C9orf72 promoter hypermethylation is reduced while hydroxymethylation is acquired during reprogramming of ALS patient cells

Rustam Esanov¹; Kinsley C. Belle²; Marka van Blitterswijk³; Veronique V. Belzil³; Rosa Rademakers³; Dennis W. Dickson³; Leonard Petrucelli³; Kevin B. Boylan⁴; Derek M. Dykxhoorn²; Joanne Wu⁵; Michael Benatar⁵; Claes Wahlestedt¹; Zane Zeier^{1*}

¹Center for Therapeutic Innovation and Department of Psychiatry & Behavioral Sciences, University of Miami Miller School of Medicine, Miami, FL, USA

²John P. Hussman Institute for Human Genomics and the Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine

³Department of Neuroscience, Mayo Clinic, Jacksonville FL, USA

⁴Department of Neurology, Mayo Clinic, Jacksonville FL, USA

⁵Department of Neurology, University of Miami Miller School of Medicine

*To whom correspondence should be addressed:

Dr. Zane Zeier

Department of Psychiatry & Behavioral Sciences

Center for Therapeutic Innovation

University of Miami Miller School of Medicine

1501 NW 10th Ave * BRB # 413

Miami, FL 33136, USA

Telephone: +1 352 246 3455

E-mail: zzeier@med.miami.edu

RUNNING TITLE: C9ORF72 HYPERMETHYLATION

KEY WORDS: C9orf72, amyotrophic lateral sclerosis, repeat expansion, cytosine hydroxymethylation

29 **Abstract:**

30 Amongst several genetic mutations known to cause amyotrophic lateral sclerosis (ALS), a
31 hexanucleotide repeat expansion in the *C9orf72* gene is the most common. In approximately 30% of
32 *C9orf72*-ALS cases, 5-methylcytosine (5mC) levels within the *C9orf72* promoter are increased, resulting
33 in a modestly attenuated phenotype. The developmental timing of *C9orf72* promoter hypermethylation
34 and the reason why it occurs in only a subset of patients remains unknown. In order to model the
35 acquisition of *C9orf72* hypermethylation and examine the potential role of 5-hydroxymethylcytosine
36 (5hmC), we generated induced pluripotent stem cells (iPSCs) from an ALS patient with *C9orf72* promoter
37 hypermethylation. Our data show that 5mC levels are reduced by reprogramming and then re-acquired
38 upon neuronal specification, while 5hmC levels increase following reprogramming and are highest in
39 iPSCs and motor neurons. We confirmed the presence of 5hmC within the *C9orf72* promoter in post-
40 mortem brain tissues of hypermethylated patients. These findings show that iPSCs are a valuable model
41 system for examining epigenetic perturbations caused by the *C9orf72* mutation and reveal a potential
42 role for cytosine demethylation.

43

44 **Introduction:**

45 A GGGGCC hexanucleotide repeat expansion (HRE) mutation near the *C9orf72* gene promoter was
46 recently identified as the most common genetic cause of ALS and frontotemporal dementia (FTD) ^{1, 2}.
47 The mutation accounts for approximately 34% of familial and 6% of sporadic cases of ALS and places it
48 among a large family of nearly 40 repeat expansion disorders, many of which specifically affect neurons
49 ³. The *C9orf72* protein product remains poorly characterized but structurally resembles *Differentially*
50 *Expressed in Normal and Neoplastic cells* (DENN) proteins, which are involved in membrane trafficking
51 processes ⁴. Alleles with expansions larger than 30 repeats are enriched with repressive epigenetic
52 marks within the gene promoter including histone 3 lysine 9 trimethylation (H3K9me3) and cytosine
53 methylation (5mC) ⁵. Such heterochromatinization corresponds to reduced transcription rates, leading to
54 the hypothesis that haploinsufficiency may contribute to the pathology of *C9orf72*-related ALS and FTD ⁶.
55 Bidirectional transcription through the HRE produces expanded RNA transcripts that form RNA foci,
56 sequester RNA binding proteins, and disrupt global RNA processing ^{1, 7}. In addition, dipeptide repeat

57 proteins (DPRs) produced by repeat-associated non-ATG (RAN) translation⁸ of *C9orf72* RNAs harboring
58 the HRE represent a third potential source of pathology⁹. Recent studies have shown that
59 nucleocytoplasmic transport is disrupted by expanded RNAs and DPRs¹⁰⁻¹². Thus *C9orf72*-ALS
60 pathology may result, in part, from the loss of gene function while toxic RNAs and DPRs play a principal
61 role.

62

63 Cytosine methylation within CpG islands near the *C9orf72* promoter occurs in approximately 30% of
64 patients across multiple tissue types^{5, 13}. This hypermethylation is hypothesized to be protective due to a
65 decreased amount of toxic products made in patient cells¹⁴, reduced neuronal loss in the brain¹⁵, and is
66 inversely correlated with duration of the disease¹⁶. Similar cytosine hypermethylation is a characteristic
67 feature of other heterochromatinized repeat expansion mutations including those responsible for Fragile
68 X Syndrome (FX) and Friedrich's ataxia^{17, 18}. In addition to the role of 5mC, the potential role for 5-
69 hydroxymethylcytosine (5hmC) in the heterochromatinization of genes harboring repeat expansion
70 mutations has recently been investigated¹⁹⁻²¹. The ten-eleven translocation (TET) family of proteins
71 converts 5mC to a stable 5hmC intermediate in a step-wise series of reactions that can culminate in the
72 removal of 5mC residues (active DNA demethylation)^{22, 23}. 5hmC levels are enriched in the mammalian
73 brain^{24, 25}, pluripotent stem cells²⁶⁻²⁹, are dynamically regulated during neurodevelopment^{22, 24}, and vary
74 across cell types^{30, 31}, including neuronal subtypes^{24, 30}. The TET proteins modulate iPSC
75 reprogramming efficiency³² and levels of TET decrease upon differentiation of pluripotent cells³³. Since
76 previous examinations of *C9orf72* hypermethylation have relied on the bisulfite conversion method,
77 which is unable to disambiguate 5mC from 5hmC, we used an induced pluripotent stem cell (iPSC)
78 model system to investigate whether 5hmC is present at the *C9orf72* promoter using a 5hmC-sensitive
79 assay.

80

81 The recent advent of iPSC technology has provided an opportunity to examine aspects of the *C9orf72*
82 HRE, including repeat instability and local epigenetic alterations, such as promoter hypermethylation³⁴.
83 Furthermore, iPSCs can be used to derive disease relevant cell-types from patients with well-
84 characterized phenotypes and provide a scalable resource for therapeutic development efforts. Given its

85 large size and high GC content, the HRE is refractory to manipulation using traditional molecular genetic
86 techniques, while primary patient cells and post-mortem tissues cannot be used to model the dynamic
87 acquisition of *C9orf72* heterochromatinization in developing neurons. Therefore, we generated iPSCs
88 from a hypermethylated cell line obtained from an individual within our patient population to examine the
89 acquisition of 5mC hypermethylation and the putative role of 5hmC. Our results demonstrate that the
90 *C9orf72* HRE is unstable through the reprogramming process, promoter hypermethylation is reduced in
91 iPSCs, and 5hmC levels are enriched in primary and iPSC-derived motor neurons.

92

93 **Results:**

94 **Generation and characterization of induced pluripotent stem cells from patient-derived**
95 **lymphocytes.** We identified an individual within our patient population whose fibroblast and immortalized
96 lymphocyte cell lines maintained stable promoter hypermethylation³⁵. While both cell lines derived from
97 this patient were hypermethylated, immortalized lymphocytes generated by the standard Epstein-Barr
98 Virus (EBV) method^{36, 37} possessed a larger HRE, as determined by southern blot analysis³⁵. Therefore,
99 we used immortalized lymphocyte cells (hence-forth referred to as lymphocytes) to generate iPSCs. The
100 observation that fibroblasts tend to have smaller repeats suggests some degree of somatic repeat
101 contraction in this cell type, or that fibroblasts with smaller repeats have a selective advantage *in vitro*³⁸.
102 Two different lymphocyte-derived iPSC clones (#1, #2) from the same *C9orf72*-ALS patient were
103 generated, one of which (#1) was used to generate motor neurons. To establish a negative control cell
104 line and mitigate variability resulting from the primary cell type, reprogramming method, or presence of a
105 repeat expansion mutation, we reprogrammed immortalized lymphocytes from an individual affected by
106 Fragile-X syndrome that does not harbor a repeat expansion at the *C9orf72* locus. Using integration-free
107 episomal plasmids to express the essential Yamanaka factors, we were able to generate iPSCs that
108 express the pluripotency markers Oct4, Nanog, TRA-1-81 and SSEA4 (Figure 1A, B). Expression of
109 pluripotency markers was also confirmed by quantitative PCR (qPCR) (Figure S2A). To ensure that EBV
110 cassettes were lost after serial passages of iPSCs, we assessed the levels of EBNA-1 RNA levels and
111 confirmed its absence in iPSCs after 7 passages (Figure S1). Similar methods and loss of EBV self-
112 replicating plasmids in iPSC cultures has been previously reported^{39, 40}.

113

114 **Motor neuron fate specification of patient-derived induced pluripotent stem cells.** To examine the
115 acquisition of *C9orf72* promoter hypermethylation during neurodevelopment, we generated neural
116 precursor cells (NPCs) and terminally differentiated motor neurons from iPSCs in a two-stage
117 differentiation protocol as we have previously reported³⁵. Expression of the NPC cell markers Nestin and
118 Sox2 were confirmed by immunodetection (Figure 1C) while motor neuron specification was confirmed
119 by the expression of the cytoskeletal neuronal marker Tuj1 and Isl1, which localizes to the nucleus and
120 perinucleus of motor neurons (Figure 1D, E). Further confirmation of motor neuron specification was
121 achieved using a lentiviral vector encoding a green fluorescent protein (GFP) reporter gene driven by the
122 motor neuron specific Homebox gene promoter, HB9, as previously reported⁴¹ (Figure 1F). Endogenous
123 expression of Tuj1, Isl1 and HB9 was increased in motor neuron cultures relative to iPSCs (Figure S2B).
124 The *C9orf72* repeat expansion was maintained throughout reprogramming and motor neuron
125 specification as determined by southern blot analysis (Figure 2) and repeat primed PCR (Figure S3).
126 Interestingly, repeat size varied across different cell types, indicating repeat instability during motor
127 neuron specification, a trend that has been previously reported⁴².

128

129 **DNA methylation of the *C9orf72* promoter is progressively acquired during neuronal**
130 **specification.** During reprogramming, the genome undergoes dramatic epigenetic reorganization
131 including the reactivation of essential developmental genes such as Oct4 and Nanog, with concurrent
132 repression of cell-type specific genes^{43,44}. Globally, DNA methylation (5mC) is dramatically reduced in
133 iPSCs⁴⁵. Thus, we hypothesized that 5mC hypermethylation of CpG islands within the *C9orf72* promoter
134 would be reduced in ALS patient-derived iPSCs and re-acquired during motor neuron specification. To
135 test our hypothesis, we isolated genomic DNA from immortalized patient lymphocytes and their iPSC,
136 NPC and motor neuron derivatives. We then used a methylation-sensitive HhaI restriction digest followed
137 by qPCR (mePCR) to assess *C9orf72* promoter methylation levels (Table 1). Since this method is limited
138 to two HhaI sites within the CpG island, we also used a more sensitive and quantitative bisulfite
139 pyrosequencing method to interrogate 5mC levels at 18 CpG sites spanning the CpG island and nearby
140 *C9orf72* transcription start site (Table 2). Both methods confirmed the reduction of 5mC methylation in

141 iPSCs and subsequent re-acquisition upon motor neuron specification in *C9orf72*-ALS patient cells,
142 whereas it remained at low levels in Fragile X cells. To further investigate the acquisition of *C9orf72*
143 promoter methylation during neurodevelopment, we performed a time-course study where DNA was
144 isolated from ALS patient cells and analyzed by mePCR at 24-hour intervals during the iPSC to NPC
145 transition (Figure S4A). As early as day 1, levels of 5mC within the *C9orf72* promoter increased above
146 baseline and continued to increase as neurospheres developed. To exclude the possibility that *C9orf72*
147 promoter methylation increases over time in iPSC cultures irrespective of neuronal differentiation, we
148 used the HhaI mePCR assay to assess DNA methylation levels over the course of 5 passages (Figure
149 S4B). We confirmed that the level of *C9orf72* promoter methylation is low and stable in iPSCs, and that
150 neuronal induction promotes its increase. This observation suggests that DNA methylation at the *C9orf72*
151 promoter is acquired during early neurodevelopment. To our knowledge, this is the first study using an
152 iPSC differentiation approach to investigate the acquisition of *C9orf72* DNA methylation.

153

154 **Levels of 5-hydroxymethylcytosine at the *C9orf72* promoter are elevated in *C9orf72*-ALS induced**
155 **pluripotent stem cells and motor neurons.** Despite an emerging role of cytosine hydroxymethylation in
156 neurodegenerative diseases ⁴⁶, no studies to date have examined the putative epigenetic role of 5hmC in
157 regulating the *C9orf72* promoter activity in ALS. To distinguish between 5mC and 5hmC residues at the
158 *C9orf72* promoter at different stages of reprogramming and differentiation, we used a 5hmC sensitive
159 MspI/HpaII restriction digest of genomic DNA followed by qPCR. Assay detection sensitivity was
160 assessed using unmethylated, 5mC, and 5hmC enriched control DNA standards (Figure S5). We found
161 that 5mC levels at the *C9orf72* promoter in *C9orf72*-ALS lymphocytes were reduced after reprogramming
162 and then progressively reacquired during motor neuron specification (Table 3), confirming our previous
163 findings. While only trace levels of 5hmC levels were detectable in *C9orf72*-ALS lymphocytes, they were
164 increased in iPSCs and terminally differentiated motor neurons, mirroring the genome-wide enrichment
165 of 5hmC in these cell types ²⁸⁻³⁰. An independent experiment across disease type showed negligible
166 levels of 5mC and no detectable levels of 5hmC in the promoter of the unexpanded *C9orf72* locus in
167 primary FX control cells at any stage of reprogramming (Figure S6). The location of 5hmC enrichment
168 near the *C9orf72* promoter is consistent with genome-wide 5hmC profiling studies that show enrichment

169 near transcriptional start sites and CpG islands and shores^{28-30, 47}. In addition, the low concentration of
170 5hmC relative to 5mC is representative of genome-wide proportions^{23, 24}, which increase during
171 neurodevelopment³⁰. Our findings suggest that DNA demethylation at the expanded *C9orf72* locus is a
172 dynamic process associated with the acquisition of heterochromatinization during neurodevelopment.

173

174 **DNA 5-hydroxymethylation at the *C9orf72* promoter is enriched in the brain of *C9orf72*-ALS**
175 **patients.** To determine whether 5hmC is present at the *C9orf72* promoter in clinical samples, we
176 analyzed genomic DNA isolated from post-mortem brain tissues of three unaffected control individuals,
177 three hypermethylated *C9orf72*-ALS patients, and three unmethylated *C9orf72*-ALS patients. Methylation
178 status was initially assessed by HhaI mePCR, which detects both 5mC and 5hmC marks (without
179 distinguishing between them) at two CpG sites located -215 and -109 base pairs from the transcriptional
180 start site (Table 1, Figure 3). Next, we used the MspI/HpaII PCR assay to distinguish between 5mC and
181 5hmC at two different CpG sites (-313, +104). We found that at one CpG dinucleotide position (-313)
182 5mC and 5hmC levels were highly enriched in hypermethylated *C9orf72*-ALS patient samples. However,
183 at the +104 position this elevation was observed only for 5hmC in one hypermethylated *C9orf72*-ALS
184 sample (Figure 4). While there is a trend of higher 5hmC enrichment in hypermethylated *C9orf72*-ALS
185 brains, more samples are needed to confirm our findings. The presence of 5hmC within the *C9orf72*
186 promoter of post-mortem brain tissues might be an indication of active demethylation in *C9orf72*-ALS
187 neurons and reduces the likelihood that such enrichment in iPSC-derived neurons is an artifact of the
188 reprogramming process.

189

190 **Discussion:**

191 Patient-derived fibroblasts obtained by dermal biopsy are used for *in vitro* studies and to generate iPSCs.
192 However, DNA isolated from fibroblast cell lines tend to harbor smaller expansions than DNA isolated
193 from the blood or post mortem brains³⁸. Furthermore, our unpublished observations indicate that
194 fibroblasts have lower levels of *C9orf72* promoter hypermethylation than blood from the same individual.
195 Therefore, lymphocytes may be more similar to neurons in that they maintain larger repeats and higher
196 methylation levels as compared to fibroblasts. As such, we propose that peripheral blood mononuclear

197 cells, or immortalized lymphocytes may be a superior source for generating iPSCs than fibroblasts.
198 Regardless of the cell type sourced for iPSC production, we and others ⁴² have observed repeat
199 instability during reprogramming. Our data show a smaller repeat size in iPSC clone #1; yet larger
200 repeats reappeared during motor neuron specification. Alternatively, there may still be some
201 heterogeneity within iPSC cultures and it is possible that the observed changes in repeat size are due to
202 selective advantage of a minor subpopulation of cells to differentiate. In addition, we observed a trend for
203 5mC levels to directly correlated with repeat size, which has also been observed in post-mortem brain
204 tissues ⁴⁸. As our study is limited to one patient, further studies are warranted in order to discern the
205 relationship between repeat size, repeat instability, DNA methylation, and hydroxymethylation levels in
206 reprogrammed cells from a larger cohort of patients. It will also be important to determine whether
207 unmethylated *C9orf72*-ALS cells will wrongly acquire *C9orf72* promoter hypermethylation after
208 reprogramming, or if methylation status is faithfully preserved in iPSC-derived neurons.

209

210 Transcription through the *C9orf72* HRE produces toxic RNAs and DPR proteins that mediate a complex
211 downstream pathobiology ⁴⁹. Hypermethylation of a CpG island within the *C9orf72* promoter has been
212 observed in approximately 30% of *C9orf72*-ALS cases ¹⁶. Such epigenetic repression is a common
213 feature conferred by other repeat expansion mutations ^{17, 18}. Since DNA methylation is closely associated
214 with gene repression, *C9orf72* promoter hypermethylation would be expected to drastically reduce
215 transcription rates, levels of toxic RNAs and DPRs, and result in a mild clinical phenotype. While there is
216 evidence of reduced transcription in the somatic cells of hypermethylated individuals ^{14, 16}, there is no
217 reported correlation between gene expression rates and the level of promoter hypermethylation in the
218 brain. Furthermore, hypermethylated *C9orf72*-ALS patients present with only a modestly attenuated
219 phenotype ¹⁵, although RNA foci and DPR levels may be reduced ¹⁴. Nevertheless, our data indicate that
220 previous assessments of methylation in brain tissues using the bisulfite technique or HhaI restriction
221 digest have overestimated the levels of 5mC since these methods cannot distinguish between 5mC and
222 5hmC ⁵⁰. This is particularly poignant since 5mC is broadly considered to be a repressive mark, while
223 5hmC is typically associated with active transcription ²⁸ and is most abundant in the brain ⁵¹. These
224 findings suggest that active demethylation, which is primarily mediated by TET1 in the brain ⁵²,

225 counteracts epigenetic repression of expanded *C9orf72* alleles. Thus, TET inhibition may represent a
226 viable therapeutic strategy for hypermethylated *C9orf72*-ALS patients, although further studies are
227 needed in order to test this hypothesis. Fortunately, our data indicate that 5hmC is present within the
228 promoter of *C9orf72*-ALS iPSC-derived neurons providing a valuable model for investigating this avenue
229 of research.

230

231 **Materials and methods:**

232 **Immortalized patient-derived lymphocytes:** Primary lymphocytes (B-cells) were isolated from blood of
233 consenting individuals within our *C9orf72*-ALS patient population under IRB approved protocols at the
234 University of Miami Miller School of Medicine. Culturing and immortalization was performed using
235 standard methods^{36, 37}. Immortalized control lymphocytes from an individual with Fragile X syndrome
236 were obtained from the Coriell Institute Biorepository (GM09237).

237

238 **Brain samples:** Protocols were approved by the Mayo Clinic Institutional Review Board and Ethics
239 Committee on human experimentation. All participants or authorized family members gave written
240 informed consent after which autopsies were performed at Mayo Clinic, Jacksonville. Postmortem
241 analyses were conducted to confirm pathological diagnosis of FTLD, and genetic analyses confirmed a
242 *C9orf72* repeat expansion in each case included in this study. 50-100mg of frozen human postmortem
243 brain tissues were sampled from the hippocampus of each cases, after which DNA was extracted using
244 QIAamp DNA Mini Kit® (Qiagen #51306) as per manufacturers' instructions.

245

246 **Reprogramming of immortalized lymphocytes:** Immortalized lymphocytes were incubated in
247 StemPro-34 medium supplemented with cytokines: SCF (100ng/mL), IL-3 (50ng/mL), and GM-CSF
248 (25ng/mL) for 3 days. Two million cells were electroporated using the Lonza Nucleofector Kit, and 2µg of
249 episomal plasmids pCXLE-hOCT3/4-shp53-F (Addgene ID 27077), pCXLE-hSK (27078), pCXLE-hUL
250 (27080), and pCDNA3.1-Nanog (28221). Following transfection using the 4D-Nucleofector™ system
251 (program EO-100), cells were seeded on a monolayer of mouse embryonic fibroblast (MEF) feeder cells.
252 The next day, medium was replaced with N2B27 medium (Life Technologies, Carlsbad CA) for a week

253 before switching to mTeSR medium (Stemcell, Vancouver Canada). Cells were cultured for another two
254 weeks; emerging colonies were manually picked and expanded on feeder-free Matrigel® coated plates
255 (Corning Life Sciences, Corning NY).

256

257 **Motor neuron differentiation:** Motor neurons were differentiated according to our previously published
258 methods ³⁵. Briefly, iPSC colonies were cultured in neural stem cell medium on ultra-low attachment
259 flasks. After neural precursor sphere formation, cells were plated on poly-L-ornithine/laminin coated
260 plates and maintained with neuronal induction medium for at least one month to obtain terminally
261 differentiated motor neurons.

262

263 **Immunocytochemistry:** Cells were fixed with 4% paraformaldehyde and blocked using a buffer (PBS,
264 1% BSA, 0.02% sodium Azide) containing goat serum (20%). For detection of intracellular markers, cells
265 were permeabilized with Triton-x100 (0.03%) for 1 hour and incubated with primary antibodies Oct4,
266 Nanog, Sox2, TRA-1-81 SSEA4 (Cell Signalling Technology 9656), Nestin (Abcam ab6142), Tuj1
267 (Sigma-Aldrich T8578) and Isl1 (Millipore Ab4326) overnight at 4°C. The next day, cells were washed
268 three times with PBS and incubated with secondary antibodies (1:1000) for 2 hours at room temperature.
269 Cells were counter stained with DAPI for 30 minutes and observed by fluorescent microscopy.

270

271 **Viral preparation and transduction:** Viral packaging was performed in LentiX293 cells (Clontech
272 632180) following transfection with the Lenti-HB9-GFP (Addgene ID 37080) construct and lentiviral
273 packaging plasmids psPAX2 (12259) and pCMV delta R8.2 (12263). Viral preparations were
274 concentrated using LentiX concentrator (Clontech 631231) and tittered by qPCR. Patient-derived
275 neuronal cultures were transduced at a MOI of 5 in the presence of 2µg/mL polybrene. Cells expressing
276 GFP were visualized 10 days after transduction by fluorescent microscopy.

277

278 **C9orf72 repeat expansion assessment:** For assessment of the *C9orf72* hexanucleotide repeat
279 expansion, we used a repeat-primed PCR assay in the presence of 5% dimethyl sulfoxide and complete
280 substitution of 7-deaza-2-deoxy GTP for dGTP as previously reported ¹. PCR fragments were analyzed

281 using ABI3730 DNA Analyzer and GeneMapper software. To determine the accurate sizing of an
282 expanded allele, Southern Blot analysis was performed using methods previously reported ³⁸. Briefly, 10
283 µg of genomic DNA was digested with XbaI and electrophoresed in 0.8% agarose gel followed by DNA
284 transfer onto nitrocellulose membrane. Hybridization was performed using digoxigenin-labelled probe,
285 followed by detection via anti-digoxigenin antibody and visualization with CDP-star substrate on
286 autoradiography film.

287

288 **Methylation-sensitive quantitative PCR:** Levels of *C9orf72* promoter methylation were assessed by
289 methylation-sensitive restriction digest followed by qPCR as we previously reported ³⁵. At each stage of
290 reprogramming and neuronal specification, genomic DNA was extracted from cells and quantified by
291 spectrophotometry. 100 ng of DNA was used for methylation-sensitive restriction digest with the HhaI
292 endonuclease. Two sets of primers within the *C9orf72* promoter were used: C9me (L/R:
293 5'CGGTAAAAACAAAATTTTCATCCA / 5'GGGCAACTTGTCTGTTCTT) spans two HhaI restriction
294 sites, and C9ec (L/R: 5'AGGAAAGAGAGGTGCGTCAA / 5'TCCTAAACCCACACCTGCTC) that is in
295 close proximity to C9me but does not span HhaI sites and is used as an endogenous control.

296

297 **Bisulfite pyrosequencing:** Genomic DNA (2µg) was subjected to bisulfite conversion and analyzed by
298 pyrosequencing (EpigenDx Inc.) to assess *C9orf72* promoter methylation levels. The analysis was
299 performed using the PSQTM96HS system and custom designed primers developed by EpigenDx
300 (ADS3232-FS1, ADS3232-FS2 and ADS3233-FS1). The assay covers 18 CpG dinucleotides within the
301 promoter and 5' untranslated region of *C9orf72* from -254 to +125 relative to the transcriptional start site
302 of the human genomic sequence (GRCh38:CM000671.2/Chr9(-)).

303

304 **5-hydroxymethylcytosine assessment:** For relative quantification of 5hmC and 5mC levels, we utilized
305 the EpiMark® 5-hmC and 5-mC Analysis Kit (New England Biolabs E3317), according to the
306 manufacturer's protocol. Briefly, 5µg of genomic DNA was used for glucosylation. Converted DNA was
307 then subjected to restriction digest with MspI or HpaII endonucleases for 8 hours at 37°C and then
308 amplified by qPCR using the C9ec primer set (sequence above) and C9epimark primer set (L/R:

309 5'AAATTGCGATGACTTTGCAG / 5'ACTGCAAACCCTGGTAGG), which span one CCGG MspI/HpaII
310 recognition site.

311

312 **Quantitative PCR:** Reactions were performed using TaqMan® assays for Lin28 (Hs00702808_s1),
313 Nanog (Hs04260366_g1), beta-III-tubulin (Hs00801390_s1), HB9 (Hs00907365_m1), ISL1
314 (Hs00158126_m1), and GAPDH (Hs02758991_g1). For EBV gene expression SYBR green qPCR,
315 EBNA-1 primers (L/R: 5'ATCGTCAAAGCTGCACACAG / 5'CCCAGGAGTCCCAGTAGTCA) and
316 GAPDH primers (L/R: 5'GGCTGTGGGCAAGGTCATCCCTGA / 5'TCCACCACCCTGTTGCTGTA) were
317 designed. For visualization, PCR products were run on 2% agarose gel.

318

319 **Acknowledgements** This work was supported by grants from the U.S. Department of Defense
320 (GRANT11188144), ALS Association (GRANT2009 and Milton Safenowitz postdoctoral fellowship),
321 NINDS/NIH (R01 MH084880-05, R01 NS080882, P50 NS071674 and P01 NS084974), Robert Packard
322 Center for ALS Research at Johns Hopkins, NIH Alzheimer's Disease Research Center, the
323 Epigenomics Translational grant, Center for Individualized Medicine, Mayo Clinic, and the CReATe
324 consortium (U54 NS090291) part of the Rare Diseases Clinical Research Network (RDCRN), an initiative
325 of the Office of Rare Diseases Research (ORDR), National Center for Advancing Translational Sciences
326 (NCATS). CReATe is funded through collaboration between NCATS and the National Institute of
327 Neurological Disorders and Stroke (NINDS).

328

329 **Literature cited:**

330

- 331 1. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC hexanucleotide
332 repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*.
333 Oct 20 2011;72(2):245-256.
- 334 2. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the
335 cause of chromosome 9p21-linked ALS-FTD. *Neuron*. Oct 20 2011;72(2):257-268.
- 336 3. van Blitterswijk M, DeJesus-Hernandez M, Rademakers R. How do C9ORF72 repeat expansions
337 cause amyotrophic lateral sclerosis and frontotemporal dementia: can we learn from other
338 noncoding repeat expansion disorders? *Curr Opin Neurol*. Dec 2012;25(6):689-700.
- 339 4. Zhang D, Iyer LM, He F, Aravind L. Discovery of Novel DENN Proteins: Implications for the
340 Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Front Genet*.
341 2012;3:283.
- 342 5. Belzil VV, Bauer PO, Gendron TF, Murray ME, Dickson D, Petrucelli L. Characterization of DNA
343 hypermethylation in the cerebellum of c9FTD/ALS patients. *Brain Res*. Oct 10 2014;1584:15-21.

- 344 6. Gendron TF, Belzil VV, Zhang YJ, Petrucelli L. Mechanisms of toxicity in C9FTLD/ALS. *Acta*
345 *Neuropathol.* Mar 2014;127(3):359-376.
- 346 7. Zu T, Liu Y, Banez-Coronel M, et al. RAN proteins and RNA foci from antisense transcripts in
347 C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A.* Dec 17
348 2013;110(51):E4968-4977.
- 349 8. Zu T, Gibbens B, Doty NS, et al. Non-ATG-initiated translation directed by microsatellite
350 expansions. *Proc Natl Acad Sci U S A.* Jan 4 2011;108(1):260-265.
- 351 9. Ash PE, Bieniek KF, Gendron TF, et al. Unconventional translation of C9ORF72 GGGGCC
352 expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron.* Feb 20
353 2013;77(4):639-646.
- 354 10. Freibaum BD, Lu Y, Lopez-Gonzalez R, et al. GGGGCC repeat expansion in C9orf72
355 compromises nucleocytoplasmic transport. *Nature.* Sep 3 2015;525(7567):129-133.
- 356 11. Jovicic A, Mertens J, Boeynaems S, et al. Modifiers of C9orf72 dipeptide repeat toxicity connect
357 nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci.* Sep 2015;18(9):1226-1229.
- 358 12. Zhang K, Donnelly CJ, Haeusler AR, et al. The C9orf72 repeat expansion disrupts
359 nucleocytoplasmic transport. *Nature.* Sep 3 2015;525(7567):56-61.
- 360 13. Xi Z, Rainero I, Rubino E, et al. Hypermethylation of the CpG-island near the C9orf72 G(4)C(2)-
361 repeat expansion in FTD patients. *Hum Mol Genet.* Nov 1 2014;23(21):5630-5637.
- 362 14. Liu EY, Russ J, Wu K, et al. C9orf72 hypermethylation protects against repeat expansion-
363 associated pathology in ALS/FTD. *Acta Neuropathol.* Oct 2014;128(4):525-541.
- 364 15. McMillan CT, Russ J, Wood EM, et al. C9orf72 promoter hypermethylation is neuroprotective:
365 Neuroimaging and neuropathologic evidence. *Neurology.* Apr 21 2015;84(16):1622-1630.
- 366 16. Xi Z, Zinman L, Moreno D, et al. Hypermethylation of the CpG island near the G4C2 repeat in
367 ALS with a C9orf72 expansion. *Am J Hum Genet.* Jun 6 2013;92(6):981-989.
- 368 17. Castaldo I, Pinelli M, Monticelli A, et al. DNA methylation in intron 1 of the frataxin gene is related
369 to GAA repeat length and age of onset in Friedreich ataxia patients. *J Med Genet.* Dec
370 2008;45(12):808-812.
- 371 18. McLennan Y, Polussa J, Tassone F, Hagerman R. Fragile x syndrome. *Curr Genomics.* May
372 2011;12(3):216-224.
- 373 19. Yao B, Lin L, Street RC, et al. Genome-wide alteration of 5-hydroxymethylcytosine in a mouse
374 model of fragile X-associated tremor/ataxia syndrome. *Hum Mol Genet.* Feb 15 2014;23(4):1095-
375 1107.
- 376 20. Al-Mahdawi S, Sandi C, Mouro Pinto R, Pook MA. Friedreich ataxia patient tissues exhibit
377 increased 5-hydroxymethylcytosine modification and decreased CTCF binding at the FXN locus.
378 *PLoS One.* 2013;8(9):e74956.
- 379 21. Wang F, Yang Y, Lin X, et al. Genome-wide loss of 5-hmC is a novel epigenetic feature of
380 Huntington's disease. *Hum Mol Genet.* Sep 15 2013;22(18):3641-3653.
- 381 22. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to
382 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature.* Aug 26
383 2010;466(7310):1129-1133.
- 384 23. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in
385 mammalian DNA by MLL partner TET1. *Science.* May 15 2009;324(5929):930-935.
- 386 24. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje
387 neurons and the brain. *Science.* May 15 2009;324(5929):929-930.
- 388 25. Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human
389 brain. *Nucleic Acids Res.* Jul 2011;39(12):5015-5024.
- 390 26. Szulwach KE, Li X, Li Y, et al. 5-hmC-mediated epigenetic dynamics during postnatal
391 neurodevelopment and aging. *Nat Neurosci.* Dec 2011;14(12):1607-1616.
- 392 27. Costa Y, Ding J, Theunissen TW, et al. NANOG-dependent function of TET1 and TET2 in
393 establishment of pluripotency. *Nature.* Mar 21 2013;495(7441):370-374.
- 394 28. Ficiz G, Branco MR, Seisenberger S, et al. Dynamic regulation of 5-hydroxymethylcytosine in
395 mouse ES cells and during differentiation. *Nature.* May 19 2011;473(7347):398-402.
- 396 29. Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE. 5-Hydroxymethylcytosine is
397 associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.*
398 2011;12(6):R54.

399 30. Wang T, Pan Q, Lin L, et al. Genome-wide DNA hydroxymethylation changes are associated with
400 neurodevelopmental genes in the developing human cerebellum. *Hum Mol Genet.* Dec 15
401 2012;21(26):5500-5510.

402 31. Kinney SM, Chin HG, Vaisvila R, et al. Tissue-specific distribution and dynamic changes of 5-
403 hydroxymethylcytosine in mammalian genomes. *J Biol Chem.* Jul 15 2011;286(28):24685-24693.

404 32. Chen J, Guo L, Zhang L, et al. Vitamin C modulates TET1 function during somatic cell
405 reprogramming. *Nat Genet.* Dec 2013;45(12):1504-1509.

406 33. Cheng Y, Xie N, Jin P, Wang T. DNA methylation and hydroxymethylation in stem cells. *Cell*
407 *Biochem Funct.* Jun 2015;33(4):161-173.

408 34. Yanovsky-Dagan S, Mor-Shaked H, Eiges R. Modeling diseases of noncoding unstable repeat
409 expansions using mutant pluripotent stem cells. *World J Stem Cells.* Jun 26 2015;7(5):823-838.

410 35. Zeier Z, Esanov R, Belle KC, et al. Bromodomain inhibitors regulate the C9ORF72 locus in ALS.
411 *Exp Neurol.* Jun 19 2015.

412 36. Amoli MM, Carthy D, Platt H, Ollier WE. EBV immortalization of human B lymphocytes separated
413 from small volumes of cryo-preserved whole blood. *Int J Epidemiol.* Apr 2008;37 Suppl 1:i41-45.

414 37. Tosato G, Cohen JL. Generation of Epstein-Barr Virus (EBV)-immortalized B cell lines. *Curr*
415 *Protoc Immunol.* Feb 2007;Chapter 7:Unit 7 22.

416 38. van Blitterswijk M, DeJesus-Hernandez M, Niemantsverdriet E, et al. Association between repeat
417 sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions
418 (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol.* Oct 2013;12(10):978-988.

419 39. Choi SM, Liu H, Chaudhari P, et al. Reprogramming of EBV-immortalized B-lymphocyte cell lines
420 into induced pluripotent stem cells. *Blood.* Aug 18 2011;118(7):1801-1805.

421 40. Rajesh D, Dickerson SJ, Yu J, Brown ME, Thomson JA, Seay NJ. Human lymphoblastoid B-cell
422 lines reprogrammed to EBV-free induced pluripotent stem cells. *Blood.* Aug 18 2011;118(7):1797-
423 1800.

424 41. Son EY, Ichida JK, Wainger BJ, et al. Conversion of mouse and human fibroblasts into functional
425 spinal motor neurons. *Cell Stem Cell.* Sep 2 2011;9(3):205-218.

426 42. Almeida S, Gascon E, Tran H, et al. Modeling key pathological features of frontotemporal
427 dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. *Acta Neuropathol.*
428 Sep 2013;126(3):385-399.

429 43. Hill PW, Amouroux R, Hajkova P. DNA demethylation, Tet proteins and 5-hydroxymethylcytosine
430 in epigenetic reprogramming: an emerging complex story. *Genomics.* Nov 2014;104(5):324-333.

431 44. Santiago M, Antunes C, Guedes M, Sousa N, Marques CJ. TET enzymes and DNA
432 hydroxymethylation in neural development and function - how critical are they? *Genomics.* Nov
433 2014;104(5):334-340.

434 45. Leitch HG, McEwen KR, Turp A, et al. Naive pluripotency is associated with global DNA
435 hypomethylation. *Nat Struct Mol Biol.* Mar 2013;20(3):311-316.

436 46. Sherwani SI, Khan HA. Role of 5-hydroxymethylcytosine in neurodegeneration. *Gene.* Oct 1
437 2015;570(1):17-24.

438 47. Wu H, D'Alessio AC, Ito S, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution
439 reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.*
440 Apr 1 2011;25(7):679-684.

441 48. Gijssels I, Van Mossevelde S, van der Zee J, et al. The C9orf72 repeat size correlates with
442 onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Mol*
443 *Psychiatry.* Oct 20 2015.

444 49. van Blitterswijk M, Rademakers R. Neurodegenerative disease: C9orf72 repeats compromise
445 nucleocytoplasmic transport. *Nat Rev Neurol.* Nov 3 2015.

446 50. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-
447 hydroxymethylcytosine in bisulfite sequencing. *PLoS one.* 2010;5(1):e8888.

448 51. Globisch D, Munzel M, Muller M, et al. Tissue distribution of 5-hydroxymethylcytosine and search
449 for active demethylation intermediates. *PLoS One.* 2010;5(12):e15367.

450 52. Kaas GA, Zhong C, Eason DE, et al. TET1 controls CNS 5-methylcytosine hydroxylation, active
451 DNA demethylation, gene transcription, and memory formation. *Neuron.* Sep 18 2013;79(6):1086-
452 1093.

453

454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481

Figure legends

Figure 1. Characterization of patient-derived iPSCs, NPCs, and motor neurons.

Immunofluorescence analysis of pluripotency markers Oct4, TRA-1-81, Nanog and SSEA4 in iPSC colonies (A, B), neural stem cell markers Sox2 and Nestin in neural precursor cells (C), neuronal marker beta-III-tubulin (Tuj1) (D), and motor neuron specific marker ISL1 in neuronal cultures (E). Motor neurons, differentiated for 4 weeks visualized on day 5 post HB9-GFP lentiviral transduction (F), scale bar = 50µm.

Figure 2. Southern blot analysis of the *C9orf72* hexanucleotide repeat expansion through

reprogramming and neuronal differentiation. Concise bands demonstrate the presence of both wild type (WT) and long expanded alleles in ALS patient lymphocytes (Lym), two iPSC clones (#1, #2), neural precursor cells (NPC), and motor neurons (MN). NPCs and MNs, were derived from iPSC clone #1.

Figure 3. Schematic of the *C9orf72* promoter region, indicating CpG sites that were subjected to

DNA methylation and hydroxymethylation assessment. The first *C9orf72* transcriptional start site (TSS) and CpG dinucleotides (lollipops) are indicated. Position of two MspI restriction sites (and its isoschizomere HpaII) used for the hydroxymethylation sensitive assay are indicated at -313 and +104 base pairs (#). Two hydroxymethylation insensitive HhaI sites used for methylation PCR are also indicated (*). Grey boxes above the line indicate regions queried by bisulfite pyrosequencing (PYRO), which includes 18 CpGs from -254 to +125. The black box below line indicates position of exon 1a. The position of the hexanucleotide repeat expansion (HRE) is also indicated. Numbers are base pair position relative to the TSS (+1).

Figure 4. DNA hydroxymethylation assessment near the *C9orf72* promoter in *C9orf72*-ALS post-mortem brain tissues. The abundance of 5mC and 5hmC at two MspI/HpaII sites located at -313 and

482 +104 base pairs from the transcriptional start site were assessed in post-mortem brain tissues of
483 unaffected control individuals (C9-), hypermethylated (C9+, meth+) and unmethylated (C9+, meth -)
484 *C9orf72*-ALS patients (N=3 per group). Black and grey horizontal lines represent mean levels of 5mC
485 and 5hmC per group, respectively.

486

487 **Table 1. Site-specific methylation assessment at the *C9orf72* promoter using DNA methylation-**
488 **sensitive PCR.** Aggregate DNA methylation (5mC+5hmC) levels across two HhaI sites (-215 & -109) in
489 patient-derived cell lines and post-mortem brain samples.

490

491 **Table 2. Pyrosequencing analysis of *C9orf72* promoter in ALS and Fragile X cells during**
492 **reprogramming and differentiation.** Heat-map illustrating quantification of cytosine methylation at 18
493 CpG sites near *C9orf72* transcription start site using bisulfite pyrosequencing.

494

495 **Table 3. Site-specific DNA hydroxymethylation assessment near the *C9orf72* promoter in ALS**
496 **patient-derived cells.** For hydroxymethylation assessment, two CpG dinucleotides at positions -313 and
497 +104 base pairs from the transcriptional start site (TSS) within MspI/HpaII restriction sites were
498 considered. Numbers indicate percent hydroxymethylcytosine (5hmC, green), methylcytosine (5mC, red)
499 and the average of two marks across both positions.

Table 1

Cell type	5mC+5hmC %	Brain sample	5mC+5hmC %
ALS-Lym	19.9	C9 brain 44	39.5
ALS-iPSC-1	7.2	C9 brain 54	13.4
ALS-iPSC-2	13.9	C9 brain 56	9.5
ALS-NPC-1	15.3	C9 brain 48	5.6
ALS-MN-1	35.1	C9 brain 32	4.8
Fx-Lym	0.3	C9 brain 43	4.2
Fx-iPSC	0.3	CTL brain 1	3.3
Fx-NPC	0.3	CTL brain 2	3.2
Fx-MN	3.8	CTL brain 3	2.2
CTL-Low	6.1	CTL-High	100

Table 2

PYRO	CpG 1	CpG 2	CpG 3	CpG 4	CpG 5	CpG 6	CpG 7	CpG 8	CpG 9	CpG 10	CpG 11	CpG 12	CpG 13	CpG 14	CpG 15	CpG 16	CpG 17	CpG 18	Mean
ALS-Lym	47.8	36.2	26.1	29.1	10.8	11.2	31.2	12.3	11.1	18.3	20.5	22.4	35.3	37.9	25	9.4	9.8	10.8	22.5
ALS-iPSC-1	8.1	10.1	8.2	8.8	4.3	5.1	8.7	7.8	9.2	9.2	8.8	8.6	10.5	8.7	8	6.6	6.4	5.4	7.9
ALS-NPC-1	15.9	16.5	9.7	9.5	4.8	5.1	9.8	12	11.7	13.8	8.5	8.3	16.2	13.1	12.4	9.9	7.3	6.9	10.6
ALS-MN-1	32.5	31.4	29.2	29.9	20.8	18.4	28.3	28.5	30.2	32.9	32	29	35.7	33.2	31.1	25.9	18.4	21.1	28.2
Fx-Lym	0	1.9	0	2.4	1.6	0	0	0	0	2.7	1.8	1.9	2.4	2.4	0	0	3.3	0	1.1
Fx-iPSC	0	2.5	1.2	2.7	0	0	0	0	0	0	1.6	2.8	2.3	2.4	1.7	0	3.1	0	1.1
Fx-NPC	0	2.7	0	2.2	0	0	0	0	0	0	3	2.1	3.4	2.5	0	0	0	0	0.9
Fx-MN	0	2.7	1.4	2.1	0	0	0	0	0	0	1.9	2.9	3.2	1.9	0	0	2.6	0	1
Low Control	1.5	2.9	3.3	3.5	0	0	0	0	2.8	0	2	3	7	6.1	1.9	0	0	0	1.9
High Control	91.8	54.5	81.2	95.7	94.6	79.2	84.8	75.1	83.6	87.3	91.5	88.2	97.6	85.7	85.7	79.7	90.8	84.7	85.1
CpG (From TSS)	-254	-215	-55	-45	-36	-26	-20	2	11	19	37	47	57	63	75	91	104	125	

Table 3

CpG	-313		+104		
Sample	5mC %	5hmC %	5mC %	5hmC %	Average 5mC+5hmC
ALS-Lym	69.3	0	4.6	0	37
ALS-iPSC-1	20.7	16.2	1.7	1.9	20.3
ALS-iPSC-2	31.4	2.7	11.3	2.4	23.9
ALS-NPC-1	37.1	13.8	3.3	0.6	27.4
ALS-MN-1	49	3.9	15.4	2	35.2

Figure's

Figure 1

Figure 2

Figure 3

Figure 4