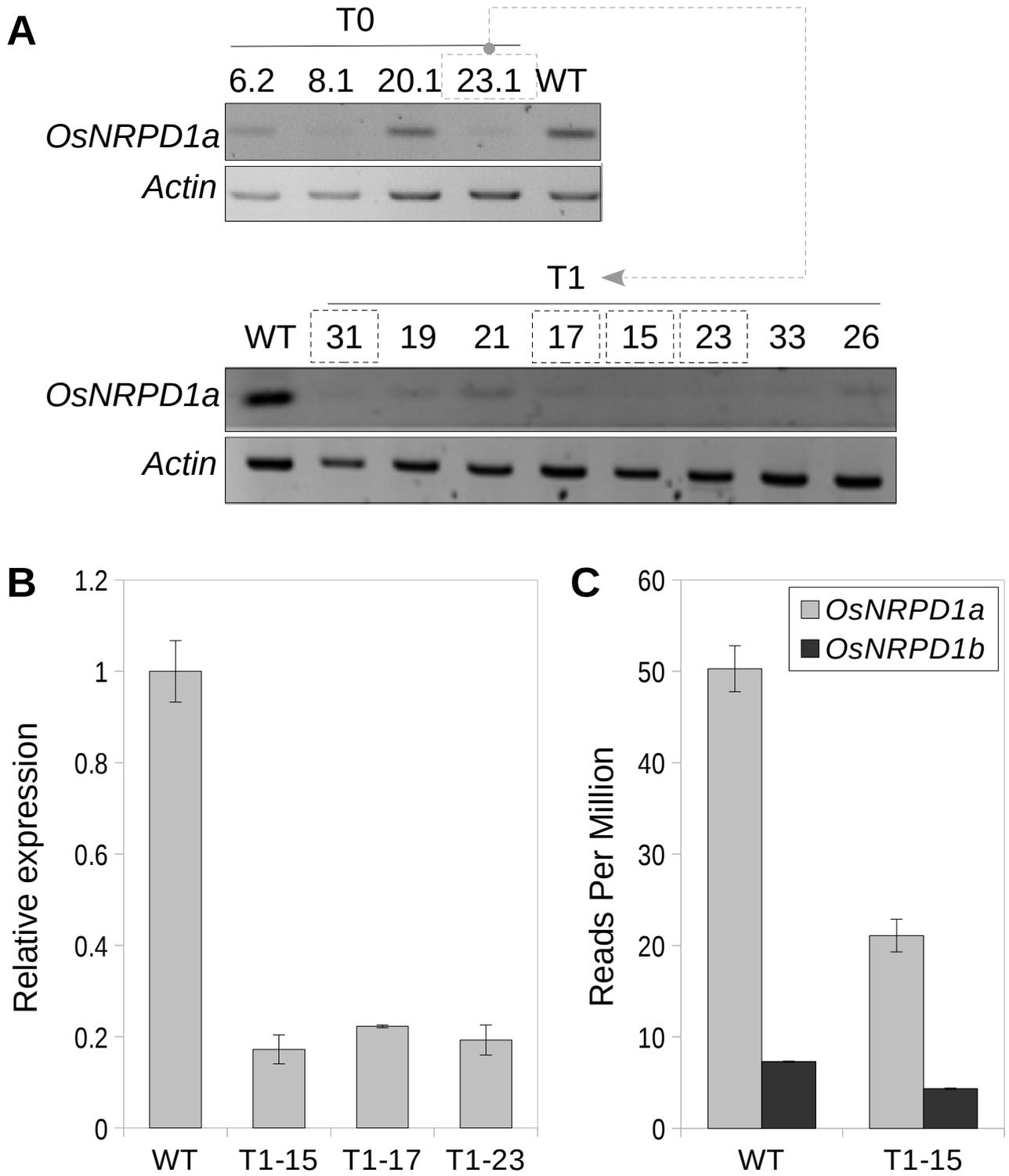


Figure 1

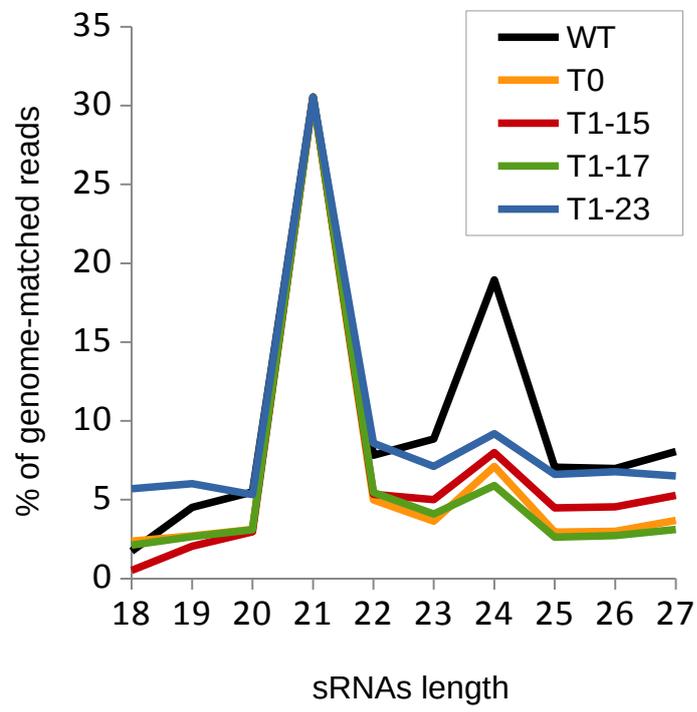


Figure 2

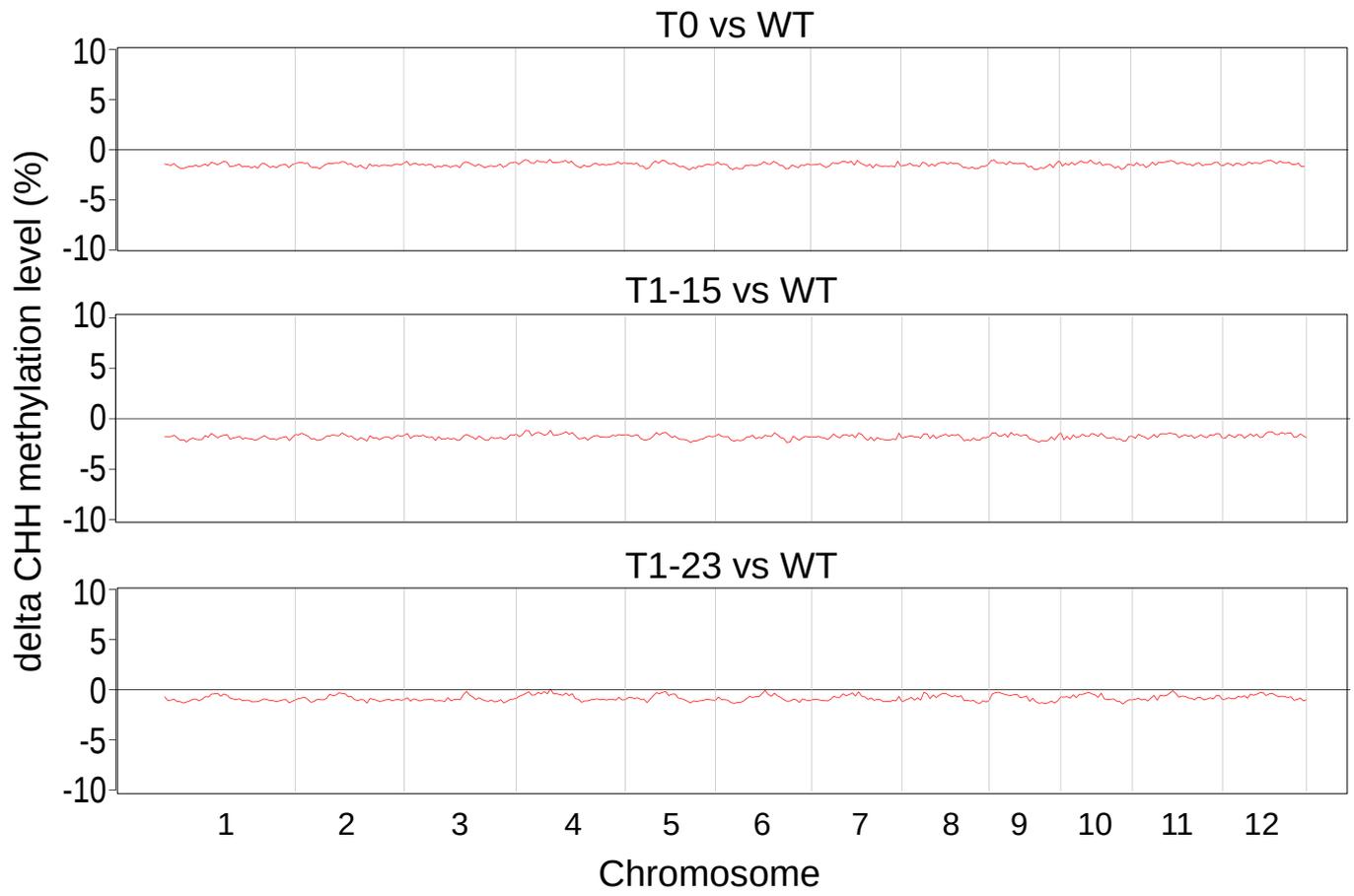


Figure 3

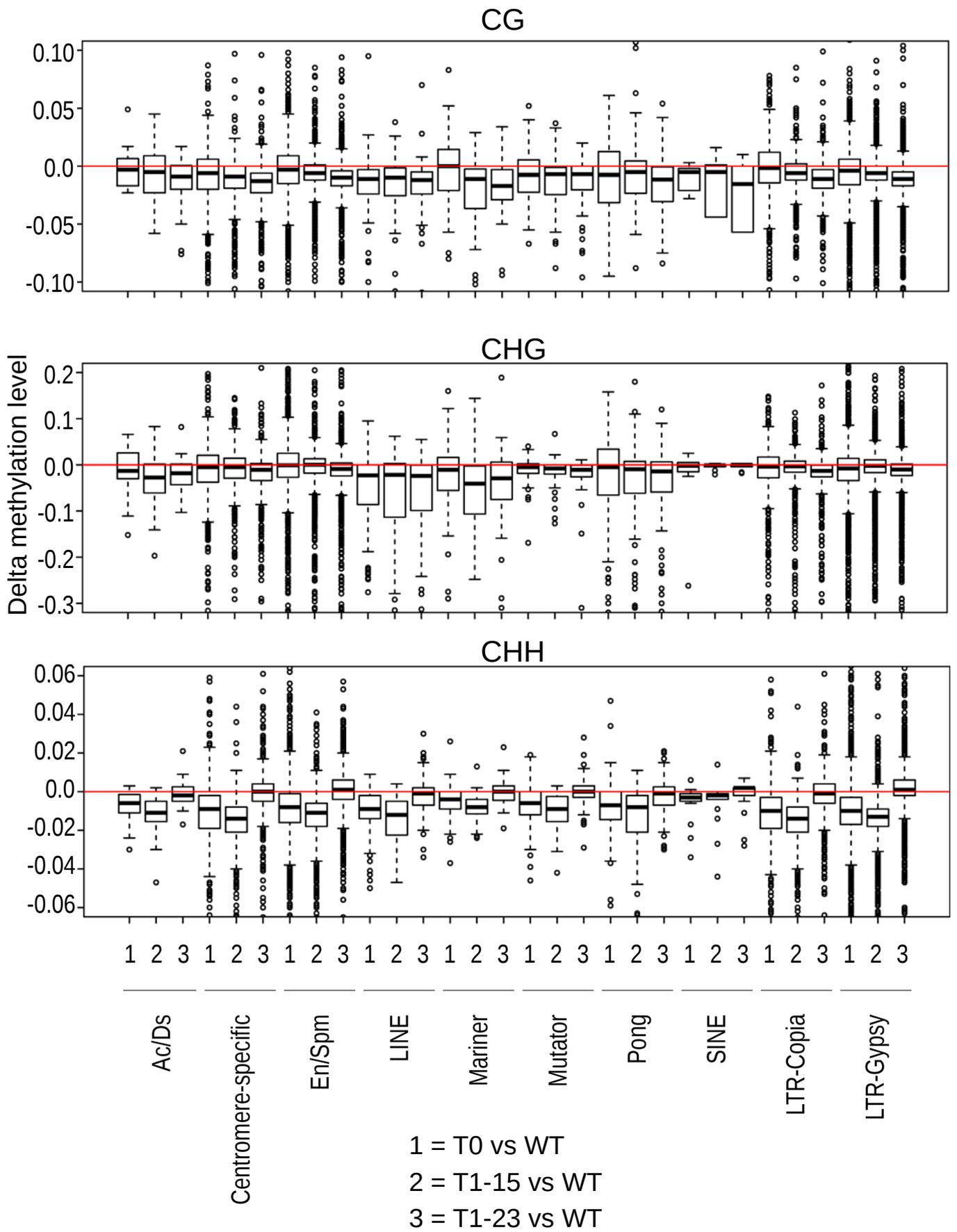


Figure 4

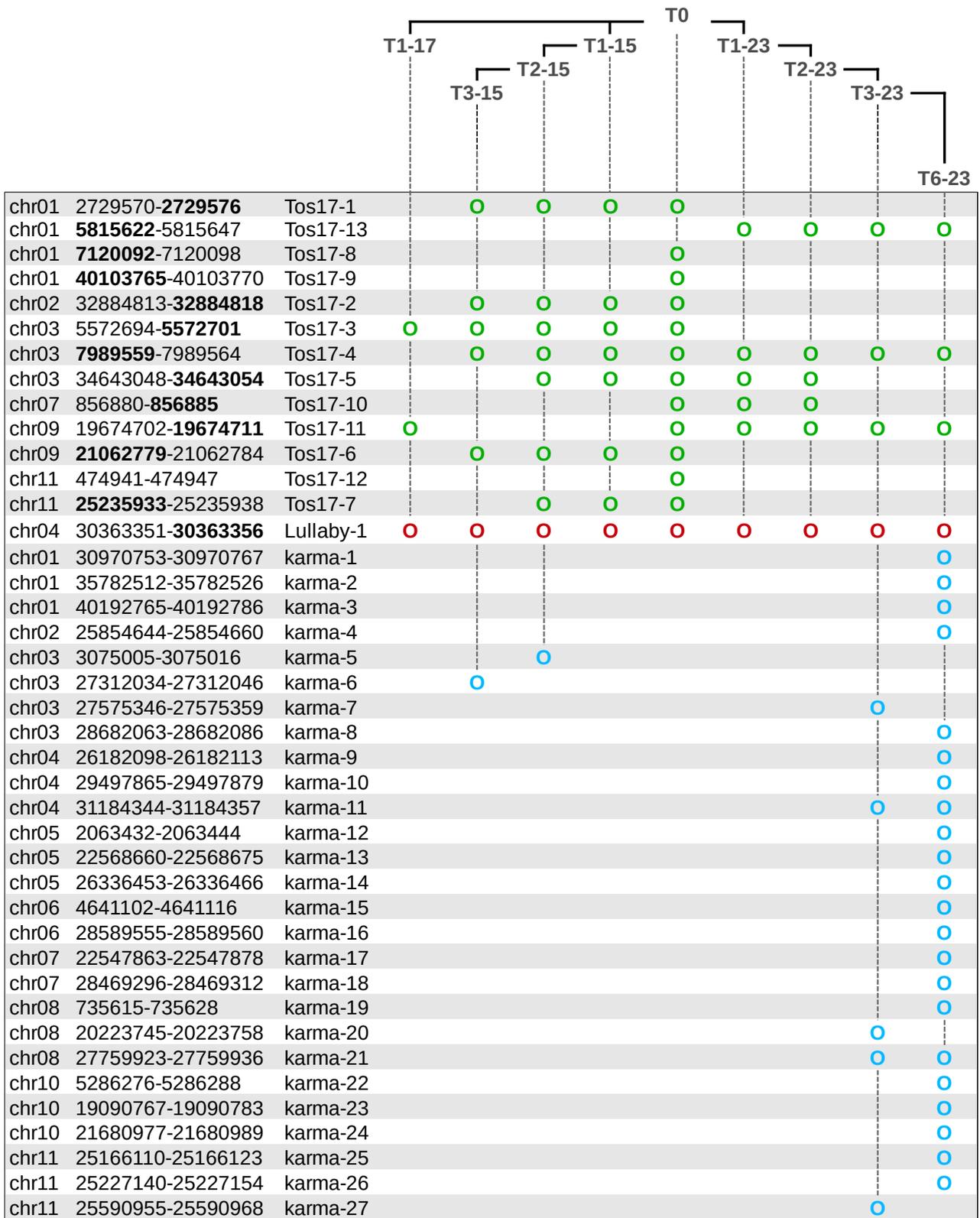


Figure 5