DNMT3A haploinsufficiency results in behavioral deficits and global epigenomic dysregulation shared across neurodevelopment disorders

Diana L. Christian[1,6], Dennis Y. Wu[1,6], Jenna R. Martin[1], J. Russell Moore[1], Yiran R. Liu[1], Adam W. Clemens[1], Sabin A. Nettles[1], Nicole M. Kirkland[2], Cheryl A. Hill[2], David F. Wozniak[3,5], Joseph D. Dougherty[3,4,5], Harrison W. Gabel[1,*]

[1]Department of Neuroscience, Washington University School of Medicine, St. Louis, MO 63110-1093, USA

[2]Department of Pathology and Anatomical Science, University of Missouri School of Medicine, Columbia, MO 65212, USA

[3]Department of Psychiatry, Washington University School of Medicine, St. Louis, MO 63110-1093, USA

[4]Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110-1093, USA

[5.]Intellectual and Developmental Disabilities Research Center, Washington University School of Medicine, St. Louis, MO, USA.

[6]These authors contributed equally: Diana L. Christian, Dennis Y. Wu.

[*]Email: gabelh@wustl.edu

Summary

Mutations in DNA methyltransferase 3A (DNMT3A) have been detected in autism and related disorders, but how these mutations disrupt nervous system function is unknown. Here we define the effects of neurodevelopmental disease-associated DNMT3A mutations. We show that diverse mutations affect different aspects of protein activity yet lead to shared deficiencies in neuronal DNA methylation. Heterozygous DNMT3A knockout mice mimicking DNMT3A disruption in disease display growth and behavioral alterations consistent with human phenotypes. Strikingly, in these mice we detect global disruption of neuron-enriched non-CG DNA methylation, a binding site for the Rett syndrome protein MeCP2. Loss of this methylation leads to enhancer and gene dysregulation that overlaps with models of Rett syndrome and autism. These findings define effects of DNMT3A haploinsufficiency in the brain and uncover

30 disruption of the non-CG methylation pathway as a convergence point across
31 neurodevelopmental disorders.
32
33 **Introduction**
34 Precise regulation of transcription through epigenetic mechanisms is critical for nervous
35 system development (Cholewa-Waclaw et al., 2016). Exome sequencing studies have revealed
36 mutations of genes encoding epigenetic modifiers of chromatin structure as a major underlying
37 cause of neurodevelopmental diseases (NDD), including autism spectrum disorder (ASD) (McRae
38 et al., 2017; Sanders et al., 2015; Satterstrom et al., 2019). A challenge emerging from these
39 discoveries is to define the cellular functions of the disrupted proteins during normal
40 development and search for shared pathways between these proteins that can potentially be
41 targeted for therapeutic development.
42 Gene regulation mediated by DNA methylation has emerged as an epigenetic mechanism
43 that plays a critical role in nervous system function (Kinde et al., 2015). In addition to the classical
44 methylation of cytosines found at CG dinucleotides (mCG), neurons contain uniquely high levels
45 of methyl-cytosine (mC) in a non-CG context, with this mark occurring primarily at CA
46 dinucleotides (mCA) (Guo et al., 2014; Lister et al., 2013a; Xie et al., 2012). mCA is deposited *de*
47 *novo* through the activation of the DNA methyltransferase 3A (DNMT3A) enzyme during the early
48 postnatal period (1-6 weeks of age in mice). Levels of mCA increase specifically in neurons until
49 the number of methylation sites in the non-CG context are nearly equivalent to mCG sites (Guo
50 et al., 2014; Lister et al., 2013a; Xie et al., 2012). A critical function for mCA is to serve as a binding
51 site for a neuron-enriched chromatin protein, Methyl-CpG binding Protein 2 (MeCP2) (Chen et
52 al., 2015; Gabel et al., 2015; Guo et al., 2014). MeCP2 was initially defined by its high affinity for
53 mCG, but biochemical and genomic studies indicate that it preferentially interacts with mCA to
54 down-regulate transcription of genes with essential functions in the brain (Boxer et al., 2019;
55 Gabel et al., 2015; Kinde et al., 2016; Lagger et al., 2017; Lyst and Bird, 2015). Loss of MeCP2
56 leads to the severe neurological disorder Rett syndrome, while duplication causes MeCP2-
57 duplication syndrome, an ASD, suggesting that read-out of mCA is critical to nervous system
58 function (Amir et al., 1999; Van Esch et al., 2005).

59 Notably, human exome sequencing studies have recently identified *de novo* mutations in
60 DNMT3A in individuals with ASD (Feliciano et al., 2019; Sanders et al., 2015; Satterstrom et al.,
61 2019). Separate studies have also defined heterozygous disruption of DNMT3A as the underlying
62 cause of Tatton-Brown Rahman syndrome (TBRS), a heterogeneous NDD characterized by
63 intellectual disability, overgrowth, craniofacial abnormalities, anxiety, and high penetrance of
64 ASD (Tatton-Brown et al., 2014, 2018). While a portion of the mutations identified in affected
65 individuals are truncations that are predicted to cause complete inactivation of the enzyme, a
66 majority of disease-associated alleles are missense mutations, raising questions about whether
67 loss-of-function effects are a primary mechanism of disruption in DNMT3A disorders (Tatton-
68 Brown et al., 2014, 2018). In addition, while heterozygous loss of DNMT3A has been studied in
69 the context of oncogenesis in the hematopoietic system (Cole et al., 2017), the effects of partial
70 loss of DNMT3A on nervous system function *in vivo* have not been examined; therefore, the
71 consequences of possible methylation changes on neuronal gene regulation and behavior are
72 unknown.
73 Here we examine the molecular effects of neurodevelopmental disease-associated
74 DNMT3A mutations and explore the consequences of heterozygous DNMT3A mutation on the
75 neuronal epigenome. Our results indicate that missense mutations across canonical domains of
76 DNMT3A disrupt different aspects of protein function, yet mutations in all domains reduce the
77 capacity of the enzyme to deposit neuronal mCA. We detect altered growth and behavior in
78 DNMT3A heterozygous deletion mice, supporting haploinsufficiency as a driver of pathology in
79 DNMT3A disorders. Through integrated epigenomic analysis, we reveal disruption of mCA
80 throughout the brain of DNMT3A mutant mice. Strikingly, we show that this loss of mCA leads to
81 disruption of distal regulatory enhancer activity and changes in gene expression that overlap with
82 models of MeCP2 disorders and other ASDs. These findings define the effects of NDD-associated
83 DNMT3A mutations for the first time and reveal disruption of mCA-mediated epigenomic
84 regulation as a convergence site across clinically distinct NDDs.
85
86 **Results**
87 **Functional analysis of disease-associated DNMT3A mutations**

88 Multiple DNMT3A mutations have been identified in individuals with ASD and TBRS.
89 However, the large number of missense mutations identified and the phenotypic heterogeneity
90 of individuals with these mutations raise the possibility that alterations of amino acids within
91 different protein domains may have distinct consequences and may dictate the nature and
92 severity of disease. We therefore sought to assess the effects of disease-associated DNMT3A
93 mutations on protein expression, cellular localization, and catalytic activity, looking for common
94 effects that may be core to the development of NDD.

95 We engineered amino-acid alterations homologous to human disease mutations into a
96 FLAG-tagged DNMT3A protein expression vector and assessed multiple mutations found within
97 each functional domain of the protein (Figure 1A). These analyses included mutations in the
98 chromatin interacting proline-tryptophan-tryptophan-proline (PWWP) domain, the auto-
99 inhibitory Histone H3 lysine 4 interacting ATRX-DNMT3-DNMT3L (ADD) domain, and the well-
100 defined methyltransferase catalytic domain (Gowher and Jeltsch, 2018). Transfection into
101 heterologous cells facilitated rapid assessment of protein expression by western blot, cellular
102 localization by immunocytochemistry, and catalytic activity using an *in vitro* methyltransferase
103 assay (Figure 1). Mutations in the PWWP domain resulted in a reduction in DNMT3A protein
104 expression and loss of nuclear localization compared to wild-type controls (Figure 1B-D, Figure
105 S1A-C). When expressed at equal levels to that of wild-type protein however, these mutations
106 exhibited substantial catalytic activity (Figure 1E,F). In contrast, mutations found in the catalytic
107 methyltransferase domain of DNMT3A showed wild-type expression and localization but
108 displayed deficits in catalytic activity in the *in vitro* methyltransferase analysis (Figure 1B-F, Figure
109 S1A-C). Mutations in the ADD domain of DNMT3A displayed normal protein localization and
110 expression levels and exhibited equal or higher methylation activity compared to wild-type
111 protein *in vitro* (Figure 1B-F, Figure S1A-C).

112 To further evaluate the functional effects of disease-associated DNMT3A mutations in the
113 context of endogenous chromatin, we tested the capacity of DNMT3A mutants to establish DNA
114 methylation across the genome in mouse cortical neurons. For this analysis, we focused on the
115 global build-up of mCA in postmitotic neurons that requires DNMT3A (Gabel et al., 2015; Lister
116 et al., 2013a). Cultured neurons isolated from the cerebral cortex at embryonic day 14.5

117 accumulate mCA *in vitro* and this build-up can be blocked by lentiviral-mediated delivery of Cre
118 recombinase to DNMT3A$^{flx/flx}$ cells at 3 days *in vitro* (DIV) (Figure 2A, Figure S1D,E). We co-
119 transduced wild-type or mutant DNMT3A lentivirus at equal levels (Figure S1F) to test the
120 capacity of each protein to rescue deposition of DNA methylation. Analysis of mutations across
121 the major domains of DNMT3A detected deficits in mCA accumulation for all disease-associated
122 mutations tested (Figure 2B). Notably, mutations in the ADD domain that exhibited robust
123 catalytic activity *in vitro* displayed moderate-to-severe deficits in mCA deposition in neurons. The
124 ADD domain has been implicated in both histone binding and auto-inhibition of the protein (Guo
125 et al., 2015), thus effects in this neuronal assay may indicate that loss of ADD function blocks the
126 capacity of the enzyme to engage with chromatin and promote DNMT3A methylation activity in
127 cells. Together our results indicate that although NDD-associated mutations in DNMT3A affect
128 different protein domains and alter distinct aspects of protein function (e.g. localization,
129 chromatin interaction, catalysis), these mutations share a common outcome of reduced
130 enzymatic activity on neuronal DNA, with many mutations resulting in functionally null proteins.
131

132 ***In vivo* effects of heterozygous DNMT3A disruption**

133 In light of our findings that multiple NDD-associated missense mutations in DNMT3A
134 result in complete or near-complete loss of function, we next sought to understand the effects
135 of heterozygous inactivation of DNMT3A *in vivo*. Previous studies have demonstrated severe
136 developmental deficits and perinatal lethality associated with complete loss of DNMT3A
137 (homozygous null mutation) in mice (Okano et al., 1999). However, the relevance of
138 heterozygous mutation of DNMT3A to neurodevelopmental disease has only recently been
139 uncovered and growth and behavioral effects of partial DNMT3A inactivation have not been
140 systematically assessed. We therefore carried out growth, behavioral, and molecular analyses of
141 mice carrying a constitutive heterozygous deletion of exon 19 of *Dnmt3a* (see *methods*) (Kaneda
142 et al., 2004). We find that this mutation leads to 50% reduction of RNA and protein expression,
143 allowing us to study the *in vivo* effects of heterozygous null mutation of DNMT3A (referred to as
144 DNMT3A$^{KO/+}$) (Figure S2A-C).

145 We first examined phenotypes with relevance to the overgrowth in individuals with
146 heterozygous DNMT3A mutations (Tatton-Brown et al., 2018), including enlarged body size and
147 obesity (body weight), tall stature (long-bone length), and macrocephaly (skull dimensions).
148 DNMT3A$^{KO/+}$ mice showed similar body weight to controls in the early postnatal period but were
149 significantly heavier than controls as mature adults (Figure 3A). This phenotype mimics a
150 maturity-associated trend toward increasing body weight observed in TBRS patients (Tatton-
151 Brown et al., 2018). Measurements of bone length indicated a small but significant increase in
152 tibia length in DNMT3A$^{KO/+}$ mice, with a trend towards longer femur length (Figure 3B, Figure
153 S3A-C). Morphometric analyses of the cranium and mandible indicated very subtle differences
154 between DNMT3A$^{KO/+}$ mice and their wild-type littermates (Figure S3D-F). One linear distance
155 spanning the rostrocaudal length of the interparietal bone is larger in DNMT3A$^{KO/+}$ mice relative
156 to wild-type littermates. Two linear distances in the facial region were significantly larger in wild-
157 type mice, while all other comparisons were not significantly different (Figure S3D). This suggests
158 very slight disruptions in growth of the facial region in DNMT3A $^{KO/+}$ mice. Together these findings
159 uncover effects on long bone length that mirror aspects of the human disorder, while skull
160 development in DNMT3A$^{KO/+}$ mice shows more limited effects. Additionally, enlarged body mass
161 in these mice appears to mimic overgrowth and obesity detected in individuals with TBRS
162 (Tatton-Brown et al., 2014).

163 To examine neurological and behavioral phenotypes in DNMT3A$^{KO/+}$ mice, we assessed
164 basic measures of sensation and motor performance such as balance (ledge test, platform test),
165 grip strength (inverted screen test), motor coordination (walking initiation, rotarod), and
166 sensorimotor gating (pre-pulse inhibition). DNMT3A$^{KO/+}$ mice were not significantly different in
167 these assays (Figure S4A-G), indicating that heterozygous loss of DNMT3A does not grossly
168 disrupt sensorimotor function. This allowed us to accurately assess more complex aspects of
169 behavior and cognition.

170 We carried out a panel of assays with relevance to neuropathology observed in humans
171 with DNMT3A mutations, including anxiety, autism, and intellectual disability. DNMT3A$^{KO/+}$ mice
172 displayed reduced exploratory behavior during open field testing, including reduced distance
173 traveled and rearing (Figure 3C-D). DNMT3A$^{KO/+}$ mice also displayed anxiety-like behaviors in this

174 assay, as they spent less time in the center of the open field arena (Figure 3E). In tests of climbing
175 behavior, DNMT3A$^{KO/+}$ mice showed longer latency to climb to the bottom of a pole and to the
176 top of mesh screens (Figure S4H-J), suggesting DNMT3A$^{KO/+}$ mice display differences in volitional
177 movement. To further assess anxiety-like behavior, we tested mice in the elevated plus maze and
178 observed that DNMT3A$^{KO/+}$ mice spent less time exploring the open arms of the maze with no
179 change in percent entries made into all arms (Figure 3F, Figure S4K). Overall, these results
180 demonstrate that the DNMT3A$^{KO/+}$ mice display changes in exploratory behavior, suggesting
181 altered emotionality and increased anxiety-like behaviors.

182 Analysis of common phenotypes examined in mouse models of autism (marble burying,
183 three chamber social approach, ultrasonic vocalizations) revealed additional changes in behavior.
184 We detected a significant reduction in marble burying activity for DNMT3A$^{KO/+}$ mice, indicating
185 alterations in repetitive digging behavior (Figure 3G). Testing of social interaction in the three
186 chamber social approach test for adult mice (Yang et al., 2011) detected trends toward similar
187 sociability and preference for social novelty in DNMT3A$^{KO/+}$ compared to controls (Figure S4L).
188 However, mutant mice showed reduced time investigating both mice and objects, as well as
189 reduced overall activity (Figure 3H,I, Figure S4M). These results may further reflect the overall
190 trend towards reduced exploration and anxiety-like phenotypes in these animals, instead of
191 changes in sociability (Nygaard et al., 2019). Reduced maternal-isolation induced ultrasonic
192 vocalizations (Barnes et al., 2017) were detected in DNMT3A$^{KO/+}$ mice at postnatal day five,
193 suggesting deficits in early pro-social behaviors or slight developmental delay in the normal
194 acquisition of this behavior (Figure 3J). Together these results indicate alterations in behaviors
195 commonly assessed in mouse models of autism (Chang et al., 2017; Simola and Granon, 2019;
196 Takumi et al., 2019), with our findings suggesting a reduction in activity and exploration, as well
197 as some changes in communication behaviors.

198 Intellectual disability is observed in patients with DNMT3A mutations, so we assessed
199 learning and memory in the DNMT3A$^{KO/+}$ mice using fear conditioning and Morris water maze
200 tests. The mutant mice displayed largely similar recall performance to that of control mice in
201 foot-shock induced fear conditioning, with similar levels of shock sensitivity (Figure 3K-M, Figure
202 S4N). However, DNMT3A$^{KO/+}$ mutants showed heightened freezing response during training, as

203 well as contextual and auditory recall phases of conditioned fear testing (Figure 3K-M). Mutant
204 mice also showed delayed extinction of freezing behavior in response to the auditory cue alone,
205 which may indicate altered emotionality or cognition (see *methods* for discussion). Assessment
206 of spatial and contextual memory by Morris water maze testing demonstrated that DNMT3A$^{KO/+}$
207 mice were slower to learn to find a visible platform and did not learn the location of the hidden
208 platform over time to the level of wild-type controls (Figure S4O-P). DNMT3A$^{KO/+}$ mice also
209 showed no differences in swimming speed compared to wild types (Figure S4Q-R). There were
210 no significant effects on distance traveled in target zone or platform crossings in the probe trial,
211 though DNMT3A$^{KO/+}$ mice trended towards fewer platform crossings (Figure S4S,T). These
212 findings suggest that DNMT3A$^{KO/+}$ mutants do not show frank deficits in learning and memory
213 but do display differences in task performance that further suggest altered emotionality or
214 cognition in these mice. Our analyses demonstrate that heterozygous deletion of DNMT3A
215 results in altered behavior in mice with relevance to anxiety and memory associated behaviors
216 observed in patients with DNMT3A mutations. These data support a model in which DNMT3A
217 haploinsufficiency can alter behavioral circuits to drive phenotypes in NDD.
218
219 **Global disruption of DNA methylation in the DNMT3A$^{KO/+}$ brain**
220 We next investigated the epigenomic defects that may underlie the altered behaviors
221 observed in DNMT3A$^{KO/+}$ mice. We first used sparse whole-genome bisulfite sequencing to
222 efficiently survey effects on global DNA methylation levels for multiple brain regions and liver
223 tissue isolated from wild type and DNMT3A$^{KO/+}$ mice. This analysis detected limited reductions in
224 genome-wide mCG levels in the DNMT3A$^{KO/+}$ brain that were not apparent in the liver, a non-
225 neural tissue (Figure 4A). In contrast, mCA levels were reduced by 30-50% across all brain regions
226 examined in DNMT3A$^{KO/+}$ mice (Figure 4B). DNA methylation across postnatal development in
227 the cerebral cortex, the brain region with the highest levels of mCA at 8 weeks, suggests that
228 deficits in mCA appear during initial accumulation of this methyl mark at 1-6 weeks (Figure 4B).
229 Thus, global mCA levels in the brain appear to be highly sensitive to heterozygous DNMT3A
230 disruption, while overall global mCG levels are largely maintained.

231 DNA methylation at specific genomic elements, including promoters, enhancers, and
232 gene bodies is thought to play an important role in regulating transcription. Alterations in
233 methylation at these regions can impact gene expression to affect the development and function
234 of the brain (Clemens et al., 2019; Nord and West, 2019; Stroud et al., 2017a). We therefore
235 assessed changes in methylation at base-pair resolution by high-depth whole-genome bisulfite
236 sequencing to identify potential changes in mCA and mCG at these important regulatory sites.
237 For this analysis we focused on the cerebral cortex, as this region is enriched for mCA (Figure 4B)
238 and is disrupted in ASD and MeCP2 disorders (Clemens et al., 2019; Satterstrom et al., 2019;
239 Sceniak et al., 2016; de la Torre-Ubieta et al., 2016).

240 High-resolution analysis of mCG confirmed the subtle reduction in mCG across all classes
241 of genomic elements (Figure 5A,D). We considered that CG dinucleotides in specific sites in the
242 neuronal genome may be more sensitive to a partial reduction in DNMT3A activity. For example,
243 in the hematopoietic system, heterozygous disruption of DNMT3A leads to reductions in DNA
244 methylation in genomic regions that can be identified as sensitive to complete loss of DNMT3A
245 (Cole et al., 2017). We therefore evaluated developmentally-regulated adult-specific CG-
246 differentially methylated regions (CG-DMRs) previously identified in the cortex (Figure S5A)
247 (Lister et al., 2013a). Because DNMT3A is the only *de novo* methyltransferase expressed in the
248 postnatal brain, we hypothesized that adult-specific CG-DMRs might be sensitive to a reduction
249 in enzyme activity. Indeed, we found that these sites build up mCG during postnatal development
250 and do not become methylated in a brain-specific DNMT3A conditional knockout mouse
251 (DNMT3A cKO) (Stroud et al., 2017a) (Figure S5B). Analysis of adult-specific CG-DMRs in the
252 DNMT3A$^{KO/+}$ model indicated that these sites are particularly sensitive to partial inactivation of
253 DNMT3A compared to other regions genome wide (Figure 5A,C,D).

254 To further search for local sites of altered mCG in the DNMT3A$^{KO/+}$, we performed *de novo*
255 calling of mCG differentially methylated regions using the BSmooth algorithm (Hansen et al.,
256 2012). We identified 843 hypo- and 71 hyper-CG-DMRs across the genome that met high
257 stringency filters for size and reproducibility (Figure 5A-C,E, Figure S5C, see *methods*). These
258 hypo-DMRs significantly overlap with the previously identified adult-specific CG-DMRs (Lister et
259 al., 2013a) (Figure 5F), further supporting the idea that DNMT3A is haploinsufficient for postnatal

260 mCG deposition at these sites. Examination of the genomic distribution of all DNMT3A$^{KO/+}$ CG-
261 DMRs revealed significant overlap of hypo-DMRs with putative enhancer regions, gene bodies,
262 and promoters (Figure 5A-C,F). DMRs were also highly enriched for overlap with CpG island
263 shores, regions disrupted in studies of DNMT3A mutation outside of the nervous system (Cole et
264 al., 2017; Spencer et al., 2017) (Figure 5F). Because a substantial percentage of mCG in neurons
265 can occur in an oxidized, hydroxymethyl form (hmCG), we further performed oxidative bisulfite
266 sequencing analysis of DNA from the cortex. This analysis revealed no clear evidence of
267 differential effects on the oxidized or unoxidized forms of mCG across the genome in the
268 DNMT3A$^{KO/+}$ (Figure S5D). Together these findings indicate that a small subset of mCG sites are
269 particularly sensitive to heterozygous loss of DNMT3A. The localization of these CG-DMRs to
270 regulatory elements suggests that these methylation changes could impact gene expression.

271 We next examined the profile of mCA at higher resolution, assessing genomic elements
272 of different scales that have relevance to gene regulation. In contrast to the limited mCG changes
273 in the DNMT3A$^{KO/+}$, analysis of mCA levels detected consistent 30-50% reductions at nearly all
274 genomic regions examined (Figure 5G). This was true of gene bodies, promoters, and CpG island
275 shores. CpG island sites, which show very low mCA levels in wild-type cortex, displayed less
276 reduction of mCA, possibly due to floor effects in bisulfite-sequencing (see *methods*).
277 Comparison of mCA changes in each class of genomic elements as a function of wild-type mCA
278 levels suggested that consistent reductions occurred across the genome independent of the
279 normal levels of mCA (Figure S5E). This suggests that changes in mCA levels in the DNMT3A$^{KO/+}$
280 do not preferentially impact specific classes of genomic elements or become substantially more
281 severe in some regions based on the level of mCA that normally is deposited.

282 Recent analysis has demonstrated that topologically-associating domains (TADs) of
283 chromatin folding are regions of organization for mCA that can impact gene regulation (Clemens
284 et al., 2019). The “set-point” level of mCA within TADs is associated with the level of mCA at
285 enhancers within TADs and high-mCA enhancers found in high-mCA TADs are particularly robust
286 targets of repression by MeCP2 (Clemens et al., 2019). We therefore specifically assessed mCA
287 levels at TADs and enhancers genome-wide. This analysis detected reductions in TAD mCA levels
288 that were similar to global reductions in mCA at other genomic elements (Figure 5G). Enhancers

289 also showed this pervasive depletion of mCA (Figure 5G). Like other genomic elements, these
290 effects were consistent for TADs and enhancers with differing wild-type levels of mCA (Figure
291 S5E). Thus, widespread loss of mCA for TADs and enhancer elements occurs in DNMT3A$^{KO/+}$ mice
292 and has the potential to impact epigenetic control of regulatory elements by MeCP2.
293
294 **Enhancer dysregulation results from methylation deficits in DNMT3A$^{KO/+}$ mice**
295 We next examined how disruption of DNA methylation can affect epigenetic regulation in
296 DNMT3A$^{KO/+}$ neurons to alter gene expression and disrupt nervous system function. Recent
297 analysis indicates that mCA serves as a binding site for MeCP2 to mediate neuron-specific gene
298 regulation, in part by controlling the activity of distal regulatory enhancer elements (Clemens et
299 al., 2019). Loss of MeCP2 in mice leads to genome-wide upregulation of the activating mark
300 Histone H3 lysine 27 acetylation (H3K27ac) at enhancers that contain high levels of mCA and mCG
301 sites, while overexpression of MeCP2 leads to reciprocal downregulation of highly methylated
302 sites. Alterations in enhancer activity in MeCP2 mutants are linked to dysregulation of genes that
303 can then drive nervous system dysfunction. These findings suggest that reduced CA methylation
304 in the DNMT3A$^{KO/+}$ would remove binding sites for MeCP2 within enhancers. This mCA reduction
305 could then result in dysregulation of enhancer activity that partially phenocopies the effects we
306 have observed in MeCP2 mutant mice.
307 To investigate this possibility directly, we quantified the change in mCA binding sites in
308 the DNMT3A$^{KO/+}$ for enhancers significantly repressed by MeCP2 (Clemens et al., 2019). These
309 enhancers contain a large number of mCA sites due to high mCA/CA levels and an enrichment of
310 CA dinucleotides within these sequences. As a result, we found that the global 30-50% reduction
311 of mCA in the DNMT3A$^{KO/+}$ leads to a larger loss in the total number of mCA sites at MeCP2-
312 repressed enhancers than at other enhancers genome-wide (Figure 6A,B). Thus MeCP2-
313 repressed enhancers are particularly susceptible to mCA binding site loss from heterozygous
314 mutation of DNMT3A.
315 To determine if the reduction of mCA sites at MeCP2-repressed enhancers affects their
316 activity, we assessed changes in enhancer activation level by H3K27ac ChIP-seq analysis of the
317 DNMT3A$^{KO/+}$ and wild-type cerebral cortex. This analysis revealed significant changes in

318 acetylation at MeCP2-repressed enhancers (Figure 6A,C). Consistent with these effects arising
319 from 30-50% loss of the mCA that normally builds up post-mitotically at enhancers, we detect
320 changes that are concordant with, but smaller than, those caused by complete loss of post-
321 mitotic mCA in the DNMT3A cKO (Clemens et al., 2019) (Figure 6C).

322 Although significantly dysregulated enhancers can be detected in MeCP2 mutants, broad
323 sub-significance-threshold effects also occur genome-wide upon MeCP2 mutation, with nearly
324 all enhancers across the genome undergoing dysregulation that is proportional to the number of
325 mC binding sites at these regions (Figure 6D) (Clemens et al., 2019). Analysis of H3K27ac changes
326 at enhancers based on the normal density of mCA sites in these sequences genome-wide
327 revealed broad mCA-associated derepression of enhancers in DNMT3A$^{KO/+}$ cortex that is similar
328 to, but smaller in magnitude than, the effects observed in DNMT3A cKO and MeCP2 knockout
329 mice (MeCP2 KO). These effects are also reciprocal to effects observed in MeCP2 overexpression
330 mice (MeCP2 OE). Consistent with the limited disruption of mCG genome wide in the DNMT3A$^{KO/+}$
331 mice, there was more limited association between changes in enhancer activity and the level of
332 mCG at these sequences. This contrasts with MeCP2 mutants in which loss of protein binding at
333 both mCG and mCA sites leads to enhancer dysregulation that is associated with both mCA and
334 mCG (Clemens et al., 2019) (Figure 6D). Together, these findings demonstrate that loss of half of
335 the normal mCA sites in the DNMT3A$^{KO/+}$ cortex results in enhancer dysregulation that overlaps
336 with MeCP2 mutant mice, uncovering a role for shared neuronal chromatin pathology between
337 DNMT3A and MeCP2 disorders.

338

339 **Overlapping transcriptional pathology between DNMT3A$^{KO/+}$, MeCP2 disorders, and ASD**

340 The epigenetic alterations we observe in DNMT3A$^{KO/+}$ cerebral cortex can have direct
341 consequences on gene expression to drive neurological dysfunction in mice. Furthermore, the
342 overlapping effects on enhancers that we observe between DNMT3A$^{KO/+}$ and MeCP2 mutant
343 mice suggests that there may be shared transcriptional pathology occurring upon loss of mCA in
344 DNMT3A disorders and through disruption of MeCP2 in Rett syndrome and MeCP2-duplication
345 syndrome. We therefore assessed changes in gene expression in DNMT3A$^{KO/+}$ mice, interrogating
346 the extent to which these effects overlap with those observed in MeCP2 mutants and upon

347 complete disruption of mCA in the DNMT3A cKO. RNA-seq of DNMT3A$^{KO/+}$ cerebral cortex
348 identified subtle changes in gene expression that are consistent in magnitude with effects
349 observed in other heterozygous NDD models (Fazel Darbandi et al., 2018; Gompers et al., 2017;
350 Katayama et al., 2016) (Figure S6A). Gene set enrichment analysis on differential expression data
351 revealed dysregulation of Gene Ontology terms relating to neuronal development and function
352 (Figure S6B), suggesting that alterations in gene expression resulting from DNMT3A heterozygous
353 disruption could drive the behavioral alterations that we observed.

354 While a limited gene set is detected as significantly dysregulated in the DNMT3A$^{KO/+}$, we
355 considered if genome-wide alterations in enhancer activity could lead to wide-spread, subtle
356 dysregulation of gene expression that is below the threshold of detection for individual genes. In
357 this way, the transcriptional pathology in the DNMT3A$^{KO/+}$ brain could overlap with the
358 subthreshold genome-wide effects observed upon loss of neuronal mCA (DNMT3A cKO) and in
359 models of Rett syndrome (MeCP2 KO) and ASD (MeCP2 OE) (Clemens et al., 2019; Gabel et al.,
360 2015). Indeed, the significantly dysregulated genes in the DNMT3A$^{KO/+}$ overlapped extensively
361 with genes identified as significantly dysregulated in DNMT3A cKO and MeCP2 mutant mice
362 (Clemens et al., 2019), supporting the notion of shared gene expression effects between these
363 mouse models (Figure 6E). To more comprehensively assess the degree to which transcriptome-
364 wide changes in the DNMT3A$^{KO/+}$ phenocopy these MeCP2 mutant and DNMT3A cKO models, we
365 performed Generally Applicable Gene-set Enrichment (GAGE) analysis (Luo et al., 2009) of all
366 genes detected as dysregulated in these mutant models. This revealed highly significant,
367 concordant changes in gene expression in the DNMT3A$^{KO/+}$ for dysregulated gene sets detected
368 upon loss of mCA in the DNMT3A cKO and in MeCP2 mutant models (Figure 6F).

369 Having detected overlap in transcriptomic pathology between models of DNMT3A and
370 MeCP2 disorders, we sought to explore if shared gene expression signatures in the DNMT3A$^{KO/+}$
371 mice extend to models of disorders that do not have as clear mechanistic links to DNMT3A
372 disorders. We therefore tested if DNMT3A$^{KO/+}$ mice show significant alterations in gene sets
373 identified as dysregulated in other mouse models of NDD and human gene sets implicated as
374 altered in the autistic brain. GAGE analysis across multiple datasets detected highly significant
375 dysregulation of gene sets identified in CHD8 and PTEN mouse models of overgrowth and ASD

376 (Gompers et al., 2017; Katayama et al., 2016; Tilot et al., 2016) as well as the SetD5 model of NDD
377 (Sessa et al., 2019) (Figure 7A). These findings support a role for overlapping gene dysregulation
378 underlying common symptomology found in affected individuals carrying mutations in distinct
379 genes.

380 Analysis of human gene sets detected as dysregulated in ASD postmortem brains (Gandal
381 et al., 2018; Voineagu et al., 2011) showed significant changes in the DNMT3A$^{KO/+}$ cortex (Figure
382 7B). This analysis also indicated upregulation of candidate genes linked to ASD from human
383 genetics studies (Abrahams et al., 2013; Banerjee-Basu and Packer, 2010) (Figure 7B). In addition,
384 analysis of co-expression modules of human brain development (Parikshak et al., 2013) showed
385 overlap with several neurodevelopmental modules including those that increase during early
386 cortical development and are enriched for ASD risk genes (M13, M16, and M17) (Figure S6C).
387 Modules involved in regulation of nucleic acids and gene regulation that are expressed early in
388 development and decrease over time are also increased upon heterozygous loss of DNMT3A (M2
389 and M3) (Figure S6C). These results indicate that important sets of genes with opposing
390 developmental trajectories and function are altered upon loss of DNMT3A regulation. Notably,
391 control resampling analysis indicated that significant dysregulation of these mouse and human
392 gene sets was not driven by enriched expression of these genes in the cortex (Figure S6D).
393 Together these findings suggest that the DNMT3A$^{KO/+}$ mouse shares overlapping transcriptional
394 pathology with gene expression changes underlying ASD.

395

396 **Discussion**

397 Our functional analysis of NDD-associated DNMT3A mutations together with our *in vivo*
398 studies provide an initial working model of molecular etiology in DNMT3A disorders. Diverse *de*
399 *novo* missense mutations that arise in affected individuals disrupt enzyme function by disabling
400 the capacity of the enzyme to localize to chromatin in the nucleus, altering the ADD-regulatory
401 domain, or disrupting the activity of the methyltransferase domain. Loss-of-function effects
402 resulting from these missense mutations, as well as early truncations or gene deletions, lead to
403 insufficient DNMT3A activity. This causes deficits in deposition of mCG at specific sites during
404 development and a massive deficit in postnatal mCA accumulation throughout the brain. These

405 changes in DNA methylation lead to alterations in epigenomic regulation, including subtle but
406 wide-spread disruption of mCA-MeCP2-mediated enhancer regulation in adult neurons, resulting
407 in gene expression changes that can drive deficits in nervous system function.

408 Our studies of DNMT3A mutations not only provide insight into the molecular etiology of
409 DNMT3A disorders, but also serve as a model for understanding the functional effects of diverse
410 *de novo* mutations underlying neurodevelopmental disorders. Exome sequencing studies have
411 identified a large and growing list of mutations in genes encoding epigenetic regulators in
412 individuals with NDD. Many of these are missense mutations and occur as heterozygous
413 disruptions (McRae et al., 2017; Satterstrom et al., 2019), leaving it unclear if simple loss-of-
414 function effects are sufficient to drive pathology through haploinsufficiency, or if more complex
415 effects play a role when individual amino acids are altered. In addition, while identification of
416 multiple mutations in a gene can implicate disruption of the gene as causative for NDD, it remains
417 possible that a subset of the mutations identified in affected individuals, particularly missense
418 mutations, are not in fact deleterious or causative. Functional testing of these variants is
419 therefore necessary to determine if they may underlie disease. Here, our analysis of DNMT3A
420 mutations in multiple functional assays has uncovered diverse mechanisms by which the protein
421 can be disrupted while pointing to a shared loss of function in the deposition of neuronal DNA
422 methylation. Notably, it is only by assessing multiple aspects of protein function (i.e. expression,
423 localization, activity, and cellular mCA levels) that we can detect deficits for each mutation tested.
424 For example, mutation of the ADD domain disrupts deposition of mCA, possibly due to loss of
425 regulation that can only be assessed in the endogenous chromatin context. Together our findings
426 establish the deleterious effects of diverse DNMT3A mutations and underscore the importance
427 of multidimensional analysis of *de novo* mutations to fully assess their potential role in NDD.

428 Our *in vivo* analyses show that heterozygous deletion of DNMT3A mirrors multiple key
429 features of DNMT3A disorders, including tall stature (increased long bone length), increased body
430 weight, and behavioral alterations. Detection of robust anxiety-like phenotypes in multiple
431 assays, deficits in pro-social communication, and alterations in repetitive behaviors align with
432 observed human phenotypes. In contrast, lack of strong deficits in learning and memory assays
433 in our mouse model may indicate that some regions and systems in humans are more susceptible

434 to DNMT3A disruption than in mice. However, we do detect alterations in behavior in these
435 assays (Figure 3K-M, Figure S4O-T) and the lack of strong deficits may also reflect insensitivity of
436 the methods used to measure specific aspects of disrupted cognition. In all, our *in vivo* analysis
437 indicates that heterozygous deletion of DNMT3A results in effects which can guide future studies
438 of molecular, cellular, and organismal dysfunction caused by mutation of DNMT3A.

439 We employed the DNMT3A$^{KO/+}$ mouse experimental system to assess how heterozygous
440 DNMT3A disruption impacts epigenetic regulation in the brain. Our analysis of DNA methylation
441 in tissues from DNMT3A$^{KO/+}$ mice detected very subtle changes in genome-wide mCG levels
442 across brain regions, with no global mCG effects in non-neural tissue (Figure 4A). Analysis of local
443 changes in mCG in the brain detected evidence of disrupted CG methylation at sites methylated
444 during postnatal development (i.e. adult hyper CG-DMRs). In addition, multiple hypo-CG-DMRs
445 can be detected at regulatory elements including enhancers. While limited, these effects have
446 the potential to alter gene expression and contribute to neurological alterations in these mice.
447 The limited nature of mCG effects is likely due to the redundant function of the other DNA
448 methyltransferases. The maintenance methyltransferase DNMT1 has the capacity to preserve
449 existing mCG patterns during cell divisions (Jeltsch et al., 2018). In addition, the *de novo*
450 methyltransferase DNMT3B is expressed with DNMT3A in many tissues during early development
451 and could provide critical redundancy for mCG patterning (Okano et al., 1999). Nonetheless, the
452 site-specific changes in mCG are also likely to occur in early development and in non-neural
453 tissues. For example, constitutive heterozygous deletion of DNMT3A has been shown to disrupt
454 mCG patterns in the blood and alter hematopoietic lineages (Cole et al., 2017). These changes in
455 mCG may contribute to changes in growth and other phenotypes observed in mice and humans.

456 In contrast to mCG, we detect a global reduction in mCA to approximately 30-50% of wild-
457 type levels in DNMT3A$^{KO/+}$ cortex, striatum, cerebellum, and hippocampus (Figure 4B). These
458 results generalize and extend findings in the hypothalamus (Sendžikaitė et al., 2019),
459 demonstrating the susceptibility of broad neuronal types and circuits to heterozygous loss of
460 DNMT3A. The susceptibility of mCA to heterozygous loss of DNMT3A is likely due to several
461 related factors. For example, DNMT3B is not expressed in postnatal neurons (Lister et al., 2013a),
462 and DNMT1 is not capable of depositing mCA (Jeltsch et al., 2018), making all mCA build-up in

463 neurons dependent on DNMT3A. In addition, the enzyme has slow kinetics for activity on CA sites
464 (Zhang et al., 2018) and deposition of mCA genome-wide by DNMT3A must take place in a
465 restricted time window (1-6 weeks) when the enzyme is highly expressed and active in neurons
466 (Clemens et al., 2019; Lister et al., 2013a; Stroud et al., 2017a). These constraints may make
467 enzyme levels limiting for mCA accumulation in neurons, providing an explanation for why global
468 mCA in the brain is sensitive to DNMT3A gene dosage. Notably, our findings suggest that
469 manipulations that activate the remaining DNMT3A, or prolong its high early postnatal
470 expression, might rescue deficits in mCA deposition. Conversely, duplication of the DNMT3A gene
471 could result in too much deposition of mCA and possibly cause significant neural dysfunction akin
472 to those effects seen in MeCP2 duplication disorder. Future studies can assess the feasibility of
473 rescue approaches and explore if DNMT3A duplication alters brain function.

474 Our analysis of chromatin changes downstream of altered DNA methylation has
475 uncovered a striking point of shared molecular disruption across models of DNMT3A disorders,
476 Rett syndrome, and MeCP2 duplication syndrome. While the clinical profile and pathophysiology
477 of DNMT3A disorders is clearly distinct from MeCP2 disorders, we have shown here that loss of
478 approximately a quarter of MeCP2 binding sites across the neuronal genome in the DNMT3A$^{KO/+}$
479 cortex results in subtle but wide-spread disruption of mCA-associated enhancer regulation that
480 partially phenocopies loss of MeCP2. This enhancer dysregulation can be linked to shared
481 alterations in gene expression across these models (Clemens et al., 2019) (Figure 6). Given the
482 critical roles of MeCP2-regulated genes for nervous system function (Gabel et al., 2015; Lagger
483 et al., 2017; Lyst and Bird, 2015), these epigenomic and transcriptomic effects likely contribute
484 to aspects of neurologic dysfunction observed in DNMT3A disorders. The persistence of many
485 mCA and mCG binding sites for MeCP2 in the DNMT3A$^{KO/+}$ may partially explain how DNMT3A
486 mutations manifest with less severe symptomology than in Rett Syndrome. In addition, absence
487 of DNMT3A early in prenatal development can contribute to overgrowth and other non-
488 overlapping aspects of DNMT3A and MeCP2 disorders. Together, our findings show that
489 disruption of mCA-MeCP2 mediated enhancer regulation likely contributes to three disorders
490 with distinct symptomology, defining a site of convergent molecular etiology underlying
491 heterogeneous clinical syndromes.

492 Our transcriptomic analysis of changes of ASD/NDD gene sets in DNMT3A mice has
493 further detected overlap with NDD beyond MeCP2 disorders, including both mouse models of
494 NDD/ASD (CHD8) and gene sets identified in human idiopathic ASD. As additional transcriptomic
495 studies of mouse models and human NDD brain emerge, systematic analyses of gene expression
496 effects can identify shared aspects of transcriptional pathology that can contribute to cognitive
497 and social deficits across diverse causes of NDD. Notably, the large number of chromatin
498 modifying enzymes mutated in these disorders raises the possibility that shared transcriptomic
499 effects emerge from common chromatin pathology. Our study has identified alterations in mCA
500 and enhancer regulation as a potential site of convergent dysfunction in MeCP2 and DNMT3A
501 disorders. Future studies may identify additional gene disruptions in which alterations in mCA
502 and enhancer dysregulation contribute to molecular pathology, expanding the role of
503 "methylopathies" in neurodevelopmental disease.
504

505 **Acknowledgements**
506 We thank the Division of Comparative Medicine at Washington University in Saint Louis for their
507 assistance with mouse husbandry and veterinary support. We are grateful for sequencing
508 support from the Genome Technology Access Center and the Center for Genome Sciences and
509 Systems Biology Spike in Cooperative at Washington University in Saint Louis. We thank members
510 of the Dougherty lab including S. Maloney, K. McCullough, and M. Rieger for assistance in USV
511 experiments. We thank J. Edwards, J. Goodman, and J. Yi for critical feedback on the experimental
512 design and manuscript. We thank A. Smith and T. Ley for helpful discussions. This work was
513 supported by NIH 5T32GM00815133 to D.L.C, by NIH 5T32GM007067 and F31NS108574 to
514 A.W.C., and by the Klingenstein-Simons Fellowship Fund, the G. Harold and Leila Y. Mathers
515 Foundation, the Brain and Behavior Research Foundation, the Simons Foundation for Autism
516 Research Initiative, and NIMH R01MH117405 to H.W.G.
517

518 **Author Contributions**
519 D.L.C and D.Y.W. are joint first authors, as each led critical components of the project and
520 analysis. D.L.C., J.R.Ma., and Y.R.L. generated and analyzed *in vitro* biochemical data. Y.R.L. and
521 S.A.N. generated primary neuronal culture samples. D.L.C. and J.R.Ma. generated skeletal
522 samples. N.M.K. and C.A.H. carried out craniofacial analysis and D.L.C. carried out long bone
523 analysis. D.L.C., J.R.Ma., D.F.W., and J.D.D. carried out behavioral tests and analysis. D.L.C.,
524 J.R.Ma., J.R.Mo., Y.R.L., and A.W.C. generated genomic data. D.Y.W. developed analysis
525 algorithms and pipelines. D.L.C., D.Y.W., J.R.Mo, and A.W.C. completing genomic analyses.
526 H.W.G. conceived the project and H.W.G., D.L.C., and D.Y.W. designed the experiments. H.W.G.,
527 D.L.C., and D.Y.W. wrote the manuscript and all authors contributed to manuscript editing and
528 revisions.

529
530 **Declaration of interests**
531 The authors declare no competing interests.
532

533 **Figure Legends**
534 **Figure 1**
535 Disease-associated DNMT3A mutations disrupt distinct aspects of protein function.
536 (**A**) Schematic of human DNMT3A protein showing canonical domains and disease-associated
537 mutations identified in previous studies (Sanders et al., 2015; Tatton-Brown et al., 2018). (**B**)
538 Example immunoblot of DNMT3A mutant protein expression. (**C**) Example images of DNMT3A
539 protein immunocytochemistry from wild type and PWWP domain mutant. Scale bar = 20µm. (**D**)
540 Quantification of DNMT3A mutant protein localization (#, *P*<0.0001; *, *P*<0.05; n=6-16 images;
541 Generalized Linear Model test of percent nuclear expression per image for mutants compared to
542 WT with Bonferroni correction). (**E**) Schematic of *in vitro* methylation assay for DNMT3A mutant
543 proteins. (**F**) Activity of DNMT3A mutant proteins in the *in vitro* methylation assay. (***, *P*<0.001;
544 **, *P*<0.01; *, *P*<0.05; n=4-19; one-sample Student's T-Test from normalized WT mean of 1 with
545 Bonferroni correction). Bar graphs indicate mean with SEM error bars.
546
547 **Figure 2**
548 Disease-associated DNMT3A mutations prevent buildup of neuronal CA methylation.
549 (**A**) Schematic of DNMT3A functional analysis in primary culture neurons. Cortical neurons are
550 collected from DNMT3A$^{flx/flx}$ mice at E14.5 and cultured. After 3 days *in vitro* (DIV), neurons are
551 virally transduced with Cre recombinase and WT or mutant FLAG-tagged DNMT3A. On DIV 12.5,
552 DNA and RNA are collected. Equal DNMT3A mRNA expression is verified by qRT-PCR (Figure S1D)
553 and DNA is used for whole genome bisulfite sequencing analysis. (**B**) Relative mCA amount
554 compared to Cre only and Cre+WT DNMT3A controls (#, *P*<0.0001; ***, *P*<0.001; **, *P*<0.01; *,
555 *P*<0.05; n=4-11; one-sample Student's T-Test from normalized WT mean of 1 with Bonferroni
556 correction). Bar graphs indicate mean with SEM error bars.
557
558 **Figure 3**
559 Heterozygous disruption of DNMT3A *in vivo* leads to growth and behavioral alterations.
560 (**A**) Body weight of DNMT3A$^{KO/+}$ and WT mice at three developmental timepoints (Male *P*=0.038
561 genotype by age interaction effect, $F_{(2,50)}$=3.494, n=6-18; Female *P*=0.0032 genotype by age
562 interaction effect, $F_{(2,48)}$=6.498; Female *P*=0.0016 genotype effect, $F_{(1,48)}$=11.18, n=5-17; two-way
563 ANOVA). (**B**) Lengths of femur and tibia bones measured by dual X-ray imaging shown as standard
564 deviations from the WT mean for the DNMT3A$^{KO/+}$ mice (**, *P*<0.01; n=12; paired Student's T-
565 Test). (**C**) Total ambulations of mice during 1-hour open-field testing, split into 10-minute bins
566 (*P*=0.0008 effect by genotype, $F_{(1,46)}$=13.02, n=21,27; two-way repeated-measures ANOVA with
567 Sidak's multiple comparison test; *, *P*<0.05; ****, *P*<0.0001). (**D**) Number of rearing events of
568 mice during 1-hour open-field testing, split into 10-minute bins (*P*=0.0103 effect by genotype,
569 $F_{(1,46)}$=7.161, n=21,27; two-way repeated-measures ANOVA with Sidak's multiple comparison
570 test; *, *P*<0.05). (**E**) Schematic of open-field testing center and edge zones (left). Total time spent
571 in the center zone of field during open-field testing (right) (*P*=0.0075, n=21,27; unpaired

572 Student's T-Test). (**F**) Schematic of the elevated plus maze indicating closed and open arms (left),
573 and percent of time mice spent in the open arms compared to all arms during first day of testing
574 (right) (*P*=0.0069; n=33,39; unpaired Student's T-Test). (**G**) Example images of marble burying
575 assay (left) and quantification of marbles buried during 30 minutes of testing split into 5-minute
576 bins (right)(*P*=0.0374 effect by genotype, $F_{(1,25)}$=4.834, n=14,13; two-way repeated-measures
577 ANOVA with Sidak's multiple comparison test; *, *P*<0.05). (**H**) Schematic of 3-chamber task in
578 which subject mouse can freely explore apparatus containing a novel mouse or an empty cup
579 (left). Quantification of time spent in zones closest to each cup (right) (Empty, *P*=0.0026; Novel,
580 *P*=0.0095; n=33,39; unpaired Student's T-Test). (**I**) Schematic of 3-chamber task in which subject
581 mouse can freely explore apparatus and interact with novel mouse or familiar mouse (left). Time
582 spent in zones closest to novel mouse and familiar mouse (right) (Familiar, *P*=0.29; Novel, *P*=0.24;
583 n=33,39; unpaired Student's T-Test). (**J**) Number of ultrasonic calls from pup isolated from the
584 nest for 3-minute testing over developmental time points (Analysis run on days 5-9 as these were
585 timepoints in which all animals tested had data; *P*=0.0378 effect by genotype, $F_{(1,285)}$=4.355, n=9-
586 46; two-way ANOVA with Sidak's multiple comparisons test; *, *P*<0.05). (**K-M**) Percent time spent
587 freezing in (**K**) Conditioned fear training (Baseline: *P*=0.0071 effect by genotype, $F_{(1,50)}$=7.897;
588 Cue: *P*=0.0013 effect by genotype, $F_{(1,50)}$=11.7; n=26; two-way repeated-measures ANOVA with
589 Sidak's multiple comparisons test; **, *P*<0.01; ***, *P*<0.001), (**L**) contextual fear trials (*P*=0.0215
590 effect by genotype, $F_{(1,50)}$=5.633, n=26; two-way repeated-measures ANOVA), and (**M**) cued fear
591 trials (Baseline: *P*=0.0606 effect by genotype, $F_{(1,50)}$=3.685; Cue: *P*<0.0001 effect by genotype,
592 $F_{(1,50)}$=17.03; n=26; two-way repeated-measures ANOVA with Sidak's multiple comparisons test;
593 *, *P*<0.05; ****, *P*<0.0001). Line graphs and bar graphs indicate mean with SEM error bars. Box
594 plots contain 10th-90th percentiles of data, with remaining data represented as individual points.
595
596 **Figure 4**
597 Global DNA methylation levels upon heterozygous loss of DNMT3A.
598 (**A**) Global mCG levels in DNA isolated from tissues of 8-week old mice (left) (*, *P*<0.05; unpaired
599 Student's T-Test with Bonferroni correction), and developmental time course of global mCG
600 (right), as measured by sparse whole genome bisulfite sequencing (WGBS). (**B**) Global mCA levels
601 in DNA isolated from tissues of 8-week old mice (left) (***, *P*<0.001; *, *P*<0.05; unpaired
602 Student's T-Test with Bonferroni correction), and developmental time course of global mCA
603 (right), as measured by sparse WGBS (*P*<0.0001 effect by genotype, $F_{(1,27)}$=1024; *P*<0.0001 effect
604 by age $F_{(5,27)}$=884.6; n=3-4; two-way ANOVA). Line graphs indicate mean with SEM error bars.
605
606
607 **Figure 5**
608 High-resolution analysis of DNA methylation changes in the $DNMT3A^{KO/+}$ cerebral cortex.
609 (**A**) Genome browser views of mCA and mCG in WT and $DNMT3A^{KO/+}$ cerebral cortex as measured
610 by high-depth WGBS. Broad view showing global reduction in mCA (left). Grey dashed line in mCA
611 plots at 0.03 to facilitate visual comparison of global mCA levels between genotypes. Zoomed-in
612 view of a $DNMT3A^{KO/+}$ CG-hypo-DMR that overlaps an enhancer (center) and a $DNMT3A^{KO/+}$ CG-
613 hypo-DMR at a CpG-island shore that overlaps with an adult-specific DMR (right). WT H3K27ac
614 ChIP-seq signal (Clemens et al., 2019), peaks of enhancer-associated H3K4me1
615 (Stamatoyannopoulos et al., 2012), peaks of promoter-associated H3K4me3

616 (Stamatoyannopoulos et al., 2012), CpG islands, and gene annotations (Haeussler et al., 2019)
617 are shown below to illustrate overlap between DMRs and functional elements in the genome.
618 (**B,C**) Overlay of mCG signal for DMR regions shown in **A**. (**D**) Mean mCG/CG level per replicate
619 (top) and percent reduction, (DNMT3A$^{KO/+}$-WT)/WT, (bottom) in WT and DNMT3A$^{KO/+}$ cerebral
620 cortex across indicated classes of genomic regions. Adult-specific DMRs were identified in the
621 cortex (Lister et al., 2013a) (*, *P*<0.05; **, *P*<0.01; n=4 per genotype; paired Student's T-Test with
622 Bonferroni correction). (**E**) Heat map of CG DMRs called in the DNMT3A$^{KO/+}$ cortex. Biological
623 replicates (B1,B2) and technical replicates (R1,R2) are indicated. (**F**) Observed and expected (see
624 *methods*) overlap between DNMT3A$^{KO/+}$ cortex CG-DMRs and various genomic regions (****,
625 *P*<0.0001; ***, *P*<0.001; Fisher's Exact Test with Bonferroni correction). (**G**) Mean mCA/CA levels
626 per replicate (top) and percent reduction, (DNMT3A$^{KO/+}$-WT)/WT, (bottom) in WT and
627 DNMT3A$^{KO/+}$ cortex across indicated classes of genomic regions (*, *P*<0.05; **, *P*<0.01; ***,
628 *P*<0.001; n=4; paired Student's T-Test with Bonferroni correction). Box plots indicate median and
629 quartiles.
630
631 **Figure 6**
632 DNMT3A$^{KO/+}$ enhancer dysregulation and transcriptomic pathology overlaps with MeCP2
633 mutants.
634 (**A**) Genome browser view of DNA methylation and H3K27ac ChIP-seq data from WT and
635 DNMT3A$^{KO/+}$ cerebral cortex (top). Overlaid H3K27ac signal and mCA/CA levels at enhancer
636 regions highlighted in blue that were identified as dysregulated enhancers upon disruption of
637 mCA or MeCP2 (Clemens et al., 2019) (bottom). (**B**) Mean mCA sites/kb in WT and DNMT3A$^{KO/+}$
638 cortex (top) and number of mCA sites/kb lost in the DNMT3A$^{KO/+}$ cortex (bottom) for enhancers
639 significantly dysregulated in MeCP2 mutants (*, *P*<0.05; **, *P*<0.01; ***, *P*<0.001; ****,
640 *P*<0.0001; n=4; paired Student's T-Test with Bonferroni correction). (**C**) Boxplot of fold-change
641 in H3K27ac signal in DNMT3A cKO and the DNMT3A$^{KO/+}$ cortex for enhancers defined as
642 significantly dysregulated in MeCP2 mutants. (***, $P<10^{-8}$; **** $P<10^{-12}$; n=5 biological replicates
643 of DNMT3A$^{KO/+}$ and WT; Wilcoxon test) (**D**) Heatmap of changes in H3K27ac signal indicated
644 mutants across deciles of enhancers sorted by wild-type mCA or mCG sites (*P*<2.2e-16,
645 DNMT3A$^{KO/+}$ mCA/kb; *P*<2.2e-16, DNMT3A$^{KO/+}$ mCG/kb; Spearman Rho correlation). (**E**)
646 Observed versus expected overlap of significantly dysregulated genes (padj. < 0.1) in the
647 DNMT3A$^{KO/+}$ and genes dysregulated in DNMT3A cKO or MeCP2 mutant mice (***, *P*<1e-5; ****,
648 *P*<1e-10; hypergeometric test). (**F**) Significance of gene set expression changes in the indicated
649 direction in the DNMT3A$^{KO/+}$ cortex for GAGE analysis of gene sets identified as dysregulated in
650 DNMT3A cKO or MeCP2 mutant mice (Clemens et al., 2019). Note: legend is shared in **B** and **C.**
651 Box plots indicate median and quartiles.
652
653 **Figure 7**
654 Gene dysregulation in the DNMT3A$^{KO/+}$ overlaps with other ASD/NDD disorders.
655 (**A**) GAGE analysis of expression changes in DNMT3A$^{KO/+}$ for dysregulated gene sets identified in
656 studies of NDD mouse models (Gompers et al., 2017; Katayama et al., 2016; Sessa et al., 2019;
657 Tilot et al., 2016) (n=7 biological replicates of DNMT3A$^{KO/+}$ and WT). (**B**) GAGE analysis of
658 expression changes in DNMT3A$^{KO/+}$ for gene sets identified in studies of human ASD. Gene sets:
659 ASD module 12 (synaptic) and 16 (immune) identified in weighted-gene coexpression analysis of

human ASD brain (Voineagu et al., 2011), and ASD-dysregulated genes previously identified (Abrahams et al., 2013; Gandal et al., 2018).

Figure S1 Related to Figures 1 and 2.
(**A**) Example images of DNMT3A mutant protein immunocytochemistry. Scale bar = 20µm. (**B**) Full example immunoblot from Figure 1b with the truncated S312fs11x DNMT3A mutant protein indicated with an asterisk. (**C**) Quantification of immunoblot signal of DNMT3A (α-FLAG) from mutant proteins (***, *P*<0.001; n=6-30; unpaired Student's T-Test with Bonferroni correction). (**D**) Genome wide mCA levels over time in neuronal cortical cultures, as measured by sparse WGBS (*P*=0.0035 effect by time, $F_{(3,4)}$=29.45, n=2; one-way ANOVA). (**E**) Genome-wide mCA levels in DNMT3A mutant add-back cortical cultures (****, *P*<0.0001; *, *P*<0.05; n=7-11; planned unpaired Student's T-Tests with Bonferroni correction). (**F**) qRT-PCR of *Dnmt3a* normalized to *Actb* for samples chosen for WGBS. (****, *P*<0.0001; n=4-11; one-sample Student's T-Test with Bonferroni correction). Bar graphs indicate mean with SEM error bars.

Figure S2 Related to Figures 3,4,5,6 and 7.
(**A**) Normalized *Dnmt3a* mRNA and protein expression from 2-week cortices of DNMT3A$^{KO/+}$ and wild-type littermates (****, *P*<0.0001; *, *P*<0.05; mRNA n=5, protein n=4, unpaired Student's T-Test). (**B**) Protein expression of DNMT3A normalized to α-Tubulin measured by western blotting for cerebral cortex of DNMT3A$^{KO/+}$ and wild-type littermates over developmental time (*P*=0.0157 effect by genotype, $F_{(1,16)}$=7.303, n=2-5; two-way ANOVA). (**C**) Protein expression of DNMT3A normalized to α-Tubulin measured by western blotting for hippocampus, striatum, and cerebellum (*P*=0.0010 effect by genotype, $F_{(1,18)}$=15.48, n=4; two-way ANOVA). Line plots indicate mean with SEM error bars.

Figure S3 Related to Figure 3.
(**A,B**) Measurements of (**A**) femur and (**B**) tibia by dual x-ray imaging in WT and DNMT3A$^{KO/+}$ mice at 8, 20, and 25 weeks of age. Lines indicate mean with SEM error bars. (**C**) Example dual x-ray image of mouse body with femur and tibia indicated. (**D**) Example of reconstructed skull from µCT imaging with landmarks used for craniofacial analysis shown. Red line indicates distance that is significantly larger in the WT compared to the DNMT3A$^{KO/+}$, while blue line indicates distance that is significantly smaller in the WT compared to the DNMT3A$^{KO/+}$ (*P*<0.05). (**E,F**) Principal component analysis of (**E**) cranial and (**F**) mandibular shape shows no clear separation between groups along PC1 or PC2.

Figure S4 Related to Figure 3.
(**A-J**) Comparison of DNMT3A$^{KO/+}$ and WT mice across a battery of sensorimotor assays. DNMT3A$^{KO/+}$ mice show no significant difference in (**A**) walking initiation, or latency to fall off (**B**) ledge or (**C**) platform. (**D-E**) DNMT3A$^{KO/+}$ mice show no difference compared to WT littermates in motor coordination as evidenced by time on a continuous (**D**) and accelerating (**E**) rotarod. (**F**) DNMT3A$^{KO/+}$ mice show no difference in grip strength compared to WT littermates evidenced by no change in time on an inverted screen. (**G**) Mean % pre-pulse inhibition shows no significant difference between genotypes. DNMT3A$^{KO/+}$ mice show a significant increase in time to (**H**) climb down a pole (*P*=0.016, n=21,27; unpaired Student's T-Test), and to the top of a (**I**) 60° inclined

screen (*P*=0.039, n=21,27; unpaired Student's T-Test) or a (**J**) 90° inclined screen (*P*=0.045, n=21,27; unpaired Student's T-Test). (**K**) DNMT3A$^{KO/+}$ mice show no deficit in elevated plus maze exploration as measured by percent entries into open arms. (**L**) Percent preference for novel conspecific in the 3-chambered social approach task for WT and DNMT3A$^{KO/+}$ mice using time spent in zones as calculated: Mouse/(Mouse+Object)x100 or Novel/(Novel+Familiar)x100 within each animal. (**M**) Total distance traveled during the 3-chambered social approach task for WT and DNMT3A$^{KO/+}$ mice shows a broad reduction of distance traveled across all trials by DNMT3A$^{KO/+}$ mice (*P*=0.018 effect by genotype, $F_{(1,70)}$=5.862, n=33,39; two-way ANOVA). (**N**) Shock sensitivity during conditioned fear test as indicated by the minimum shock needed to exhibit a behavioral response in mice shows no significant difference between genotypes. (**O-R**) Path distance to escape platform and swim speeds in the Morris water maze task. DNMT3A$^{KO/+}$ mice show increased path distance to escape platform in both (**O**) cued trials (*P*=0.0012 effect by genotype, $F_{(1,46)}$=11.93; *P*=0.0433 interaction effect of genotype and trial block, $F_{(3,138)}$=2.784; n=21,27; two-way repeated-measures ANOVA) and (**P**) place trials (*P*=0.0408 interaction effect of genotype and trial block, $F_{(4,184)}$=2.55 n=21,27; two-way repeated-measures ANOVA). No significant difference is seen in swimming speed during (**Q**) cued trials (*P*=0.0634 effect by genotype, $F_{(1,46)}$=3.619 n=21,27; two-way repeated-measures ANOVA) and (**R**) place trials (*P*=0.098 effect by genotype, $F_{(1,46)}$=2.845, n=21,27; two-way repeated-measures ANOVA). (**S**) DNMT3A$^{KO/+}$ mice show no significant difference in time spent in the target quadrant of a Morris water maze compared to WT littermates. (**T**) DNMT3A$^{KO/+}$ mice show a trend towards a reduction in platform crossings in the probe trial (*P*=0.0609; unpaired Student's T-Test). Bar graphs and line plots indicate mean with SEM error bars. Box plots contain 10th-90th percentiles of data, with remaining data represented as individual points.

Figure S5 Related to Figures 4,5, and 6.

(**A**) Boxplots of CG methylation over postnatal development at regions called as having higher methylation in the frontal cortex of fetal versus adult tissue (Fetal-Specific DMR), or called as having higher methylation in the frontal cortex of the adult versus fetal tissue (Adult-specific DMR). DMRs from Lister et al. 2013, methylation data from Stroud et al. 2017. (#, *P*<2.2e-16; Wilcoxon rank sum test with Bonferroni correction) (**B**) Boxplots of cortical methylation of WT and DNMT3A cKO at 8 weeks postnatal within developmental DMRs shown in **A** (#, *P*<2.2e-16; Wilcoxon rank sum test with Bonferroni correction). (**C**) Boxplots of cortical methylation of WT and DNMT3A$^{KO/+}$ at 8 weeks postnatally within DMRs defined in the DNMT3A$^{KO/+}$ model. DMRs are called on this data set, so no additional statistics were run on genotype differences. (**D**) mCG/CG (left) and hmCG/CG (right) from WT and DNMT3A$^{KO/+}$ cortices in various genomic contexts as measured by oxidative bisulfite sequencing. CpG islands were obtained from the UCSC table browser (Haeussler et al., 2019), and CpG Shores were defined as the 8kb surrounding them (*, *P*<0.05; paired Student's T-Test with Bonferroni correction). (**E**) Smooth scatter plots of WT and DNMT3A$^{KO/+}$ mCA/CA for classes of genomic regions. Box plots indicate median and quartiles.

Figure S6 Related to Figures 6 and 7.

(**A**) Volcano plot of DESeq $\log_2$ fold changes of the DNMT3A$^{KO/+}$ versus WT. Genes reaching a significance of $p_{adj.}$<0.1 are colored in red. (**B**) Top ten up- and down-regulated Gene Ontology

748 terms from Broad GSEA Molecular Signatures Database version 7.0 (Subramanian et al., 2005).
749 All terms are significant at an FDR<0.1. (**C**) GAGE analysis of developmental expression modules
750 (Parikshak et al., 2013). Significant modules (q-value<0.1) are colored in red (left). Expression
751 matched resampling of each gene set was performed 1,000 times and analyzed using GAGE for
752 enrichment in DNMT3A$^{KO/+}$ fold-change data (gray violin). This was compared with the true gene
753 set p-value (red point) to test for significance (right). Only the direction of dysregulation in which
754 the gene sets showed significance (i.e. DNMT3A$^{KO/+}$ greater or less) is shown. (**D**) Expression
755 matched resampling of GAGE analysis for gene sets displayed in Figures 6 and 7. Only the
756 direction of dysregulation in which the gene set showed significance (i.e. DNMT3A$^{KO/+}$ greater or
757 less) is shown.
758
759
760 **STAR Methods**

761 **Key Resources Table**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-Histone H3 (acetyl K27)	Abcam	Cat# ab4729; RRID:AB_2118291
Rabbit Anti-DDDDK tag Polyclonal Antibody	Abcam	Cat#: ab1162; RRID:AB_298215
Mouse Anti-Dnmt3a Monoclonal Antibody, Clone 64B1446	Abcam	Cat#: ab13888 RRID: AB_300714
alpha Tubulin antibody [EP1332Y]	Abcam	Cat#: ab52866 RRID: AB_869989
Goat Anti-Rabbit IgG (H+L) Antibody, Alexa Fluor 568	ThermoFisher	Cat#: A-11011 RRID: AB_143157
IRDye 800CW Donkey anti-Rabbit IgG antibody	LI-COR Biosciences	Cat#: 926-32213 RRID: AB_621848
IRDye 800CW Goat anti-Rabbit IgG antibody	LI-COR Biosciences	Cat#: 926-32211 RRID: AB_621843
IRDye 800CW Goat anti-Mouse IgG antibody	LI-COR Biosciences	Cat#: 926-32210 RRID: AB_621842
Bacterial and Virus Strains		
Biological Samples		
Chemicals, Peptides, and Recombinant Proteins		
Critical Commercial Assays		
RNeasy Mini Kit	Qiagen	Cat#: 74104
DNeasy Kit	Qiagen	Cat#: 69504
Epitect Bisulfite Kit	Qiagen	Cat#: 59824
AllPrep DNA/RNA Kit	Qiagen	Cat#: 80204
NEBNext Ultra Directional RNA Library Prep Kit for Illumina	NEB	Cat#: E7420S
NEBNext Mulitplex Oligos for Illumina (Index Primers Set 1)	NEB	Cat#: E7335S
NEBNext rRNA Depletion Kit (Human/Mouse/Rat)	NEB	Cat#: E6310L
Ovation Ultralow Library System V2	NuGEN	Cat#: 0344-32
Ovation Ultralow Methyl-Seq Kit	NuGEN	Cat#: 0335-32

TrueMethyl oxBS plugin	NuGEN	Cat#: 0414-32
High-Capacity cDNA Reverse Transcription Kit	ThermoFisher	Cat#: 4368814
Power SYBR™ Green PCR Master Mix	ThermoFisher	Cat#: 4368577
Deposited Data		
RNA-sequencing data	This paper	GEO:
ChIP-sequencing data (H3K27ac)	This paper	GEO:
Bisulfite-sequencing data	This paper	GEO:
RNA, ChIP, Bisulfite-sequencing data	(Clemens et al., 2019)	GEO: GSE123373
DNMT3A ChIP-seq, Bisulfite-sequencing data	(Stroud et al., 2017b)	GEO: GSE104298
Dixon Hi-C contact matrices	(Dixon et al., 2012)	http://chromosome.sdsc.edu/mouse/hi-c/cortex.norm.tar.gz
Bisulfite-sequencing data	(Lister et al., 2013b)	GEO: GSE47966
Mus musculus mm9 genome assembly	UCSC	http://hgdownload.soe.ucsc.edu/goldenPath/mm9/
Ensembl gene models	UCSC	https://genome.ucsc.edu/cgi-bin/hgTables
Experimental Models: Cell Lines		
HEK293T	ATCC	Cat#: HEK293T
Neuro-2a	ATCC	Cat#: CCL-131
Experimental Models: Organisms/Strains		
Mouse: DNMT3A$^{KO/+}$	Derived from DNMT3A fl/fl strain (described in (Kaneda et al., 2004) provided by M. Goodell) and CMV-Cre (IMSR Cat# JAX:006054, RRID:IMSR_JAX:006054) mice obtained from the Jackson Laboratory. After DNMT3A heterozygous deletion, Cre negative mice were backcrossed to pure C57BL6 mice (IMSR Cat# JAX:000664, RRID:IMSR_JAX:000664) obtained from the Jackson Laboratory.	JAX: 006054; RRID:IMSR_JAX:006054 JAX: 000664; RRID:IMSR_JAX:000664
Mouse: *Dnmt3a*$^{fl/fl}$	(Kaneda et al., 2004)	Provided by M. Goodell
Oligonucleotides		
Actb Forward	IDT	AAGGCCAACCGTGAAAAGAT
Actb Reverse	IDT	GTGGTACGACCAGAGGCATAC
Dnmt3a Forward	IDT	GGCCTTCTCGACTCCAGATG
Dnmt3a Reverse	IDT	TTCCTCTTCTCAGCTGGCAC

Recombinant DNA		
Software and Algorithms		
DESeq2 (v1.14.1)	(Love et al., 2014)	http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html
edgeR (v3.16.5)	(Robinson et al., 2010)	https://bioconductor.org/packages/release/bioc/html/edgeR.html
SAMtools (v1.3)	(Li and Durbin, 2009)	https://sourceforge.net/projects/samtools/files/
BEDtools2 (v2.25.0)	(Quinlan and Hall, 2010a)	https://github.com/arq5x/bedtools2
Bowtie2 (v2.2.5)	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
STAR	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
fastQC		https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
MACS2 (v2.1.0)	(Zhang et al., 2008)	https://github.com/taoliu/MACS
Trim galore		https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
Tadtree	(Weinreb and Raphael, 2016)	http://compbio.cs.brown.edu/projects/tadtree/
BS-seeker2	(Guo et al., 2013)	https://github.com/BSSeeker/BSseeker2
GREAT	(McLean et al., 2010)	http://great.stanford.edu
BSmooth	(McLean et al., 2010)	https://www.bioconductor.org/packages/release/bioc/html/bsseq.html
Gene Set Enrichment Analysis	Broad Institute	http://software.broadinstitute.org/gsea/downloads.jsp
Molecular Signatures Database	Broad Institute	http://software.broadinstitute.org/gsea/msigdb
ImageJ		https://imagej.nih.gov/ij/
Avizo		http://www.vsg3d.com/

Stratovan Checkpoint		https://www.stratovan.com/products/checkpoint
MorphoJ	(Dobin et al., 2013)	https://morphometrics.uk/MorphoJ_page.html
GraphPad Prism 7.03a		https://www.graphpad.com/

762

763 **Contact for Reagent and Resource Sharing**

764

765 Requests for reagents and resources should be directed towards the Lead Contact, Harrison
766 Gabel (gabelh@wustl.edu). This study did not generate new unique reagents.

767

768 **Experimental Model and Subject Details**

769

770 **Animal Husbandry**

771 All animal protocols were approved by the Institutional Animal Care and Use Committee and the
772 Animal Studies Committee of Washington University in St. Louis, and in accordance with
773 guidelines from the National Institutes of Health (NIH). Mice were housed in a room on a 12:12
774 hour light/dark cycle, with controlled room temperature (20-22°C) and relative humidity (50%).
775 Home cages measured 28.5 cm x 17.5 cm x 12 cm and were supplied with corncob bedding and
776 standard laboratory chow and water. All mice were group-housed and adequate measures were
777 taken to minimize animal pain or discomfort.

778

779 **Transgenic animals**

780 Male and female homozygous *Dnmt3a*$^{flx/flx}$ mice (Kaneda et al., 2004) were bred together for
781 viral-mediated DNMT3A replacement assay culture experiments. To generate the DNMT3A
782 heterozygous mouse model, *Dnmt3a*$^{flx/flx}$ mice were crossed to CMV:Cre (B6.C-Tg(CMV-
783 cre)1Cgn/J) to generate *Dnmt3a*$^{KO/+}$Cre:CMV$^{+/-}$ offspring. *Dnmt3a*$^{KO/+}$Cre:CMV$^{+/-}$ progeny were
784 bred to C57BL/6J to outcross the cre recombinase and generate experimental genotype
785 (DNMT3A$^{KO/+}$). Subsequent experimental animals were generated from *Dnmt3a*$^{KO/+}$ males mated
786 to C57BL/6J females to generate *Dnmt3a*$^{KO/+}$ and *Dnmt3a*$^{+/+}$ experimental and control animals
787 for experiments. *Dnmt3a*$^{KO/+}$ females were not used for breeding to avoid social differences in
788 mothering from mutant dams. Mice were genotyped with ear-DNA by PCR for *Dnmt3a* and *Cre*,
789 and recombination was tested. Mice were weighed at a variety of timepoints to assess growth.

790

791 **Method Details**

792 **Immunocytochemistry**

793 *Staining.* Neuro-2a cells (ATCC, CCL-131) were grown on coverslips and transfected with FLAG-
794 tagged WT or mutant mouse DNMT3A plasmids and GFP plasmid. Coverslips were fixed with 4%
795 paraformaldehyde in PBS for 20 minutes at room temperature, permeabilized with 0.1% Triton
796 X-100 in PBS for 10 minutes at room temperature, and blocked with 1% BSA in PBS for 1 hour at
797 room temperature. Coverslips were incubated overnight at 4˚C in anti-DDDDK tag (FLAG-tag)
798 primary antibody (Abcam, 1:5000, ab1162). Coverslips were then washed in PBS and incubated

799 for 1 hour at room temperature with fluorescent secondary antibody (ThermoFisher, 1:500, A-
800 11011) and counterstained with DAPI. *Imaging.* Images were captured using a Nikon A1Rsi
801 confocal microscope with a 20x air objective. Laser settings were kept constant for each image.
802 *Analysis/Quantification.* Cells were counted using an automatic threshold in FIJI and manually
803 classified as displaying nuclear or non-nuclear signal by a blinded observer. This was determined
804 by evaluating the overlap of FLAG signal (DNMT3A) with DAPI signal (nucleus). For mutants that
805 did not reach expression levels comparable to the WT or for images that had too few positive
806 cells, cell number was counted manually. 8 separate transfections were run, with each mutant
807 being counted over 3 or more independent experiments. Sample sizes are as follows: WT, 15
808 images, 880 cells; W297del, 8 images, 435 cells; I310N, 8 images, 492 cells; S312fs11x, 12 images,
809 321 cells; G532S, 9 images, 695 cells; M548K, 7 images, 333 cells; V665L, 6 images, 635 cells;
810 Y735C, 16 images, 613 cells; R749C, 8 images, 667 cells. P904L, 7 images, 692 cells. Percent
811 nuclear was assessed per image and a generalized linear model was run comparing each mutant
812 to WT. *P* values for each mutant were then Bonferroni corrected. We chose to use a generalized
813 linear model with Bonferroni correction to allow for us to compare ratios of percent nuclear
814 signal while taking into account experimental and biological replicates.
815
816 **Modeling of DNMT3A disease mutations**
817 HEK293T (ATCC, ACS-4500) or Neuro-2a cells (ATCC, CCL-131) were transfected with GFP and
818 FLAG-tagged WT or mutant mouse DNMT3A plasmids. Collected cell lysates were ruptured by 3
819 freeze/thaw cycles using liquid nitrogen, or sonication ~42 hours after transfection. Samples
820 were then either used for western blotting, the *in vitro* radioactive methyltransferase assay, or
821 RNA isolation for qRT-PCR.
822
823 **qRT-PCR**
824 RNA was reverse transcribed using the High-Capacity cDNA Reverse Transcription Kit
825 (Applied Biosystems). *Dnmt3a* and *Actb* were measured by qPCR using the Power SYBR™ Green
826 PCR Master Mix and primers *Actb* (F:AAGGCCAACCGTGAAAAGAT,
827 R:GTGGTACGACCAGAGGCATAC) and *Dnmt3a* (F:GGCCTTCTCGACTCCAGATG,
828 R:TTCCTCTTCTCAGCTGGCAC). Relative quantity of *Actb* and *Dnmt3a* cDNA was determined by
829 comparing the Ct of each primer set in each sample to a standard curve and then normalizing the
830 DNMT3A signal by the ACTB signal. We chose to compare experimental conditions to WT samples
831 using Student's T-Tests with Bonferroni correction, as these are two normally distributed groups
832 (visually checked) with similar variability.
833
834 ***In vitro* radioactive methyltransferase assay**
835 30 µl of cell lysate was used in the methyltransferase reaction previously described (Russler-
836 Germain et al., 2014). Lysates were incubated at 37˚C for 20 hours in 5 µl reaction buffer of 20
837 mM HEPES, 30 mM NaCl, 0.5 mM DTT, 1 mM EDTA, 0.2 mg/ml BSA, 5 mM 3 H-labeled SAM
838 (PerkinElmer, NET155050UC) and 500 ng/µl Poly(dI-dC) substrate (Sigma P4929). Substrate was
839 purified (Macherey-Nagel NucleoSpin Gel and PCR Clean-up) and radioactivity measured using a
840 scintillation counter. In instances where DNMT3A mutant showed altered protein expression, cell
841 lysate was re-balanced to match protein expression of WT DNMT3A. Only experimental replicates
842 where WT DNMT3A showed a 1.5-fold increase compared to GFP alone were used for

843 subsequent analysis. Outliers beyond 2 standard deviations above or below the mean were
844 removed. Number of independent replicates are as follows: W297del, 18; I310N, 19; S312fs11x,
845 4; G532S, 10; M548K, 15; V665L, 11; Y375C, 13; R749C, 7; P904L, 14. Significance was assessed
846 using a one-sample student's t-test, as we are comparing groups normalized to WT and GFP back
847 to the normalized value of 1.
848
849 **Viral-mediated DNMT3A replacement assay**
850 Functional activity of DNMT3A mutants in cortical neurons was determined by measuring
851 methylation build-up *in vitro*. Timed mating of DNMT3A$^{flx/flx}$ females and DNMT3A$^{flx/flx}$ males was
852 performed to collect embryonic cortical DNMT3A$^{flx/flx}$ neurons at embryonic day 14.5. At E14.5,
853 DNMT3A$^{flx/flx}$ cortical neurons were isolated and plated (DIV 0). On DIV 3, neurons were either
854 not perturbed or virally transduced with one of three conditions: 1) Cre only, 2) Cre and WT
855 DNMT3A, or 3) Cre and mutant DNMT3A. DNA and RNA were isolated on DIV 12.5 using the
856 AllPrep DNA/RNA Kit (Qiagen, 80204). DNA was used for whole genome bisulfite sequencing, and
857 RNA was used for qRT-PCR for DNMT3A. Number of independent replicates are as follows:
858 W297del, 5; S312fs11x, 4; G532S, 7; M548K, 6; V665L, 10; Y375C, 11; R749C, 6; P904L, 8. Buildup
859 of methylation over development was done using timed mating of C57BL/6J mice and collected
860 and prepared as described above. Significance was assessed using a one-sample student's t-test,
861 as we are comparing groups normalized to WT and GFP back to the normalized value of 1.
862
863 **Ultrasonic vocalization and analysis**
864 A total of 76 DNMT3A$^{KO/+}$ (n=30, 16 male and 14 female) and litter-matched WT (n=46, 25 male
865 and 21 female) mice were used for ultrasonic pup vocalization (USV) recording and analysis as
866 previously described (Barnes et al., 2017; Dougherty et al., 2013; Holy and Guo, 2005). Dams
867 were removed from the nest for a 10-minute acclimation, and individual pups had their body
868 temperature measured using an infrared laser thermometer. Pups were then removed from their
869 nest and placed in a dark, enclosed chamber. Ultrasonic vocalizations were recorded for 3
870 minutes with an Avisoft UltraSoundGate CM16 microphone and 416H amplifier using Avisoft
871 Recorder software (gain = 6 dB, 16 bits, sampling rate = 250 kHz). Following this, pups were
872 weighed and returned to their nest and littermates. All mice were recorded at postnatal days 5,
873 7, and 9, and on either day 11 or 15. Because not all animals were recorded from on day 11 or
874 15, only days 5, 7, and 9 were used for repeated measures ANOVA analysis. Frequency sonograms
875 were prepared and analyzed in MATLAB (frequency range = 40 kHz to 120 kHz, FFT size = 256,
876 overlap = 50%) with individual syllables identified and counted according to previously published
877 methods (Dougherty et al., 2013; Holy and Guo, 2005). Significance was assessed using a within-
878 subjects repeated measures ANOVA over timepoints 5-9, as these were when there was data
879 from all experimental subjects, and these are optimal testing times where number of calls was
880 highest.
881
882 **Marble burying**
883 A total of 27 DNMT3A$^{KO/+}$ (n=13, 8 male and 5 female) and litter-matched WT (n=14, 7 male and
884 7 female) mice were used for marble burying. Marble burying is a natural murine behavior and
885 has been used to indicate repetitive digging as well as anxiety-related behaviors. Protocol was
886 adapted from previously published methods (Lazic, 2015; Maloney et al., 2019a). In brief, 8-week

old mice were placed in a transparent enclosure (28.5 cm x 17.5 cm x 12 cm) with clean aspen bedding and 20 dark blue marbles evenly spaced in a 4 x 5 grid on top of the bedding. Animals were allowed to explore freely for 30 minutes. The number of buried marbles were counted every 5 minutes by two independent blinded observers. Marbles were considered "buried" if they were at least two-thirds covered by bedding. Enclosure and marbles were cleaned thoroughly between animals. Significance was assessed using a within-subjects repeated measures ANOVA to determine if rate of burying marbles is different between genotypes. These statistical methods are more appropriate than a simple t-test at 30 minutes, as mice may have buried all marbles before this timepoint, and significant changes in marble burying behavior may have occurred at an earlier timepoint in the assay.

Adult behavioral battery

A total of 72 DNMT3A$^{KO/+}$ (n=39, 18 male and 21 female) and litter-matched WT (n=33, 15 male and 18 female) mice were used for adult behavioral testing. Mice were housed in mixed genotype home cages with 2-5 animals per cage, and all tests were performed during the light cycle. All experimenters were blinded to genotype during testing. For increased experimental rigor and reproducibility, we used three separate cohorts of mice to ensure quality and consistency in any observed phenotypes.

Testing started when mice were 3-4 months of age. The sequence of behavioral testing was designed to minimize carry-over effects across behavioral tests. Cohorts, ages, and testing order are in Supplementary Table 4. The majority of the assays were performed on cohorts 1 and 2 with cohort 3 being performed to test for reproducibility in some assays (Supplementary Table 4). Because of differences in testing sequences and exposure of mice to prior tests between cohorts, we examined separate cohorts individually and looked at combined cohorts (Supplementary Table 4). Testing was performed by the Washington University in St. Louis Animal Behavior Core.

One-hour locomotor activity

Locomotor activity was evaluated by computerized photobeam instrumentation in transparent polystyrene enclosures (47.6 cm x 25.4 cm x 20.6 cm) as previously described (Wozniak et al., 2004). Activity variables such as ambulations and vertical rearings were measured in addition to time spent in a 33 cm x 11 cm central zone.

Sensorimotor battery

Mice were assayed in walking initiation, balance (ledge and platform tests), volitional movement (pole and inclined screens), and strength (inverted screen) as previously described (Grady et al., 2006; Wozniak et al., 2004). For the walking initiation test, mice were placed on the surface in the center of a 21 cm x 21 cm square marked with tape and the time for the mouse to leave the square was recorded. During the balance tests, the time the mouse remained on an elevated plexiglass ledge (0.75 cm wide) or small circular wooden platform (3.0 cm in diameter) was recorded. During the Pole test, mice were placed at the top of a vertical pole with nose pointing upwards. The time for the mouse to turn and climb down the pole was recorded. For the inclined screen tests, a mouse was placed (oriented head-down) in the middle of an elevated mesh grid

measuring 16 squares per 10 cm angled at 60° or 90°. Time for the mouse to turn 180° and climb to the top was recorded. For the inverted screen test, a mouse was placed on a similar screen and when the mouse appeared to have a secure grasp of the screen, the screen was inverted 180° and the latency for the mouse to fall was recorded. All tests had a duration of 60 seconds, except for the pole test which was 120 seconds. Two separate trials were done on subsequent days and averaged time of both trials was used for analysis. Data from the walking initiation, ledge, and platform tests were not normally distributed and therefore analyzed using Mann-Whitney U tests.

Continuous and accelerating rotarod

Motor coordination and balance were assessed using the rotarod test (Rotamex-5, Columbus Instruments, Columbus, OH) with three conditions: a stationary rod (60-second maximum), a rotating rod at constant 5 rpm (60-second maximum), and a rod with accelerating rotational speed (5 – 20 rpm, 180-second maximum) as previously described (Grady et al., 2006). This protocol is designed to minimize learning and instead measure motor coordination, so testing sessions were separated by 4 days to allow for extinction. Testing included one trial on stationary rod, and two trials on both the constant-speed rotarod and accelerating rotarod. Later timepoints in the constant speed rotarod test failed tests of normality, as the majority of mice stayed on the rotating rod for all 60 seconds. However, data were analyzed with two-way repeated-measures ANOVA.

Morris water maze

Spatial learning was assessed as previously described (Wozniak et al., 2004). Cued trials (visible platform, variable location) and place trials (submerged, hidden platform, consistent location) were conducted in which escape path latency, length, and swimming speeds were recorded. Animal tracking was done using a computerized system (ANY-maze, Stoelting). During cued trials, animals underwent 4 trials per day over 2 consecutive days with the platform being moved to a different location for each trial with few distal spatial cues available. Each trial lasted no longer than 60 seconds, with a 30-minute interval between each trial. Performance was analyzed across four blocks of trials (2 trials/block). After a three-day rest period, animals were tested on place trials, in which mice were required to learn the single location of a submerged platform with several salient distal spatial cues. Place trials occurred over 5 consecutive days of training, with 2 blocks of 2 consecutive trials (60-second trial maximum, 30-second inter-trial-interval after the mouse has reached the platform) with each block separated by 2 hours. Mice were released into different quadrants over different trials. Place trials were averaged over each of the five consecutive days (4 trials/block). One hour after the final block, a probe trial occurred (60-second trial maximum) in which the platform is removed, and the mouse is released from the quadrant opposite where the platform had been located. The time spent in pool quadrants, and the number of crossings over the exact platform location were recorded. DNMT3A$^{KO/+}$ mice showed a small, but significant reduction in target zone time in cohort 2, though there was no difference in cohort 1 (Supplementary Table 4). Additionally, female mice had significantly faster swimming speeds than male mice across both genotypes (Supplementary Table 4). These findings and the observation that the DNMT3A$^{KO/+}$ were slower moving on some of the sensorimotor tests, make

path length a more appropriate variable than escape latency for evaluating performance, as escape latency can be confounded by the differences in swimming speeds.

3-Chamber social approach

Sociability was assayed as previously described (Moy et al., 2004; Silverman et al., 2011). Mice were tested in a rectangular all-Plexiglas apparatus (each chamber measuring 19.5 cm x 39 cm x 22cm) divided into three chambers with walls containing rectangular openings (5 cm x 8 cm) and sliding doors. The apparatus was in a room with indirect light and was cleaned between tests with Nolvasan solution. Stimulus mice were contained within a small stainless-steel withholding cage (10 cm height x 10 cm diameter; Galaxy Pencil/Utility Cup, Spectrum Diversified Designs), allowing minimal contact between mice without allowing fighting. Between tests, withholding cages were cleaned with 75% ethanol solution. A digital video camera recorded movement of the mouse within the apparatus and allowed for tracking with ANY-maze (Stoelting). Distance and time spent in each chamber and investigation zones surrounding the withholding cages were recorded. Zones were defined as 12 cm in diameter from the center of withholding cages.

The test sequence consisted of 4 consecutive 10-minute trials in which the test mouse is placed in the middle chamber and allowed to freely explore the environment. In the first trial, the mouse is placed in the middle chamber with the doors to other chambers shut. In the second trial, the mouse is placed in the middle chamber and can explore all three chambers of the task, allowing it to acclimate to the environment. Neither genotype tested showed a preference towards a side of the chamber during this habituation. For the third trial, a sex-matched novel conspecific was placed within a withholding cage with the other cage remaining empty. For the fourth trial, the same sex-matched conspecific was in one withholding cage, while a new unfamiliar sex-matched stimulus mouse was placed in the other withholding cage. The locations of stimuli mice were counterbalanced across groups for the third trial and randomized novel or familiar for the fourth trial.

Elevated plus maze

Anxiety-like behaviors were examined using the elevated plus maze as previously described (Boyle, 2006). The apparatus contains a central platform (5.5 cm x 5.5 cm) with two opposing open arms and two opposing closed arms (each 36 cm x 6.1 cm x 15 cm) constructed of black Plexiglas. Mouse position is measured using beam-breaks from pairs of photocells configured in a 16 x 16 matrix and outputs are recorded using an interface assembly (Kinder Scientific) and analyzed using software (MotoMonitor, Kinder Scientific) to determine time spent, distance traveled, and entries made into open arms, closed arms, and the center area. Test sessions were conducted in a dimly lit room with each session lasting 5 minutes and each mouse tested over 3 consecutive days. Data shown are from day 1. All mice showed a decrease in time, distance, and entries into open arms on days 2 and 3. There was no significant difference between genotypes in percent entries into open arms (Figure S4K; P=0.137; unpaired Student's T-Test) or total entries into arms (data not shown), indicating that both genotypes explored the maze. Percent distance traveled in open arms showed similar effects to percent time in open arms (Percent distance traveled: P=0.027; unpaired Student's T-Test). Analysis of these data in individual cohorts detected DNMT3A$^{KO/+}$ significant effects for the percent of open arm time on the first day in

cohorts 1 and 3, with no evidence of an effect in cohort 2 (Supplementary Table 4). Individual cohorts also showed no significant difference between genotypes in percent open arm entries (Supplementary Table 4) suggesting that mice explored the elevated plus maze sufficiently to detect anxiety-like behaviors.

Acoustic startle/prepulse inhibition

Sensorimotor gating was evaluated as previously described (Dougherty et al., 2013; Gallitano-Mendel et al., 2008; Hartman et al., 2001). In short, mice were presented with an acoustic startle response (120 dB auditory stimulus pulse, 40 ms broadband burst) and a pre-pulse (response to pre-pulse plus startle pulse). Stimulus onset began at 65 seconds, and 1ms force readings were obtained and averaged to produce an animal's startle amplitude. 20 startle trials were presented in 20 minutes. The first 5 minutes were an acclimation period where no stimuli above the 65 dB background were presented. The session started and finished with 5 consecutive startle (120 dB pulse) trials. The middle 10 trials were interspersed with pre-pulse trials, consisting of an additional 30 presentations of 120 dB startle stimuli preceded by pre-pulse stimuli of 4, 12, or 20 dB above background (10 trials for each PPI trial type). To calculate percent pre-pulse inhibition, we used $\%PPI = 100 \times (ASR_{\text{startle pulse alone}} - ASR_{\text{prepulse + startle pulse}})/ASR_{\text{startle pulse alone}}$.

Conditioned fear

Fear conditioning was done as previously described (Maloney et al., 2019a, 2019b). Mice were habituated to an acrylic chamber (26 cm x 18 cm x 18 cm) containing a metal grid floor and an odorant and was illuminated by LED light which remained on for the duration of the trial. Day 1 testing lasted 5 minutes in which an 80 dB tone sounded for 20 seconds at trial timepoints 100, 160, and 220 seconds. A 1.0 mA shock (unconditioned stimulus) occurred within the last 2 seconds of the tone (conditioned stimulus). Baseline freezing behavior during the first 2 minutes and the freezing behavior during the last 3 minutes was quantified using image analysis (Actimetrics, Evanston, Illinois). On Day 2, testing lasted for 8 minutes in which the light was illuminated but no tones or shocks were presented. On Day 3, testing lasted for 10 minutes in which the mouse was placed in an opaque chamber with a different odorant than the original test chamber. The 80 dB tone began at 120 seconds and lasted for the remainder of the trial and freezing behavior to the conditioned auditory stimulus was quantified for the remaining 8 minutes. Small elevated freezing levels of the DNMT3A$^{KO/+}$ mice for the contextual fear and auditory cue data could be interpreted as evidence for an increased baseline propensity to freeze or stronger fear conditioning in the mutant mice. However, an alternative hypothesis is that the exaggerated freezing levels displayed by the DNMT3A$^{KO/+}$ mice may reflect an emotional hypersensitivity to the footshock as was documented by their freezing levels during tone-shock training. In support of the latter hypothesis, evaluation of baseline freezing levels in individual cohorts showed that they were only significantly different in one of the two cohorts tested.

Statistical analysis for behavioral tests

Behavioral data were analyzed with R v3.3.2 (including the ANOVA function from the Car package in R (Fox and Weisberg, 2011)) and plots were made using GraphPad Prism 7.03a. Normality was assessed using the Shapiro-Wilkes test and visually confirmed. Data not normally distributed were analyzed using non-parametric tests, with the exception of continuous rotarod data. Sexes

1062 were considered separately with genotype and times/block as fixed factors with no consistent
1063 sex effects observed, therefore data were collapsed across sex. Statistical testing was performed
1064 using planned assay-specific methods, such as using Student's T-Tests for single parameter
1065 comparisons between genotypes, and within-subjects two-way repeated-measures ANOVA for
1066 comparisons across timepoints. Individual timepoints within repeated measures tests were
1067 evaluated using Sidak's multiple comparisons test. Individual cohorts were analyzed separately
1068 and in aggregate with similar trends seen across cohorts (Supplementary Table 4), therefore data
1069 from all cohorts were included together.
1070
1071 **Tissue**
1072 Brain tissue was dissected from DNMT3A$^{KO/+}$ and WT littermate mice in ice-cold PBS, flash-frozen
1073 in liquid nitrogen, and stored at -80°C.
1074
1075 **Western blotting**
1076 *Western blotting from cell culture*. Neuro-2a or HEK293T cells were collected and combined with
1077 2x laemmli buffer with 5% β-mercaptoethanol. Samples were passed through a Wizard Column
1078 (Fisher, Wizard Minipreps Mini Columns, PRA7211), boiled for 5 minutes, and run on a BioRad 4-
1079 12% acrylamide gel at 125 V for 60 minutes. Samples were then transferred to a nitrocellulose
1080 membrane, which was bisected between 37kDa and 50kDa bands. Membranes were blocked
1081 with 3% bovine serum albumin in TBS-T for 1 hour at room temperature and then the lower
1082 membrane was immunostained with anti-GFP (ThermoFisher, 1:2000, A-11122) while the upper
1083 membrane was immunostained with anti-DDDDK (Abcam, 1:1000, ab1162) for 12-16 hours at
1084 4°C. All primary and secondary antibodies were diluted in 3% Bovine Serum Albumin in TBS-T.
1085 Membranes were then washed with TBS-T then incubated for 1 hour at room temperature with
1086 IR-dye secondary antibody (IRDye 800CW Donkey anti-Rabbit, LI-COR Biosciences, 1:15,000,
1087 product number: 926-32213). Blots were then washed in PBS, and imaged using the LiCOR
1088 Odyssey XCL system, and quantified using Image Studio Lite software (LI-COR Biosciences). FLAG
1089 (DDDDK) and GFP levels were normalized to a standard curve, and protein levels are expressed
1090 as normalized DDDDK values divided by normalized GFP values to enable comparison of FLAG
1091 (DDDDK) levels between blots. Each blot included a standard curve and WT samples. Outliers
1092 beyond 2 standard deviations above or below the mean were removed. Number of independent
1093 replicates are as follows: WT, 29; W297del, 7; I310N, 7; S312fs11x, 12; G532S, 7; M548K, 9;
1094 V665L, 7; Y375C, 8; R749C, 6; P904L, 7. Significance was assessed using a one sample T-Test, as
1095 protein expression levels were normalized to GFP and WT, and mutant protein expression was
1096 compared to the normalized WT value of 1.
1097
1098 *Western blotting from tissue.* Brain tissue samples were homogenized with a dounce
1099 homogenizer in buffer with protease inhibitors (10mM HEPES pH 7.9, 10mM KCl, 1.5mM $MgCl_2$,
1100 1mM DTT, 10mM EDTA). A portion of the lysate was removed and 1% SDS was added. Samples
1101 were boiled for 10 minutes, followed by a 10-minute spin at 15,000g. Supernatant was collected
1102 and run through a Wizard Column (Fisher, Wizard Minipreps Mini Columns, PRA7211), then
1103 diluted in LDS sample buffer with 5% β-mercaptoethanol. Samples were boiled for 5 minutes, run
1104 on an 8% acrylamide gel for 60 minutes at 125 V, and transferred to a nitrocellulose membrane.
1105 Membrane was bisected between 75kDa and 100kDa. Membranes were blocked with 3% bovine

1106 serum albumin in TBS-T for 1 hour at room temperature, and the upper membrane was
1107 immunostained with anti-DNMT3A (Abcam, 1:1000, ab13888) while the lower membrane was
1108 immunostained with anti-α-Tubulin (Abcam, 1:1000, ab52866) for 12-16 hours at 4°C. All primary
1109 and secondary antibodies were diluted in 3% Bovine Serum Albumin in TBS-T. Membranes were
1110 then washed with TBS-T then incubated for 1 hour at room temperature with IR-dye secondary
1111 antibody (IRDye 800CW Goat anti-Rabbit, or IRDye 800CW Goat anti-Mouse, LI-COR Biosciences,
1112 1:15,000, product numbers: 926-32211 and 926-32210 respectively). Blots were then washed in
1113 PBS, imaged using the LiCOR Odyssey XCL system, and quantified using Image Studio Lite
1114 software (LI-COR Biosciences). DNMT3A and α-Tubulin levels were normalized to a standard
1115 curve, and protein levels are expressed as normalized DNMT3A values divided by normalized α-
1116 Tubulin values to enable comparison of DNMT3A levels between blots. For brain region analysis,
1117 sample sizes of n=4 per genotype (2 male and 2 female pairs) were used. For time course analysis,
1118 sample sizes of n=2 per genotype (1 male and 1 female pairs) were used for all time points except
1119 the 2-week timepoint in which n=6 (3 male and 3 female pairs) was used. Significance was
1120 assessed using a two way ANOVA considering genotype and time to determine if there was a
1121 detectable difference in protein expression over time.
1122
1123 **Bisulfite sequencing**
1124 *Whole genome bisulfite sequencing from cortical cultures.* Samples were chosen for whole
1125 genome bisulfite sequencing if mutant and WT samples expressed equal amounts of *Dnmt3a*
1126 mRNA as measured by qRT-PCR (Figure S2D). DNA from cortical cultures was bisulfite converted
1127 and prepared for sequencing using the Tecan Ovation Ultralow Methyl-Seq Kit (Tecan, 0335-32)
1128 and the Epitect Bisulfite Kit (Qiagen, 59824) was used for bisulfite conversion. We used alternate
1129 bisulfite conversion cycling conditions ([95°C, 5 min; 60°C, 20 min] x 4 cycles, 20°C hold) to ensure
1130 lowest possible bisulfite non-conversion rate. Libraries were PCR-amplified for 10-11 cycles.
1131 Libraries were then pooled and sequenced at a depth of 0.01-0.03x genomic coverage using an
1132 Illumina MiSeq 2x150 through the Spike-In Cooperative at Washington University in St. Louis.
1133 Significance was assessed using a one-sample student's t-test, as we are comparing groups
1134 normalized to WT and GFP back to the normalized value of 1.
1135
1136 *Whole genome bisulfite sequencing from tissue.* DNA was isolated from tissue using the DNEasy
1137 Kit (Qiagen). 300 ng of DNA was prepared for sequencing using the Ovation Ultralow Methyl-Seq
1138 Kit (Tecan, 0335-32) with and the Epitect Bisulfite Kit (Qiagen, 59824) was used for bisulfite
1139 conversion. For these samples, 300 ng of DNA was fragmented for 45 seconds with the Covaris
1140 E220 sonicator (10% Duty Factory, 175 Peak Incidence Power, 200 cycles per burst, milliTUBE
1141 200µL AFA Fiber). DNA was then purified using 0.7 volumes of Agencourt Beads to select for long
1142 DNA inserts for sequencing. We used alternate bisulfite conversion cycling conditions ([95°C, 5
1143 min; 60°C, 20 min] x 4 cycles, 20°C hold) to ensure lowest possible bisulfite non-conversion rate.
1144 Libraries were PCR-amplified for 12 cycles. Libraries were then pooled and sequenced using an
1145 Illumina MiSeq 2x150 through the Spike-In Cooperative at Washington University in St. Louis.
1146 Samples for shallow-depth sequencing (Figure 4A,B) were sequenced at 0.01-0.03x genomic
1147 coverage. For brain region and liver methylation, n=2 per genotype per region (one male pair,
1148 one female pair). For developmental time course methylation, n=3-4 per genotype per timepoint,
1149 with at least one male and one female pair. 8-week cortex samples for deep sequencing (Figure

5, Figure 6A,B,D, Figure S5C-E) were sequenced at 6.4-7.8x coverage per biological rep (two technical reps per biological rep, two biological reps per genotype) using only male genotype pairs. For shallow sequencing experiments, significance was assessed using either a two-sample Student's T-Test to compare global methylation values of the cortex between two genotypes, or using a two way ANOVA to compare broad methylation changes across a variety of brain regions between genotypes. Genomic element comparisons were done using two-sample Student's T-Tests with Bonferroni correction.

Oxidative bisulfite sequencing from tissue. DNA was isolated from tissue using the DNEasy Kit (Qiagen, 69504). 450 ng of DNA was prepared for sequencing using the Ovation Ultralow Methyl-Seq Kit (Tecan, 0335-32) with TrueMethyl oxBS plugin (Tecan, 0414-32). For these samples, 450 ng of DNA was fragmented for 45 seconds with the Covaris E220 sonicator (10% Duty Factory, 175 Peak Incidence Power, 200 cycles per burst, milliTUBE 200µL AFA Fiber). DNA was then purified using 0.7 volumes of Agencourt Beads to select for long DNA inserts for sequencing. 300 ng of DNA was used for OxBS libraries, whereas the remaining 150 ng of DNA was used for bisulfite libraries. We used alternate bisulfite conversion cycling conditions ([95°C, 5 min; 60°C, 20 min] x 2 cycles; 95°C, 5 min; 60°C, 40 min; 95°C, 5 min; 60°C, 45 min; 20°C hold) to ensure lowest possible bisulfite non-conversion rate. Bisulfite and oxidative bisulfite libraries were PCR-amplified for 11 and 13 cycles respectively. Libraries were then pooled and sequenced using an Illumina MiSeq 2x150 through the Spike-In Cooperative at Washington University in St. Louis. Samples were sequenced at 0.8-2.2x genomic coverage per replicate (two replicates per genotype). Genomic element comparisons were done using two-sample Student's T-Tests with Bonferroni correction.

Whole-genome bisulfite analysis

Bisulfite sequencing analysis was performed as previously described (Clemens et al., 2019). Briefly, data were adapter-trimmed, mapped to mm9, then deduplicated and called for methylation using BS-seeker2. Methylation levels across regions were assessed using bedtools map -o sum, summing the number of reads mapping to Cs (supporting mC) and the amount of coverage in the region, then dividing those two numbers(Quinlan and Hall, 2010b). Hydroxymethylation was calculated as the percent methylation found in the BS-seq minus the percent methylation found in the matching oxBS-seq. Due to count noise, this occasionally resulted in apparent negative hydroxymethylation. During bisulfite sequencing not all DNA can be efficiently bisulfite converted. Though our methods should maximize the amount of converted unmethylated C, there is still a small percentage of unmethylated cytosines that are called as methylated due to non-conversion (0.2-0.3%). Due to this non-conversion, very lowly methylated regions (e.g. mCA at CpG islands) may not show the same percent reduction in mCA as highly methylated regions. Data were visualized using the UCSC genome browser (http://genome.ucsc.edu) (Kent et al., 2002). Average methylation per-sample is normally distributed in all regions examined, and variance between genotypes is similar, fitting the assumptions of a 2-sample t-test. Methylation levels for individual elements are not necessarily normally distributed, so non-parametric tests were used instead.

Differentially methylated region detection

BSmooth (Hansen et al., 2012) was used to call differentially CpG methylated regions between DNMT3A$^{KO/+}$ and WT mice, using two technical replicates each of two biological replicates. CG sites were filtered for requiring at least 2x coverage in all replicates and differentially methylated regions were called with a statistical threshold of t-stat >2.0. These regions were further filtered for a length >100 bp and a requirement that the smoothed per-rep methylation values were consistent. For hypomethylated regions, all WT mCG/CG values needed to be greater than any KO mCG/CG value and all KO methylation values needed to be higher than all WT methylation values for hypermethylated regions. Data fit the assumptions and requirements of BSmooth. Data were distributed evenly between chromosomes, and the overlap between DMRs and regions of interest fit a hypergeometric distribution, making a fisher's exact test appropriate.

RNA sequencing

Total RNA isolation was carried out as previously described (Clemens et al., 2019). In brief, cerebral cortex was dissected in ice-cold PBS from DNMT3A$^{KO/+}$ and WT littermates at 8 weeks of age (n=7 pairs, 3 male, 4 female). Cortex was lysed in RLT buffer following the RNeasy Mini Kit (Qiagen, 74104). RNA libraries were generated from 250 ng of RNA with NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB) using a modified amplification protocol (37°C, 15 minutes; 98°C, 30 seconds; [98°C, 10 seconds; 65°C, 30 seconds; 72°C, 30 seconds]x13; 72°C, 5 minutes; 4°C hold). RNA libraries were pooled at a final concentration of 10nM and sequenced using Illumina HiSeq3000 1x50bp with the Genome Technology Access Center at Washington University in St. Louis, typically yielding 15-30 million single-end reads per sample.

RNA sequencing analysis

RNA sequencing analysis was performed as previously described (Clemens et al., 2019). Briefly, raw FASTQ files were trimmed with Trim Galore and rRNA sequences were filtered out with Bowtie. Remaining reads were aligned to mm9 using STAR (Dobin et al., 2013) with the default parameters. Reads mapping to multiple regions in the genome were then filtered out, and uniquely mapping reads were converted to BED files and separated into intronic and exonic reads. Finally, reads were assigned to genes using bedtools coverage -counts (Quinlan and Hall, 2010b).

For gene annotation we defined a "flattened" list of longest transcript forms for each gene, generated on Ensgene annotations and obtained from the UCSC table browser. For each gene, Ensembl IDs were matched up to MGI gene names. Then, for each unique MGI gene name, the most upstream Ensgene TSS and the most downstream TES were taken as that gene's start and stop. Based on these Ensembl gene models, we defined TSS regions and gene bodies. Differentially expressed genes were identified using a Wald test through DESeq2, running using default parameters on exonic reads from the DNMT3A$^{KO/+}$ and WT.

Chromatin immunoprecipitation protocol

Chromatin immunoprecipitation was performed as previously described (Clemens et al., 2019; Cohen et al., 2011). Cerebral cortex was dissected on ice in PBS from DNMT3A$^{KO/+}$ and WT

1238 littermates at 8-weeks old (n=5 pairs, 3 male, 2 female). The tissue was flash-frozen in liquid
1239 nitrogen and stored at -80°C. Chromatin were fragmented with the Covaris E220 sonicator (5%
1240 Duty Factory, 140 Peak Incidence Power, 200 cycles per burst, milliTUBE 1mL AFA Fiber). ChIP
1241 was performed with H3K27ac antibody (0.025-0.1µg; Abcam, ab4729) and libraries were
1242 generated using Ovation Ultralow Library System V2 (Tecan, 0344NB-32). Libraries were pooled
1243 to a final concentration of 8-10nM and sequenced using Illumina HiSeq 3000 with the Genome
1244 Technology Access Center at Washington University in St. Louis, typically yielding 15-40 million
1245 single-end reads per sample.
1246
1247 **Chromatin immunoprecipitation analysis**
1248 ChIP sequencing analysis was performed as previously described (Clemens et al., 2019). Briefly,
1249 reads were mapped to mm9 using bowtie2 and reads were extended based on library sizes and
1250 deduplicated. Bedtools coverage –counts was used to quantify ChIP signal at the transcriptional
1251 start site (TSS), gene body (GB), and transcriptional end site (TES) (Quinlan and Hall, 2010b).
1252 edgeR was then used to determine differential ChIP-signal across genotypes. Data were
1253 visualized using the UCSC genome browser (http://genome.ucsc.edu) (Kent et al., 2002).
1254
1255 **Controlled resampling**
1256 A similar resampling approach was used as previously described (Clemens et al., 2019). Briefly,
1257 for every entry in a sample set (e.g. DNMT3A-dysregulated genes), an entry in the control set
1258 (e.g. all other genes) with a similar desired characteristic (e.g. expression) was selected,
1259 generating a control set of the same size and variable distribution as the sample set.
1260
1261 **Identification of dysregulated enhancers**
1262 Enhancer regions from Clemens et al. 2019 were used, and enhancers dysregulated in the
1263 DNMT3A$^{KO/+}$ were called using the same method. Briefly, H3K27ac ChIP-seq reads were
1264 quantified in all acetyl peak regions, and edgeR was used to identify peaks with significantly
1265 different amounts of H3K27ac signal. Peak regions were then divided into promoters, enhancers,
1266 and non-identified peaks. Data fits the assumptions of BSmooth. Overlap between misregulated
1267 enhancers in different genotypes fit a hypergeometric distribution.
1268
1269
1270 **GAGE**
1271 Gene set enrichment analysis for the gene sets described was performed using the Generally
1272 Applicable Gene-set Enrichment (GAGE) program (Luo et al., 2009). Analysis was performed
1273 directionally on input of shrunken, log-normalized exonic fold changes output from DESeq2
1274 analysis of DNMT3A$^{KO/+}$ versus WT RNA-seq data. Gene sets with an FDR q-value below 0.1 and
1275 an adjusted p-value below 0.5 following expression matched resampling repeated 1,000 times
1276 were considered statistically significant. Gene sets were selected for analysis from both human
1277 and mouse studies of autism associated genes. SFARI genes (Abrahams et al., 2013) with scores
1278 of equal to or less than 3 were considered. Date accessed: 6/20/2019.
1279
1280 **GSEA**

Gene Set Enrichment Analysis (GSEA) (version 7.0, the Broad Institute of MIT and Harvard, http://software.broadinstitute.org/gsea/downloads.jsp) was performed on shrunken, log-normalized exonic fold changes from DESeq2 between DNMT3A^{KO+} and WT RNA-seq data. GSEA calculated a gene set Enrichment Score (ES) that analyzed genes were enriched in the biological signal conduction on the MsigDB (Molecular Signatures Database, http://software.broadinstitute.org/gsea/msigdb). Background was set to all expressed genes in this study and 1,000 permutations were set to generate a null distribution for enrichment score in the hallmark gene sets and functional annotation gene sets. The gene sets database used for enrichment analysis were 'c5.all.v7.0.symbols.gmt', 'c5.bp.v7.0.symbols.gmt', 'c5.cc.v7.0.symbols.gmt'and 'c5.mf.v7.0.symbols.gmt' and FDR <0.1 was defined as the cut-off criteria for significance.

Craniofacial morphological analyses

A total of 24 sex-matched littermate paired mice (DNMT3A$^{KO/+}$ n=12, 7 male, 5 female; WT n=12, 7 male, 5 female) across 3 time-points (8 weeks DNMT3A$^{KO/+}$ n=4, WT n=4; 20 weeks DNMT3A$^{KO/+}$ n=4, WT n=4; 25 weeks DNMT3A$^{KO/+}$ n=4, WT n=4) were fixed in 4% paraformaldehyde through intracardiac perfusions. Whole mouse heads were scanned at the Musculoskeletal Research Center at Washington University in St. Louis using a Scanco µCT40 machine. CT images had voxel dimensions of 0.018 millimeters and were reconstructed on a 2048x2048 pixel grid. The CT images were converted to 8bit images using ImageJ (https://imagej.nih.gov/ij/) and surface reconstructions were acquired in Avizo (http://www.vsg3d.com/). Thirty-five three-dimensional landmarks were collected from surface reconstructions of the cranium and mandible using Stratovan Checkpoint (https://www.stratovan.com/products/checkpoint).

Generalized Procrustes Analysis in MorphoJ software was used to explore the differences and similarities of shape between the DNMT3A$^{KO/+}$ mice and their WT littermates as previously described (Hill et al., 2013). To control for possible differences in size, the landmark coordinate data were natural log-transformed and analyzed with a linear regression model. Additionally, to localize differences in form to specific linear distances, landmark data were analyzed using Euclidean Distance Matrix Analysis (EDMA).

Bone length measurements

A total of 24 sex-matched littermate paired mice (DNMT3A$^{KO/+}$ n=12, 7 male, 5 female; WT n=12, 7 male, 5 female) across 3 time-points (8 weeks DNMT3A$^{KO/+}$ n=4, WT n=4; 20 weeks DNMT3A$^{KO/+}$ n=4, WT n=4; 25 weeks DNMT3A$^{KO/+}$ n=4, WT n=4) were fixed in 4% paraformaldehyde through intracardiac perfusions. Decapitated mouse bodies were scanned at the Musculoskeletal Research Center at Washington University in St. Louis using a Faxitron Model UltraFocus100 Dual X-Ray machine. Bone lengths were measured using ImageJ. Data were taken over three age time-points: 8 weeks, 20 weeks, and 25 weeks of age for male and female mice. There was no significant difference in bone lengths based upon sex, but there was a difference based by age. To normalize for this age effect, data were expressed as DNMT3A$^{KO/+}$ bone lengths normalized to the WT lengths within groups. Left and right bones were measured and the larger was used for analysis.

1325 **Experimental design**
1326 Authenticated cell lines from ATCC (HEK293T, NEURO2A) were used, and no mycoplasma
1327 contamination testing was needed. Sample sizes were chosen based upon previously published
1328 studies using similar techniques. Statistical tests and exclusion criteria (values beyond 2 standard
1329 deviations of the group mean) were similar to that of previously published studies and indicated
1330 in the appropriate methods. For all animal experiments, experimenters were blinded to genotype
1331 during data collection. No treatment conditions were used, so no samples or animals were
1332 allocated to experimental groups and no randomization was needed. Tests that assume equal
1333 variance were only run if group variances were similar, otherwise alternative tests were used.
1334
1335 **Data availability statement**
1336 The data that support the findings of this study are available from the corresponding author upon
1337 request. DOIs for all published gene sets used in comparison and enrichment analysis:
1338 Lister et al. 2013: https://doi.org/10.1126/science.1237905;
1339 Clemens et al. 2019: https://doi.org/10.1016/j.molcel.2019.10.033;
1340 Stroud et al. 2017: https://doi.org/10.1016/j.cell.2017.09.047;
1341 Gompers et al. 2017: https://doi.org/10.1038/nn.4592;
1342 Katayama et al. 2016: https://dx.doi.org/10.1038/nature19357;
1343 Tilot et al. 2016: https://doi.org/10.1038/mp.2015.17;
1344 Sessa et al. 2019: https://doi.org/10.1016/j.neuron.2019.07.013;
1345 Gandal et al. 2018: https://doi.org/10.1126/science.aat8127;
1346 Voineagu et al. 2011: https://doi.org/10.1038/nature10110;
1347 Abrahams et al. 2013: https://doi.org/10.1186/2040-2392-4-36;
1348 Parikshak et al. 2013: https://dx.doi.org/10.1038/nature20612;
1349
1350 Raw and aggregate bisulfite-seq, raw and gene-count data for RNA-seq, and raw and peak call
1351 data for ChIP seq will be available on GEO.
1352
1353 **Supplemental Information**
1354 Table S1. DNMT3A Mutation Table, Related to Figure 1, Figure 2, Figure S1
1355 Table S2. BSsmooth-defined Differentially Methylated Regions, Related to Figure 4, Figure 5,
1356 Figure S5
1357 Table S3. Table of differentially expressed genes, Related to Figure 6, Figure 7, Figure S6
1358 Table S4. Behavioral Test Table, Related to Figure 3, Figure S3, Figure S4
1359
1360 **References**
1361 Abrahams, B.S., Arking, D.E., Campbell, D.B., Mefford, H.C., Morrow, E.M., Weiss, L.A.,
1362 Menashe, I., Wadkins, T., Banerjee-Basu, S., and Packer, A. (2013). SFARI Gene 2.0: a
1363 community-driven knowledgebase for the autism spectrum disorders (ASDs). Mol. Autism *4*, 36.

1364 Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett
1365 syndrome is caused by mutations in X-linked MECP2 , encoding methyl-CpG-binding protein 2.
1366 Nat. Genet. *23*, 185–188.

1367 Banerjee-Basu, S., and Packer, A. (2010). SFARI Gene: an evolving database for the autism
1368 research community. Dis. Model. Mech. *3*, 133–135.

1369 Barnes, T.D., Rieger, M.A., Dougherty, J.D., and Holy, T.E. (2017). Group and Individual
1370 Variability in Mouse Pup Isolation Calls Recorded on the Same Day Show Stability. Front. Behav.
1371 Neurosci. *11*.

1372 Boxer, L.D., Renthal, W., Greben, A.W., Whitwam, T., Silberfeld, A., Stroud, H., Li, E., Yang, M.G.,
1373 Kinde, B., Griffith, E.C., et al. (2019). MeCP2 Represses the Rate of Transcriptional Initiation of
1374 Highly Methylated Long Genes. Mol. Cell.

1375 Boyle, M.P. (2006). Forebrain Glucocorticoid Receptors Modulate Anxiety-Associated
1376 Locomotor Activation and Adrenal Responsiveness. J. Neurosci. *26*, 1971–1978.

1377 Chang, Y., Cole, T.B., and Costa, L.G. (2017). Behavioral Phenotyping for Autism Spectrum
1378 Disorders in Mice. Curr. Protoc. Toxicol. *72*.

1379 Chen, L., Chen, K., Lavery, L.A., Baker, S.A., Shaw, C.A., Li, W., and Zoghbi, H.Y. (2015). MeCP2
1380 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing
1381 of onset for Rett syndrome. Proc. Natl. Acad. Sci. *112*, 5509–5514.

1382 Cholewa-Waclaw, J., Bird, A., von Schimmelmann, M., Schaefer, A., Yu, H., Song, H.,
1383 Madabhushi, R., and Tsai, L.-H. (2016). The Role of Epigenetic Mechanisms in the Regulation of
1384 Gene Expression in the Nervous System. J. Neurosci. *36*, 11427–11434.

1385 Clemens, A.W., Wu, D.Y., Moore, J.R., Christian, D.L., Zhao, G., and Gabel, H.W. (2019). MeCP2
1386 Represses Enhancers through Chromosome Topology-Associated DNA Methylation. Mol. Cell.

1387 Cohen, S., Gabel, H.W., Hemberg, M., Hutchinson, A.N., Sadacca, L.A., Ebert, D.H., Harmin, D.A.,
1388 Greenberg, R.S., Verdine, V.K., Zhou, Z., et al. (2011). Genome-Wide Activity-Dependent MeCP2
1389 Phosphorylation Regulates Nervous System Development and Function. Neuron *72*, 72–85.

1390 Cole, C.B., Russler-Germain, D.A., Ketkar, S., Verdoni, A.M., Smith, A.M., Bangert, C.V., Helton,
1391 N.M., Guo, M., Klco, J.M., O'Laughlin, S., et al. (2017). Haploinsufficiency for DNA
1392 methyltransferase 3A predisposes hematopoietic cells to myeloid malignancies. J. Clin. Invest.
1393 *127*, 3657–3674.

1394 Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012).
1395 Topological domains in mammalian genomes identified by analysis of chromatin interactions.
1396 Nature *485*, 376–380.

1397 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M.,
1398 and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

1399 Dougherty, J.D., Maloney, S.E., Wozniak, D.F., Rieger, M.A., Sonnenblick, L., Coppola, G.,
1400 Mahieu, N.G., Zhang, J., Cai, J., Patti, G.J., et al. (2013). The Disruption of *Celf6*, a Gene

1401 Identified by Translational Profiling of Serotonergic Neurons, Results in Autism-Related
1402 Behaviors. J. Neurosci. *33*, 2732 LP – 2753.

1403 Fazel Darbandi, S., Robinson Schwartz, S.E., Qi, Q., Catta-Preta, R., Pai, E.L.-L., Mandell, J.D.,
1404 Everitt, A., Rubin, A., Krasnoff, R.A., Katzman, S., et al. (2018). Neonatal Tbr1 Dosage Controls
1405 Cortical Layer 6 Connectivity. Neuron *100*, 831-845.e7.

1406 Feliciano, P., Zhou, X., Astrovskaya, I., Turner, T.N., Wang, T., Brueggeman, L., Barnard, R.,
1407 Hsieh, A., Snyder, L.G., Muzny, D.M., et al. (2019). Exome sequencing of 457 autism families
1408 recruited online provides evidence for autism risk genes. Npj Genomic Med. *4*, 1–14.

1409 Fox, J., and Weisberg, S. (2011). An R Companion to Applied Regression (SAGE Publications).

1410 Gabel, H.W., Kinde, B., Stroud, H., Gilbert, C.S., Harmin, D.A., Kastan, N.R., Hemberg, M., Ebert,
1411 D.H., and Greenberg, M.E. (2015). Disruption of DNA-methylation-dependent long gene
1412 repression in Rett syndrome. Nature *522*, 89–93.

1413 Gallitano-Mendel, A., Wozniak, D.F., Pehek, E.A., and Milbrandt, J. (2008). Mice Lacking the
1414 Immediate Early Gene Egr3 Respond to the Anti-Aggressive Effects of Clozapine Yet are
1415 Relatively Resistant to its Sedating Effects. Neuropsychopharmacol. Off. Publ. Am. Coll.
1416 Neuropsychopharmacol. *33*, 1266–1275.

1417 Gandal, M.J., Zhang, P., Hadjimichael, E., Walker, R.L., Chen, C., Liu, S., Won, H., van Bakel, H.,
1418 Varghese, M., Wang, Y., et al. (2018). Transcriptome-wide isoform-level dysregulation in ASD,
1419 schizophrenia, and bipolar disorder. Science *362*, eaat8127.

1420 Gompers, A.L., Su-Feher, L., Ellegood, J., Copping, N.A., Riyadh, M.A., Stradleigh, T.W., Pride,
1421 M.C., Schaffler, M.D., Wade, A.A., Catta-Preta, R., et al. (2017). Germline Chd8
1422 haploinsufficiency alters brain development in mouse. Nat. Neurosci. *20*, 1062–1073.

1423 Gowher, H., and Jeltsch, A. (2018). Mammalian DNA methyltransferases: new discoveries and
1424 open questions. Biochem. Soc. Trans. *46*, 1191–1202.

1425 Grady, R.M., Wozniak, D.F., Ohlemiller, K.K., and Sanes, J.R. (2006). Cerebellar Synaptic Defects
1426 and Abnormal Motor Behavior in Mice Lacking α- and β-Dystrobrevin. J. Neurosci. *26*, 2841–
1427 2851.

1428 Guo, J.U., Su, Y., Shin, J.H., Shin, J., Li, H., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G., et al. (2014).
1429 Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain.
1430 Nat. Neurosci. *17*, 215–222.

1431 Guo, W., Fiziev, P., Yan, W., Cokus, S., Sun, X., Zhang, M.Q., Chen, P.-Y., and Pellegrini, M.
1432 (2013). BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. BMC Genomics
1433 *14*, 774.

1434 Guo, X., Wang, L., Li, J., Ding, Z., Xiao, J., Yin, X., He, S., Shi, P., Dong, L., Li, G., et al. (2015).
1435 Structural insight into autoinhibition and histone H3-induced activation of DNMT3A. Nature
1436 *517*, 640–644.

1437 Haeussler, M., Zweig, A.S., Tyner, C., Speir, M.L., Rosenbloom, K.R., Raney, B.J., Lee, C.M., Lee,
1438 B.T., Hinrichs, A.S., Gonzalez, J.N., et al. (2019). The UCSC Genome Browser database: 2019
1439 update. Nucleic Acids Res. *47*, D853–D858.

1440 Hansen, K.D., Langmead, B., and Irizarry, R.A. (2012). BSmooth: from whole genome bisulfite
1441 sequencing reads to differentially methylated regions. Genome Biol. *13*, R83.

1442 Hartman, R.E., Wozniak, D.F., Nardi, A., Olney, J.W., Sartorius, L., and Holtzman, D.M. (2001).
1443 Behavioral Phenotyping of GFAP-ApoE3 and -ApoE4 Transgenic Mice: ApoE4 Mice Show
1444 Profound Working Memory Impairments in the Absence of Alzheimer's-like Neuropathology.
1445 Exp. Neurol. *170*, 326–344.

1446 Hill, C.A., Martínez-Abadías, N., Motch, S.M., Austin, J.R., Wang, Y., Jabs, E.W., Richtsmeier, J.T.,
1447 and Aldridge, K. (2013). Postnatal brain and skull growth in an Apert syndrome mouse model.
1448 Am. J. Med. Genet. A. *161*, 745–757.

1449 Holy, T.E., and Guo, Z. (2005). Ultrasonic Songs of Male Mice. PLoS Biol. *3*, e386.

1450 Jeltsch, A., Broche, J., and Bashtrykov, P. (2018). Molecular Processes Connecting DNA
1451 Methylation Patterns with DNA Methyltransferases and Histone Modifications in Mammalian
1452 Genomes. Genes *9*, 566.

1453 Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004). Essential
1454 role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. Nature
1455 *429*, 900–903.

1456 Katayama, Y., Nishiyama, M., Shoji, H., Ohkawa, Y., Kawamura, A., Sato, T., Suyama, M., Takumi,
1457 T., Miyakawa, T., and Nakayama, K.I. (2016). CHD8 haploinsufficiency results in autistic-like
1458 phenotypes in mice. Nature *537*, 675–679.

1459 Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, and
1460 D. (2002). The Human Genome Browser at UCSC. Genome Res. *12*, 996–1006.

1461 Kinde, B., Gabel, H.W., Gilbert, C.S., Griffith, E.C., and Greenberg, M.E. (2015). Reading the
1462 unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation,
1463 and MeCP2. Proc. Natl. Acad. Sci. *112*, 6800–6806.

1464 Kinde, B., Wu, D.Y., Greenberg, M.E., and Gabel, H.W. (2016). DNA methylation in the gene
1465 body influences MeCP2-mediated gene repression. Proc. Natl. Acad. Sci. *113*, 15114–15119.

1466 Lagger, S., Connelly, J.C., Schweikert, G., Webb, S., Selfridge, J., Ramsahoye, B.H., Yu, M., He, C.,
1467 Sanguinetti, G., Sowers, L.C., et al. (2017). MeCP2 recognizes cytosine methylated tri-nucleotide

1468 and di-nucleotide sequences to tune transcription in the mammalian brain. PLOS Genet. *13*,
1469 e1006793.

1470 Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat.
1471 Methods *9*, 357–359.

1472 Lazic, S.E. (2015). Analytical strategies for the marble burying test: avoiding impossible
1473 predictions and invalid p-values. BMC Res. Notes *8*, 141.

1474 Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler
1475 transform. Bioinformatics *25*, 1754–1760.

1476 Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang,
1477 Y., Dwork, A.J., Schultz, M.D., et al. (2013a). Global epigenomic reconfiguration during
1478 mammalian brain development. Science *341*.

1479 Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang,
1480 Y., Dwork, A.J., Schultz, M.D., et al. (2013b). Global Epigenomic Reconfiguration During
1481 Mammalian Brain Development. Science *341*, 1237905.

1482 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
1483 dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

1484 Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D., and Woolf, P.J. (2009). GAGE: generally
1485 applicable gene set enrichment for pathway analysis. BMC Bioinformatics *10*, 161.

1486 Lyst, M.J., and Bird, A. (2015). Rett syndrome: a complex disorder with simple roots. Nat. Rev.
1487 Genet. *16*, 261–275.

1488 Maloney, S.E., Rieger, M.A., Al-Hasani, R., Bruchas, M.R., Wozniak, D.F., and Dougherty, J.D.
1489 (2019a). Loss of CELF6 RNA binding protein impairs cocaine conditioned place preference and
1490 contextual fear conditioning. Genes Brain Behav. *18*, e12593.

1491 Maloney, S.E., Yuede, C.M., Creeley, C.E., Williams, S.L., Huffman, J.N., Taylor, G.T., Noguchi,
1492 K.N., and Wozniak, D.F. (2019b). Repeated neonatal isoflurane exposures in the mouse induce
1493 apoptotic degenerative changes in the brain and relatively mild long-term behavioral deficits.
1494 Sci. Rep. *9*, 1–14.

1495 McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and
1496 Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat.
1497 Biotechnol. *28*, 495–501.

1498 McRae, J.F., Clayton, S., Fitzgerald, T.W., Kaplanis, J., Prigmore, E., Rajan, D., Sifrim, A., Aitken,
1499 S., Akawi, N., Alvi, M., et al. (2017). Prevalence and architecture of de novo mutations in
1500 developmental disorders. Nature *542*, 433–438.

1501 Moy, S.S., Nadler, J.J., Perez, A., Barbaro, R.P., Johns, J.M., Magnuson, T.R., Piven, J., and
1502 Crawley, J.N. (2004). Sociability and preference for social novelty in five inbred strains: an
1503 approach to assess autistic-like behavior in mice. Genes Brain Behav. *3*, 287–302.

1504 Nord, A.S., and West, A.E. (2019). Neurobiological functions of transcriptional enhancers. Nat.
1505 Neurosci.

1506 Nygaard, K.R., Maloney, S.E., and Dougherty, J.D. (2019). Erroneous inference based on a lack of
1507 preference within one group: Autism, mice, and the social approach task. Autism Res. *12*, 1171–
1508 1183.

1509 Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and
1510 Dnmt3b are essential for de novo methylation and mammalian development. Cell *99*, 247–257.

1511 Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and
1512 Geschwind, D.H. (2013). Integrative Functional Genomic Analyses Implicate Specific Molecular
1513 Pathways and Circuits in Autism. Cell *155*, 1008–1021.

1514 Quinlan, A.R., and Hall, I.M. (2010a). BEDTools: a flexible suite of utilities for comparing
1515 genomic features. Bioinformatics *26*, 841–842.

1516 Quinlan, A.R., and Hall, I.M. (2010b). BEDTools: a flexible suite of utilities for comparing
1517 genomic features. Bioinforma. Oxf. Engl. *26*, 841–842.

1518 Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for
1519 differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

1520 Russler-Germain, D.A., Spencer, D.H., Young, M.A., Lamprecht, T.L., Miller, C.A., Fulton, R.,
1521 Meyer, M.R., Erdmann-Gilmore, P., Townsend, R.R., Wilson, R.K., et al. (2014). The R882H
1522 DNMT3A Mutation Associated with AML Dominantly Inhibits Wild-Type DNMT3A by Blocking Its
1523 Ability to Form Active Tetramers. Cancer Cell *25*, 442–454.

1524 Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E., Murtha, M.T.,
1525 Bal, V.H., Bishop, S.L., Dong, S., et al. (2015). Insights into Autism Spectrum Disorder Genomic
1526 Architecture and Biology from 71 Risk Loci. Neuron *87*, 1215–1233.

1527 Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.-Y., Peng, M., Collins,
1528 R., Grove, J., Klei, L., et al. (2019). Large-scale exome sequencing study implicates both
1529 developmental and functional changes in the neurobiology of autism. BioRxiv 484113.

1530 Sceniak, M.P., Lang, M., Enomoto, A.C., James Howell, C., Hermes, D.J., and Katz, D.M. (2016).
1531 Mechanisms of Functional Hypoconnectivity in the Medial Prefrontal Cortex of Mecp2 Null
1532 Mice. Cereb. Cortex N. Y. NY *26*, 1938–1956.

1533 Sendžikaitė, G., Hanna, C.W., Stewart-Morgan, K.R., Ivanova, E., and Kelsey, G. (2019). A
1534 DNMT3A PWWP mutation leads to methylation of bivalent chromatin and growth retardation
1535 in mice. Nat. Commun. *10*, 1–16.

1536 Sessa, A., Fagnocchi, L., Mastrototaro, G., Massimino, L., Zaghi, M., Indrigo, M., Cattaneo, S.,
1537 Martini, D., Gabellini, C., Pucci, C., et al. (2019). SETD5 Regulates Chromatin Methylation State
1538 and Preserves Global Transcriptional Fidelity during Brain Development and Neuronal Wiring.
1539 Neuron *104*, 271-289.e13.

1540 Silverman, J.L., Turner, S.M., Barkan, C.L., Tolu, S.S., Saxena, R., Hung, A.Y., Sheng, M., and
1541 Crawley, J.N. (2011). Sociability and motor functions in Shank1 mutant mice. Brain Res. *1380*,
1542 120–137.

1543 Simola, N., and Granon, S. (2019). Ultrasonic vocalizations as a tool in studying emotional states
1544 in rodent models of social behavior and brain disease. Neuropharmacology *159*, 107420.

1545 Spencer, D.H., Russler-Germain, D.A., Ketkar, S., Helton, N.M., Lamprecht, T.L., Fulton, R.S.,
1546 Fronick, C.C., O'Laughlin, M., Heath, S.E., Shinawi, M., et al. (2017). CpG Island
1547 Hypermethylation Mediated by DNMT3A Is a Consequence of AML Progression. Cell *168*, 801-
1548 816.e13.

1549 Stamatoyannopoulos, J.A., Snyder, M., Hardison, R., Ren, B., Gingeras, T., Gilbert, D.M.,
1550 Groudine, M., Bender, M., Kaul, R., Canfield, T., et al. (2012). An encyclopedia of mouse DNA
1551 elements (Mouse ENCODE). Genome Biol. *13*, 418.

1552 Stroud, H., Su, S.C., Hrvatin, S., Greben, A.W., Renthal, W., Boxer, L.D., Nagy, M.A., Hochbaum,
1553 D.R., Kinde, B., Gabel, H.W., et al. (2017a). Early-Life Gene Expression in Neurons Modulates
1554 Lasting Epigenetic States. Cell *171*, 1151-1164.e16.

1555 Stroud, H., Su, S.C., Hrvatin, S., Greben, A.W., Renthal, W., Boxer, L.D., Nagy, M.A., Hochbaum,
1556 D.R., Kinde, B., Gabel, H.W., et al. (2017b). Early-Life Gene Expression in Neurons Modulates
1557 Lasting Epigenetic States. Cell *171*, 1151-1164.e16.

1558 Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich,
1559 A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A
1560 knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad.
1561 Sci. *102*, 15545–15550.

1562 Takumi, T., Tamada, K., Hatanaka, F., Nakai, N., and Bolton, P.F. (2019). Behavioral neuroscience
1563 of autism. Neurosci. Biobehav. Rev. S0149763418303725.

1564 Tatton-Brown, K., Seal, S., Ruark, E., Harmer, J., Ramsay, E., Del Vecchio Duarte, S., Zachariou,
1565 A., Hanks, S., O'Brien, E., Aksglaede, L., et al. (2014). Mutations in the DNA methyltransferase
1566 gene DNMT3A cause an overgrowth syndrome with intellectual disability. Nat. Genet. *46*, 385–
1567 388.

1568 Tatton-Brown, K., Zachariou, A., Loveday, C., Renwick, A., Mahamdallie, S., Aksglaede, L.,
1569 Baralle, D., Barge-Schaapveld, D., Blyth, M., Bouma, M., et al. (2018). The Tatton-Brown-
1570 Rahman Syndrome: A clinical study of 55 individuals with de novo constitutive DNMT3A
1571 variants. Wellcome Open Res. *3*, 1–16.

1572 Tilot, A.K., Bebek, G., Niazi, F., Altemus, J.B., Romigh, T., Frazier, T.W., and Eng, C. (2016). Neural
1573 transcriptome of constitutional Pten dysfunction in mice and its relevance to human idiopathic
1574 autism spectrum disorder. Mol. Psychiatry *21*, 118–125.

1575 de la Torre-Ubieta, L., Won, H., Stein, J.L., and Geschwind, D.H. (2016). Advancing the
1576 understanding of autism disease mechanisms through genetics. Nat. Med. *22*, 345–361.

1577 Van Esch, H., Bauters, M., Ignatius, J., Jansen, M., Raynaud, M., Hollanders, K., Lugtenberg, D.,
1578 Bienvenu, T., Jensen, L.R., Gécz, J., et al. (2005). Duplication of the MECP2 Region Is a Frequent
1579 Cause of Severe Mental Retardation and Progressive Neurological Symptoms in Males. Am. J.
1580 Hum. Genet. *77*, 442–453.

1581 Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R.M.,
1582 Blencowe, B.J., and Geschwind, D.H. (2011). Transcriptomic analysis of autistic brain reveals
1583 convergent molecular pathology. Nature *474*, 380–384.

1584 Weinreb, C., and Raphael, B.J. (2016). Identification of hierarchical chromatin domains.
1585 Bioinformatics *32*, 1601–1609.

1586 Wozniak, D.F., Hartman, R.E., Boyle, M.P., Vogt, S.K., Brooks, A.R., Tenkova, T., Young, C., Olney,
1587 J.W., and Muglia, L.J. (2004). Apoptotic neurodegeneration induced by ethanol in neonatal mice
1588 is associated with profound learning/memory deficits in juveniles followed by progressive
1589 functional recovery in adults. Neurobiol. Dis. *17*, 403–414.

1590 Xie, W., Barr, C.L., Kim, A., Yue, F., Lee, A.Y., Eubanks, J., Dempster, E.L., and Ren, B. (2012).
1591 Base-Resolution Analyses of Sequence and Parent-of-Origin Dependent DNA Methylation in the
1592 Mouse Genome. Cell *148*, 816–831.

1593 Yang, M., Silverman, J.L., and Crawley, J.N. (2011). Automated Three-Chambered Social
1594 Approach Task for Mice. Curr. Protoc. Neurosci. *56*, 8.26.1-8.26.16.

1595 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers,
1596 R.M., Brown, M., Li, W., et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biol.
1597 *9*, R137.

1598 Zhang, Z.M., Lu, R., Wang, P., Yu, Y., Chen, D., Gao, L., Liu, S., Ji, D., Rothbart, S.B., Wang, Y., et
1599 al. (2018). Structural basis for DNMT3A-mediated de novo DNA methylation. Nature *554*, 387–
1600 391.

1601

Figure 1

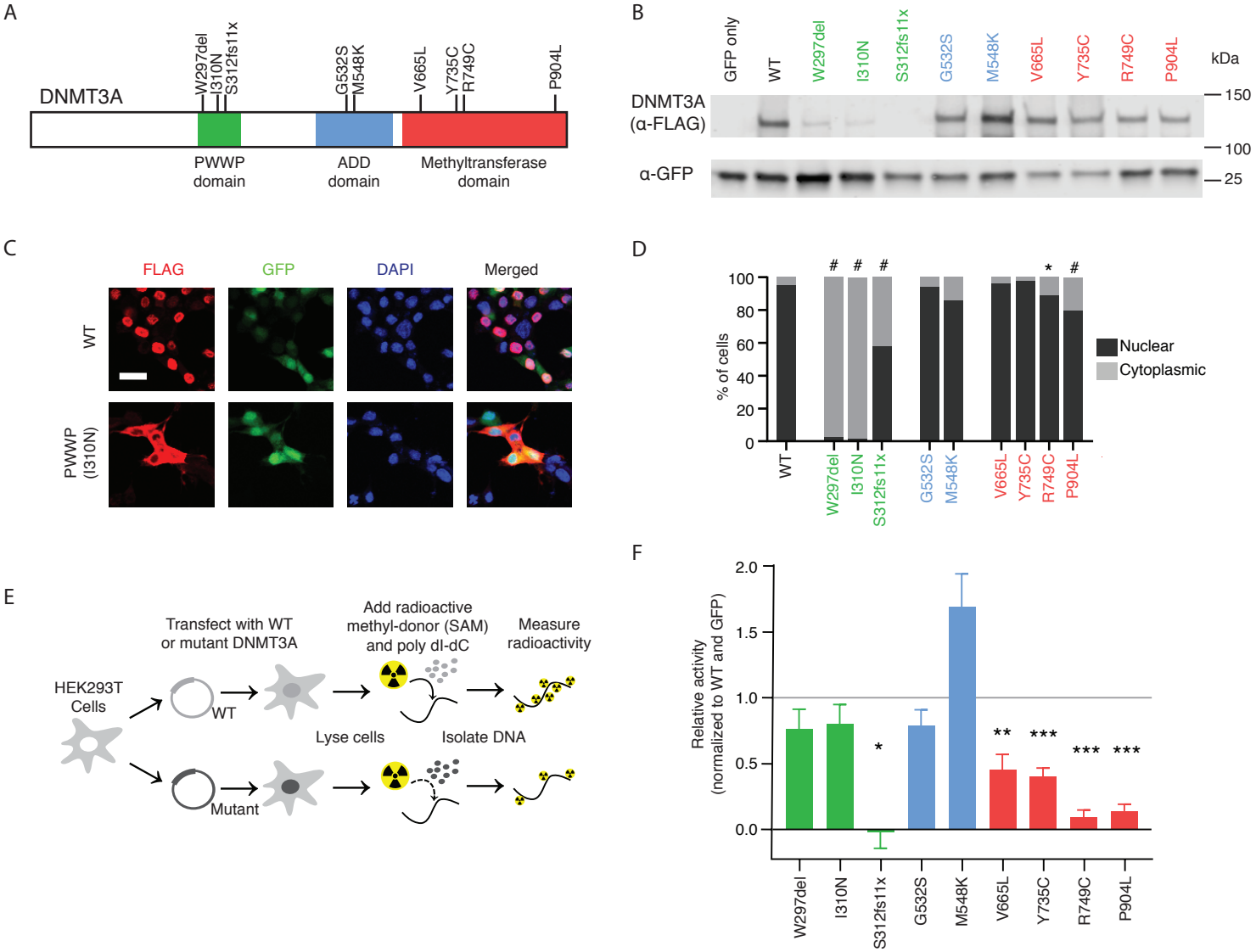

A
DNMT3A
W297del
I310N
S312fs11x
G532S
M548K
V665L
Y735C
R749C
P904L
PWWP domain
ADD domain
Methyltransferase domain
B
GFP only
WT
W297del
I310N
S312fs11x
G532S
M548K
V665L
Y735C
R749C
P904L
kDa
DNMT3A (α-FLAG)
α-GFP
150
100
25
C
FLAG
GFP
DAPI
Merged
WT
PWWP (I310N)
D
% of cells
Nuclear
Cytoplasmic
E
HEK293T Cells
Transfect with WT or mutant DNMT3A
WT
Mutant
Add radioactive methyl-donor (SAM) and poly dI-dC
Measure radioactivity
Lyse cells
Isolate DNA
F
Relative activity (normalized to WT and GFP)

Figure 2

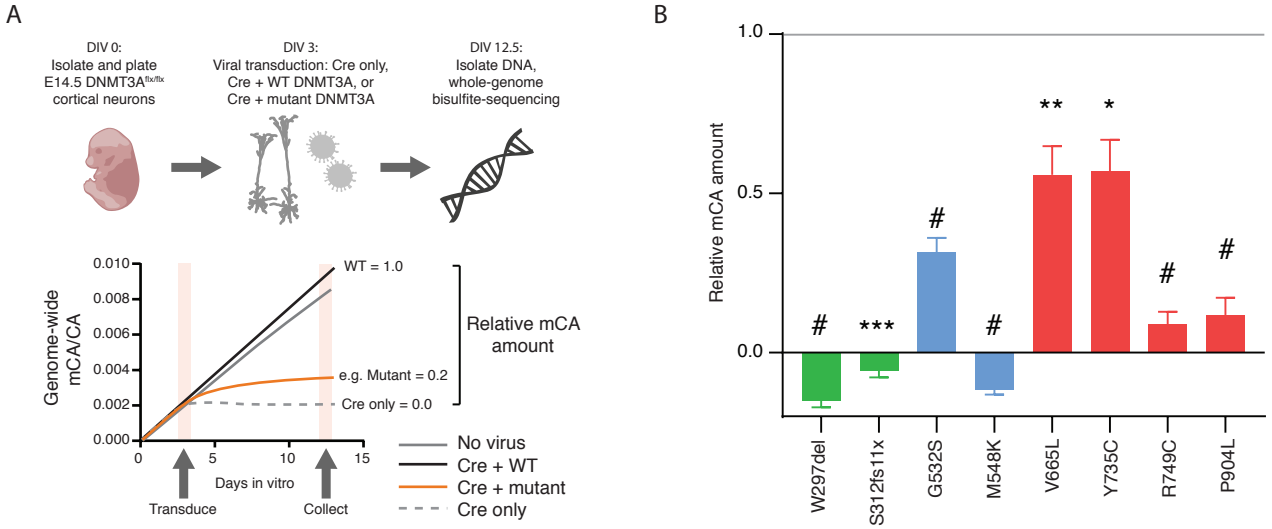

A
DIV 0:
Isolate and plate
E14.5 DNMT3A flx/flx
cortical neurons
DIV 3:
Viral transduction: Cre only,
Cre + WT DNMT3A, or
Cre + mutant DNMT3A
DIV 12.5:
Isolate DNA,
whole-genome
bisulfite-sequencing
Genome-wide
mCA/CA
0.010
0.008
0.006
0.004
0.002
0.000
0
5
10
15
Days in vitro
Transduce
Collect
WT = 1.0
e.g. Mutant = 0.2
Cre only = 0.0
Relative mCA
amount
No virus
Cre + WT
Cre + mutant
Cre only
B
Relative mCA amount
1.0
0.5
0.0
#
***
#
#
**
*
#
#
W297del
S312fs11x
G532S
M548K
V665L
Y735C
R749C
P904L

Figure 3

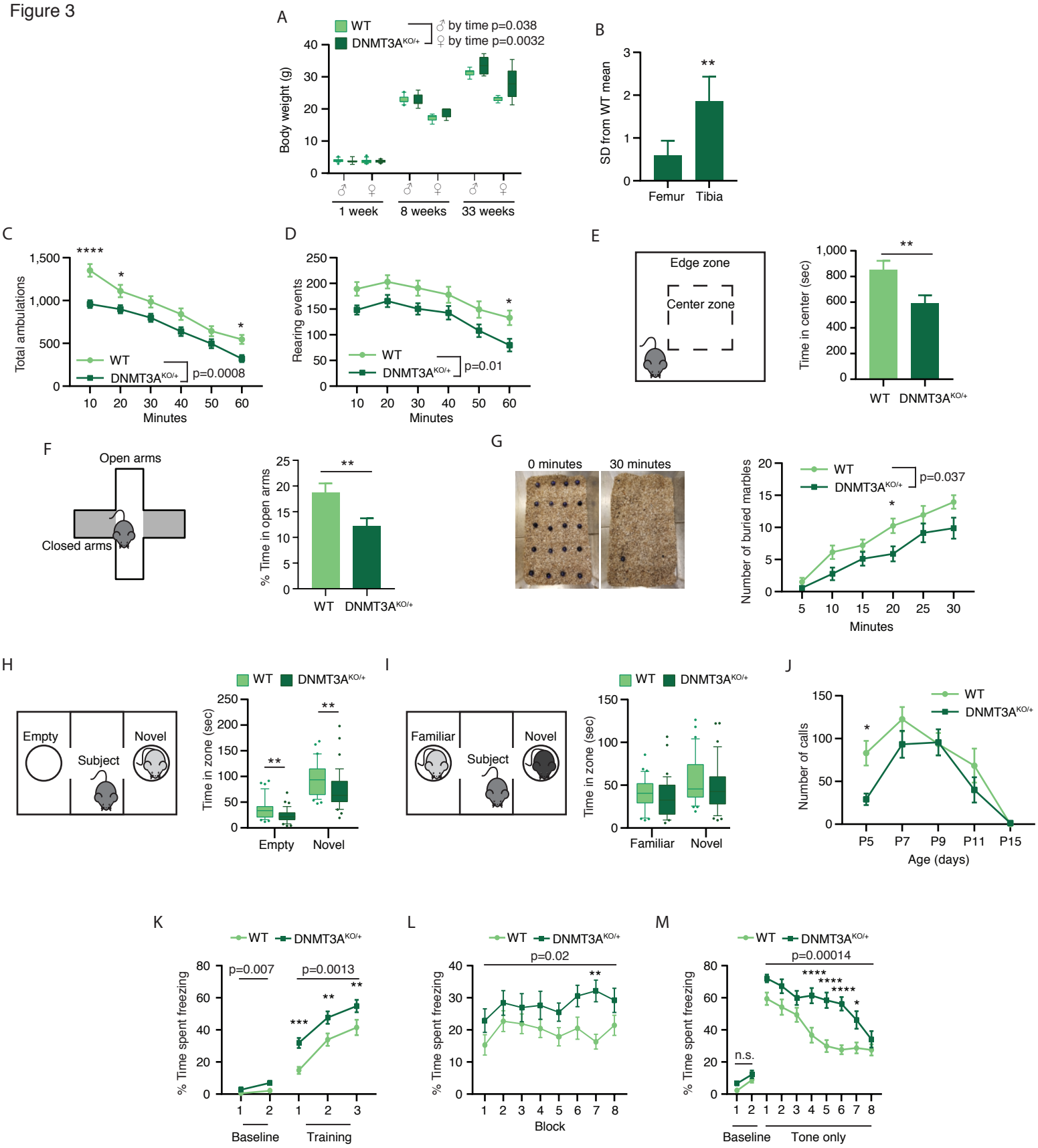

A
WT
DNMT3A KO/+
♂ by time p=0.038
♀ by time p=0.0032
Body weight (g)
1 week
8 weeks
33 weeks
B
SD from WT mean
Femur
Tibia
C
Total ambulations
WT
DNMT3A KO/+
p=0.0008
Minutes
D
Rearing events
WT
DNMT3A KO/+
p=0.01
Minutes
E
Edge zone
Center zone
Time in center (sec)
WT
DNMT3A KO/+
F
Open arms
Closed arms
% Time in open arms
WT
DNMT3A KO/+
G
0 minutes
30 minutes
Number of buried marbles
WT
DNMT3A KO/+
p=0.037
Minutes
H
Empty
Subject
Novel
Time in zone (sec)
WT
DNMT3A KO/+
Empty
Novel
I
Familiar
Subject
Novel
Time in zone (sec)
WT
DNMT3A KO/+
Familiar
Novel
J
Number of calls
WT
DNMT3A KO/+
P5
P7
P9
P11
P15
Age (days)
K
WT
DNMT3A KO/+
% Time spent freezing
p=0.007
p=0.0013
Baseline
Training
L
WT
DNMT3A KO/+
% Time spent freezing
p=0.02
Block
M
WT
DNMT3A KO/+
% Time spent freezing
p=0.00014
n.s.
Baseline
Tone only

Figure 4

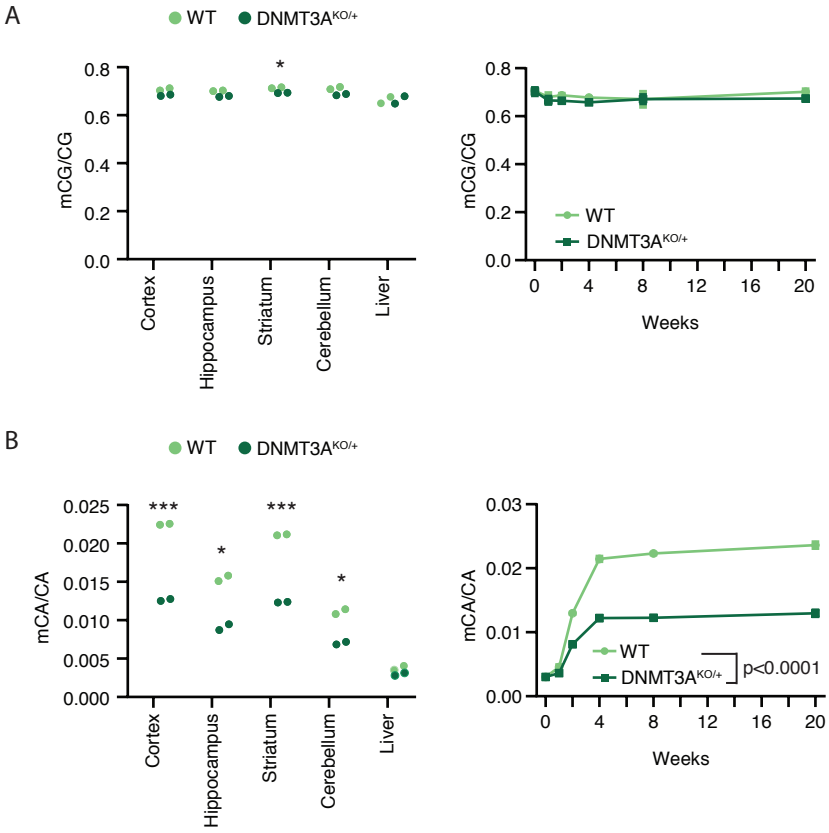

A
WT
DNMT3A^KO/+
mCG/CG
0.8
0.6
0.4
0.2
0.0
*
Cortex
Hippocampus
Striatum
Cerebellum
Liver
Weeks
0
4
8
12
16
20
B
mCA/CA
0.025
0.020
0.015
0.010
0.005
0.000
***
*
***
*
0.03
0.02
0.01
0.00
p<0.0001

Figure 5

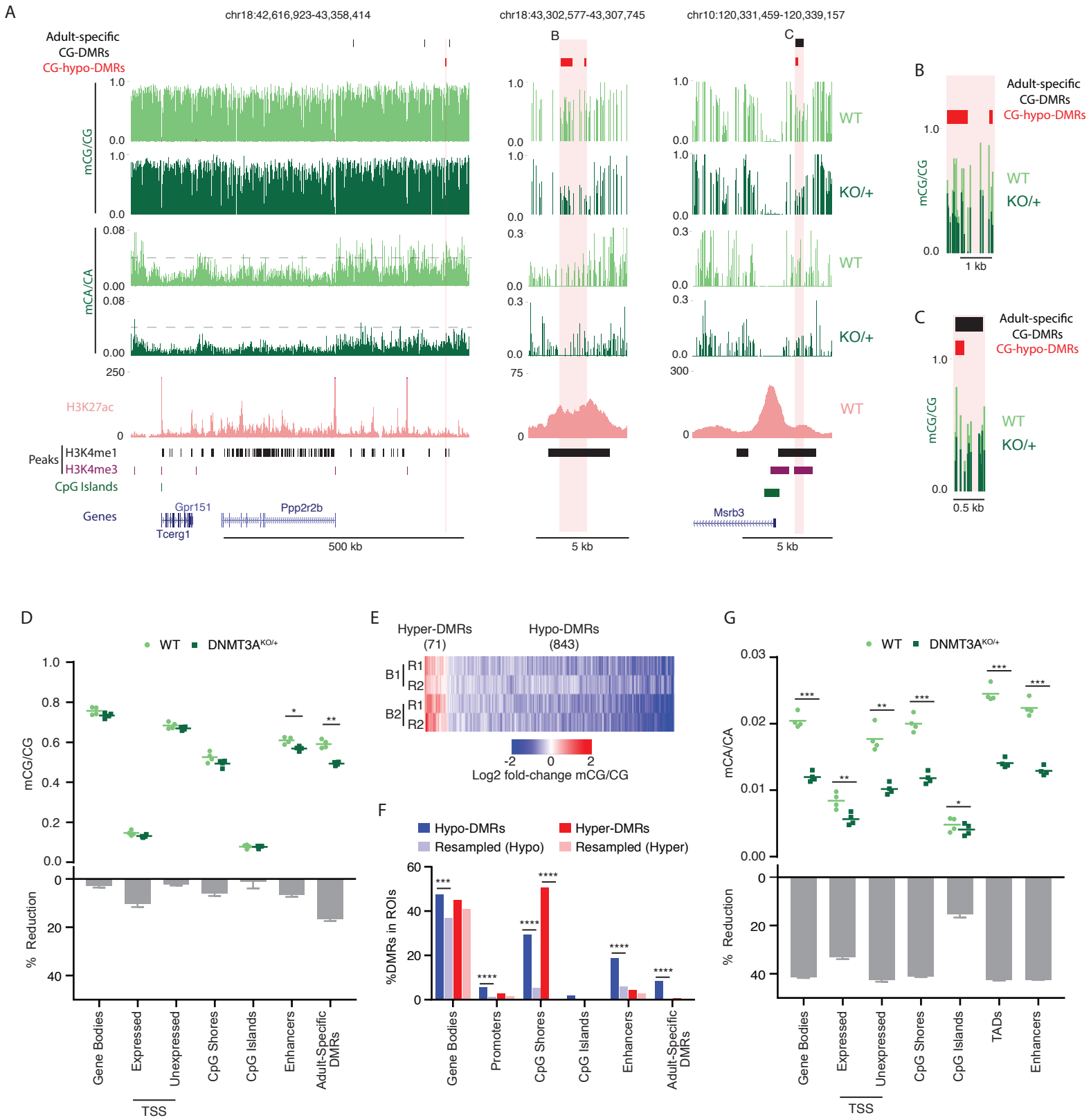

A
chr18:42,616,923-43,358,414
chr18:43,302,577-43,307,745
chr10:120,331,459-120,339,157
Adult-specific CG-DMRs
CG-hypo-DMRs
mCG/CG
mCA/CA
WT
KO/+
H3K27ac
Peaks
H3K4me1
H3K4me3
CpG Islands
Genes
Gpr151
Ppp2r2b
Tcerg1
Msrb3
500 kb
5 kb
1 kb
0.5 kb
B
C
D
WT
DNMT3A^KO/+
% Reduction
Gene Bodies
Expressed
Unexpressed
CpG Shores
CpG Islands
Enhancers
Adult-Specific DMRs
TSS
E
Hyper-DMRs (71)
Hypo-DMRs (843)
B1
B2
R1
R2
Log2 fold-change mCG/CG
F
Hypo-DMRs
Hyper-DMRs
Resampled (Hypo)
Resampled (Hyper)
%DMRs in ROIs
Promoters
G
TADs

Figure 6

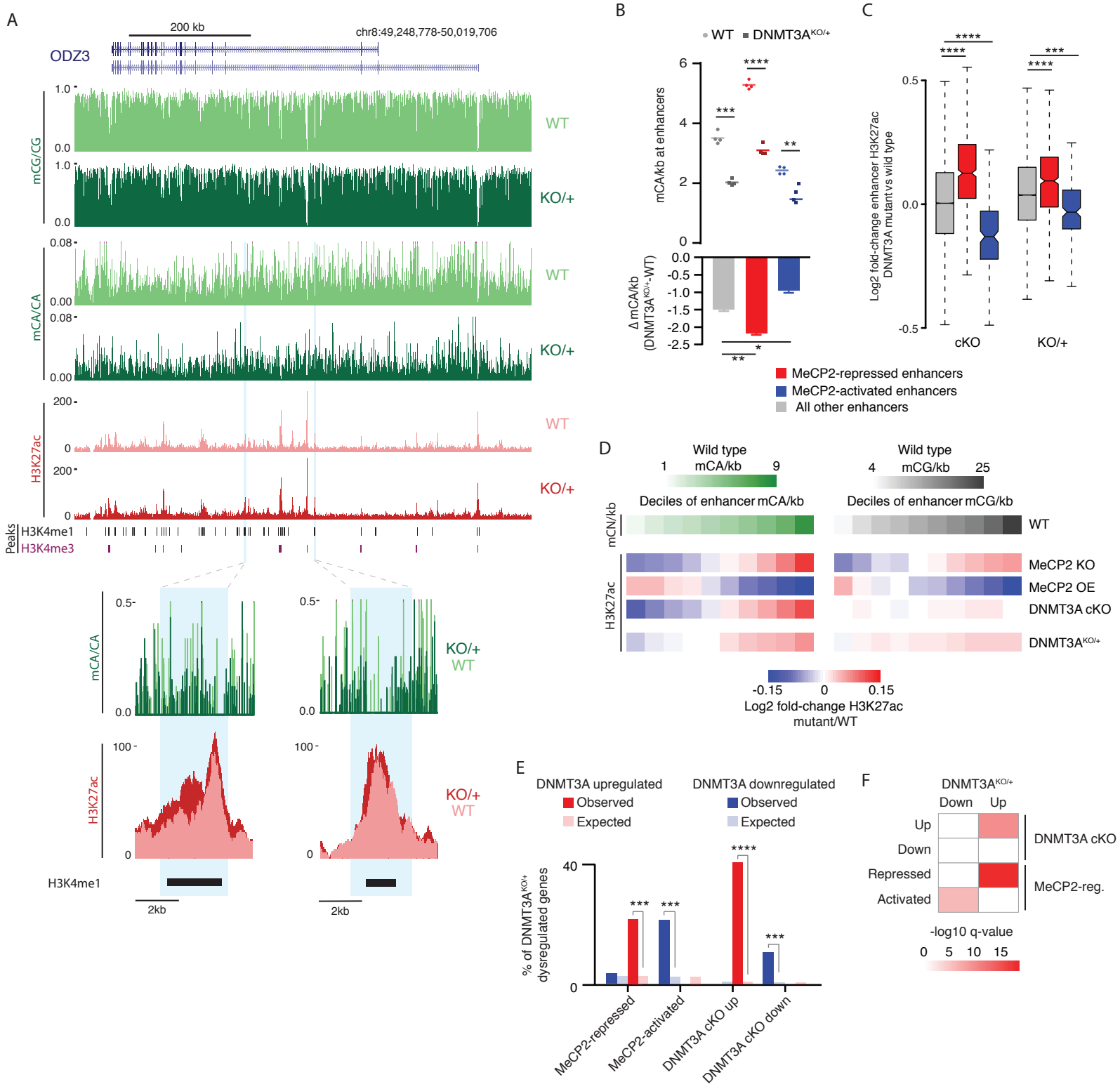
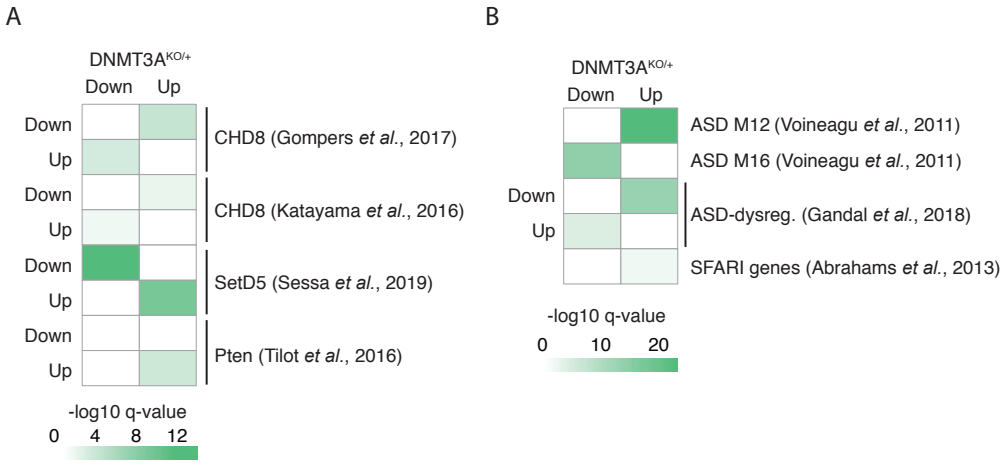
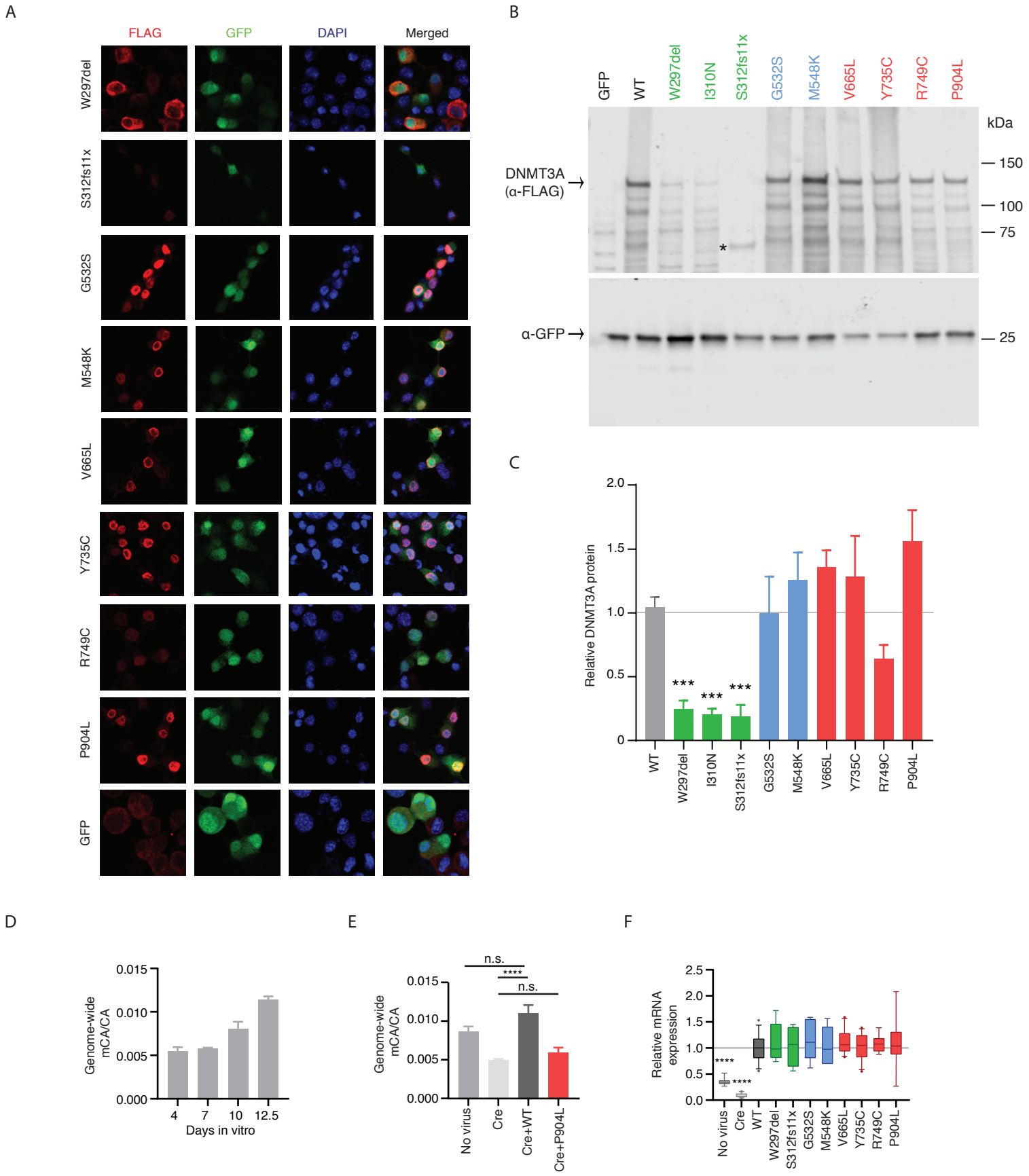

A
200 kb
chr8:49,248,778-50,019,706
ODZ3
mCG/CG
WT
KO/+
mCA/CA
H3K27ac
Peaks
H3K4me1
H3K4me3
2kb
B
WT
DNMT3A KO/+
mCA/kb at enhancers
Δ mCA/kb (DNMT3A KO/+-WT)
MeCP2-repressed enhancers
MeCP2-activated enhancers
All other enhancers
C
Log2 fold-change enhancer H3K27ac DNMT3A mutant vs wild type
cKO
KO/+
D
Wild type mCA/kb
Wild type mCG/kb
Deciles of enhancer mCA/kb
Deciles of enhancer mCG/kb
mCN/kb
H3K27ac
MeCP2 KO
MeCP2 OE
DNMT3A cKO
DNMT3A KO/+
Log2 fold-change H3K27ac mutant/WT
E
DNMT3A upregulated
DNMT3A downregulated
Observed
Expected
% of DNMT3A KO/+ dysregulated genes
MeCP2-repressed
MeCP2-activated
DNMT3A cKO up
DNMT3A cKO down
F
DNMT3A KO/+
Down
Up
Repressed
Activated
DNMT3A cKO
MeCP2-reg.
-log10 q-value

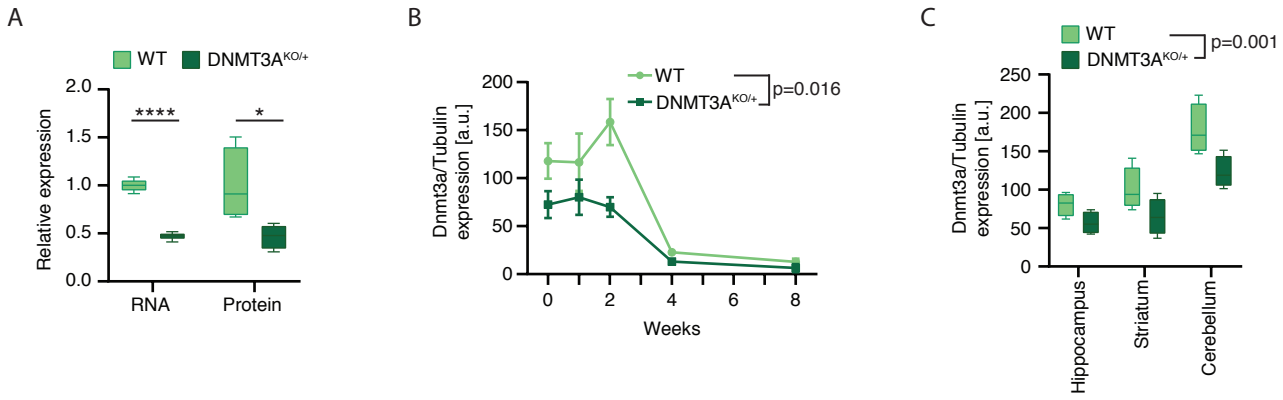
Figure 7


A
DNMT3A KO/+
Down
Up
Down
Up
Down
Up
Down
Up
Down
Up
CHD8 (Gompers et al., 2017)
CHD8 (Katayama et al., 2016)
SetD5 (Sessa et al., 2019)
Pten (Tilot et al., 2016)
-log10 q-value
0
4
8
12
B
DNMT3A KO/+
Down
Up
Down
Up
ASD M12 (Voineagu et al., 2011)
ASD M16 (Voineagu et al., 2011)
ASD-dysreg. (Gandal et al., 2018)
SFARI genes (Abrahams et al., 2013)
-log10 q-value
0
10
20

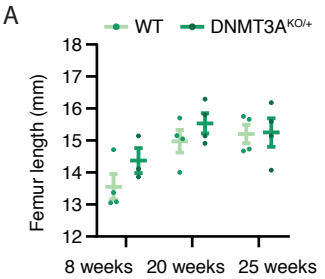
Supplementary Figure 1

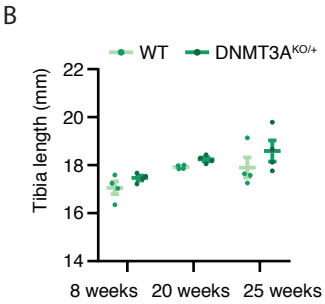


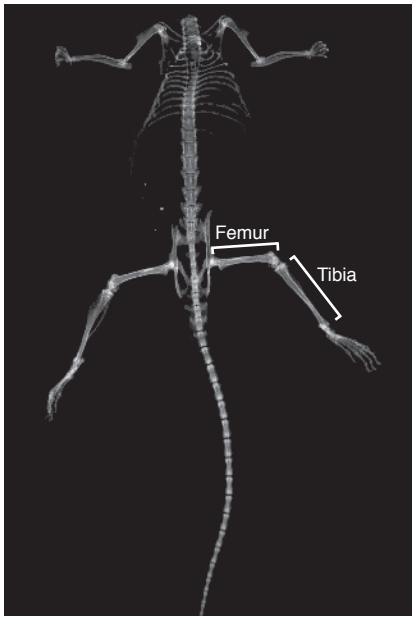
Supplementary Figure 2

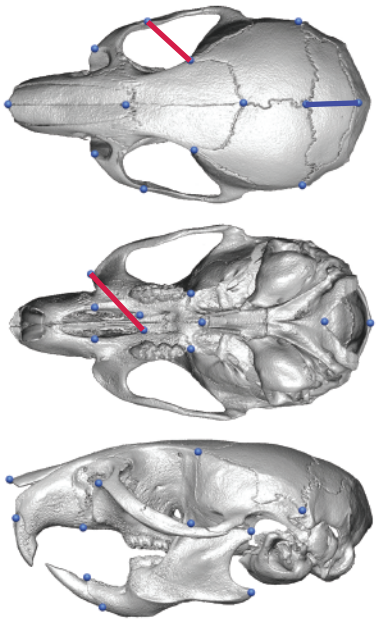


Supplementary Figure 3

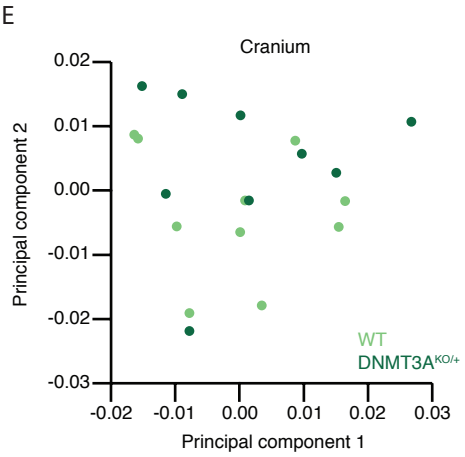


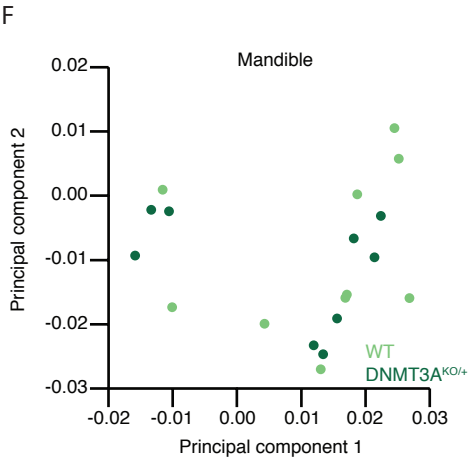




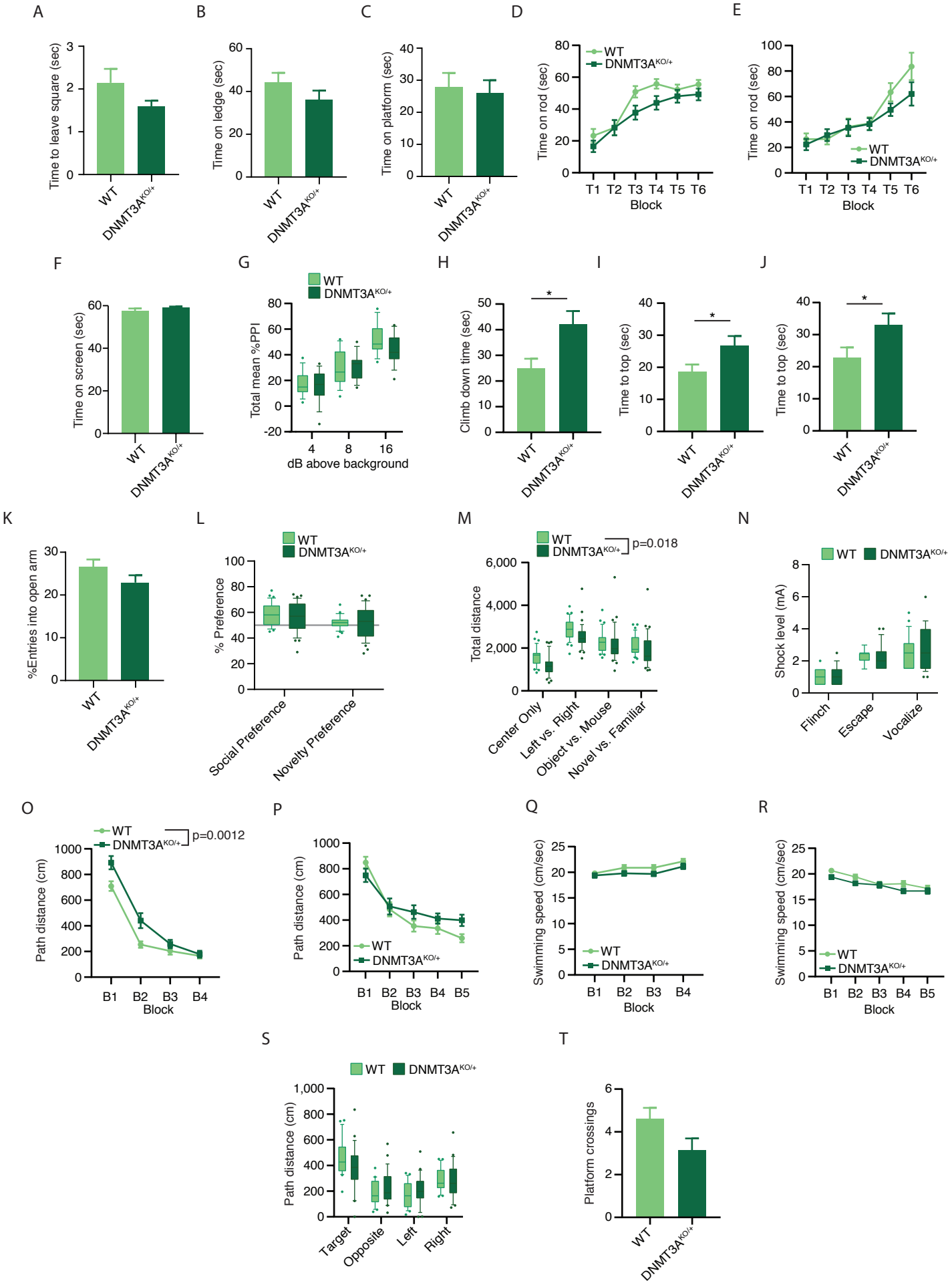


Key: WT significantly larger than DNMT3A mice
WT significantly smaller than DNMT3A mice

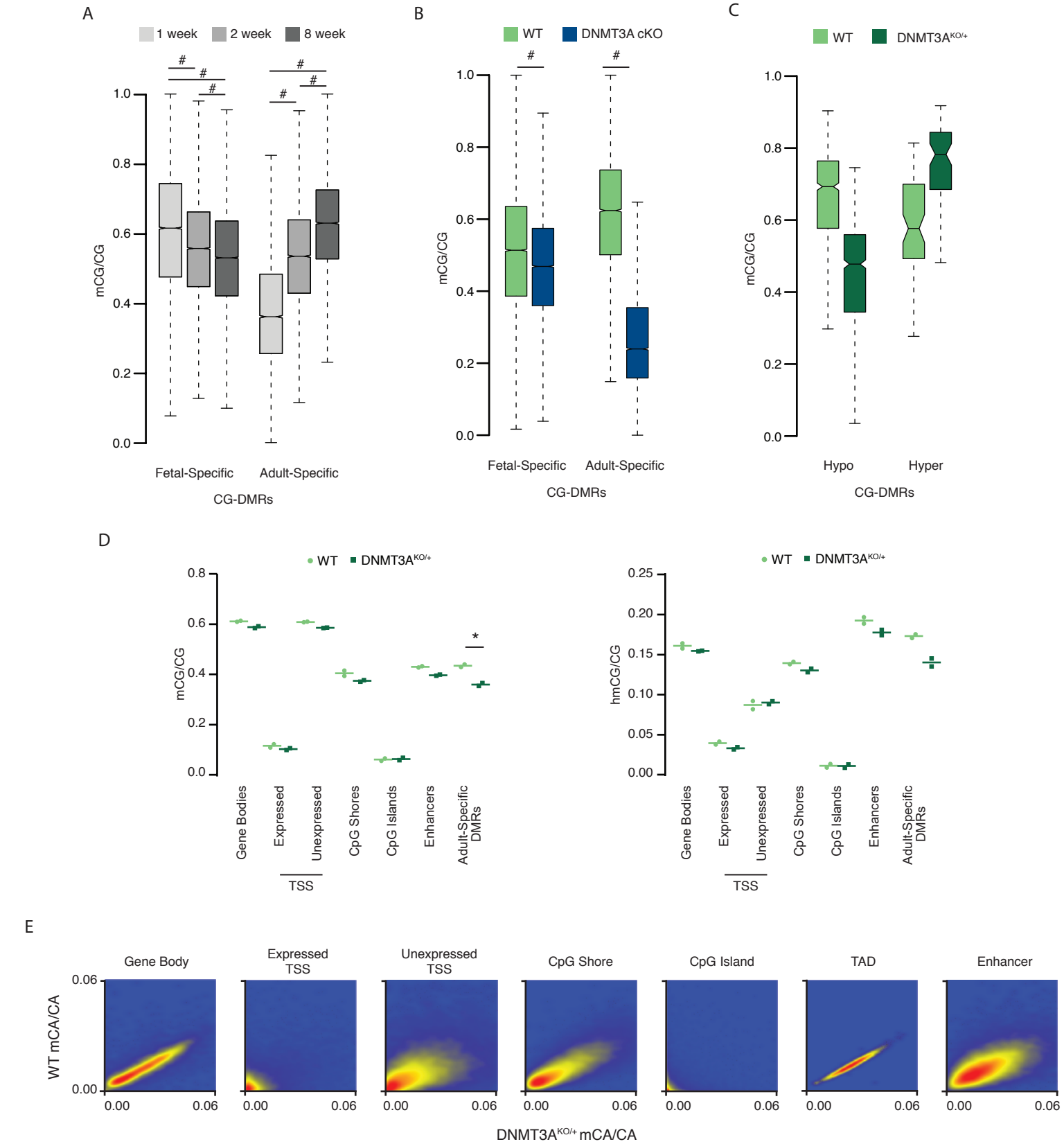
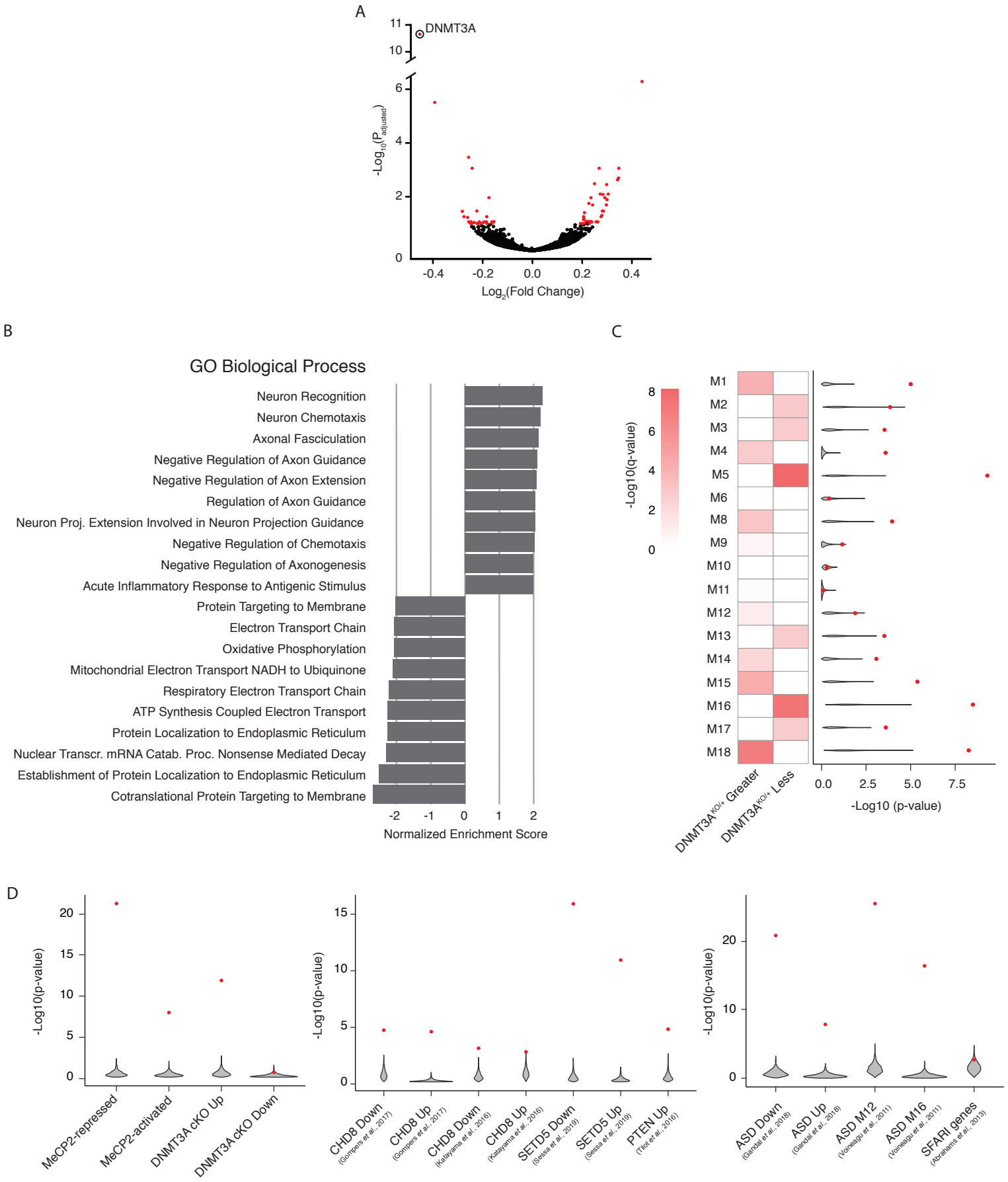




Supplementary Figure 4



Supplementary Figure 5



Supplementary Figure 6