MORC2 directs transcription-dependent CpG methylation of human LINE-1 transposons in early neurodevelopment

Fereshteh Dorazehi,[1] Caterina Francesconi,[1]# Ninoslav Pandiloski,[1,2]# Laura Castilla-Vallmanya,[2]# Anna Albecka,[3,4] Symela Koutounidou,[1] Sivani Bala Mohan,[1] Anita Adami,[2] Angelos Katsikas,[1] Jana Matijević,[1] Ofelia Karlsson,[2] Carrie Davis-Hansson,[2] Jenny G. Johansson,[2] Yorgo Modis,[3] Johan Jakobsson,[2] Christopher H. Douse[1]*

# these authors contributed equally

* correspondence: christopher.douse@med.lu.se

1. Lab of Epigenetics and Chromatin Dynamics, Lund Stem Cell Center and Section of Neurobiology, Department of Experimental Medical Science, Lund University, Sweden.
2. Lab of Molecular Neurogenetics, Lund Stem Cell Center and Section of Neurobiology, Department of Experimental Medical Science, Lund University, Sweden.
3. Molecular Immunity Unit, Department of Medicine, University of Cambridge, MRC Laboratory of Molecular Biology, Cambridge, U.K.
4. Present address: MRC Laboratory of Molecular Biology, Cambridge, U.K.

ABSTRACT.

Methylation of CpG dinucleotides is essential for silencing genomic repeats such as LINE-1 retrotransposons (L1s) in the germline and soma. Evolutionarily-young L1s are transcribed in human pluripotent stem cells, but how CpG methylation is patterned to these L1s upon exit of pluripotency is unknown. Here we investigate the critical functions of chromatin regulator MORC2 in epigenome reprogramming of the repetitive genome in early human neurodevelopment. We find that reversible ATP-dependent dimerization is required for MORC2 accumulation over L1s but not gene promoters. Engineered mutations of the MORC2 ATPase module severely disrupts the distribution of MORC2 chromatin binding, leading to simultaneous loss of L1 transcriptional control and hyper-repression of clustered ZNF genes in human pluripotent stem cells. Upon neural differentiation these phenotypes persist due to striking, targeted defects in CpG methylation patterning. Together our results define the vital role of MORC2 in safeguarding the somatic human genome upon exit of pluripotency by directing CpG methylation patterning over transcriptionally-active retrotransposons in a manner analogous to the piRNA pathway in the germline.

33 **INTRODUCTION.**

34 Transposable elements (TEs) – in particular, retrotransposons such as long interspersed
35 nuclear element 1 (L1) – have accumulated in mammalian genomes and account for at least
36 half of human DNA.[1] Out of more than half a million fixed L1s in the reference genome, around
37 two thousand hominoid-specific copies retain an intact internal promoter and are thus
38 competent for transcription. These elements are a source of novel transcripts in human
39 tissues.[2–5] A smaller subset of around 100 human-specific L1s encode proteins that enable
40 transposition of their own and other TE transcripts to new sites, resulting in polymorphisms
41 within and between individuals.[6,7] L1 fragments lacking their own promoter often reside in
42 introns of transcribed genes where they can influence mRNA processing and gene
43 expression.[8–11] L1 activity is therefore responsible for substantial genetic and transcriptional
44 variation in the human population.

45 Unrestricted mobilization of L1s and other TEs is thought to be deleterious and transcription
46 of intact sequences is controlled by promoter CpG methylation in somatic mammalian cells.[12–
47 15] As part of the epigenetic reprogramming that happens during germ cell and embryonic
48 development, when TE transcription is transiently tolerated and in certain cases functionally
49 important,[16–20] DNA methylation patterns are reestablished after periods of global erasure. In
50 the germline, the piRNA pathway licenses transcription-dependent TE methylation.[21] In
51 embryonic stem cells, the most studied mechanism of restriction works via the KRAB zinc-
52 finger (KRAB-ZNF) family of transcription factors, which bind to specific TE DNA sequences
53 and recruit the TRIM28 corepressor that promotes histone 3 lysine 9 trimethylation
54 (H3K9me3) and CpG methylation.[22,23] KRAB-ZNF genes have rapidly diversified by segmental
55 duplication during vertebrate evolution to recognize TE and repeat family sequences.[24,25]
56 However, in the human setting, it is thought the youngest L1 subfamilies have evaded this
57 DNA-directed silencing system: changes in the ancestral L1PA3 promoter allowed subsequent,
58 evolutionarily-younger L1 copies to escape silencing by ZNF93.[26,27] In human pluripotent stem
59 cells, L1 expression is indeed mostly restricted to human-specific (L1HS) and hominoid-
60 specific (L1PA2/3) elements, which correlates with relative promoter hypomethylation and
61 varies between individual integrants.[28–30] Although lineage-specific L1-driven transcripts exist,
62 and intronic L1s are passively transcribed in pre-mRNA of thousands of genes, CpG
63 methylation mostly silences the promoters of the youngest, intact L1s in somatic human

64 tissues.[13,31] How methylation is patterned to young L1 promoters upon exit of pluripotency
65 and differentiation has been an open question, but is not thought to depend solely on the
66 KRAB-ZNF/TRIM28 system.

67 A key player in the regulatory relationship (sometimes referred to as an 'arms race'[26])
68 between L1s and KRAB-ZNF genes is the chromatin remodelling ATPase MORC2.[8,32] In
69 collaboration with the human silencing hub (HUSH) complex, MORC2 is recruited to repress
70 transcriptionally-active L1s and repetitive gene clusters, including those encoding KRAB-ZNFs,
71 at H3K9me3-marked sites.[8,32–35] ChIP-seq experiments in cancer cell lines also identified a
72 second class of MORC2 peaks over transcription start sites, the functional significance of
73 which is unknown.[32] MORC2 contains an N-terminal ATPase module that reversibly dimerizes
74 upon ATP binding and hydrolysis, and a C-terminal coiled coil domain that forms constitutive
75 dimers.[36–38] Elegant recent experiments with purified components have shown that MORC2
76 binds and entraps free and nucleosomal DNA between these dimerization domains in a length
77 but not sequence-dependent manner, forming loops and multimeric assemblies that compact
78 the underlying chromatin.[36] This mechanism of MORC chromatin remodelling appears to be
79 conserved at least across the animal kingdom[39] and perhaps more deeply – MORCs are
80 encoded in plant, basal eukaryote and even prokaryote genomes.[40–43]

81 MORC2 functions are vital to human health: predicted loss of function alleles are strongly
82 selected against in the population[44] and heterozygous missense mutations in the MORC2
83 ATPase module cause severe neurodevelopmental disorders with symptoms including
84 intellectual disability and cerebellar ataxia.[45–47] Patient mutations cause changes to the
85 protein's enzymatic activity and hyperrepression of clustered KRAB-ZNF genes in cancer cell
86 lines.[32] A recent paper profiling peripheral blood and fibroblasts from MORC2 disorder
87 patients revealed an 'episignature' corresponding to hypermethylation and repression of
88 genes including KRAB-ZNFs, showing that MORC2 neurodevelopmental disorders are
89 characterized by striking epigenetic changes in cells outside the nervous system.[47] Notably,
90 the functional consequences of removing MORC2 depends substantially on the cellular
91 context. For example, loss of MORC2 in the globally-hypomethylated K562 line causes
92 massive upregulation of young L1 subfamilies, together with changes in KRAB-ZNF
93 expression.[8] In contrast, we recently found that removal of MORC2 in a robustly-methylated
94 neural progenitor cell line had only mild effects on L1 expression – indeed, only when DNA

95 methylation was experimentally removed did MORC2 bind to the activated L1s and dampen
96 their expression.[33] Thus, when young L1 promoters are methylated and silenced, MORC2
97 ignores these sequences and is bound to either older, intronic L1s (transcribed as part of genic
98 pre-mRNA) or other transcribed lineage-specific repeats – in the nervous system, data from
99 our lab and others suggest these to be primarily alpha-satellite repeats and clustered
100 protocadherin genes.[33,34]

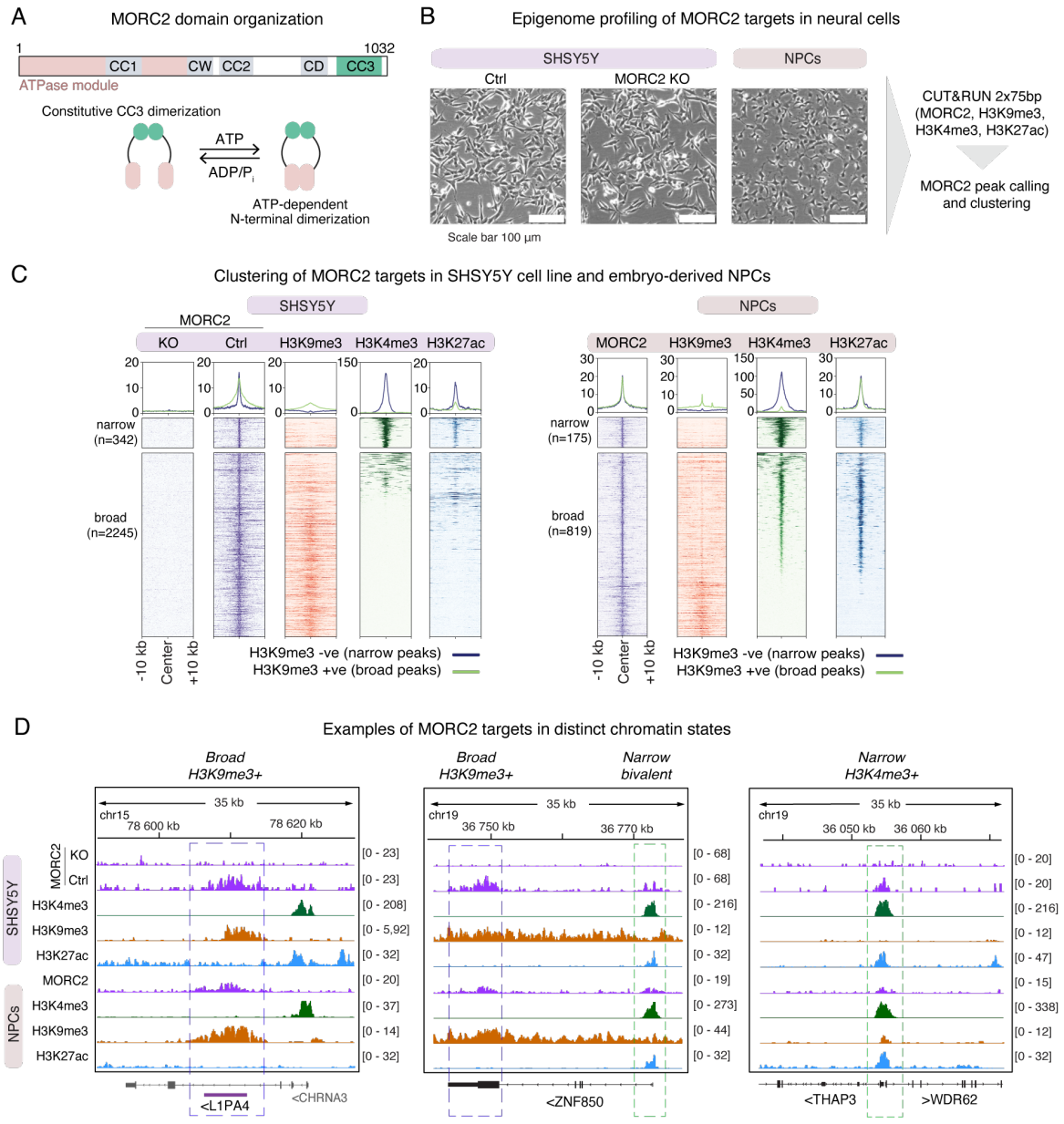
101 Together these observations underscore the layered regulation of repeats in somatic cells and
102 suggest critical functions of MORC2 early in development when the genome is relatively
103 hypomethylated. In this study we use a combination of chromatin profiling, long-read DNA
104 methylation analysis and iPSC modelling to demonstrate (i) how ATPase mutations affect the
105 distribution of MORC2 chromatin binding in the neural lineage, and (ii) the critical function of
106 MORC2 chromatin binding in directing CpG methylation patterns over hominoid-specific L1s
107 and repetitive gene clusters in early neurodevelopment. Together our results define the vital
108 role of MORC2 in safeguarding the somatic human genome upon exit of pluripotency by
109 ensuring CpG methylation patterning over repeats in a transcription-dependent manner,
110 analogous to the piRNA pathway in the germline.

111 **RESULTS.**

112 **MORC2 binds to distinct chromatin states in neural cell lines**

113 MORC2 is a potent and avid dsDNA binding protein.[36,37] The prevailing biochemical model is
114 that reversible ATP-dependent dimerization of the N-terminal ATPase module coupled to
115 constitutive C-terminal dimerization via coiled coil domain 3 enables chromatin compaction
116 by trapping DNA between these domains (**Figure 1A)**.[36,39] We first profiled MORC2 chromatin
117 binding using an optimized CUT&RUN (cleavage under targets and release using nuclease)
118 protocol in SHSY5Y cells and embryo-derived human neural progenitor cells (**Figure 1B, Supp**
119 **Fig 1A,B**).[48] In each cell type we clustered MORC2 binding regions by assessing peak shape
120 and co-occupancy with three histone modifications in control cells: H3K9me3, H3K4me3 and
121 H3K27Ac (**Figure 1C**). We observed broad MORC2 peaks in transcriptionally active, H3K9me3-
122 positive regions such as intronic full-length (>6 kb) L1 retrotransposons and 3' exons of
123 expressed zinc finger (ZNF) genes (**Figure 1D, Supp Fig 1C)**. We also observed widespread
124 evidence of MORC2 binding to bivalent chromatin i.e. H3K4me3/H3K27Ac-marked promoters
125 of expressed genes embedded in H3K9me3-positive chromatin, such as clustered ZNF and

126 protocadherin genes. In our hands a minority of narrow MORC2 peaks were identified over
127 gene regulatory elements marked by H3K27Ac and/or H3K4me3 in H3K9me3-negative
128 chromatin contexts. Taken together, in neural cells MORC2 localizes to distinct chromatin
129 states characterized by transcriptional activity and either broad or narrow occupancy,
130 exemplified by L1 retrotransposons and repetitive clustered genes such as ZNFs.



131

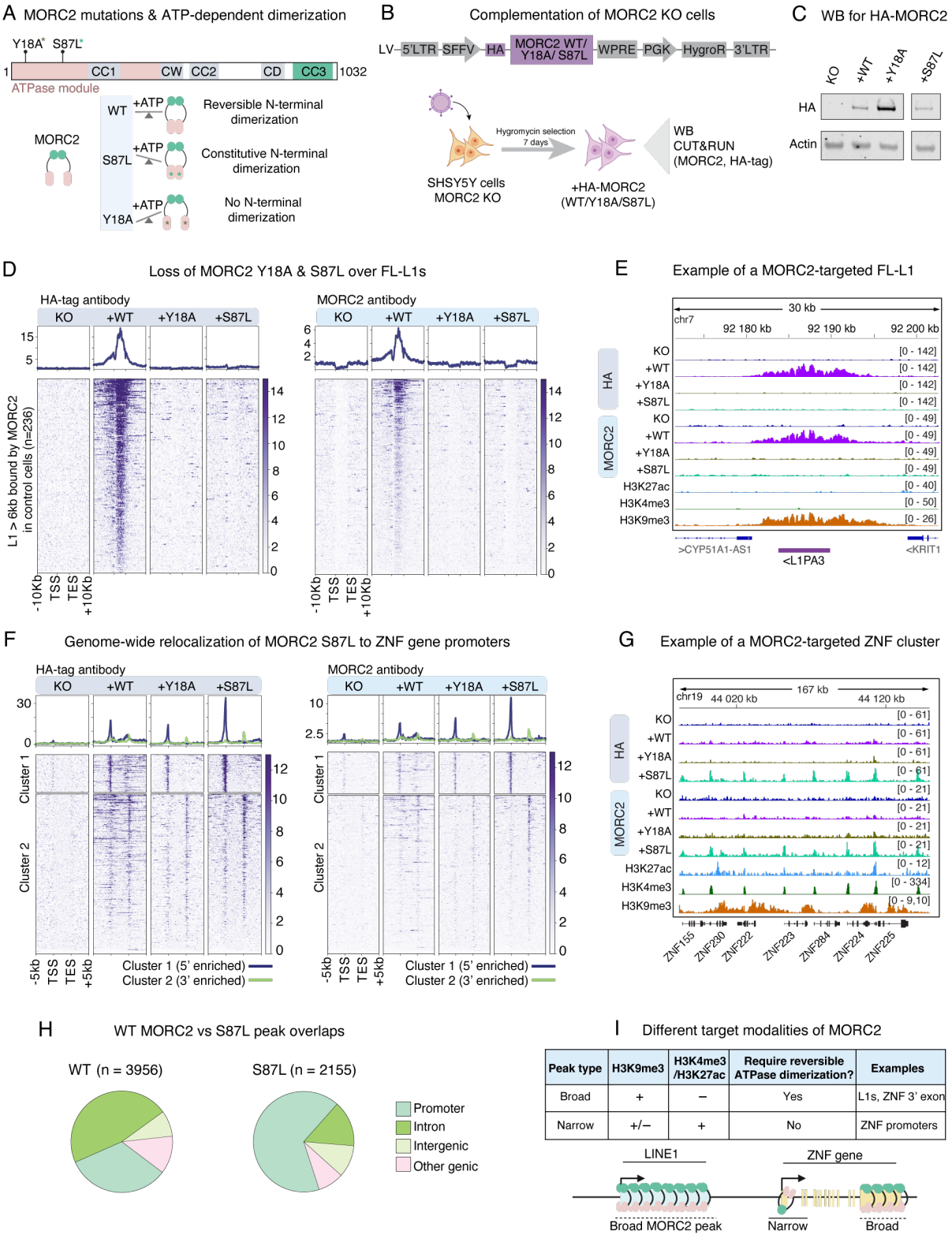
132 **Figure 1. MORC2 binds to distinct chromatin states in neural cell lines. (A)** MORC2 domain structure (top) and
133 schematic of ATP-dependent N-terminal dimerization (bottom). **(B)** CUT&RUN workflow for epigenome profiling
134 of SHSY5Y and hNPC cell lines. **(C)** Heatmap illustrating the epigenomic context of distinct MORC2 targets by
135 CUT&RUN signal in SHSY5Y (left) and NPC (right) cell lines with indicated antibodies. Displayed are the genomic
136 regions spanning ± 10 kb from the center. K-means clustering defined H3K9me3-negative, MORC2-narrow peaks
137 (top) and H3K9me3-positive, MORC2-broad peaks (bottom) in each cell type. **(D)** Genome browser tracks
138 showing CUT&RUN enrichment profiles of MORC2 and histone modification over examples MORC2 targets with

139 distinct epigenomic associations. Data are representative of at least two replicates for MORC2, H3K9me3 and
140 H3K4me3 in each line, while H3K27Ac profiling was conducted once in SHSY5Y cells and twice in hNPCs.

141

142 **ATPase dimerization mutants abolish L1 binding and disrupt the balance of MORC2**
143 **chromatin occupancy**

144 Mutations associated with MORC2 neurodevelopmental disorders all cluster to the N-
145 terminal ATPase module.[46] To interrogate how reversible ATP-dependent dimerization
146 impacts MORC2 chromatin binding we made use of two mutants that we previously found
147 shift this equilibrium in opposite directions: Y18A, a structure-guided mutation that prevents
148 ATPase dimer formation, and S87L, a patient mutation that leads to constitutive ATPase
149 dimerization (**Figure 2A**).[37] We reintroduced HA-tagged proteins to clonal MORC2-KO SHSY5Y
150 cells (**Figure 2B**) and first confirmed that WT, Y18A and S87L MORC2 variants were expressed
151 and nuclear as assessed by Western blotting (**Figure 2C**) and imaging (**Supp Fig 2**). CUT&RUN
152 analysis revealed that both Y18A and S87L variants were completely lost over L1s bound by
153 the WT protein (**Figure 2D,E**). Importantly, these effects were reproducible using antibodies
154 raised against the MORC2 protein itself and its HA epitope tag. Both variants were lost from
155 3' exons of ZNF genes but remained bound to the promoters of those genes, suggesting that
156 ATPase dimerization is not required for promoter occupancy. Notably, the binding of S87L
157 MORC2 to ZNF promoters was reproducibly enhanced relative to WT, suggesting that this
158 mutation causes relocalization of the protein away from L1s and onto gene promoters (**Figure**
159 **2F,G**). Indeed, peak calling revealed a substantially greater overlap of S87L MORC2 with
160 promoters when compared to the WT distribution (**Figure 2H**). Taken together, these results
161 demonstrate that ATPase dimerization mutants disrupt the balance of MORC2 chromatin
162 occupancy. Reversible ATP-dependent dimerization is required for broad MORC2
163 accumulation over L1s but not promoters of target genes (**Figure 2I**). The S87L mutation
164 directs the protein away from L1s and other broad MORC2 targets and increases binding over
165 gene promoters.



Peak type	H3K9me3	H3K4me3 /H3K27ac	Require reversible ATPase dimerization?	Examples
Broad	+	–	Yes	L1s, ZNF 3' exon
Narrow	+/–	+	No	ZNF promoters

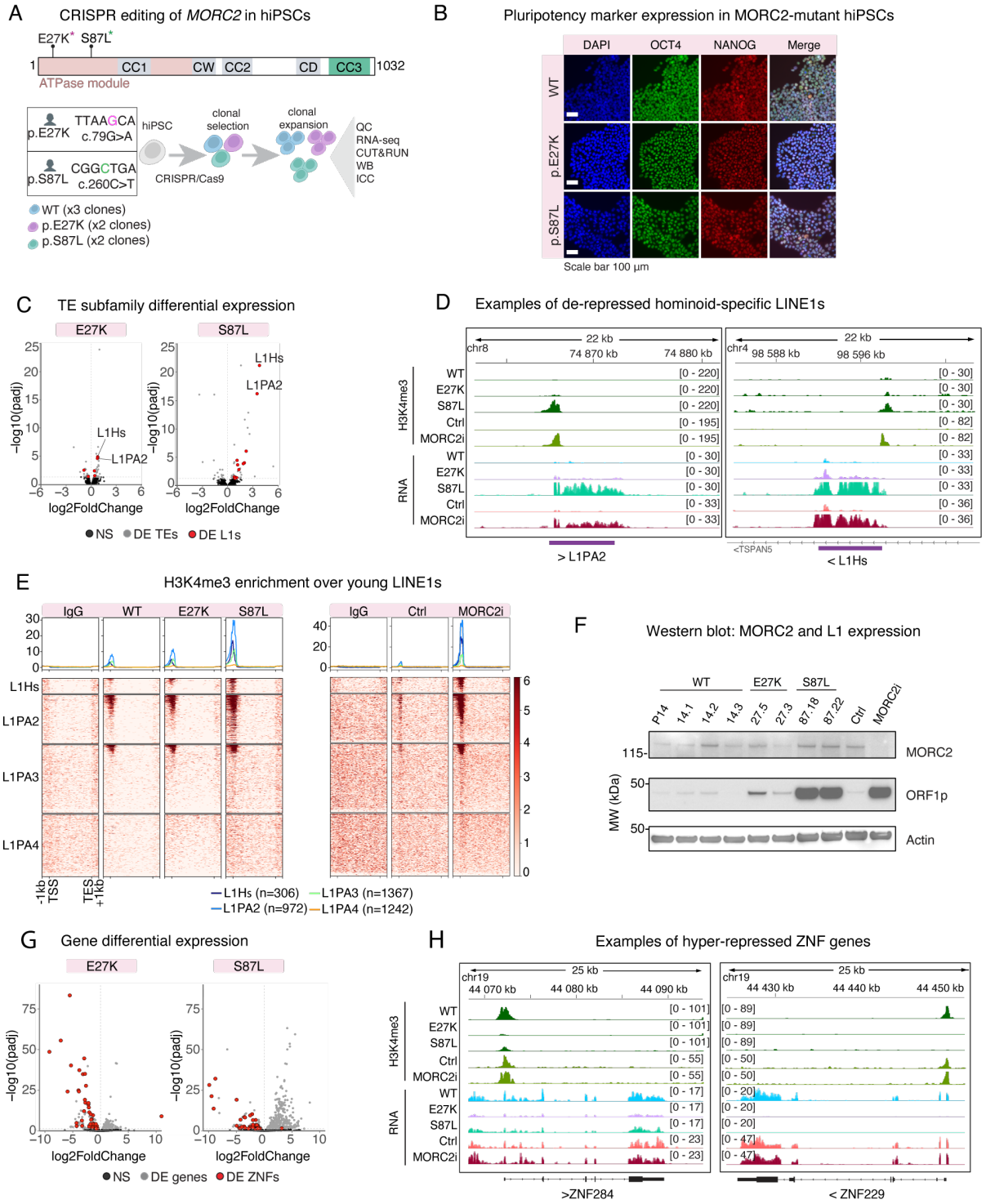
Figure 2. ATPase dimerization mutants abolish L1 binding and disrupt the balance of MORC2 chromatin occupancy. (A) Schematic representation of MORC2 protein domains and the position of ATPase disease mutations (top) and the effect of the mutations on the N-terminal dimerization equilibrium (bottom). **(B)** Schematic workflow of MORC2 complementation experiments and subsequent analyses. **(C)** Western blot validating the expression of complemented HA-tagged MORC2 variants. **(D)** Heatmap showing CUT&RUN enrichment of MORC2 binding over full length (FL, >6 kb) LINE1 elements bound by MORC2 WT in KO and complemented cell lines (+WT MORC2, +Y18A MORC2, +S87L MORC2) as detected using HA-tag antibody (left)

and MORC2 antibody (right). Shown are the LINE1 loci scaled to 6 kb with 10 kb up- and down-stream of the elements. Unique mapping was enforced and experiments were repeated at least once with similar results. **(E)** Genome browser tracks of CUT&RUN enrichment signal for MORC2 binding (KO and complemented cell lines) and histone modifications (H3K27ac, H3K4me3, H3K9me3) over an example LINE1 element. Unique mapping was enforced. **(F)** Heatmap showing CUT&RUN enrichment of MORC2 binding over all ZNF genes (n=503, hg38) in KO and complemented cell lines (+WT MORC2, +Y18A MORC2, +S87L MORC2) as detected using HA-tag antibody (left) and MORC2 antibody (right). Shown are the scaled ZNF genes ± 5 kb. K-means clustering on the S87L MORC2 sample identified 5'-enriched (top) and 3'-enriched (bottom) ZNF genes. **(G)** Genome browser tracks of CUT&RUN enrichment signal for MORC2 binding (KO and complemented cell lines) and histone modifications (H3K27ac, H3K4me3, H3K9me3) over a ZNF gene cluster on chr19. **(H)** Genome overlaps of WT and S87L HA-MORC2 peaks. **(I)** Schematic of the requirements for reversible N-terminal dimerization over different target classes of MORC2. MORC2 accumulation over broad peaks is found over regions enriched in H3K9me3 signal, such as LINE1s and ZNF 3´exons, and requires reversible N-terminal dimerization. Narrow MORC2 binding to gene regulatory elements such as ZNF gene promoters is found in regions enriched in H3K4me3 and H3K27ac signal, and does not require reversible N-terminal dimerization.

MORC2 ATPase mutations cause loss of L1 silencing and ZNF hyperrepression in iPSCs

Next we sought to explore the functional impact of chromatin binding on transcriptional regulation in a relevant developmental cell model. We used CRISPR editing to derive induced pluripotent stem cell (iPSC) clones bearing the S87L mutation described above and a second ATPase mutation, E27K, which also causes a severe neurodevelopmental phenotype and would be predicted to similarly disrupt ATPase activity based on its structural position at the N-terminal ATP-induced dimerization interface.[37,46] We note here that though patients carry heterozygous mutations, our study sought to provide mechanistic insight into how ATPase mutations affect MORC2 function over a developmental timecourse and as such we selected homozygous *MORC2* edited clones (**Figure 3A, Supp Fig 3A**). We subjected unedited controls to the same clonal expansion procedures and confirmed that each iPSC clone expressed pluripotency markers and that the copy number and genomic integrity at predicted off-targets was not affected by the editing (**Figure 3B, Supp Fig 3B,C**). Transcriptomic analysis using paired-end (150x150-bp) polyA+ RNA sequencing (RNA-seq) showed that the ATPase mutations both cause significant derepression of evolutionarily-young, hominoid-specific L1 subfamilies, with the S87L mutation having a more pronounced effect (**Figure 3C, D**). CUT&RUN profiling of the H3K4me3 mark mapped to individual full-length L1 elements provided orthogonal validation of young L1 expression **(Figure 3E)**. In line with chromatin binding data above illustrating complete loss of L1 binding, the S87L mutation caused L1 transcriptional activation at comparable levels to a loss of function control in which we used a validated CRISPRi vector [33] to robustly knock down MORC2 expression in iPSCs (**Figure 3E, Supp. Fig 3D-G**). L1 derepression was further confirmed in Western blots and

212 immunofluorescence experiments as measured by L1-encoded protein orf1p in the mutant
213 and CRISPRi conditions relative to controls (**Figure 3F, Supp Fig 3H**). We also observed
214 pronounced derepression of L1 fusion genes where evolutionarily-young L1s act as alternative
215 promoters (**Supp Fig 4A**). Several of these genes have documented or predicted functions in
216 the brain. For example, *GABRR1*, whose promoter is shared with a L1PA2 element, encodes a
217 key factor involved in GABAergic neurotransmission in the central nervous system[49] and
218 *DYNC1I1*, which contains an intronic L1PA3, encodes a neuron-enriched dynein intermediate
219 chain (**Supp Fig 4B**). Notably, and in line with the chromatin binding data, alongside L1 loss of
220 function we observed simultaneous hyperrepression of zinc finger genes in E27K (n=22), S87L
221 (n=11) or both (n=19) MORC2 mutants (log2 fold-change < -2, padj < 0.05, DESeq2) (**Figure**
222 **3G,H**). This effect was a gain of function associated with the S87L and E27K mutants that was
223 not observed in the MORC2-CRISPRi cells. Taken together these data show how MORC2
224 ATPase mutations change the dynamics of chromatin target engagement, leading to both loss
225 and gain of function effects on transcriptional control of its distinct targets.



226

227 **Figure 3. MORC2 ATPase mutations cause loss of L1 silencing and ZNF hyperrepression in iPSCs. (A)** Schematic
228 of CRISPR editing and subsequent analyses. Induced homozygous mutations were c79G>A and c260C>T for E27K
229 and S87L mutants, respectively. **(B)** Immunocytochemestry of pluripotency markers (NANOG, OCT4, SOX2) for a
230 representative hiPSC clone of each MORC2 genotype. **(C)** Volcano plot showing the results of the differential
231 expression analysis (DESeq2) on a TE subfamily level using TEtranscripts. Differentially expressed LINE1
232 subfamilies are highlighted in red, other differentially expressed subfamilies are presented in grey
233 (log2FoldChange $\neq 0$, padj <0.05). Two experimental replicates per clone were sequenced and data from each
234 clone pooled prior to differential analysis (n=4 WT vs n=4 E27K vs n=4 S87L). **(D)** Genome browser tracks showing
235 H3K4me3 and bulk RNA-seq data over two example loci of derepressed hominoid-specific full-length L1s. Unique

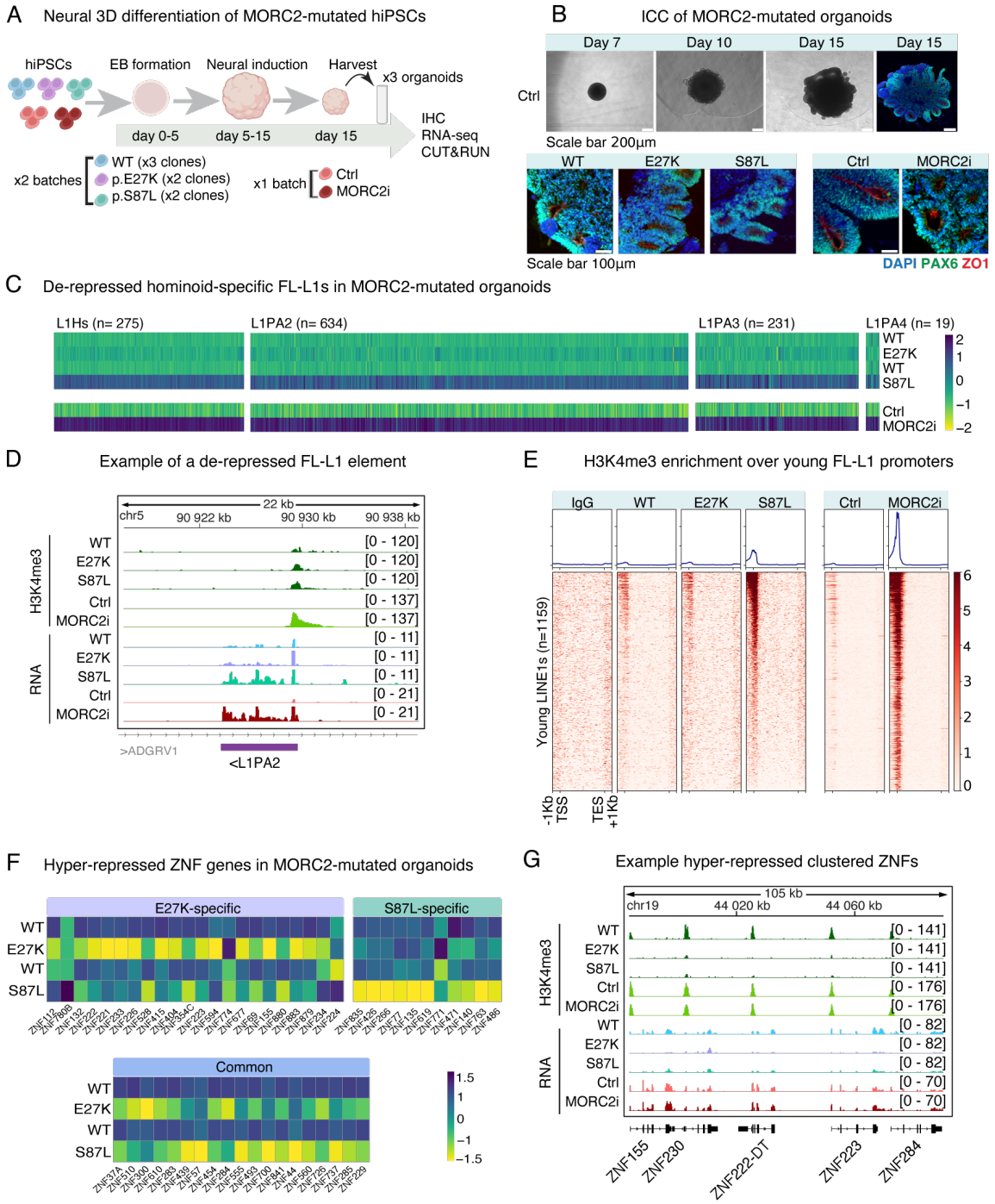
236 mapping was enforced for all data over individual loci. Also shown are representative data from MORC2 CRISPRi
237 (n=3) and Control (n=3) hiPSCs, see also Supp Fig 3. **(E)** Heatmap of the CUT&RUN signal enrichment of H3K4me3
238 and non-targeting IgG controls over reference L1 insertions belonging to the evolutionarily-youngest subfamilies
239 (L1HS, L1PA2-4) in MORC2 mutant and CRISPRi cell lines. For each WT and mutant cell line, cells from two clones
240 were pooled prior to CUT&RUN, which was conducted at least twice with similar results. Experiments with
241 control and MORC2 CRISPRi cells were repeated twice with similar results. Unique mapping was enforced for
242 CUT&RUN data shown. **(F)** Western Blot showing expression of MORC2 and of L1 ORF1p in MORC2 mutant and
243 CRISPRi cell lines. The experiment was repeated once with similar results. **(G)** Volcano plot showing the results
244 of the differential gene expression analysis (DESeq2). Differentially expressed ZNF genes are highlighted in red,
245 other differentially expressed genes are presented in grey (padj <0.05). Two experimental replicates per clone
246 were sequenced and data from each clone pooled prior to differential analysis (n=4 WT vs n=4 E27K vs n=4 S87L).
247 **(H)** Genome browser tracks showing H3K4me3 and bulk RNA-seq data over example loci of two hyperrepressed
248 ZNF genes. Unique mapping was enforced for all data over individual loci. Also shown are data from MORC2
249 CRISPRi (n=3) and Control (n=3) hiPSCs, see also Supp Fig 3.

250

251 **Transcriptional phenotypes of MORC2 mislocalization persist upon neural differentiation**

252 To model how these transcriptional phenotypes change upon exit from pluripotency and
253 neural differentiation, we set out to differentiate the iPSC lines to neural progenitor cells
254 (NPCs). Using a previously-described 2D NPC differentiation protocol[50] we observed
255 widespread cell death in the MORC2-CRISPRi and S87L lines after pluripotency exit, precluding
256 further analysis. As an alternative we made use of an undirected 15-day cerebral organoid
257 protocol as a model for human brain development in 3D (**Figure 4A**). At this timepoint our
258 previous data showed that organoids are largely composed of NPCs, enabling detailed
259 analysis of L1 and gene regulation in bulk.[28] 3D differentiation improved cell survival and cells
260 bearing both MORC2 mutations or carrying the MORC2 CRISPRi lentiviral vector could form
261 characteristic neural rosettes after 15 days, albeit with morphological and size changes
262 (**Figure 4B**). Single nucleus RNA sequencing confirmed that NPCs and early neurons
263 constituted the major cell populations in these preparations across multiple batches and
264 clones (**Supp Fig 4A-D**). Differential expression analysis on day-15 organoids from at least two
265 clones and three experimental batches showed that the transcriptional phenotype observed
266 in hiPSCs – that is, L1 derepression in mutant and CRISPRi lines and ZNF hyperrepression
267 specifically in the mutants – persisted upon differentiation (**Figure 4C-G**). We found that the
268 promoters of young L1 integrants identified to be derepressed in hiPSCs were highly enriched
269 with the H3K4me3 active mark in MORC2-S87L and MORC2-CRISPRi day 15 organoids (**Figure
270 4F**). We also observed upregulation of lineage-specific clustered protocadherin genes in
271 MORC2 mutant organoids, while this effect was stronger in MORC2-CRISPRi samples (**Supp**

272 **Fig 5A,B**). Taken together, our data illustrate that the transcriptional phenotypes of MORC2
273 targets in the pluripotent state persist upon exit from pluripotency and early
274 neurodevelopment.



275

276 **Figure 4. Transcriptional phenotypes of MORC2 mislocalization persist upon neural differentiation. (A)**
277 Schematic of differentiation of MORC2 mutant and CRISPRi cell lines from hiPSCs to cerebral organoids. Two
278 batches of organoids were produced for both S87L and E27K clones and three for WT clones. One batch was
279 produced for MORC2 CRISPRi and its corresponding control cell line in which hiPSCs were transduced with a non-
280 targeting lentiviral CRISPRi vector. **(B)** Representative bright field images and immunohistochemistry of cerebral
281 organoid formation in a control line at different time points (top). Immunohistochemistry of PAX6 (green) and
282 ZO1 (red) in day 15 organoid sections (bottom). **(C)** Scaled heatmap showing the normalized read counts of

individual full length (FL) LINE1 from youngest subfamilies (L1HS, L1PA2-4) that had been derepressed in MORC2 E27K and S87L mutant and MORC2 CRISPRi hiPSCs (log2FoldChange >0, padj <0.05). Data are representative of three replicates of two clone across two batches (n=9 WT vs n=9 E27K vs n=9 S87L). Unique mapping was enforced. **(D)** Genome browser tracks showing CUT&RUN H3K4me3 enrichment and bulk RNA-seq data over an example derepressed full-length L1. Unique mapping was enforced. **(E)** Heatmap of the CUT&RUN signal enrichment of H3K4me3 and non-targeting IgG control over reference full-length L1 insertions belonging to the to the youngest subfamilies (L1HS, L1PA2-4) that were expressed in MORC2 mutants and CRISPRi hiPSCs. For WT and mutant genotypes, organoids from two clones were pooled prior to CUT&RUN, which was conducted at least twice with similar results. Unique mapping was enforced **(F)** Scaled heatmap showing the normalized read counts over ZNF genes hyperrepressed in MORC2 E27K and S87L mutants hiPSCs (log2FoldChange <0, padj <0.05). **(G)** Genome browser tracks showing CUT&RUN H3K4me3 enrichment and bulk RNA-seq data over a hyper-repressed ZNF cluster on chr19. Unique mapping was enforced.

MORC2 directs CpG methylation patterning over the repetitive human genome

As the observed transcriptional phenotypes persisted from the pluripotent state to day-15 organoids, we hypothesized a defect in CpG methylation patterning over MORC2 targets. To investigate methylation changes over MORC2-targeted repetitive elements we used whole-genome Oxford Nanopore long-read sequencing (**Figure 5A**). While genome-wide methylation was robust and stable across all conditions (**Figure 5B**), the S87L mutation in MORC2 caused a catastrophic methylation loss over the promoter of the youngest L1Hs and L1PA2 integrants (**Figure 5C-E**). Loss of methylation over young L1 promoters was mirrored in the MORC2-CRISPRi organoids and across organoid preparations derived from independent clones, confirming the effect over these targets was reproducible and a direct loss of function effect (**Figure 5C-E, Supp Fig 6A**). Leveraging the Nanopore data to call polymorphic L1 insertions absent from the hg38 reference, we also observed striking loss of methylation over these youngest copies in both genetic backgrounds tested (**Supp Fig 6B**). Importantly, we found that the effect was most striking to the L1 elements transcribed in the pluripotent state, illustrating a transcription dependence of the methylation patterning directed by MORC2 (**Figure 5F**). We also observed hypermethylation over zinc finger gene clusters that were specific to the MORC2 mutant but not CRISPRi organoids – in particular the ZNF genes hyperrepressed in pluripotent cells **(Figure 5G,H).** As major alterations of the methylation are present across the ZNF gene clusters (**Supp Fig 6C**), including gene bodies, we speculate that this could be driven by effects on 3D chromatin structure in MORC2 mutant cells, though this remains to be tested systematically. Together our observations demonstrate the critical role of MORC2 in directing transcription dependent CpG methylation patterning over its repetitive targets in early development.

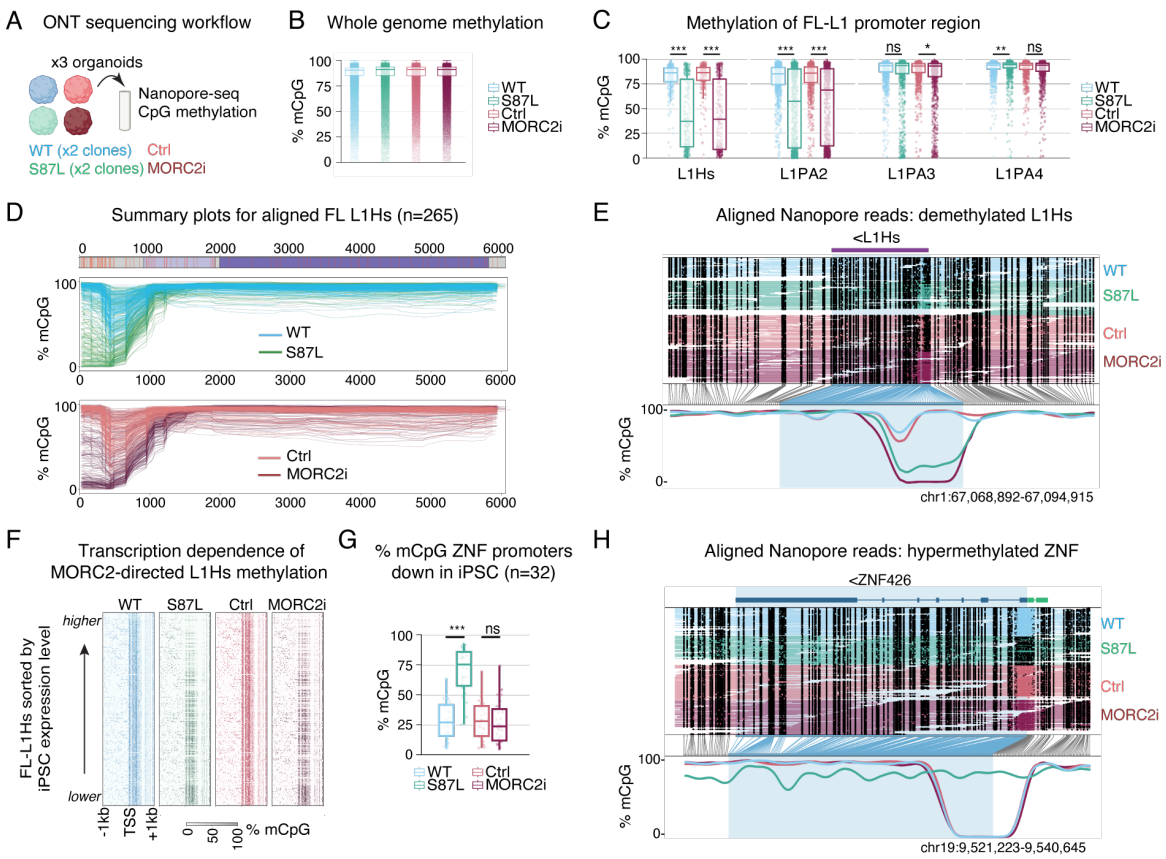


Figure 5. MORC2 directs CpG methylation patterning over the repetitive human genome. (A) Schematic of DNA methylation analysis workflow. For each clone and cell line, DNA from three organoids was extracted and pooled for Oxford Nanopore sequencing (ONT). **(B)** Boxplots showing the CpG methylation levels across the whole genome and **(C)** over the first 900-bp of full length (FL) evolutionary young LINE1 elements (L1HS, L1PA2-4), which corresponds to the 5' UTR promoter region. P-value as calculated by a two-sided Wilcoxon rank sum test with continuity correction (p-values adjusted following Benjamini–Hochberg correction). **(D)** Composite plot showing CpG methylation levels of all reference FL-L1HS elements greater than 6 kb (hg38) over the entire element in the given conditions. **(E)** Aligned Nanopore reads showing loss of CpG methylation over the promoter of an example L1HS locus in S87L and MORC2i organoids. Methylated and non-methylated CpGs are represented by filled black circles and open (coloured) circles, respectively. **(F)** Heatmap showing the CpG methylation level over the promoter of FL-L1HS, sorted by expression level in hiPSC S87L mutants as assessed by promoter H3K4me3 level. Methylation changes are anti-correlated to their expression in hiPSCs. **(G)** CpG methylation level of the promoter of ZNF genes downregulated in MORC2-mutant iPSCs. **(H)** Aligned Nanopore reads showing hypermethylation over the promoter and general methylation alteration over the body of a hyperrepressed ZNF gene. ns, non-significant; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$.

L1 derepression upon exit of pluripotency causes loss of NPC proliferation

Finally we tested the impact of young L1 derepression on NPC differentiation. We returned to a directed 2D NPC protocol where both MORC2-CRISPRi and S87L mutant cells – i.e. those lines with high L1 expression – had undergone cell death in differentiation trials. Leveraging our recently-developed L1 CRISPRi lentiviral vector (**Fig. 6A,B**), which efficiently and specifically silences expression from full-length L1Hs, L1PA2 and L1PA3 subfamilies,[28] we

342 asked whether re-silencing L1s on the MORC2 S87L genetic background could rescue this
343 apparent loss of cellular fitness. We generated six iPSC lines where WT or S87L clones were
344 transduced with the L1 CRISPRi vector or a non-targeting control (**Fig. 6B**). We confirmed high
345 young L1 expression in S87L clones by Western blot analysis of L1-orf1p, which was robustly
346 lost upon L1 CRISPRi (**Fig. 6C**). Upon dual-SMAD inhibition to direct neural induction, we found
347 that all lines exited pluripotency as assessed by RT-qPCR analysis of marker genes (**Supp Fig**
348 **8A**) and characteristic morphological changes of cellular polarization between days 4-7 of
349 differentiation (**Fig. 6D, Supp Fig 8B**).[50] By day 14, surviving cells from all lines expressed PAX6
350 and FOXG1 according to RT-qPCR and immunocytochemistry analysis (**Supp Fig 8A,B**).
351 However, between day 9-14 of differentiation we observed striking cell death and 6-16 fold
352 lower final cell yields in S87L clones transduced with the control CRISPRi vector, but not in
353 those same clones transduced with the L1 CRISPRi vector, an effect that was reproducible
354 across multiple blinded differentiation batches (**Fig 6E,F**). Thus, excessive L1 expression may
355 be incompatible with rapid proliferation and/or early maturation of NPCs leading to loss of
356 fitness, at least in this protocol. Taken together, our data indicate that MORC2-directed CpG
357 methylation safeguards the somatic human genome against deleterious L1 expression upon
358 exit of pluripotency and early neurodevelopment.

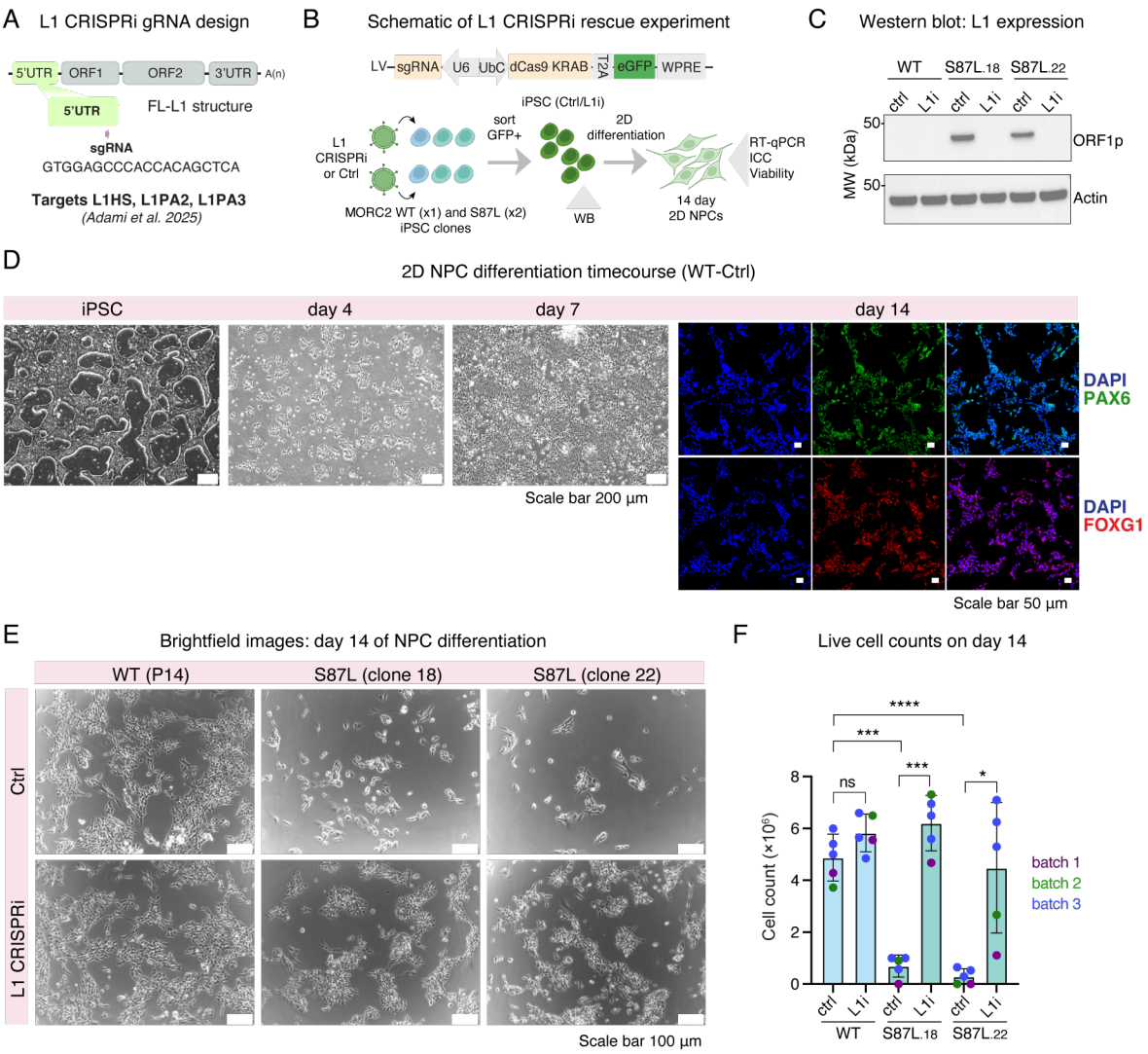
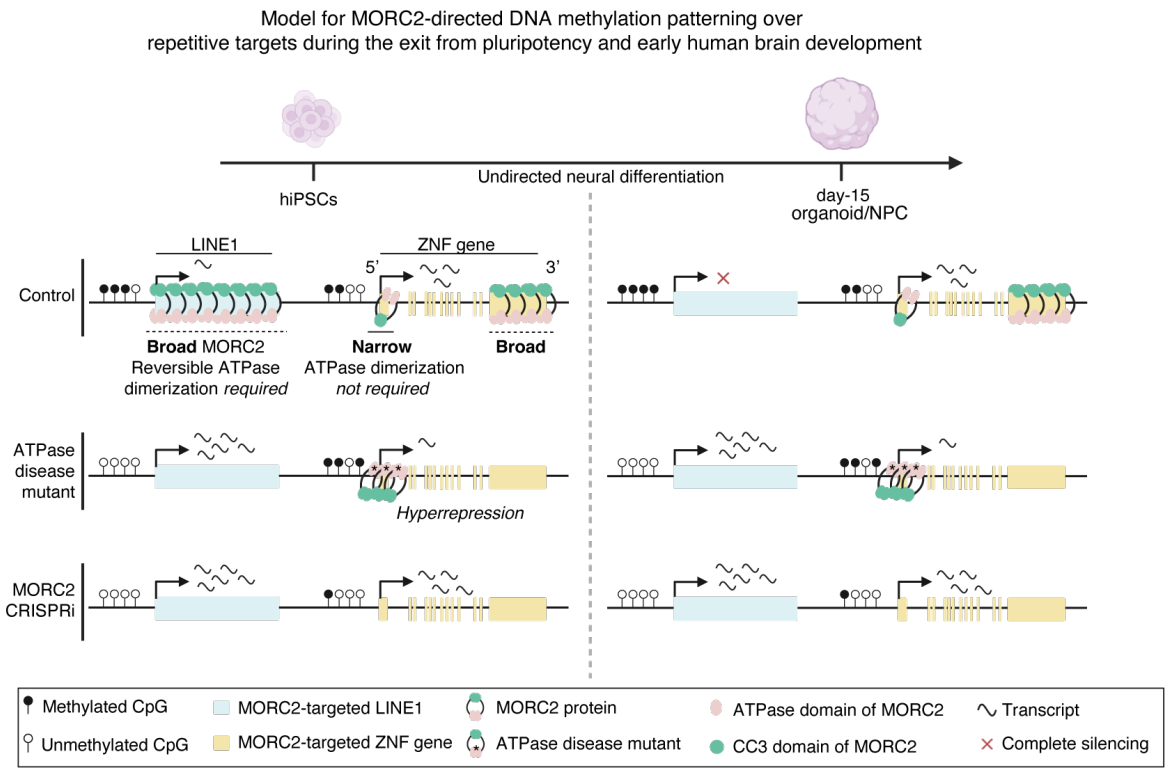


Figure 6. L1 derepression upon exit of pluripotency causes loss of NPC proliferation. (A) Schematic of L1 CRISPRi gRNA design to ectopically silence young full-length L1s. For full validation see Adami et al.[28] **(B)** Experimental outline for L1 CRISPRi experiment in WT and S87L MORC2 iPSC lines, and subsequent 2D NPC differentiation. **(C)** Validation of young L1 silencing by Western blot of orf1p in the indicated iPSC lines. Experiment was repeated twice with similar results. **(D)** Representative brightfield (left) and ICC (right) images of WT Ctrl line during 2D NPC differentiation at the indicated timepoints. See also Supp Fig 8. **(E)** Representative brightfield images of the six indicated lines at the end of the 14-day 2D NPC differentiation protocol. **(F)** Quantification of live cell counts at the end of NPC differentiations from five replicates across three experimental batches (cells seeded on different days). Batch 3 was done in triplicate (cells seeded independently in different wells) to account for in-batch variation. Cell counts given are the mean value of two counts from each tube of cells. Bars indicate the mean (n=5) and error bars indicate the standard deviation. In batch 2, a small subset of cells were removed from all conditions at day 11 for subsequent immunocytochemistry. Experimenters were blinded to the identity of the cell lines during differentiation experiments. Significance was assessed using paired t-tests between indicated conditions. *, $p < 0.05$; ***, $p < 0.001$; ****, $p < 0.0001$; ns, not significant.

DISCUSSION.

In this study we report and relate two key mechanistic insights on the role of the MORC2 ATPase in epigenetic reprogramming of the repetitive genome. First, we find that reversible

378 ATP-dependent dimerization is required for broad MORC2 accumulation over L1
379 retrotransposons, but not required for its binding to gene promoters, and that mutation of
380 the ATPase module titrates MORC2 occupancy between these target classes. Second, we find
381 that MORC2 chromatin binding directs CpG methylation patterns over its targets upon exit of
382 pluripotency and early human neurodevelopment, which protects against the otherwise
383 deleterious effects of unrestricted L1 expression. The model is summarized in **Figure 7**.



384

385 **Figure 7.** Model for MORC2-directed CpG methylation patterning over repetitive targets during the exit from
386 pluripotency and early human brain development.

387

388 The physiological relevance of MORC2 ATP-dependent dimerization is evident from clinical
389 genetics reports over the past ten years, where missense mutations in patients with severe
390 neurodevelopmental disorders all map to the ATPase module.[45–47] Our previous biochemical
391 and structural studies identified that these mutations change the ATPase activity and
392 dimerization dynamics of the purified protein,[37] and here we provide a framework to explain
393 how this translates at the level of chromatin binding. MORC2 has two main chromatin binding
394 modes – broad accumulation over H3K9me3-marked sites, and narrow binding to gene
395 regulatory elements marked by H3K4me3/H3K27Ac.[32] When reversible ATP-dependent

396 dimerization is perturbed by mutation of the ATPase module, MORC2 fails to accumulate over
397 broad sites but binding is maintained or enhanced at narrow sites. Crucially, mutations that
398 shift the equilibrium of N-terminal dimerization in either direction cause loss of function over
399 L1s – supporting the notion that the lifetime of ATP-dependent MORC dimers is a critical
400 parameter in driving chromatinization of certain targets.[36,37,39] The observation that binding
401 of MORC2 mutants to gene promoters is maintained or enhanced explains the hyper-
402 repression phenotype observed in engineered cell lines or patient samples,[32,47] and in the
403 stem cell models of early brain development that we characterize here. That many of the
404 genes hyperrepressed in MORC2 mutant cells encode KRAB-ZNFs, themselves associated with
405 epigenetic TE silencing, highlights the central and multifaceted role of MORC2 in TE biology.
406 Many of the hyperrepressed ZNFs in MORC2 mutant lines are transcription factors or
407 repressors that each target hundreds of genomic locations, illustrating enormous potential
408 for pleiotropic, indirect effects on chromatin and transcription of disrupting the balance of
409 MORC2 chromatin binding sites via ATPase mutation. These complex gene regulatory
410 networks require further investigation, but our data suggest that TE transcription and TE-
411 induced chromatin structure are likely important molecular phenotypes to consider in MORC2
412 disorders.

413 Work by us and others has established that the HUSH-MORC2 corepressor targets transcribed
414 and relatively hypomethylated repeats with little apparent preference for particular DNA
415 sequence motifs.[8,33,51,52] The prevailing model is that the complex is recruited to long,
416 intronless units with some A-rich bias in the sense strand.[51] In human pluripotent stem cells,
417 the primary targets of MORC2 are the hominoid-specific LINE-1 retrotransposons that we
418 recently reported to be transcribed and relatively hypomethylated at steady-state.[28] This is
419 likely the critical way that the HUSH-MORC2 axis preserves genome stability in early
420 development: a subset of the youngest L1s remain competent for autonomous transposition
421 and their encoded proteins provide the molecular apparatus for mobilization of other TEs.
422 Transcription-dependent TE targeting is in contrast to the KRAB-ZNF/TRIM28 pathway, which
423 is recruited to particular DNA sequence motifs in repeats via rapidly-evolving ZNF domains.[22–
424 26] Such complementary modes of targeting help explain why HUSH-MORC2 and KRAB-ZNFs
425 may regulate overlapping yet distinct sets of genomic repeats and genes in different cellular
426 contexts.[30,35] The situation at the youngest human L1s is of particular interest: while

427 numerous KRAB-ZNFs (and TRIM28) are competent to bind L1HS and L1PA2 subfamilies
428 through conserved sequences in the L1 ORFs,[24,25] deletion of 129 bp in the ancestral L1PA3
429 promoter region within the ZNF93 binding site may have evaded promoter methylation by
430 this factor in evolutionarily-younger copies.[26,27] Our data supports the idea that the HUSH-
431 MORC2 corepressor is recruited to these youngest, transcriptionally-active elements for
432 which a new KRAB-ZNF has yet to evolve.[51]

433 The interplay of these epigenetic silencing systems is of particular interest in early
434 development, when methylation patterns of the germline are globally erased and re-
435 established.[12] The KRAB-ZNF/TRIM28 system is thought to protect against erasure of germline
436 methylation at certain targets: for example, ZFP57 and ZFP445 bind DNA sequences at
437 imprinting control regions, recruiting TRIM28 and DNMTs to maintain methylation at paternal
438 or maternal imprints.[53,54] Many TEs are also resistant to demethylation, but it is thought the
439 youngest human L1s are demethylated early post-fertilization, maintaining relative
440 hypomethylation throughout pre-implantation development.[27] Although our primed iPSC
441 models reflect a post-implantation developmental stage, we note that *MORC2* is also robustly
442 expressed in the pre-implantation human embryo[55] and our data suggest that the protein is
443 a key factor that directs methylation over these youngest L1 copies. Whether MORC2 protects
444 against erasure of germline methylation or promotes re-establishment of human L1
445 methylation following erasure is challenging to answer definitively, but the transcription-
446 dependence of HUSH-MORC2 recruitment to L1s strongly favours a reestablishment model.
447 Nevertheless, the underlying mechanism and timing of human L1 promoter methylation
448 warrants further investigation, for example leveraging new blastoid models.[56]

449 Taken together, our study reveals how ATP-dependent dimerization of MORC2 is coupled to
450 chromatin binding and DNA methylation patterning. The precise sequence of how
451 establishment and/or maintenance DNA methyltransferases (DNMTs) are recruited to
452 MORC2 targets in early development remains an interesting open question. MORC2-driven
453 chromatin compaction is known to correlate with H3K9me3 deposition at certain sites,[8,32,33]
454 which may directly recruit maintenance methyltransferase DNMT1.[57,58] However, since
455 aberrant MORC2 mutant binding to certain (H3K9me3-negative) gene promoters also results
456 in promoter hypermethylation, we suggest that MORC2 may be directly involved in licensing
457 DNA methylation its target sites. This is somewhat reminiscent of a recent study in mouse

458 gonocytes where Morc1, a germline-restricted TE repressor,[59] was suggested to make distinct
459 complexes leading to piRNA-directed H3K9me3 and DNA methylation at different TEs.[60] Our
460 data reinforce the model that the transcriptional activity of a locus determines its sensitivity
461 to MORC2 binding and, we show here, targeted CpG methylation upon differentiation. Such
462 a model could explain how the transcription factor yin-yang 1 (YY1) functions as both
463 activator[61] and repressor[31] of human L1s: YY1 drive transcription initiation of these young
464 elements in the pluripotent state, which marks the locus for MORC2-directed methylation. It
465 is also noteworthy that certain young L1s are demethylated and transcriptionally-active in
466 both organoids and human brain tissue.[3,28,31] Further work is required to determine how
467 MORC2-directed methylation may be avoided over these elements – or whether the elements
468 are subject to active demethylation upon neural differentiation. How the mechanisms we
469 describe here play out in terminally differentiated cells such as neurons also promises to be a
470 fruitful area of research.

471 **METHODS.**

472 **Cell culture.**

473 *hiPSCs.* Two human iPSC lines were used in this study: Ctrl-14 hiPSCs, which were
474 reprogrammed by the Cell and Gene Therapy Core at Lund Stem Cell Center from human
475 fibroblasts via mRNA mediated reprogramming as described elsewhere,[62] and a
476 commercially-available line from RIKEN (RBRC-HPS0328 606A1). hiPSCs were cultured based
477 on previously developed protocols.[50,63] Briefly, cells were maintained in multi-well plates
478 coated with Biolaminin 521 (0.7 mg cm^{-2}), and fed daily with StemMACS iPS-Brew XF
479 containing 0.5% penicillin–streptomycin (Gibco). hiPSCs were passaged upon reaching 70-
480 80% confluency. Firstly, cells were rinsed with DPBS (Gibco), dissociated using Accutase
481 (Gibco) for 5-7 min at 37°C then washed off by DMEM-F12 (Gibco) and KnockOut Serum
482 Replacement (KSR) (Gibco) and centrifuged at 400g for 5 min. Rock inhibitor (10 μM, Miltenyi)
483 was added to the media following cell thawing, lentiviral transduction, FACS and CUT&RUN
484 experiments.

485 *SH-SY5Y cells.* SH-SY5Y human neuroblastoma cells (ATCC CRL-2266) were a gift from Harvey
486 McMahon (MRC-LMB). Cells were cultured in flasks and fed every 3-4 days with complete
487 culture medium containing DMEM-F12, 10% FBS, MEM non-essential amino acids solution
488 and 1x penicillin/streptomycin (Gibco). When 80% confluent, cells were rinsed with DPBS

489 (Gibco), incubated with trypsin (Gibco) for 5-7 minutes at 37°C to detach, and washed off with
490 the media to quench the enzyme. To maintain a Mycoplasma-free culture, cells were routinely
491 tested and returned negative.
492 *Unguided cerebral organoid differentiation*. Human cerebral brain organoids were generated
493 using the STEMdiff Cerebral organoid kit (Stem Cell Technologies) following the
494 manufacturer's instructions and based on a previously established protocol.[64] First, embryoid
495 bodies (EB) were generated following disassociation and resuspension of hiPSCs in EB seeding
496 medium (STEMdiff™ Cerebral Organoids Basal Medium 1 and STEMdiff™ Cerebral Organoids
497 Supplement A, supplemented with 10 μM of Y-27632) in round bottom ultra-low attachment
498 96-well plates. On day 5, to avoid EB fusion individual EBs were distributed into single wells
499 of ultra-low attachment 24-well plates and incubated with Induction Medium (STEMdiff™
500 Cerebral Organoids Basal Medium 1 and STEMdiff™ Cerebral Organoids Supplement B) for
501 48 h at 37°C. On day 7, organoids were then embedded into Matrigel (Stem Sell Technologies)
502 and cultured in ultra-low attachment 6-well plates (10-12 organoids/well) with Expansion
503 medium (STEMdiff™ Cerebral Organoids Basal Medium 2, STEMdiff™ Cerebral Organoids
504 Supplement C, STEMdiff™ Cerebral Organoids Supplement D). Up until day 15, organoids
505 were fed twice (on day 10 &13) with Maturation media (STEMdiff™ Cerebral Organoids Basal
506 Medium 2, STEMdiff™ Cerebral Organoids Supplement E). Organoids were harvested on day
507 15 for downstream analysis such as bulk RNA-seq (3 snap-frozen organoids per replicate),
508 snRNA-seq (4-5 organoids per replicate), whole genome Oxford Nanopore sequencing (3-4
509 organoids per replicate), immunostaining (3-4 organoids) and CUT&RUN (3-4 organoids per
510 replicate).
511 *2D NPC differentiation*. Transduced iPSC cell lines vectors were differentiated to neural
512 progenitor cells (NPCs) using a 2D-protocol based on dual inhibition of the SMAD-signalling
513 with Noggin and SB431542 from the start of the differentiation as described elsewhere.[50] A
514 two-day coating protocol was followed. On the first day, 6-well plate was coated with 100 µL
515 of poly-L-ornithine (PLO, Sigma-Aldrich) diluted in 10 mL of PBS +/+ (1.2 mL per well) and
516 incubated at 37℃ overnight. The following day, PLO solution was aspirated and the wells were
517 washed three times with PBS +/+. PBS +/+ was aspirated and the plate was coated with 50 µL
518 of Laminin-23017-015 (Fisher Scientific) diluted in 10 mL of PBS +/+ (1.2 mL per well). The
519 plate was wrapped in parafilm and placed in the fridge to incubate at 4℃ overnight. Before
520 seeding the cells, Laminin-23017-015 was aspirated, and the wells were washed once with

521 PBS +/+. hiPSC lines were counted and 90,000 cells were transferred to a tube, pelleted then
522 resuspended in N2 medium supplemented with SB431542, Noggin and Rock Inhibitor, before
523 being seeded into the wells. On days 2, 4 and 7 of the differentiation, N2 media supplemented
524 with SB431542 and Noggin was supplied to the cells. On the 9th and 10th day of
525 differentiation N2 media without supplements was used. On day 11 cells were passaged and
526 replated in B27 media (Table 3) supplemented with brain-derived neurotrophic factor (BDNF,
527 R&D systems), ascorbic acid (AA, Sigma-Aldrich) and Revita (Thermo Fisher Scientific). B27
528 media supplemented with BDNF and AA was supplied to the cells every day until the day of
529 the harvest.
530
531 **Lentiviral vectors.**
532 Lentiviral vectors were produced according to a standard protocol described previously.[65]
533 Briefly, HEK293T cells with a confluency of ~80% were co-transfected with three packaging
534 plasmids (pMDL, psRev, and pMD2G) and the relevant transfer plasmid using
535 polyethyleneimine (PEI; Polysciences), in DPBS. Cell culture media containing lentiviral
536 particles was harvested 48 h after transfection, filtered, and centrifuged at 25,000xg for 1.5h
537 at 4 °C. After removing the supernatant, pellets were gently resuspended in cold PBS and
538 incubated overnight at 4°C. Finally, the lentivirus was aliquoted and stored at -80 °C until use.
539 Lentiviral titration was determined using qPCR targeting the viral WPRE sequence and
540 comparison to a reference virus batch with known titre.
541
542 **CRISPR approaches.**
543 *CRISPR KO.* To generate MORC2 KO SHSY5Y cells two sgRNA were cloned into pSpCas9(BB)-
544 2A-Puro vector (Addgene #48139). sg21 (GGGATCAATATAGAGCACAG) targets exon 9 and
545 sg22 (GCGGACATTAGAAGTACGCCT) exon 11 of *MORC2*. 8 mg of the plasmids was
546 electroporated into $2x10^6$ cells using 100ul tips from Neon electroporation system and placed
547 into 25 cm flasks. After 24 h, 4 mg/mL of puromycin was added to select for electroporated
548 cells. After 10 days single cells were sorted into 96-well plates. Chosen clones were first
549 analysed for the presence of the MORC2 band using Western blotting. Cells were lysed for 30
550 min at 4°C in buffer containing 300 mM NaCl, 100 mM Tris pH 8.0, 0.2 mM EDTA, 0.1% NP40,
551 10% glycerol supplemented with protease inhibitors. Lysates were cleared by centrifugation
552 and supernatant was used for Western blotting using MORC2 (anti-rabbit MORC2, Bethyl),

553 and actin antibodies (Mouse α-β-actin, ab8226). DyLight-680 or 800-conjugated secondary
554 antibodies (Thermo Fisher Scientific) were used at 1:10,000 dilution and blots were imaged
555 on the Odyssey near-infrared system (LI-COR). Following isolation of genomic DNA (Qiagen),
556 the knockout was confirmed using Sanger sequencing and TIDE analysis.
557 *CRISPR-editing.* To introduce MORC2 S87L and E27K mutations in Ctrl-14 hiPSCs, HiFi Cas9v3,
558 relevant gRNAs and desired sequences to target the MORC2 gene were delivered to cells as
559 Ribonucleoprotein (RNP) and single-stranded oligodeoxynucleotide (ssODN) via
560 nucleofection (P3 Primary Cell 4D-Nucleofector X Kit, Lonza) (**Table 1**). CRISPR editing
561 efficiency was estimated in bulk using genomic PCR and restriction enzyme digestion followed
562 by clonal expansion. Individual clones were collected and primarily tested for the presence of
563 the edit using genomic PCR and restriction enzyme digestion followed by Sanger sequencing.
564 The selected clones with the desired edits underwent karyotyping and sequencing of the 5
565 top-predicted off-targets. Once these quality controls were passed, immunostaining of
566 pluripotency markers OCT4, NANOG and SOX2 was performed on selected clones.
567

Genotype	gRNA	ssODN	hiPSC clone numbers
WT	-	-	P14, 14.1, 14.2, 14.3
MORC2 E27K	ACCAAACAAGAATTCGTGAG	TTATCAACCAGTTCAGCAAGGGCACCAA ACAAGAATTTGTGAGTGGTTCTGAAATG ATAAATACTAAAAGTCAGCAAAAG	27.3, 27.5
MORC2 S87L	CTTCCCAAACTGGATCACAC	GACTCAGGTGTTCGCTTGGCCAACTTCC CAAACTGGATCACACTAGCAGCATCACC TGAAAGGGCAGA CACAAGAGAAAG	87.18, 87.22

568 **Table 1. CRISPR editing of *MORC2*.**
569
570 *CRISPR interference*. A previously-described protocol for CRISPR interference was used to
571 silence *MORC2* or young L1 transcription in hiPSCs.[66] gRNAs (**Table 2**) were cloned into a
572 dCas9-KRAB-T2A-GFP lentiviral backbone containing both the gRNA under the U6 promoter
573 and dCas9-KRAB-T2A-GFP under the Ubiquitin C promoter (pLV hU6-sgRNA hUbC-dCas9-
574 KRAB-T2a-GFP) using annealed oligos and the BsmBI cloning site. An identical CRISPRi vector
575 containing a gRNA sequence absent from the human genome (LacZ) was used as a non-
576 targeting control. hiPSCs were transduced with MOI 5 and GFP-positive cells were sorted
577 using FACS (FACSAria, BD sciences) on day 10 at 10°C. Sorted cells were centrifuged at 400x*g*
578 for 7 min and pellets were snap frozen on dry ice and kept at −80°C for downstream analysis,

579 or maintained in culture. MORC2 CRISPRi was validated in Pandiloski et al.[33]. Young LINE-1
580 silencing was validated in Adami et al.[28]
581

CRISPRi sgRNA target	20-bp targeting sequence (5'-3')
LacZ non-targeting control	TGCGAATACGCCCACGCGAT
MORC2	GCTTCCAAGGACCGGATCGA
LINE-1	GTGGAGCCCACCACAGCTCA

582 **Table 2. CRISPRi gRNA sequences.**
583

584 **RT-qPCR.**

585 RT-qPCR was used to check the efficiency of MORC2 CRISPRi. 500 ng RNA was extracted from
586 hiPSCs and used for reverse transcription to generate cDNA with random hexamer primers
587 and the Superscript III reverse transcriptase (Invitrogen). RT-qPCR was performed using SYBR
588 Green I master (Roche) on a LightCycler 480 instrument (Roche). Data were analysed with the
589 ΔΔCt method normalized to housekeeping gene *ACTB*. Primer sequences are given below.

Target	Forward (5'-3')	Reverse (5'-3')
MORC2	ACATGAAGACGCAGGAAGAG	ACTTCCAAGGGCAATTTCTT
Beta-Actin	CCTTGCACATGCCGGAG	GCACAGAGCCTCGCCTT
NANOG	TTGGGACTGGTGGAAGAATC	GATTTGTGGGCCTGAAGAAA
OCT4	TCTCCAGGTTGCCTCTCACT	GTGGAGGAAGCTGACAACAA
FOXG1	TGGCCCATGTCGCCCTTCCT	GCCGACGTGGTGCCGTTGTA
PAX6	TGGTATTCTCTCCCCCTCCT	TAAGGATGTTGAACGGGCAG

590 **Table 3. RT-qPCR primer sequences.**
591

592 **Lentiviral expression of HA-MORC2 variants.**

593 To express WT and mutant versions of MORC2 on the SHSY5Y MORC2-KO background, we
594 cloned MORC2 into the lentiviral pHRSIN-pSFFV-WPRE-pPGK-Hygro vector (a gift from Paul
595 Lehner, University of Cambridge) with an N-terminal HA-tag. For lentiviral production 293T
596 cells were co-transfected with lentiviral MORC2 vector, pCMVΔR8.91 and pMD.G using the
597 TransIT-293 transfection reagent (Mirus). Supernatants were collected two days post-
598 transfection and cleared by filtering through 0.45 mm membranes. $1x10^6$ of cells were
599 transduced by centrifugation for 90 min at 1000xg in the presence of 300 μL lentivirus-
600 containing supernatant and 5 μg/mL of polybrene. The following day, cells were transferred
601 into 10cm dishes and hygromycin at 400 μg/mL added for selection.

602

603 **Quantitative genotyping qPCR.**

604 Quantitative genotyping PCR (qgPCR) was performed to rule out unintended CRISPR-induced
605 on-target effects (onTEs) at the edited sites. We followed the protocol described in Weisheit
606 et al., 2021.[67] In brief, genomic DNA was extracted from iPSC cells using the DNeasy Blood &
607 Tissue kit (QIAGEN). IDTPrimerQuest was used to design primer pairs to amplify the regions
608 surrounding the edits and FAM-labeled target-specific probes and purchased from IDT (see
609 **Table 4**). We use the TERT Taqman Copy Number Reference assay (Thermo Fisher Scientific)
610 labeled with VIC as a control assay. The qgPCR reaction mix was prepared using the Prime-
611 Time Gene Expression Master Mix (IDT). Each sample was run in technical triplicates on a
612 LightCycler 480 instrument (Roche). Protocol guidelines were followed to perform the Ct
613 values analysis and total allele number calculation.

Primers	Sequence (5'-3')
MORC2_E27K-gqPCR-F	AGTTGGGGAGCTCTGGAAAG
MORC2_E27K-gqPCR-F	AGGGAAGTGGTTTTGGAGGT
MORC2_E27K-FAM-probe	56-FAM/TTGTTTGGT/ZEN/GCCCTTGCT
MORC2_S87L-gqPCR-F	ACCCTGGGAAAGAAATGCAAA
MORC2_S87L -gqPCR-F	GCTACTCGAGAAACCACCACT
MORC2_ S87L -FAM-probe	56-FAM/TGGGCAGTA/ZEN/CGGGAATGG

614 **Table 4. Primers for quantitative genotyping qPCR.**

615

616 **Bulk RNA sequencing and analysis.**

617 RNA was extracted from cells and organoids using the RNeasy mini and QIAshredder kits
618 (Qiagen) respectively. Sequencing libraries were prepared using the Illumina TruSeq Stranded
619 mRNA Library prep kit optimized for long fragments and subsequently sequenced on a
620 NovoSeq X Plus platform with paired end reads (2x150bp) sequencing.

621 *Genes and repetitive element subfamilies quantification.* The generated reads were mapped
622 to the human reference genome hg38 (GRCh38) with STAR aligner (v 2.6.0)[68] with modified
623 parameters (--outFilterMultimapNmax 100 and --winAnchorMultimapNmax 200). BAM files
624 were sorted using SAMtools (v 1.16.1).[69] For visualisation, BigWig files were created from the
625 BAM files using bamCoverage (v 2.5.4) applying RPKM normalization. Gene quantification was
626 conducted with Subread featureCounts (v1.6.3; flags -p and -s 2 for paired-end reads and
627 strand-specific reads, respectively) using the reference annotation GENCODE version 38.

628 Repetitive elements quantification at a subfamily level was conducted with TEtranscripts (v
629 2.2.3) using default options and using a curated annotation file.[70]
630 *Individual repetitive elements quantification.* To obtain an accurate quantification of
631 individual repetitive elements, alignment of the reads to the reference genome with STAR
632 aligner was conducted to allow only for uniquely mapped reads. To this end, the reads were
633 allowed to map only to a single locus (--outFilterMultimapNmax 1) and the allowed
634 mismatched ratio was decreased (--outFilterMismatchNoverLmax 0.03) compared to default
635 settings. After sorting the obtained BAM files with SAMtools, the quantification was
636 conducted with featureCounts (v1.6.3) using the curated annotation file provided within the
637 TEtranscripts package. Generation of the BigWig files was conducted as described above.
638 *Differential expression analysis.* Differential expression analysis was conducted with DESeq2
639 (v 1.40.2)[71] using as input the count matrices obtained from featureCounts and TEtranscripts.
640 Raw reads were normalized with DESeq2's median-of-ratios normalization method and fold
641 changes were shrunk using DESeq2's lfcShrink. For creating the heatmaps, the mean of
642 replicates for each condition + 0.5 was used and further visualized with pheatmap.
643
644 **Nuclear isolation**.
645 Nuclei extraction was performed from day-15 organoids according to previously described
646 protocols.[72,73] Briefly, organoids were dissociated in cold lysis buffer (0.32 M sucrose, 5 mM
647 $CaCl_2$, 3 mM MgAc, 0.1 mM Na_2EDTA, 10 mM Tris-HCl, pH 8.0, 1 mM DTT) using a 1ml tissue
648 douncer (Wheaton). Isolation was sucrose gradient-based, thus the lysate was layered
649 carefully on top of a sucrose cusion (1.8 M sucrose, 3 mM MgAc, 10 mM Tris-HCl pH 8.0, and
650 1 mM DTT diluted in milliQ water) and then centrifuged for 2h and 15 minutes at 30,000xg 4
651 ^{o}C. Supernatant was thoroughly aspirated, pellet was soften for 10 minutes by adding nuclear
652 storage buffer (15% sucrose, 10 mM Tris-HCl pH 7.2, 70 mM KCl, and 2 mM $MgCl_2$, all diluted
653 in milliQ water), resuspended in dilution buffer (10 mM Tris-HCl pH 7.2, 70 mM KCl, and 2 mM
654 $MgCl_2$, diluted in milliQ water) and filtered using a 70um cell strainer. Finally, nuclei were
655 sorted using FACS (Aria, BD Biosciences) at low flow rate and 4° C.
656
657 **Single-nuclei RNA-sequencing and analysis.**
658 Approximately 8,500 to 10,000 nuclei per sample were prepared for single-nucleus RNA
659 sequencing and loaded onto the Chromium Next GEM Chip G using the Single Cell Kit, along

660 with the reverse transcription mastermix, following the manufacturer's protocol for the
661 Chromium Next GEM Single Cell 3' Kit (10X Genomics, PN-1000268). Single-cell gel beads in
662 emulsion (GEMs) were generated during this process. cDNA was then amplified according to
663 guidelines from 10X Genomics, using 13 PCR cycles for the 3' libraries. Sequencing libraries
664 were constructed with unique dual indices (TT set A), pooled, and sequenced on a NovaSeq X
665 Plus system using a 100-cycle kit with 28-10-10-90 read lengths. Cell Ranger count (v 6.0.0)
666 was run with default settings,[74] with an mRNA and pre-mRNA reference for single-cell samples
667 (generated using 10x Genomics Cell Ranger 6.0.0 guidelines) for single-nucleus samples.
668 Samples clustering was analyzed using Seurat (version 5.1.0).[75] Each sample was filtered
669 according to two criteria, I) nuclei with percentage of mitochondrial content over 2%
670 (perc_mitochondrial) and II) nuclei with number of features (n_Feature) lower than 600 and
671 greater than 2 standard deviations were filtered out. The counts were normalized using
672 Seurat's LogNormalize (NormalizeData) and clusters were defined with a resolution of 0.1
673 (FindClusters). Standard Seurat functions (e.g. DimPlot and DotPlot) were used for all
674 downstream visualization of the data.
675
676 **Oxford Nanopore sequencing and analysis.**
677 For each sample, high molecular weight DNA was extracted from pooled organoids (n = 3)
678 using the Nanobind HMW DNA extraction kit (PacBio) according to manufacturer´s
679 instructions. Extracted DNA was then depleted of short DNA fragments (< 10 kb) and sheared
680 to 10 kb using the Short Reads Eliminator (SRE) XS kit (PacBio) and the g-TUBE (Covaris),
681 respectively. The sequencing libraries for LacZ control and MORC2 CRISPRi samples were
682 prepared using the Ligation Sequencing Kit V14 (SQK-LSK114 - ONT), while for the WT control
683 (clones P14 and 14.3) and S87L MORC2 mutant (clones 87.18 and 87.22) the library was
684 prepared, and the clones were barcoded and using the Native Barcoding Kit 24 V14 (SQK-
685 NBD114.24 - ONT). All libraries were sequenced on FLO-PRO114M flow cells on a PromethION
686 device (ONT) for 72 h. Basecalling and alignment to the human reference genome hg38 was
687 conducted with Dorado (v0.7.1) using the DNA model dna_r10.4.1_e8.2_400bps_sup@v4.3.0
688 and the DNA modification model dna_r10.4.1_e8.2_400bps_sup@v4.3.0_5mCG_5hmCG@v1
689 for the detection of CpG methylation. Resulting BAM files were sorted with SAMtools (v1.18).
690 Methylartist (v1.3.0) was used for the subsequent analysis of the methylation data visualised
691 through Methylartist and R (v4.4.2).[76] Modkit (v0.3.3) and BedGraphToBigWig (v4) were used

692 to produce bedGraph from the BAM files and BigWigs from the bedGraph files, respectively.
693 The BigWig files were then used for visualisation of the methylation with computeMatrix and
694 plotHeatmap (DeepTools v2.5.4).
695
696 **CUT&RUN.**
697 Two different CUT&RUN approaches were followed. For profiling histone modifications
698 (H3K4me3, H3K9me3, H3K27Ac), a standard CUT&RUN protocol was used. To profile MORC2
699 chromatin binding, a crosslinked protocol was used.
700 *Standard CUT&RUN.* The CUT&RUN protocol used here has previously been described.[77,78]
701 Briefly, 500k cells were harvested as described above (see cell culture: hiPSCs or SHSY5Y) and
702 washed three times using Wash buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl, 0.5
703 mM spermidine, 1x Roche cOmplete protease inhibitors. Cells were then incubated and
704 attached to 10 uL pre-activated ConA-coated magnetic beads (Bangs Laboratories) on an end-
705 to-end rotator for 7-10 min. To activate ConA-coated magnetic beads, they were washed
706 twice with binding buffer containing 20 mM HEPES pH 7.9, 10 mM KCl, 1 mM $CaCl_2$, 1 mM
707 $MnCl_2$. Bead-bound cells then were incubated with primary antibody solution (20 mM HEPES
708 pH 7.5, 0.15 M NaCl, 0.5 mM Spermidine, 1x Roche complete protease inhibitors, 0.05% w/v
709 digitonin, 2 mM EDTA and 1:100 primary antibody) overnight at 4 ^{o}C. On the second day,
710 bead-bound cells were washed twice using digitonin buffer (20 mM HEPES pH 7.5, 150 mM
711 NaCl, 0.5 mM Spermidine, 1x Roche cOmplete protease inhibitors, 0.05% digitonin) then
712 incubated with pAG-MNase enzyme resuspended in digitonin buffer at 4 ^{o}C for 1 h with a
713 gentle mixing. Bead-bound cells were then washed twice with digitonin buffer, resuspended
714 in 100μL digitonin buffer, chilled to 0 ^{o}C for 5 min then incubated for 30 min at 0 ^{o}C with 2 mM
715 CaCl2 to activate genome cleavage. Finally, the reaction was quenched by adding 100 mL 2x
716 stop buffer (0.35 M NaCl, 20 mM EDTA, 4mM EGTA, 0.05% digitonin, 50 ng/mL glycogen, 50
717 ng/mL RNase A) and gentle vortexing followed by a 30 min incubation at 37 ^{o}C to release the
718 genomic fragments. The supernatant was collected and purified using PCR clean-up spin
719 columns (Macherey-Nagel).
720 *Day-15 organoid CUT&RUN.* 3-6 organoids per cell line was used. Organoids were collected
721 on day-15 of differentiation, washed with PBS, and incubated 3 minutes with Accutase on a
722 shaker at 37 ^{o}C to dissociate cells. They were then gently resuspended in Accutase and
723 incubated on the shaker for another 4 minutes followed by a 5-minute centrifugation at 400xg.

724 Finally, the cell suspension was filtered to remove Matrigel. The standard *CUT&RUN* protocol
725 was followed as described above.
726 *Crosslinked CUT&RUN.* The approach was same as above with addition of the followings: a
727 cross-linking step, an increase in cell number and a slightly higher centrifugation force. 900K-
728 1M cells were collected by centrifugation at 600xg for 3 minutes and crosslinked for 1 min
729 using 1% formaldehyde (in cell culture media, diluted from a 16% solution, Thermo Scientific).
730 Cross-linked cells were then washed three times with wash buffer containing 20 mM HEPES
731 pH 7.5, 150 mM NaCl, 0.5 mM spermidine, 1x Roche complete protease inhibitors, 1% Triton
732 X-100, 0.05% SDS. To avoid loss of the crosslinked cell/nuclear pellet during washes, a slight
733 increase in centrifugation speed (1000-1300xg) was applied. Cells were incubated with ConA
734 beads for 7-10 minutes as above and then resuspended in primary antibody buffer (20 mM
735 HEPES pH 7.5, 150 mM NaCl, 0.5 mM spermidine, 1x Roche complete protease inhibitors, 1%
736 Triton X-100, 0.05% SDS, 0.05% digitonin, 2 mM EDTA). After an overnight incubation at 4 ^{o}C
737 with the primary antibody, bead-bound cells were thoroughly washed two times with
738 digitonin buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 1x Roche complete
739 protease inhibitors, 1% Triton X-100, 0.05% SDS, 0.05% digitonin). Following the 37 ^{o}C
740 incubation to release genomic fragments, 0.09% SDS and 0.22 mg/ml proteinase K (Thermo
741 Scientific) was added to supernatant to reverse the crosslinking reaction. Finally, DNA was
742 purified using PCR clean-up spin column (Macherey-Nagel).
743
744 **CUT&RUN sequencing and analysis.**
745 Hyperprep kit (KAPA) with unique dual indexed adaptors were used to prepare sequencing
746 libraries which were then pooled and sequenced on NovaSeqX plus machine
747 (Illumina). Paired-end reads (2×75) were aligned to the human genome (GRCh38) using
748 bowtie2 (v 2.4.5) (–local –very-sensitive-local –no-mixed –no-unal –no-discordant –phred33
749 -I 10 -X 700),[79] converted to BAM files and further sorted with SAMtools (v 1.16.1). Peaks were
750 called using HOMER (v 4.10). Tag directories were created from the BAM files using
751 makeTagDirectory, and peaks were called with findPeaks (-style histone -minDist 1000 and -
752 size 300). When available, MORC2 CRISPR KO samples were used as control for peak calling
753 within the findPeaks command (-i). For the NPC lines, where the MORC2 CRISPR KO was not
754 available, the IgG sample was used as control, and MORC2 peaks with signal in the IgG
755 samples were removed. Peaks were filtered based on number of tags (> 10) and on fold-

756 change compared to control (> 8).Reads per kilobase per million mapped reads (RPKM)-
757 normalized BigWig coverage tracks were made with bamCoverage (DeepTools, v 2.5.4).
758 Heatmaps matrices were created by computeMatrix (DeepTools v2.5.4) using called peaks as
759 the input regions and RPKM normalized bigwigs as input signal. Matrices were then visualized
760 using plotHeatmap (DeepTools, v2.5.4), adjusting the upstream and downstream visualization
761 parameters (-a and -b) and splitting the signal in heatmaps (--kmeans 2).
762
763 **Immunostaining**.
764 *hiPSCs immunostaining.* Cells cultured on 24-well ibidi plates were rinsed three times with
765 DPBS and fixed in 4% formaldehyde for 15 min at room temperature. Before blocking with
766 TKPBS (KPBS with 0.25% Triton X-100) and 5% Normal Donkey Serum (NDS), cells were
767 washed another three times with DPBS. After 1h blocking, they were incubated overnight at
768 4 °C with primary antibody or TKPBS + 5% NDS for negative controls. The next day cells were
769 rinsed in TKPBS twice for 5 minutes each and once in TKBPS/ 5% NDS. They were then
770 incubated with secondary antibody solution (TKPBS + 5% NDS) at room temperature for 2 h
771 followed by a 5-minute incubation with DAPI (1:1000, Sigma D817) diluted in TKPBS. After
772 two washes with KPBS, cells were imaged using a Zeiss 780 confocal microscope.
773 *Day-15 organoid immunostaining.* Organoids harvested at day 15 of differentiation were
774 gently rinsed once with DPBS (Gibco) in 24-well plates then fixed with 4% paraformaldehyde
775 (PFA) for 90-120 minutes at room temperature. After fixation, organoids were washed three
776 times with KPBS for 10 minutes each time and incubated overnight with 30% sucrose at 4 °C
777 overnight on gentle shake. The following day, organoids were immersed in OCT (catalog no.
778 45830, HistoLab), snap frozen on dry ice and stored in -20°C until use. Before staining,
779 organoids were sliced into 20-μm sections using a cryostat (Thermo Scientific Cryostar NX70)
780 at -20°C and placed onto Superfrost Plus microscope slides. Borders of each slide were
781 defined using a hydrophobic PAP Pen before starting the immunostaining. After two washes
782 with KPBS, slides were incubated with blocking solution (TKPBS & 5% NDS) for 1 h at room
783 temperature. They were then incubated with primary antibody solution (relevant primary
784 antibody diluted in TKPBS & 5% NDS) overnight at 4 °C. The following day slides were washed
785 three times with KPBS and incubated with secondary antibody solution for 1 h at room
786 temperature. Finally, after three more washes, slides were incubated with DAPI, rinsed off

787 with KPBS, mounted using FluorSave mounting medium and imaged with a Zeiss 780 Confocal
788 microscope. Further image adjustments were performed in batch on Image J.
789
790 **STORM.**
791 Cells were seeded on glass 8-well Ibidi dishes and next day fixed with 4% formaldehyde in PBS
792 for 15 min. Membranes were permeabilised with 0.1% Triton-X100 in PBS for 5 min. For
793 staining, wells were first blocked with 10% FBS in PBS then incubated with primary anti-HA
794 antibody at 1/500 followed by secondary Alexa Fluor 647 at 1/1000. After washing samples
795 were fixed again for 5 min, washed and kept in PBS at 4°C till the imaging. STORM data was
796 acquired using Nikon N-STORM inverted microscope with 647 nm laser line, Apochromat TIRF
797 100×/1.49 NA oil immersion objective and single-photon detection iXon Ultra DU897 EMCCD
798 camera (Andor). Samples were washed with PBS and then STORM buffer (7mM NaOH, 100
799 mM mercaptoethylamine (MEA) in PBS) was added for the imaging. Samples were illuminated
800 with the 647 nm laser at 2–3 $kW \cdot cm^{-2}$ and 100 000 frames were collected for each field of
801 view at a frame rate of 50–70 Hz. Data was processed in NIS elements Advanced Research
802 with N-STORM module. Further image processing and cropping was performed using ImageJ.
803
804 **Western blot.**
805 *MORC2 KO SHSY5Y cells and complementation.* Cells were lysed for 30 min at 4°C in buffer:
806 300 mM NaCl, 100 mM Tris pH 8.0, 0.2 mM EDTA, 0.1% NP40, 10% glycerol supplemented
807 with protease inhibitors. Lysates were cleared by centrifugation and supernatant used for
808 Western blotting, probed with indicated antibodies and scanned with a near-infrared system
809 (LI-COR Odyssey) after incubation with DyLight 680- or 800-conjugated secondary antibodies
810 (Thermo Fisher Scientific, 1:10,000 dilution).
811 *Assessment of MORC2 and L1 expression in iPSCs.* Cell lysis was performed using either RIPA
812 (Sigma-Aldrich) containing complete protease inhibitor cocktail or 1:100 benzonase in 1% SDS
813 on ice for 30 minutes and pelleted for 20 minutes at 4 °C and 12,000x*g*. Before running on a
814 4–12% Tris-glycine SDS-PAGE gel (200 V, 40-45 min, RT), the supernatant was aspirated into
815 new tubes and mixed with loading dye (Novex LDS 4x) and reducing agent (Thermo) then
816 boiled for 10 minutes at 70 °C. To transfer proteins from gel to membrane, premade transfer
817 stacks (iBlot™ 2 NC Mini Stacks, Invitrogen™) were used in the iBlot™ 2 gel transfer device,
818 using the P0 preset template, according to manufacturer's instructions. The membrane was

819 briefly rinsed with TBS-T (Tris-buffered saline with 0.1% Tween) and blocked in 5% skimmed
820 milk diluted in TBS-T (MTBST) for 1h at RT on a shaker. Membrane was rinsed once in TBS-T
821 and incubated at 4 °C overnight with primary antibody diluted in MTBST. The following day
822 membrane was washed three times in TBS-T each time for 5-10 minutes with shaking. It was
823 then incubated with HRP-conjugated anti-mouse (Cell signalling) or HRP-conjugated anti-
824 rabbit secondary antibody (Santa Cruz Biotechnology, 1:5,000) diluted in MTBST. Before
825 protein detection, membrane was washed three times in TBS-T each time for 5-10 minutes
826 with shaking and once with TBS. Protein was detected by chemiluminescence using ECL Select
827 reagents (Cytiva) according to manufacturer's instructions and imaged on a Chemi-Doc
828 system (BioRad). Finally, β-actin staining was performed by stripping the membrane using the
829 Restore PLUS Western Blot Stripping Buffer (Thermo) as per instructions, re-blocking for 1 h
830 in MTBST, followed by incubation with HRP-472 conjugated anti-β-actin (1:50000 dilution
831 Sigma A3854). For uncropped blots, see **Supp Fig 9.**
832
833 **Antibodies.**
834 *CUT&RUN*: Goat anti-rabbit IgG (abcam ab97047), Rabbit anti-H3K4me3 (Active Motif 39159),
835 Rabbit anti-MORC2(A300-149A, Bethyl), Rabbit anti HA-Tag (C29F4) mAb 3724, Rabbit anti-
836 H3K9me3 (abcam 8898), Rabbit anti-H3K27ac (MABE647 clone RM172) (1:50-1:100 dilution).
837 *Immunostaining*: Rabbit anti-NANOG (1:600 dilution, Abcam, RRID: AB446437), Mouse anti-
838 OCT3/4 (1:300 dilution,Santa Cruz, RRID: AB628051), Rabbit anti-Sox2 1:500 (Merck AB5603,
839 RRID: AB2286686), DAPI (1:1000 dilution, Sigma-Aldrich), Mouse anti-ORF1p, clone 4H1
840 (1:300 dilution, Sigma-Aldrich, RRID: RRID:AB2941775), Rabbit anti-PAX6 (1:300, Biolegend,
841 901301), Mouse anti-ZO1 (1:300, Invitrogen, 339100), Mouse anti-HA11 (Biolegend, 901501),
842 Donkey anti-mouse Cy3 (1:500 dilution, Jackson Lab), Donkey anti-rabbit Alexa647 (1:500
843 dilution,Jackson Lab), Alexa-488 anti-mouse (Thermofisher, A21202, RRID AB141607), Alexa-
844 568 anti-rabbit (Thermofisher A10042, RRID AB2534017).
845 *Western Blot*: Rabbit anti-MORC2 (A300-149A, Bethyl), Mouse anti-HA (HA.11 Biolegend
846 901501, clone 16B12), Anti LINE-1 ORF1p clone 4H1 (Merck Millipore, MABC1152), HRP-472
847 conjugated anti-β-actin (1:50000 dilution Sigma A3854), HRP-conjugated anti-mouse IgG
848 (1:10000, Cell signalling 7076P2), HRP-conjugated anti-rabbit IgG (1:5000, Santa Cruz
849 Biotechnology), Mouse α-β-actin (abcam ab8226).

850 **DATA AVAILABILITY.**

851 There are no restrictions on data availability. The RNA and DNA sequencing data presented in
852 this study have been deposited at GEOs: GSEXXXXX.

853

854 **REFERENCES.**
855

856 1. Nurk, S. *et al.* The complete sequence of a human genome. *Science (1979)* **376**, (2022).

857 2. Ewing, A. D. *et al.* Nanopore Sequencing Enables Comprehensive Transposable Element
858 Epigenomic Profiling. *Mol Cell* (2020) doi:10.1016/j.molcel.2020.10.024.

859 3. Garza, R. *et al.* LINE-1 retrotransposons drive human neuronal transcriptome
860 complexity and functional diversification. *Sci Adv* **9**, (2023).

861 4. Criscione, S. W. *et al.* Genome-wide characterization of human L1 antisense promoter-
862 driven transcripts. *BMC Genomics* **17**, (2016).

863 5. Rangwala, S. H., Zhang, L. & Kazazian, H. H. Many LINE1 elements contribute to the
864 transcriptome of human somatic cells. *Genome Biol* **10**, (2009).

865 6. Beck, C. R. *et al.* LINE-1 retrotransposition activity in human genomes. *Cell* (2010)
866 doi:10.1016/j.cell.2010.05.021.

867 7. Brouha, B. *et al.* Hot L1s account for the bulk of retrotransposition in the human
868 population. *Proc Natl Acad Sci U S A* **100**, (2003).

869 8. Liu, N. *et al.* Selective silencing of euchromatic L1s revealed by genome-wide screens
870 for L1 regulators. *Nature* **553**, 228–232 (2018).

871 9. Ilık, İ. A. *et al.* Autonomous transposons tune their sequences to ensure somatic
872 suppression. *Nature* **626**, (2024).

873 10. Marasca, F. *et al.* LINE1 are spliced in non-canonical transcript variants to regulate T
874 cell quiescence and exhaustion. *Nat Genet* **54**, (2022).

875 11. Attig, J. *et al.* Heteromeric RNP Assembly at LINEs Controls Lineage-Specific RNA
876 Processing. *Cell* (2018) doi:10.1016/j.cell.2018.07.001.

877 12. Greenberg, M. V. C. & Bourc'his, D. The diverse roles of DNA methylation in mammalian
878 development and disease. *Nature Reviews Molecular Cell Biology* vol. 20
879 https://doi.org/10.1038/s41580-019-0159-6 (2019).

880 13. Jönsson, M. E. *et al.* Activation of neuronal genes via LINE-1 elements upon global DNA
881 demethylation in human neural progenitors. *Nat Commun* (2019) doi:10.1038/s41467-
882 019-11150-8.

883 14. Deniz, Ö., Frost, J. M. & Branco, M. R. Regulation of transposable elements by DNA
884 modifications. *Nature Reviews Genetics* vol. 20 https://doi.org/10.1038/s41576-019-
885 0106-6 (2019).

886 15. Walsh, C. P., Chaillet, J. R. & Bestor, T. H. Transcription of IAP endogenous retroviruses
887 is constrained by cytosine methylation. *Nature Genetics* vol. 20
888 https://doi.org/10.1038/2413 (1998).

889 16. Oomen, M. E. *et al.* An atlas of transcription initiation reveals regulatory principles of
890 gene and transposable element expression in early mammalian development. *Cell* **188**,
891 1156-1174.e20 (2025).

892 17. Grow, E. J. *et al.* Intrinsic retroviral reactivation in human preimplantation embryos and
893 pluripotent cells. *Nature* **522**, (2015).

894 18. Jachowicz, J. W. *et al.* LINE-1 activation after fertilization regulates global chromatin
895 accessibility in the early mouse embryo. *Nat Genet* **49**, (2017).

896 19. Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous
897 retrovirus activity. *Nature* **487**, (2012).

898 20. Percharde, M. *et al.* A LINE1-Nucleolin Partnership Regulates Early Development and
899 ESC Identity. *Cell* (2018) doi:10.1016/j.cell.2018.05.043.

900 21. Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D. & Zamore, P. D. PIWI-interacting
901 RNAs: small RNAs with big functions. *Nature Reviews Genetics* vol. 20
902 https://doi.org/10.1038/s41576-018-0073-3 (2019).

903 22. Rowe, H. M. *et al.* KAP1 controls endogenous retroviruses in embryonic stem cells.
904 *Nature* (2010) doi:10.1038/nature08674.

905 23. Turelli, P. *et al.* Interplay of TRIM28 and DNA methylation in controlling human
906 endogenous retroelements. *Genome Res* (2014) doi:10.1101/gr.172833.114.

907 24. Imbeault, M., Helleboid, P. Y. & Trono, D. KRAB zinc-finger proteins contribute to the
908 evolution of gene regulatory networks. *Nature* (2017) doi:10.1038/nature21683.

909 25. de Tribolet-Hardy, J. *et al.* Genetic features and genomic targets of human KRAB-zinc
910 finger proteins. *Genome Res* **33**, (2023).

911 26. Jacobs, F. M. J. *et al.* An evolutionary arms race between KRAB zinc-finger genes
912 ZNF91/93 and SVA/L1 retrotransposons. *Nature* (2014) doi:10.1038/nature13760.

913 27. Smith, Z. D. *et al.* DNA methylation dynamics of the human preimplantation embryo.
914 *Nature* **511**, (2014).

915 28. Adami, A. *et al.* LINE-1 retrotransposons regulate the exit of human pluripotency and
916 early brain development. *bioRxiv* 2025.01.17.633315 (2025)
917 doi:10.1101/2025.01.17.633315.

918 29. Klawitter, S. *et al.* Reprogramming triggers endogenous L1 and Alu retrotransposition
919 in human induced pluripotent stem cells. *Nat Commun* **7**, (2016).

920 30. Lanciano, S. *et al.* Locus-level L1 DNA methylation profiling reveals the epigenetic and
921 transcriptional interplay between L1s and their integration sites. *Cell Genomics* **4**,
922 (2024).

923 31. Sanchez-Luque, F. J. *et al.* LINE-1 Evasion of Epigenetic Repression in Humans. *Mol Cell*
924 (2019) doi:10.1016/j.molcel.2019.05.024.

925 32. Tchasovnikarova, I. A. *et al.* Hyperactivation of HUSH complex function by Charcot-
926 Marie-Tooth disease mutation in MORC2. *Nat Genet* **49**, 1035–1044 (2017).

927 33. Pandiloski, N. *et al.* DNA methylation governs the sensitivity of repeats to restriction by
928 the HUSH-MORC2 corepressor. *Nat Commun* **15**, 7534 (2024).

929 34. Hagelkruys, A. *et al.* The HUSH complex controls brain architecture and protocadherin
930 fidelity. *Sci Adv* **8**, (2022).

931 35. Robbez-Masson, L. *et al.* The hush complex cooperates with trim28 to repress young
932 retrotransposons and new genes. *Genome Res* (2018) doi:10.1101/gr.228171.117.

933 36. Fendler, N. L. *et al.* Identification and characterization of a human MORC2 DNA binding
934 region that is required for gene silencing. *Nucleic Acids Res* **53**, gkae1273 (2025).

935 37. Douse, C. H. *et al.* Neuropathic MORC2 mutations perturb GHKL ATPase dimerization
936 dynamics and epigenetic silencing by multiple structural mechanisms. *Nat Commun*
937 (2018) doi:10.1038/s41467-018-03045-x.

938 38. Xie, H. Y., Zhang, T. M., Hu, S. Y., Shao, Z. M. & Li, D. Q. Dimerization of MORC2 through
939 its C-terminal coiled-coil domain enhances chromatin dynamics and promotes DNA
940 repair. *Cell Communication and Signaling* **17**, (2019).

941 39. Kim, H. J. *et al.* The Gene-Silencing Protein MORC-1 Topologically Entraps DNA and
942 Forms Multimeric Assemblies to Cause DNA Compaction. *Mol Cell* (2019)
943 doi:10.1016/j.molcel.2019.07.032.

944 40. Harris, C. J. *et al.* Arabidopsis AtMORC4 and AtMORC7 Form Nuclear Bodies and
945 Repress a Large Number of Protein-Coding Genes. *PLoS Genet* **12**, (2016).

946 41. Weiser, N. E. *et al.* MORC-1 Integrates Nuclear RNAi and Transgenerational Chromatin
947 Architecture to Promote Germline Immortality. *Dev Cell* **41**, 408–423 (2017).

948 42. Moissiard, G. *et al.* MORC family ATPases required for heterochromatin condensation
949 and gene silencing. *Science* **336**, 1448–51 (2012).

950 43. Iyer, L. M., Abhiman, S. & Aravind, L. MutL homologs in restriction-modification
951 systems and the origin of eukaryotic MORC ATPases. *Biol Direct* **3**, 8 (2008).

952 44. Chen, S. *et al.* A genomic mutational constraint map using variation in 76,156 human
953 genomes. *Nature* **625**, (2024).

954 45. Sevilla, T. *et al.* Mutations in the MORC2 gene cause axonal Charcot-Marie-Tooth
955 disease. *Brain* **139**, 62–72 (2016).

956 46. Guillen Sacoto, M. J. *et al.* De Novo Variants in the ATPase Module of MORC2 Cause a
957 Neurodevelopmental Disorder with Growth Retardation and Variable Craniofacial
958 Dysmorphism. *Am J Hum Genet* (2020) doi:10.1016/j.ajhg.2020.06.013.

959 47. Peymani, F. *et al.* Pleiotropic effects of MORC2 derive from its epigenetic signature.
960 *Brain* awaf159 (2025) doi:10.1093/brain/awaf159.

961 48. Tailor, J. *et al.* Stem cells expanded from the human embryonic hindbrain stably retain
962 regional specification and high neurogenic potency. *Journal of Neuroscience* **33**, (2013).

963 49. Hanley, J. G., Koulen, P., Bedford, F., Gordon-Weeks, P. R. & Moss, S. J. The protein
964 MAP-1B links GABAC receptors to the cytoskeleton at retinal synapses. *Nature* **397**, 66–
965 69 (1999).

966 50. Grassi, D. A. *et al.* Profiling of lincRNAs in human pluripotent stem cell derived forebrain
967 neural progenitor cells. *Heliyon* (2020) doi:10.1016/j.heliyon.2019.e03067.

968 51. Seczynska, M., Bloor, S., Cuesta, S. M. & Lehner, P. J. Genome surveillance by HUSH-
969 mediated silencing of intronless mobile elements. *Nature* (2022) doi:10.1038/s41586-
970 021-04228-1.

971 52. Douse, C. H. *et al.* TASOR is a pseudo-PARP that directs HUSH complex assembly and
972 epigenetic transposon control. *Nat Commun* (2020) doi:10.1038/s41467-020-18761-6.

973 53. Li, X. *et al.* A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and
974 Paternal Imprints. *Dev Cell* **15**, (2008).

975 54. Takahashi, N. *et al.* ZNF445 is a primary regulator of genomic imprinting. *Genes Dev* **33**,
976 (2019).

977 55. Petropoulos, S. *et al.* Single-Cell RNA-Seq Reveals Lineage and X Chromosome
978 Dynamics in Human Preimplantation Embryos. *Cell* **165**, (2016).

979 56. Kagawa, H. *et al.* Human blastoids model blastocyst development and implantation.
980 *Nature* **601**, (2022).

981 57. Janssen, S. M. & Lorincz, M. C. Interplay between chromatin marks in development and
982 disease. *Nature Reviews Genetics* vol. 23 https://doi.org/10.1038/s41576-021-00416-
983 x (2022).

984 58. Ren, W. *et al.* Direct readout of heterochromatic H3K9me3 regulates DNMT1-mediated
985 maintenance DNA methylation. *Proc Natl Acad Sci U S A* **117**, (2020).

986 59. Pastor, W. A. *et al.* MORC1 represses transposable elements in the mouse male
987 germline. *Nat Commun* **5**, (2014).

988 60. Uneme, Y. *et al.* Morc1 reestablishes H3K9me3 heterochromatin on piRNA-targeted
989 transposons in gonocytes. *Proceedings of the National Academy of Sciences* **121**,
990 e2317095121 (2024).

991 61. Athanikar, J. N., Badge, R. M. & Moran, J. V. A YY1-binding site is required for accurate
992 human LINE-1 transcription initiation. *Nucleic Acids Res* **32**, (2004).

993 62. Uhlin, E. *et al.* Derivation of human iPS cell lines from monozygotic twins in defined and
994 xeno free conditions. *Stem Cell Res* **18**, (2017).

995 63. Nolbrant, S., Heuer, A., Parmar, M. & Kirkeby, A. Generation of high-purity human
996 ventral midbrain dopaminergic progenitors for in vitro maturation and intracerebral
997 transplantation. *Nat Protoc* (2017) doi:10.1038/nprot.2017.078.

998 64. Lancaster, M. A. & Knoblich, J. A. Generation of cerebral organoids from human
999 pluripotent stem cells. *Nat Protoc* **9**, (2014).

1000 65. Zufferey, R., Nagy, D., Mandel, R. J., Naldini, L. & Trono, D. Multiply attenuated lentiviral
1001 vector achieves efficient gene delivery in vivo. *Nat Biotechnol* (1997)
1002 doi:10.1038/nbt0997-871.

1003 66. Johansson, P. A., Adami, A. & Jakobsson, J. CRISPRi-mediated transcriptional silencing
1004 in iPSCs for the study of human brain development. *STAR Protoc* **3**, (2022).

1005 67. Weisheit, I. *et al.* Simple and reliable detection of CRISPR-induced on-target effects by
1006 qgPCR and SNP genotyping. *Nat Protoc* **16**, (2021).

1007 68. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* (2013)
1008 doi:10.1093/bioinformatics/bts635.

1009 69. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (2009)
1010 doi:10.1093/bioinformatics/btp352.

1011 70. Jin, Y., Tam, O. H., Paniagua, E. & Hammell, M. TEtranscripts: A package for including
1012 transposable elements in differential expression analysis of RNA-seq datasets.
1013 *Bioinformatics* (2015) doi:10.1093/bioinformatics/btv422.

1014 71. Love, M. I., Anders, S. & Huber, W. *Differential Analysis of Count Data - the DESeq2*
1015 *Package*. *Genome Biology* (2014). doi:110.1186/s13059-014-0550-8.

1016 72. Södersten, E. *et al.* A comprehensive map coupling histone modifications with gene
1017 regulation in adult dopaminergic and serotonergic neurons. *Nat Commun* **9**, (2018).

1018 73. Jönsson, M. E. *et al.* Activation of endogenous retroviruses during brain development
1019 causes an inflammatory response. *EMBO J* (2021) doi:10.15252/embj.2020106423.

1020 74. Zheng, G. X. Y. *et al.* Massively parallel digital transcriptional profiling of single cells.
1021 *Nat Commun* **8**, (2017).

1022 75. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, (2019).

1023 76. Cheetham, S. W., Kindlova, M. & Ewing, A. D. Methylartist: tools for visualizing modified
1024 bases from nanopore sequence data. *Bioinformatics* **38**, (2022).

1025 77. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with
1026 high efficiency for low cell numbers. *Nat Protoc* (2018) doi:10.1038/nprot.2018.015.

1027 78. Meers, M. P., Bryson, T. D., Henikoff, J. G. & Henikoff, S. Improved cut&run chromatin
1028 profiling tools. *Elife* **8**, (2019).

1029 79. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods*
1030 (2012) doi:10.1038/nmeth.1923.

1031

1032 **ACKNOWLEDGEMENTS.**

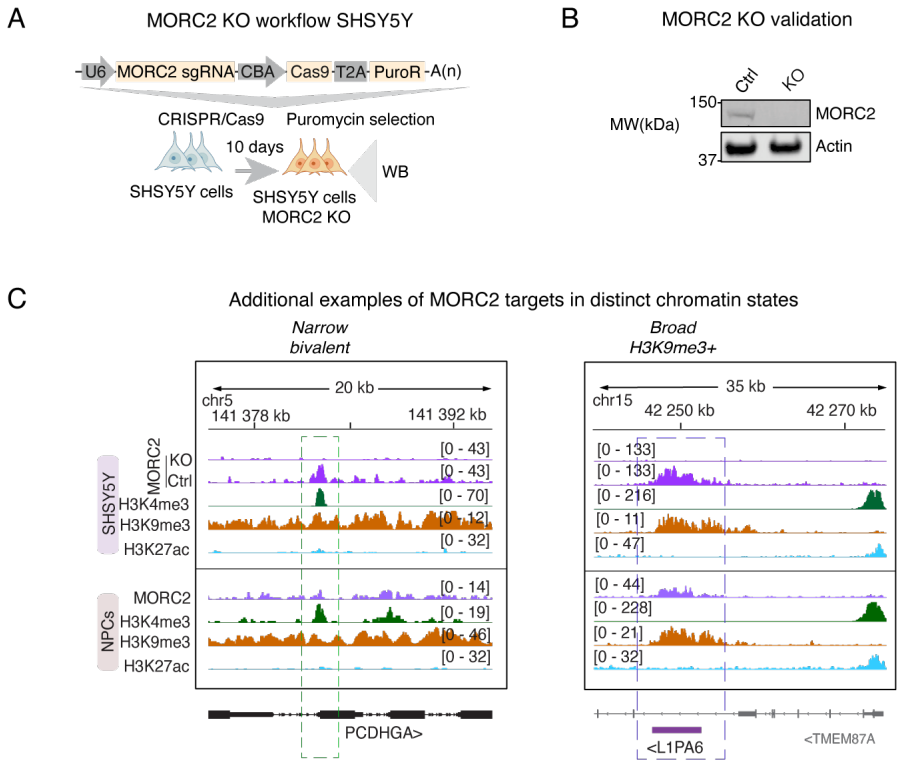
1033 We thank A. Hammarberg, S. da Rocha Baez, E. Monni, M. Persson-Vejgården, J. Nelander and
1034 S. Tesson for their technical assistance, S. Henikoff for the gift of the pA-MNase protein used
1035 in initial hNPC CUT&RUN experiments and the pAG-MNase plasmid (Addgene #123461) for in
1036 house production of the pAG-MNase used in all subsequent experiments. We thank P. Lehner
1037 for providing lentiviral vectors to knockout MORC2 and re-express MORC2 variants, the SCC
1038 Cell and Gene Therapy Core at Lund University and iPSC core at Karolinska Institute for
1039 generation of MORC2 edited iPSC clones. We acknowledge Clinical Genomics Lund and the
1040 Center for Translational Genomics (CTG) at Lund University for providing expertise and service
1041 with sequencing and analysis. This work was supported by the Swedish Government Initiative
1042 for Strategic Research Areas (MultiPark & StemTherapy) and grants from the Swedish
1043 Research Council (VR, 2022-02673 to J.J., 2021-03494 to C.H.D.), the Swedish Brain
1044 Foundation (Hjärnfonden, FO2023-0232, FO2023-0229 to C.H.D.), the Swedish Society for
1045 Medical Research (SSMF, S19-0100 to C.H.D.), Cancerfonden (222185Pj to J.J.),

Barncancerfonden (PR2023-0099 to J.J.) and the Wellcome Trust (101908/Z/13/Z and 217191/Z/19/Z to Y.M.). This project has been made possible in part by grant 2023-331773 to C.H.D. from the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation. Finally, we also acknowledge The Royal Physiographic Society of Lund (project grant to F.D.), the Crafoord Foundation and Bente-Rexed-Gerstedt Foundation (project grants to C.H.D.).

AUTHOR CONTRIBUTIONS.

AA made and validated the SHSY5Y MORC2 KO cell line and complemented cell lines. FD and LC-V designed and performed iPSC culture and cerebral organoid differentiations. FD, SK, SB and AK performed CUT&RUN experiments. FD, LC-V and CD-H performed quality control analysis of iPSC lines. FD and JM conducted L1 CRISPRi experiments, validations and 2D NPC differentiations, with expert advice from AAd. CF prepared DNA for Nanopore sequencing and performed bioinformatic analysis of the data. CF, NP and SK processed Illumina sequencing experiments and performed bioinformatic analysis of the data. OEK, CD-H and JM tested 2D neural differentiation protocols. FD, LC-V and AAd performed nuclear isolation of organoids. JGJ prepared all Illumina sequencing libraries for bulk and single-nuclei RNA-seq and CUT&RUN experiments. CHD conceived and designed the study. YM, JJ and CHD supervised the study and acquired funding. FD, CF and CHD wrote the manuscript and made figures with assistance from NP. All authors reviewed the final version.

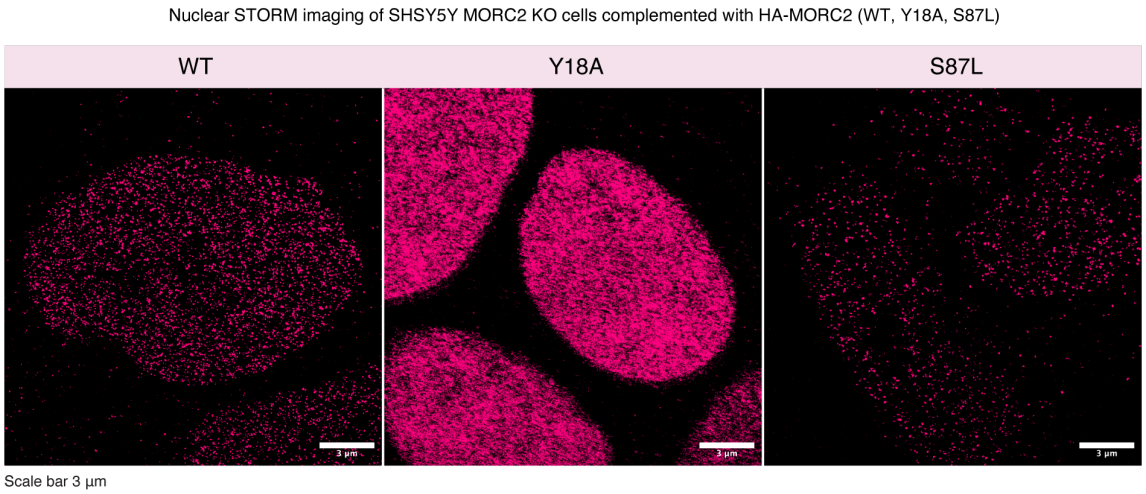
1074 **SUPPLEMENTARY FIGURES.**



1075

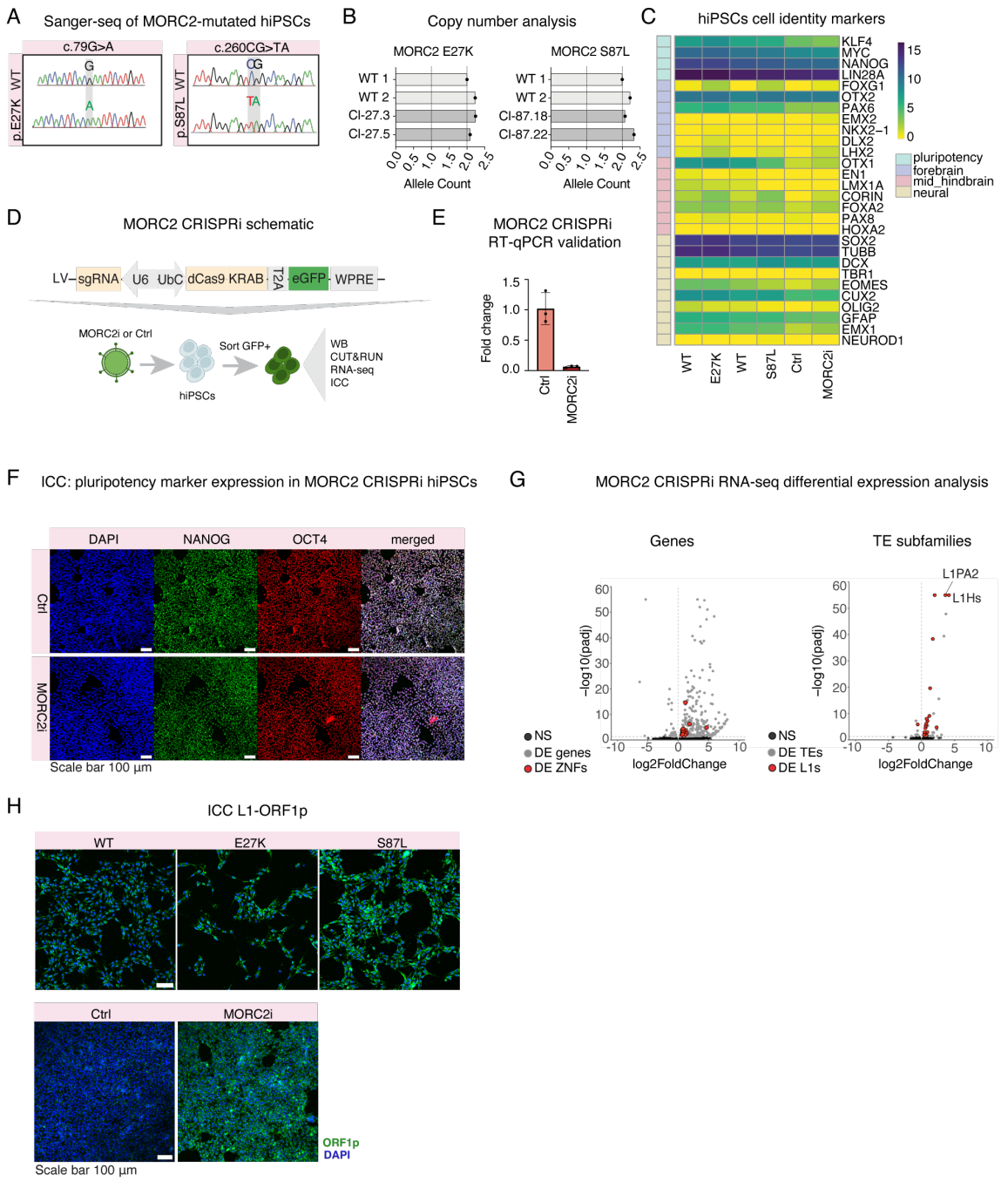
1076 **Supplementary Figure 1. (A)** Workflow for generating the MORC2 KO SHSY5Y cell line. **(B)** Validation of MORC2
1077 KO by Western blot. **(C)** Further examples of the different chromatin states characterizing MORC2 targets, see
1078 also Fig 1.

1079

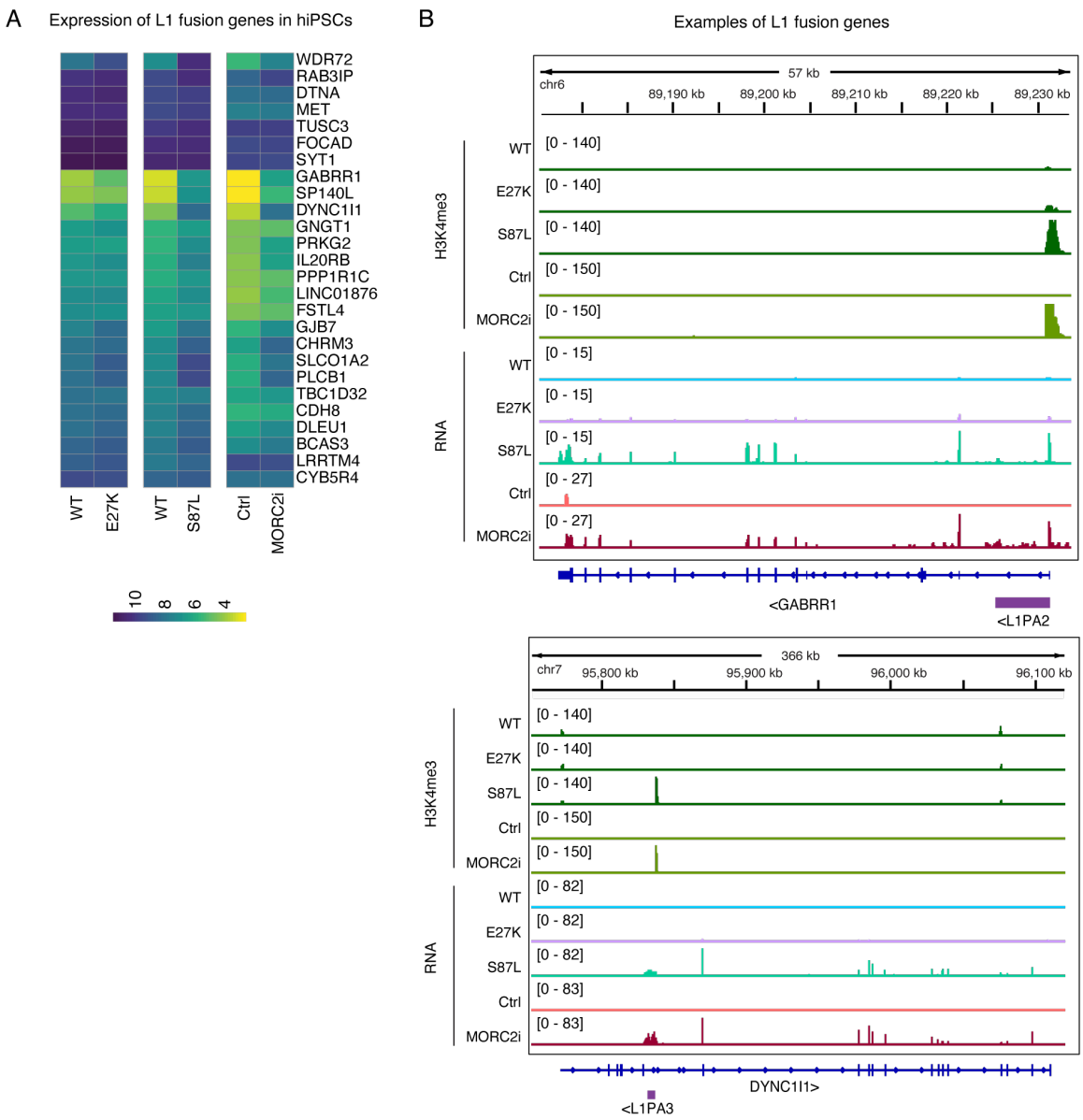


1080

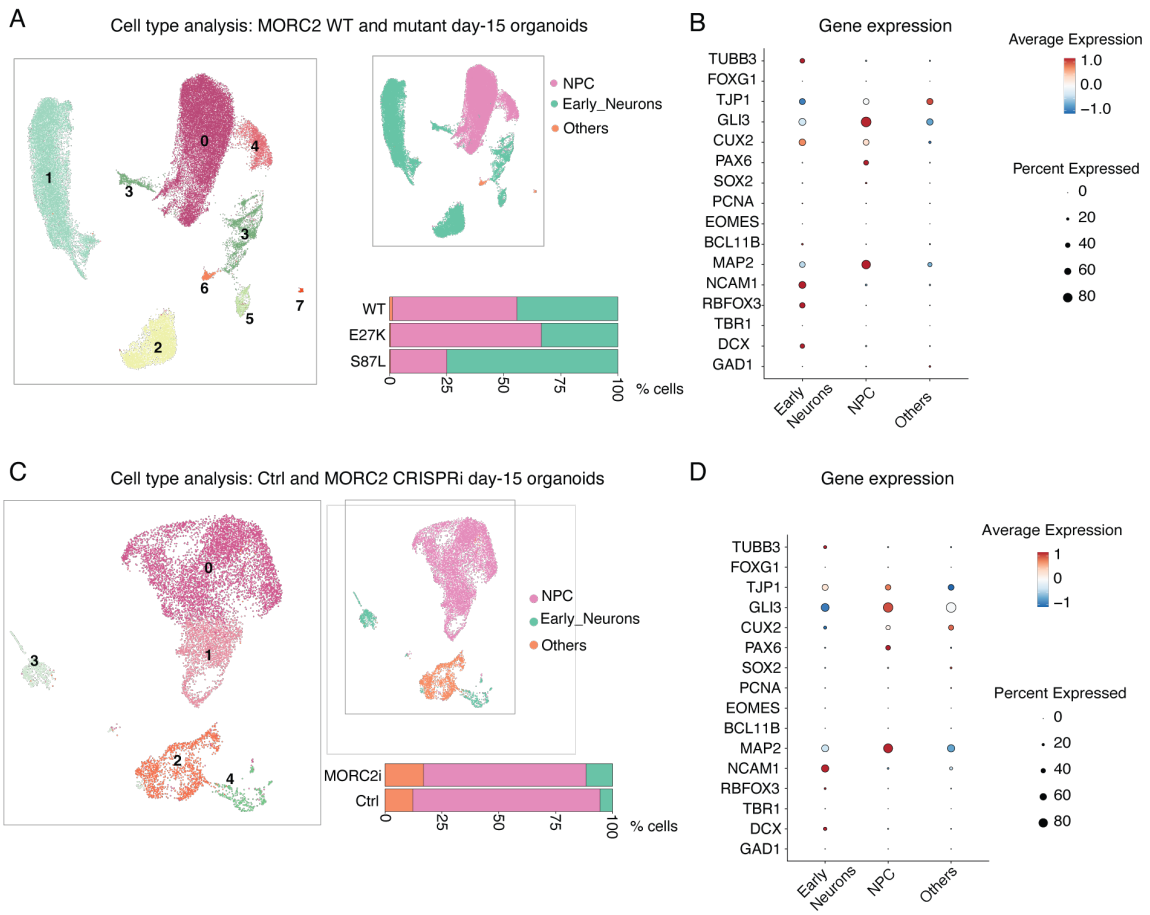
1081 **Supplementary Figure 2.** Nuclear STORM imaging of SHSY5Y MORC2 KO cells complemented with HA-MORC2
1082 (WT, Y18A, S87L) targeting the HA epitope tag.



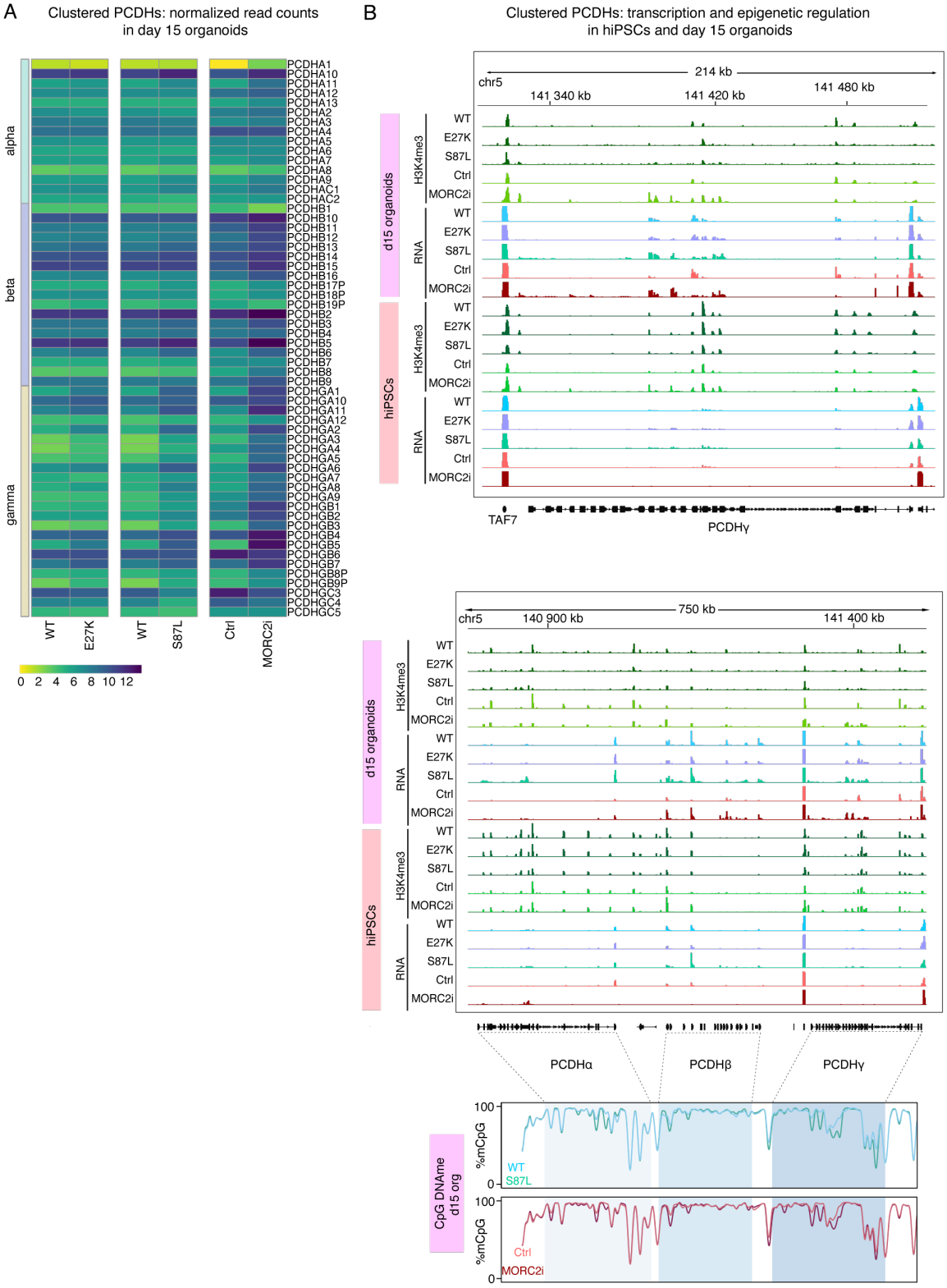
Supplementary Figure 3. (A) Electropherogram of the Sanger sequencing confirming the CRISPR-directed MORC2 mutations (E27K and S87L) in representative clones **(B)** Copy number analysis showing allele count for each MORC2 mutant clone and WT hiPSC lines. **(C)** Unscaled heatmap showing normalized read counts of pluripotency and neural marker genes as calculated by DESeq2 in the hiPSCs used in the study. Two batches of RNA from each clone was used for WT (n=4), E27K (n=4) and S87L (n=4) genotypes. **(D)** Schematic workflow of MORC2 CRISPRi experiment in hiPSCs and subsequent analyses. **(E)** Validation of CRISPRi through RT-qPCR analysis of MORC2 expression in control (n=3) and CRISPRi (n=3) hiPSCs. **(F)** Immunohistochemistry of pluripotency markers (NANOG, OCT4) for MORC2 control and CRISPRi cell lines. **(G)** Volcano plot showing the results of the differential expression analysis (DESeq2) for genes (left) and TE subfamilies using TEtranscripts (right) in the MORC2 CRISPRi cell line (n=3) compared to control (n=3). Differentially expressed genes and LINE1 subfamilies are highlighted in red, other differentially expressed genes and subfamilies are presented in grey (padj <0.05). **(H)** Immunohistochemistry of L1-ORF1p (green) in E27K and S87L MORC2 mutant and CRISPRi cell lines with their respective controls. Experiments were performed at least twice with similar results.



Supplementary Figure 4. (A) Unscaled heatmap showing the normalized read counts of L1 fusion genes in the E27K (n=4) and S87L (n=4) MORC2 mutant iPSC lines as compared to WT controls (n=4) and MORC2 CRISPRi (n=3) iPSCs compared to its control (n=3). **(B)** Genome browser snapshots showing two examples of L1 fusion genes derepressed in MORC2 mutant and CRISPRi treatments.



Supplementary Figure 5. Cell composition of day 15, unguided cerebral organoids. **(A)** Left/Middle: UMAP showing the cell types identified in MORC2-WT (4-5 organoids from n=4 batches/clones), MORC2-E27K (4-5 organoids from n=2 batches/clones) and MORC2-S87L mutant (4-5 organoids from n=2 batches/clones) unguided cerebral organoids. Bottom: bar plot showing percentage of the cell type composition of the MORC2 WT and MORC2-S87L mutant cerebral organoids. **(B)** Dot plot showing neuronal and NPC gene markers used to characterize the cell clusters. Dot size corresponds to the proportion of cells expressing the gene. Dot color indicates average expression in each cell type. **(C)** Left/Middle: UMAP showing the cell types identified in control (4-5 organoids from n=1 batch) and MORC2 CRISPRi (4-5 organoids from n=1 batch) unguided cerebral organoids. Bottom: bar plot showing percentage of the cell type composition of the control and MORC2 CRISPRi cerebral organoids. **(D)** Dot plot showing neuronal and NPC gene markers used to characterize the cell clusters. Dot size corresponds to the proportion of cells expressing the gene. Dot color indicates average expression in each cell type.

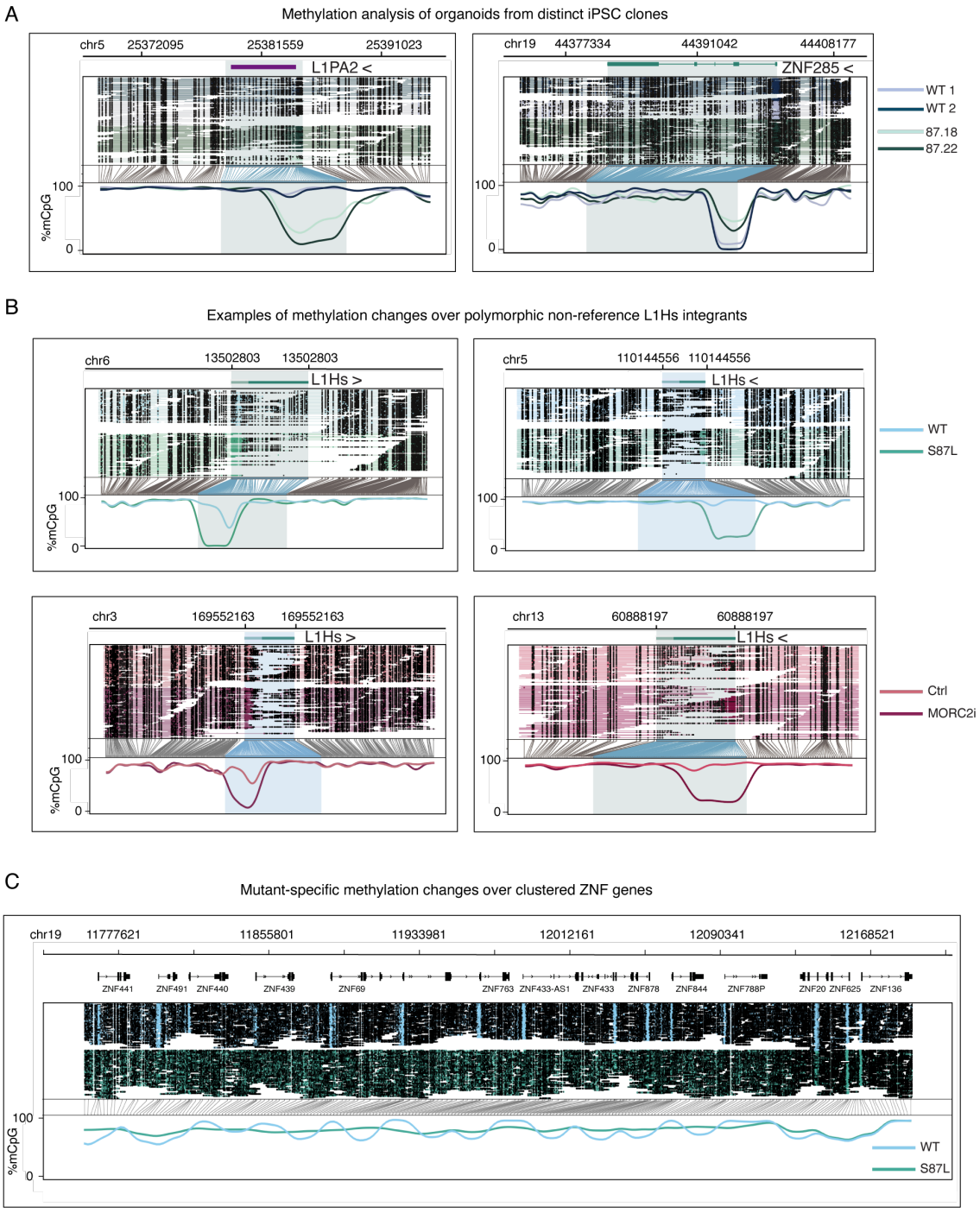


Supplementary Figure 6. (A) Unscaled heatmap showing normalized read counts of bulk RNA-seq data for clustered protocadherin genes (PCDHα, PCDHβ and PCDHγ) as calculated by DESeq2 in WT (n=9 samples from a total of 27 organoids differentiated from 3 iPSC lines across 3 batches,) E27K (n=9 samples from a total of 27 organoids differentiated from 2 iPSC lines across 2 batches) and S87L (n=9 samples from a total of 27 organoids differentiated from 2 iPSC lines across 2 batches) cell lines, and from Control (n=3 samples derived from 9 organoids differentiated from 1 bulk transduced cell line across 1 batch) and MORC2 CRISPRi (n=3 samples derived from 9 organoids differentiated from 1 bulk transduced cell line across 1 batch). **(B)** Genome browser

1124 view of bulk RNA-seq and H3K4me3 CUT&RUN from organoid experiments of the PCDHγ cluster (top) and PCDHα,
1125 PCDHβ and PCDHγ clusters (center), and their mean percentage methylation profile (bottom). Data are
1126 representative of all replicates performed.

1127

1128

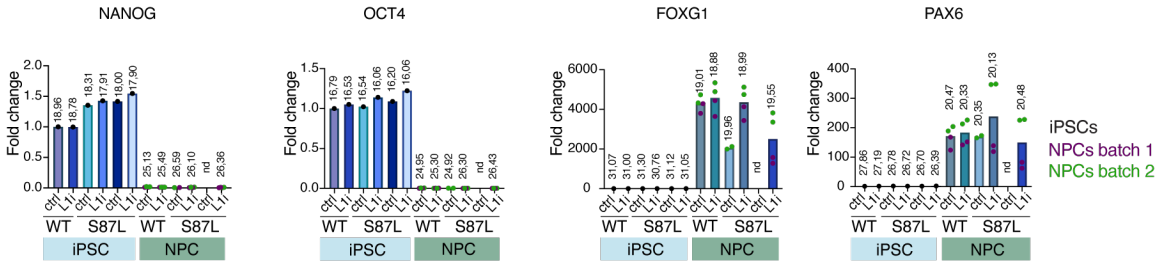


1129

1130 **Supplementary Figure 7. (A)** Aligned Nanopore reads showing CpG methylation across clones belonging to the
1131 WT (n=2) and S87L (n=2) cell lines over example L1 and ZNF gene loci. **(B)** Aligned Nanopore reads showing CpG
1132 methylation over examples of non-reference LINE1 insertions in WT and S87L cell lines, and Ctrl and MORC2
1133 CRISPRi cell lines. **(C)** Aligned Nanopore reads showing long-range dysregulated CpG methylation over a ZNF
1134 gene cluster.

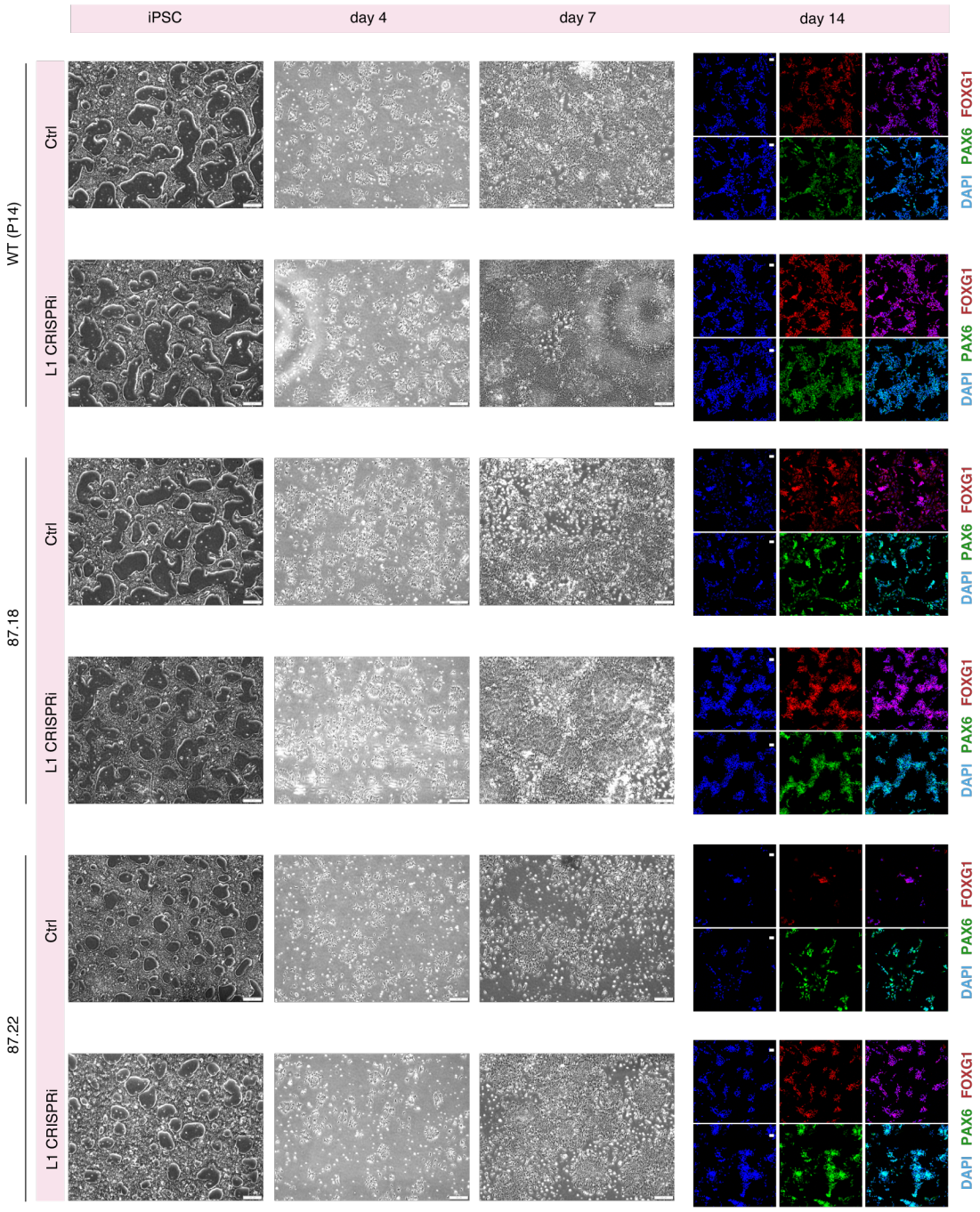
A

RT-qPCR: iPSC and NPC marker expression



B

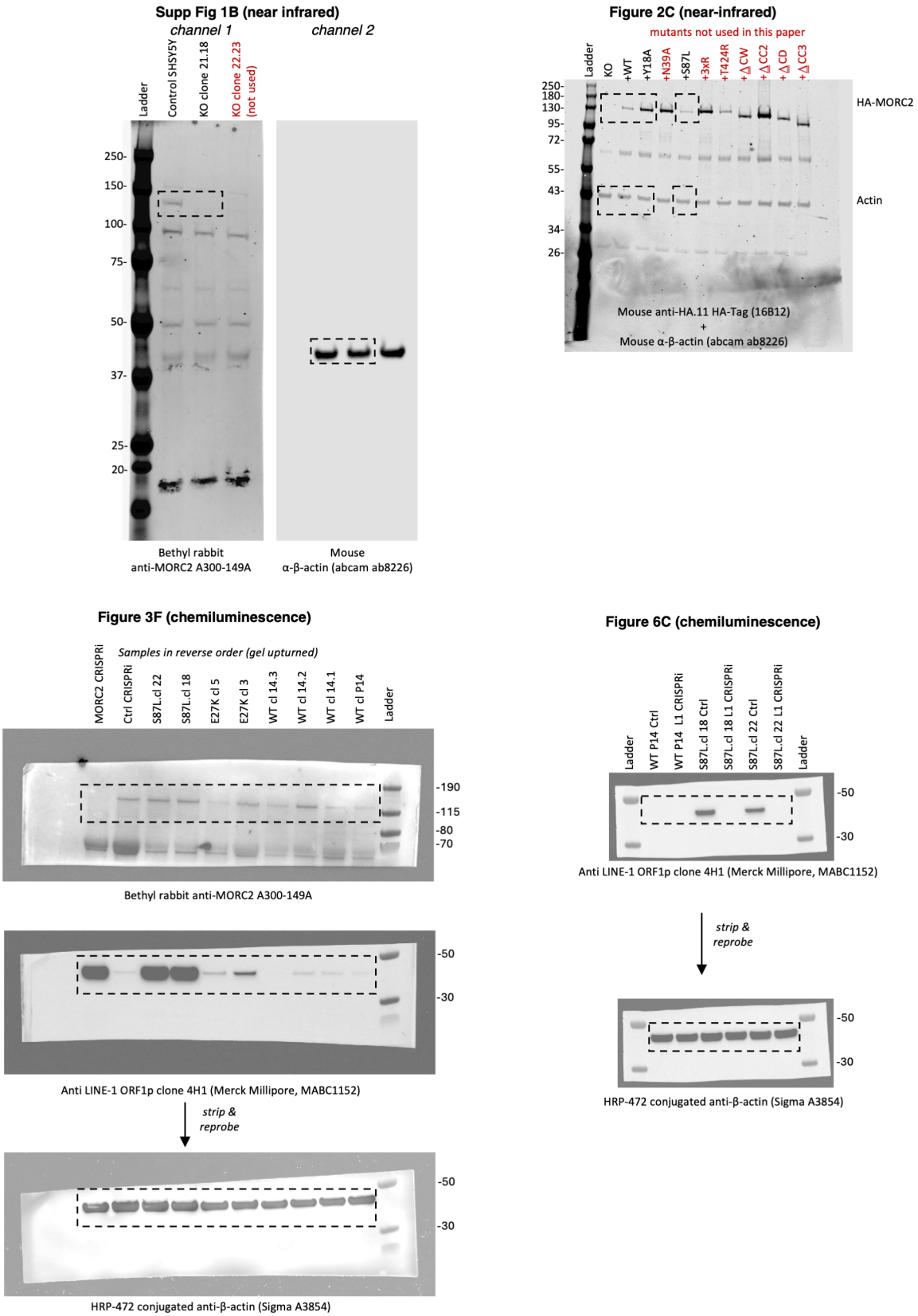
2D NPC differentiation timecourse: brightfield and ICC



1136 **Supplementary Figure 8. (A)** RT-qPCR data from indicated iPSC (n=1) and NPCs derived therefrom (n=2 from two
1137 separate differentiation batches) to assess expression of pluripotency and neural progenitor markers. Each value
1138 is the mean of three technical replicates. nd, not determined due to insufficient RNA yield from the 87.22-Ctrl
1139 condition. In all cases Ct values were normalized to the *ACTB* housekeeping gene and fold-change calculated
1140 relative to the WT-Ctrl condition. **(B)** Representative brightfield and immunocytochemistry analysis of NPC
1141 differentiations from six cell lines at the indicated timepoints. Note, data shown for the WT-Ctrl line is the same
1142 as that shown in figure 6.

1143

Uncropped Western blots



Supplementary Figure 9. Uncropped Western blots used in the given figure panels. Where samples were not used in the study the labels are written in red. See also Methods.