

1           **Heterozygous mutation to *Chd8* causes macrocephaly and widespread**  
2           **alteration of neurodevelopmental transcriptional networks in mouse**

3

4 Andrea L. Gompers<sup>1,2^</sup>, Linda Su-Feher<sup>1,2^</sup>, Jacob Ellegood<sup>3</sup>, Tyler W. Stradleigh<sup>1,2</sup>, Iva  
5 Zdilar<sup>1,2</sup>, Nycole A. Copping<sup>1,4</sup>, Michael C. Pride<sup>1,4</sup>, Melanie D. Schaffler<sup>1,4</sup>, M. Asrafuzzaman  
6 Riyadh<sup>5</sup>, Gaurav Kaushik<sup>5</sup>, Brandon Mannion<sup>6</sup>, Ingrid Plajzer-Frick<sup>6</sup>, Veena Afzal<sup>6</sup>, Axel  
7 Visel<sup>6,7,8</sup>, Len A. Pennacchio<sup>6,7</sup>, Diane Dickel<sup>6</sup>, Jason P. Lerch<sup>3</sup>, Jacqueline N. Crawley<sup>1,4</sup>,  
8 Konstantinos S. Zarbalis<sup>5</sup>, Jill L. Silverman<sup>1,4</sup>, Alex S. Nord<sup>1,2\*</sup>

9

10 <sup>1</sup>University of California, Davis, Department of Psychiatry and Behavioral Sciences, <sup>2</sup>University  
11 of California, Davis, Department of Neurobiology, Physiology and Behavior, <sup>3</sup>The Hospital for  
12 Sick Children, Mouse Imaging Centre, Toronto, Canada, <sup>4</sup>University of California, Davis, MIND  
13 Institute, School of Medicine, Davis, CA, <sup>5</sup>University of California, Davis, Department of  
14 Pathology and Laboratory Medicine, Shriners Hospitals for Children, Institute for Pediatric  
15 Regenerative Medicine, Sacramento, CA <sup>6</sup>Lawrence Berkeley National Laboratory, Berkeley,  
16 CA, <sup>7</sup>Department of Energy Joint Genome Institute, Walnut Creek, CA, <sup>8</sup>School of Natural  
17 Sciences, University of California, Merced, CA.

18

19

20 ^ These authors contributed equally

21 \* Corresponding author

22

## 23 Summary

24 The chromatin remodeling gene *CHD8* represents a central node in early neurodevelopmental  
25 gene networks implicated in autism. We examined the impact of heterozygous germline *Chd8*  
26 mutation on neurodevelopment in mice. Network analysis of neurodevelopmental gene  
27 expression revealed subtle yet strongly significant widespread transcriptional changes in *Chd8*<sup>+/-</sup>  
28 mice across autism-relevant networks from neurogenesis to synapse function. *Chd8*<sup>+/-</sup> expression  
29 signatures included enrichment of RNA processing genes and a *Chd8*-regulated module featuring  
30 altered transcription of chromatin remodeling, splicing, and cell cycle genes. *Chd8*<sup>+/-</sup> mice  
31 exhibited increased proliferation during brain development and neonatal increase in cortical  
32 length and volume. Structural MRI confirmed regional brain volume increase in adult *Chd8*<sup>+/-</sup>  
33 mice, consistent with clinical macrocephaly. Adult *Chd8*<sup>+/-</sup> mice displayed normal social  
34 interactions, and repetitive behaviors were not evident. Our results show that *Chd8*<sup>+/-</sup> mice  
35 exhibit neurodevelopmental changes paralleling *CHD8*<sup>+/-</sup> humans and show that *Chd8* is a global  
36 genomic regulator of pathways disrupted in neurodevelopmental disorders.

37

## 38 **Introduction**

39 DNA packaging determines the transcriptional potential of a cell and is central to the  
40 development and function of metazoan cell types. Chromatin remodeling complexes control  
41 local chromatin state, yielding either transcriptional activation or repression. Pluripotency,  
42 proliferation, and differentiation are dependent on genomic regulation at the chromatin level, and  
43 proteins that control chromatin packaging are critical in development and cancer<sup>1</sup>. Although  
44 many chromatin remodeling factors function across systems, case sequencing efforts have linked  
45 mutations of chromatin genes with specific, causal roles in neurodevelopmental disorders<sup>2-5</sup>. This  
46 finding is particularly strong for rare and *de novo* mutations in autism spectrum disorder  
47 (ASD)<sup>6,7</sup>. Understanding how mutations to chromatin remodeling genes affect transcriptional  
48 regulation during brain development may reveal developmental and cellular mechanisms driving  
49 neurodevelopmental disorders.

50 A key gene that has emerged from studies profiling rare and *de novo* coding variation in  
51 ASD is the chromatin remodeler, *CHD8* (*Chromodomain helicase DNA binding protein 8*)<sup>8</sup>. In  
52 addition to ASD, *CHD8*<sup>+/-</sup> individuals exhibit macrocephaly, distinct craniofacial morphology,  
53 mild-to-severe cognitive impairment, and gastrointestinal problems<sup>8</sup>. *CHD8* mutation also has  
54 been linked to attention deficit hyperactivity disorder, seizures, and schizophrenia<sup>4,8</sup>, as well as  
55 cancer<sup>9,10</sup>. Homozygous deletion of *Chd8* in mice is early embryonic lethal<sup>11</sup>. *Chd8* knockdown  
56 in zebrafish recapitulated macrocephaly and gastrointestinal phenotypes<sup>8,12</sup>, suggesting a high  
57 degree of evolutionary conservation of Chd8 function in brain development.

58 Studies of genetic and protein networks have raised the possibility that CHD8 is a central  
59 node and master regulator of early neurodevelopmental networks implicated in autism<sup>12-14</sup>.  
60 CHD8 has been proposed to achieve this regulatory function in brain development by binding to

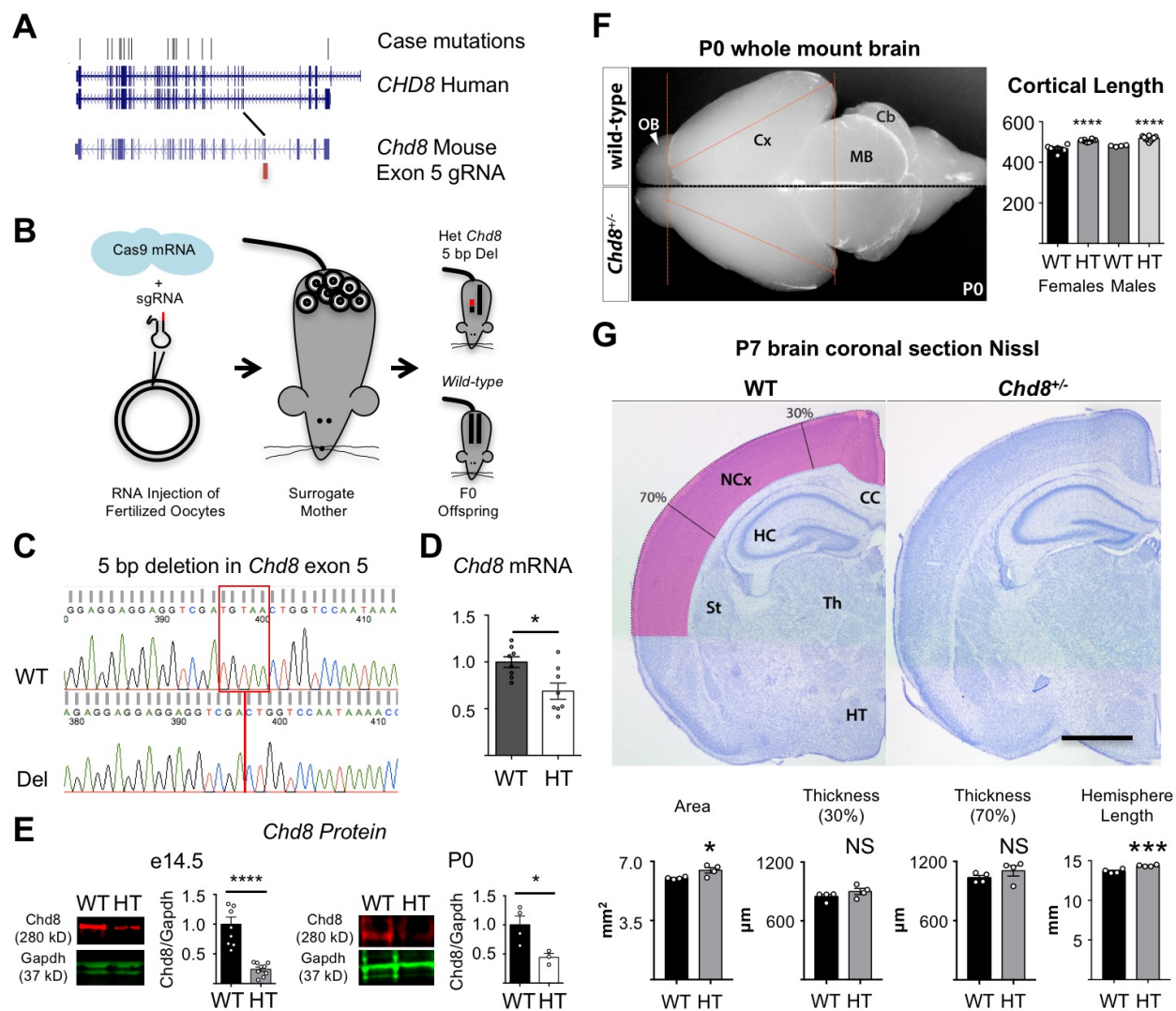
61 relevant gene promoters and enhancers<sup>12,14</sup>. CHD8 DNA binding and knockdown studies in  
62 human and mouse tissues and cells have revealed a multitude of genes directly and indirectly  
63 activated or repressed by CHD8 during neurodevelopment<sup>12,14</sup>.

64 Consequently, characterizing the functional impact of heterozygous *CHD8* mutation on  
65 brain development could reveal generalizable mechanisms linking chromatin biology to  
66 pathology. Towards this goal, we generated two new *Chd8* mutant mouse lines and performed  
67 analyses to characterize neuroanatomic, transcriptional, and behavioral phenotypes of *Chd8*<sup>+/-</sup>  
68 mice. Our interrogations identified changes in structural and developmental neuroanatomy and  
69 subtle but highly significant changes to developmental gene expression. These results provide  
70 insight into *in vivo* pathological changes, showing that germline *Chd8* haploinsufficiency results  
71 in altered gene expression across neurodevelopment and produces increased regional brain  
72 volume. The results from this study indicate the presence of genomic and neuroanatomic  
73 phenotypes that parallel the clinical signature of human *CHD8* mutations, suggesting similar  
74 neurodevelopmental pathology between human and mouse.

## 75 **Results**

### 76 ***Mice harboring heterozygous germline Chd8 mutation exhibit megalencephaly***

77 We used CRISPR/Cas9 targeting to generate two mouse lines harboring short deletions in  
78 *Chd8*, upstream of the majority of human mutations identified in autism cohorts<sup>8</sup> (Figure 1A-  
79 1C). Consistent with an earlier study<sup>11</sup>, our two newly generated *Chd8* alleles (5 and 14 bp  
80 deletions within exon 5) resulted in embryonic lethality in homozygous mutants, but  
81 heterozygous (*Chd8*<sup>+/-</sup>) mice were viable, reached a normal lifespan, and were fertile irrespective  
82 of sex. Quantitative PCR (qPCR) and western blot analysis on brain lysates from embryonic day



**Figure 1. *Chd8*<sup>+/-</sup> mouse model.** **A.** Location of case mutations in human *CHD8* and corresponding guideRNA sequence homology for Cas9-targeting of mouse *Chd8*. **B.** Schematic of mouse line generation. **C.** Sequence trace showing 5 bp indel in exon 5. **D.** qPCR showing reduction of RNA in *Chd8*<sup>+/-</sup> forebrain at P0 (n WT = 7, n *Chd8*<sup>+/-</sup> = 5). **E.** Western blot of *Chd8*<sup>+/-</sup> mice showing reduction of Chd8 protein (ab114126; Abcam) in *Chd8*<sup>+/-</sup> forebrain at E14.5 (n WT = 9, *Chd8*<sup>+/-</sup> = 9) and P0 (n WT = 4, *Chd8*<sup>+/-</sup> = 3). **F.** Whole mount brain of *Chd8*<sup>+/-</sup> mice at P0 reveals increased cortical length, indicative of megalencephaly. OB: olfactory bulb, Cx: cortex, MB: midbrain, Cb: cerebellum. WT n male = 10, female = 4; *Chd8*<sup>+/-</sup> n male = 10, female = 4. **G.** (Upper) Representative coronal sections of wild-type and *Chd8*<sup>+/-</sup> brains at P7 visualized with Nissl, n = 4 for both genotypes. Scale bar 1000  $\mu$ m. (Lower) Plots (mean  $\pm$  SEM with dots representing individual samples) of cortical area, thickness at 30% and 70% distance from the dorsal midline, and cortical hemispheric circumference. P-values derived using Student's t-test for D, E and G and using ANOVA for F, \*P < 0.05 \*\*\*\*P < 0.005 \*\*\*\*P < 0.001.

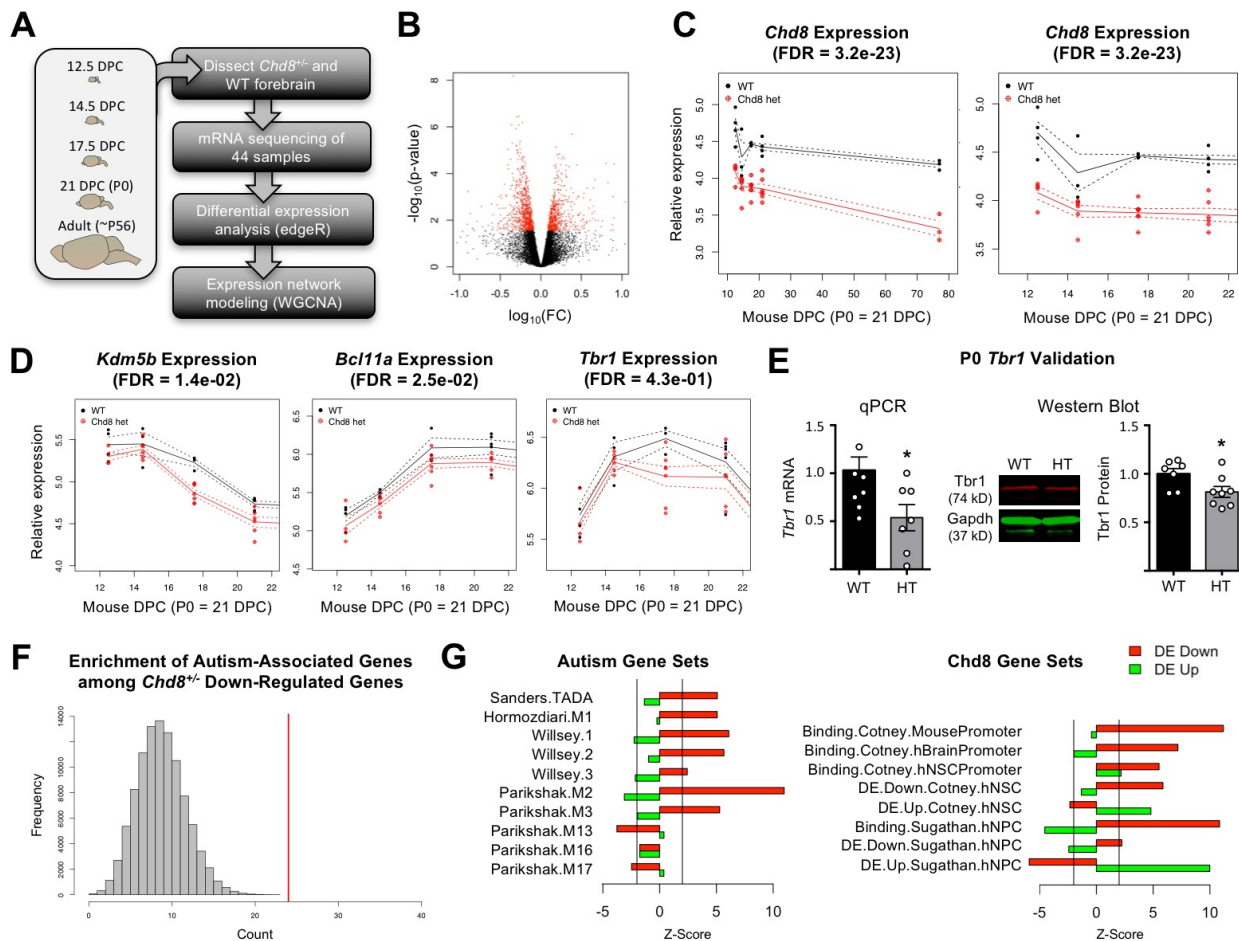
83 14.5 (E14.5), postnatal day 0 (P0), and adult mice (>P56) showed that heterozygous mutations to  
 84 *Chd8* resulted in decreased *Chd8* transcript and protein levels (Figure 1D-1E, S1C; Chd8  
 85 antibody: ab114126, Abcam). For the majority of the following studies, we analyzed mice

86 harboring a 5 bp deletion in *Chd8* exon 5. Male *Chd8*<sup>+/-</sup> mice were bred to wild-type females for  
87 at least four generations before further experiments, and multiple litters were used for all  
88 experiments to eliminate the potential impact of Cas9 off-target mutation. We tested for  
89 differences in brain size in *Chd8*<sup>+/-</sup> mice at birth (postnatal day 0, P0), as macrocephaly is a  
90 hallmark trait in human *CHD8*<sup>+/-</sup> individuals<sup>8</sup>. Maximal cortical anteroposterior length of *Chd8*<sup>+/-</sup>  
91 brains was ~7% longer than matched wild-type (WT) littermates (ANOVA, p-value = 0.0440)  
92 with no significant differences between sexes (Figure 1F). These results show that *Chd8*<sup>+/-</sup> mice  
93 are megalencephalic, suggesting neuropathological phenotypes that parallel *CHD8*<sup>+/-</sup> humans.

94 To further test the parameters of developmental megalencephaly in *Chd8*<sup>+/-</sup> mutants, we  
95 examined brains in whole-mount and Nissl-stained coronal brain sections at P7 (Figure 1G), a  
96 time point after the conclusion of developmental neurogenesis and gliogenesis. At this stage, the  
97 mutant brains did not present overt neuropathological anomalies other than size increase. Total  
98 hemispheric circumference in *Chd8*<sup>+/-</sup> brains was 4.7% longer (Student's t-test, p = 0.0001) than  
99 in WT littermates (Figure 1G). We measured cortical thickness in two positions across the  
100 neocortex, at 30% and 70% distance from the dorsal midline. No significant differences between  
101 the genotypes were observed in either location (Student's t-test, p = 0.2242; 0.2678), though  
102 *Chd8*<sup>+/-</sup> mice trended larger across both measurements. Overall neocortical section area was ~8%  
103 higher in *Chd8*<sup>+/-</sup> brains (Student's t-test, p = 0.0009) compared to WT controls, confirming  
104 cerebral megalencephaly at this stage.

### 105 ***Differential gene expression across neurodevelopment in Chd8<sup>+/-</sup> mice***

106 Having established neuroanatomical changes in *Chd8*<sup>+/-</sup> mice that parallel clinical  
107 phenotypes described in *CHD8*<sup>+/-</sup> humans and considering the role of *Chd8* in global  
108 transcriptional regulation, we profiled mRNA utilizing RNA-sequencing in forebrain dissected



**Figure 2. Differential gene expression in *Chd8*<sup>+/-</sup> neurodevelopment.** **A.** Schematic of our experimental pipeline: dissections of forebrain at five stages (E12.5, E14.5, E17.5, P0, ~P56) followed by RNA isolation, library preparation, sequencing and expression networking analysis (WGCNA). **B.** Volcano plot showing that most DE genes (plotted in red) exhibit relatively subtle fold changes. **C.** *Chd8* is the top differentially expressed gene, with panels showing expression log (RPKM) in *Chd8*<sup>+/-</sup> and WT littermates across brain development (right panel shows only developmental stages). FDR of DE p-value shown. **D.** Example expression patterns in *Chd8*<sup>+/-</sup> forebrain of three DE autism risk loci across developmental stages (FDR of DE p-value shown). **E.** Validation of DE expression of *Tbr1* RNA (left; n WT = 7, *Chd8*<sup>+/-</sup> = 5) and protein (right; n WT = 7, *Chd8*<sup>+/-</sup> = 8) in *Chd8*<sup>+/-</sup> forebrain at P0 (Student's t-test, \*P = 0.0264, 0.0338). **G.** Distribution of expected number of autism-associated genes among based on 100,000 randomly sampled gene sets (grey bars) versus observed number of autism-associated DE genes in our dataset (red bar). **H.** Comparison of DE down- and up-regulated genes identified here with autism- and *Chd8*-binding genes. Z-score generated via permutation test comparing enrichment of test gene set to randomly sampled genes.

109 from four early developmental stages (embryonic days E12.5, E14.5, E17.5, and P0) and adult  
 110 mice (>P56) (Figure 2A). After quality filtering, we individually analyzed 26 *Chd8*<sup>+/-</sup> and 18 WT  
 111 littermates, with full sample details reported in Table S1. Sample developmental stage  
 112 represented the major components of gene expression variation, as expected given the large

113 changes to transcription that occur across brain development (Figure S1). We observed reads  
114 overlapping the *Chd8* deletion sequence in all but one *Chd8*<sup>+/-</sup> library and in no WT libraries.  
115 Decreased expression and corresponding decreased Chd8 protein levels were present in mice  
116 harboring the 5 bp exon 5 deletion as well as in the second line of mutant *Chd8*<sup>+/-</sup> mice harboring  
117 an overlapping 14 bp deletion (Figure S1). For both mutant lines, *Chd8* deletion reads occurred  
118 at lower frequency relative to WT allele reads, suggesting that the frameshift transcript  
119 undergoes degradation (Figure S1).

120         Using a statistical model that accounted for sex, developmental stage, and sequencing  
121 batch, we tested for differential expression across 14,163 genes that were robustly expressed in  
122 our datasets. At a significance cutoff corresponding to FDR < 0.05 (p-value < 0.0006), FDR <  
123 0.1 (p-value < 0.0029) or FDR < 0.25 (p-value < 0.0272), 178, 418, and 1,536 genes,  
124 respectively, were differentially expressed (DE) (Table S2). While our full model is best suited  
125 for identification of genes where DE extends across developmental stages, we also examined  
126 stage-specific expression changes and full results are reported in Table S2. DE genes identified  
127 in our full model exhibited a range of expression trajectories across development, and the  
128 majority of significant expression changes in *Chd8*<sup>+/-</sup> were relatively small (99.3% < 1.0 absolute  
129 log fold change, Figure 2B). These findings suggest that changes in neurodevelopmental gene  
130 expression are widespread yet subtle. We validated change in expression at P0 for a set of DE  
131 genes via qPCR (Figure S2, primers used reported in Table S3). Confirming our model, the top  
132 DE gene was *Chd8*, with decreased expression in the *Chd8*<sup>+/-</sup> mice (log fold change = 0.59, p-  
133 value = 2.20E-27, FDR = 3.18E-23, Figure 2B). Irrespective of genotype, *Chd8* expression  
134 gradually declined across development in mouse forebrain, and significant reduction in  
135 expression was observed in *Chd8*<sup>+/-</sup> mice at each stage (Figure 2C). No obvious isoform-specific



136 changes in *Chd8* expression were present in the *Chd8*<sup>+/-</sup> mice based on exon coverage (Figure  
137 S3). DE genes were significantly overrepresented among a number of Reactome<sup>15</sup> pathways  
138 (Table S4). Example pathways with strong enrichment include RNA processing (e.g. Processing  
139 of Capped Intron-Containing Pre-mRNA, FDR = 1.59E-09), gene expression (FDR = 1.66E-04),  
140 cell cycle (FDR = 0.002), Regulation of TP53 Activity (FDR = 0.009), and Axon Guidance  
141 (FDR = 0.009). Similar overall pathway enrichment was observed for DE genes at FDR cutoff of  
142 0.05, 0.1, and 0.25. The strong signature of differential expression enabled us to map the  
143 perturbation of biological pathways and processes caused by *Chd8* haploinsufficiency across  
144 neurodevelopment.

#### 145 ***Overlap of autism-relevant genes and Chd8 binding targets with Chd8*<sup>+/-</sup> DE genes**

146 First we examined whether *Chd8* acts as a regulator of autism-linked genes during brain  
147 development. In agreement with previous *in vitro* knockdown models<sup>12,14</sup>, autism risk genes were  
148 among DE genes in *Chd8*<sup>+/-</sup>. Figure 2D shows example DE high-confidence autism risk loci  
149 (*Kdm5b*, *Bcl11a*, and *Tbr1*) with different developmental expression patterns. We validated  
150 decrease in mRNA expression and protein level for *Tbr1* at P0 in *Chd8*<sup>+/-</sup> forebrain via qPCR  
151 analysis and western blot (Figure 2E, Figure S4). We next tested for overlap between DE down-  
152 regulated and up-regulated genes (FDR < 0.25) and published gene sets of relevance to autism  
153 genetics and *Chd8* regulation. For example, of the 143 genes implicated by presence of  
154 mutations in autism cases<sup>16</sup> that were detected in our expression data, 24 were DE at FDR < 0.25  
155 and down-regulated (permutation test p-value = 3.08E-07) (Figure 2F). We also observed  
156 significant enrichment among down-regulated DE genes with autism risk genes identified by  
157 other studies<sup>13,17</sup> (Figure 2G). We examined global gene co-expression networks relevant to  
158 autism as identified via network analysis of human neurodevelopmental gene expression<sup>18</sup>,

159 including two early developmental autism-associated networks (Parikshak.M2 and  
160 Parikshak.M3) as well as three autism-relevant networks expressed later in human brain  
161 development (Parikshak.M13, Parikshak.M16, Parikshak.M17). We observed strong enrichment  
162 specific to the early developmental networks among the global set of *Chd8*<sup>+/-</sup> down-regulated  
163 genes (Parikshak.M2 p-value = 5.30E-28; Parikshak.M3 p-value = 1.20E-07). We observed no  
164 global enrichment among DE down-regulated genes for the later developmental modules or  
165 among FMRP targets<sup>19</sup> and gene networks identified in postmortem autism case brains<sup>20</sup>.

166       Next we asked whether our DE data is consistent with CHD8 binding and differential  
167 expression after *CHD8* knockdown in human *in vitro* models, as reported in previous studies<sup>12,14</sup>.  
168 While these studies show that CHD8 directly regulates a large number of promoters across  
169 mouse and human systems, we nonetheless observed consistent enrichment among down-  
170 regulated DE genes in *Chd8*<sup>+/-</sup> forebrain for CHD8 target genes identified in these earlier studies  
171 (Figure 2G). For example, down-regulated DE genes from our study were enriched for genes  
172 with *Chd8* promoter binding in E17.5 mouse frontal or occipital cortex<sup>14</sup> (p-value = 4.70E-29).  
173 There was no enrichment among our up-regulated DE sets for genes targeted by CHD8,  
174 suggesting that up-regulation is indirect or occurs at earlier time points. Comparing DE genes  
175 that exhibited down- or up-regulation in our study with genes that show DE in the matched  
176 direction in the independent knockdown studies discussed above, we observed strong direction-  
177 specific enrichment between our study and these two previous studies (Figure 2G).

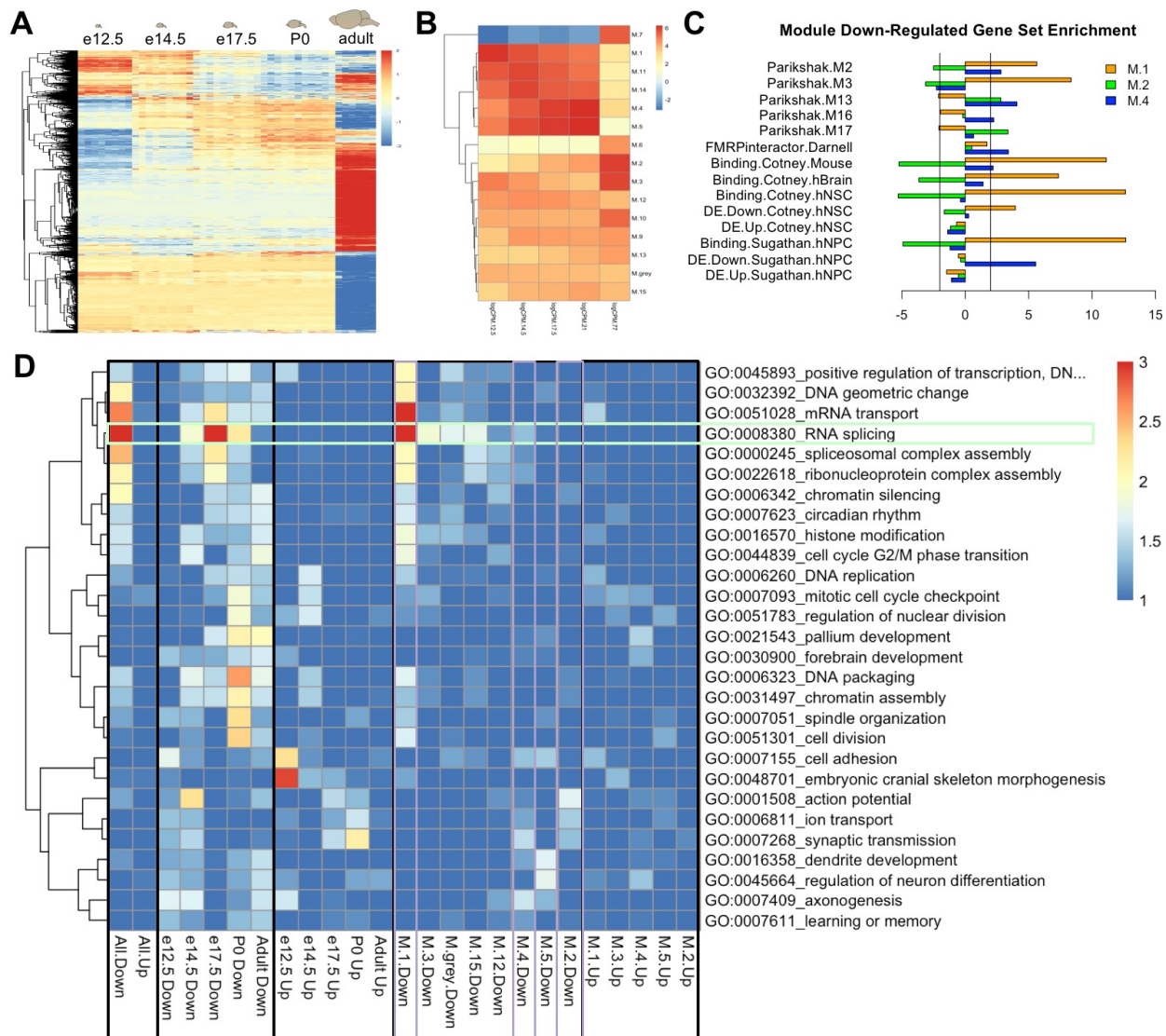
178       Finally, we see that differential expression is not limited to early developmental effects.  
179 For example, many synaptic genes are also impacted (Figure S5). Consistent with previous  
180 studies, we observed that *Chd8* regulates TP53 and, to a lesser degree, *Wnt* signaling pathways  
181 that control processes from early neurodevelopment to synaptic function<sup>10,21,22</sup>. As such, *CHD8*

182 haploinsufficiency may drive autism-associated pathology via multiple neurodevelopmental  
183 mechanisms. This analysis confirms that *Chd8* is required either directly or indirectly for typical  
184 expression of autism-relevant gene networks during neurodevelopment.

### 185 ***Gene expression network analysis to identify perturbations to *Chd8*<sup>+/-</sup> neurodevelopment***

186 We next explored how DE genes are organized into networks that follow parallel  
187 expression trajectories during brain development with a goal of identifying stage-specific  
188 neurodevelopmental processes that are perturbed in *Chd8*<sup>+/-</sup> mice. We used weighted gene  
189 correlation network analysis (WGCNA<sup>23</sup>) to identify co-regulated gene modules from our  
190 developmental transcriptomic data. Fifteen discrete modules were identified that exhibited  
191 specific trajectories of expression and covariation across forebrain development (Figure 3A-3B,  
192 Figure S6, Table S5). DE genes assigned to specific modules were enriched for relevant gene  
193 sets (Figure 3C) and stage-specific Gene Ontology Biological Process annotation terms (Figure  
194 3D, Table S7), confirming that the modules captured by this approach are developmentally and  
195 biologically relevant.

196 First, we looked at module representation among high-confidence autism genes identified  
197 in autism genome sequencing efforts<sup>16</sup>. While these genes were strongly enriched overall among  
198 DE down-regulated genes (Figure 2F), autism risk genes are present across modules rather than  
199 exhibiting strong module-specific clustering. For example, M.1 is characterized by decreasing  
200 expression across neurodevelopment, is associated with chromatin, RNA processing, and cell  
201 cycle, and includes the largest number of DE down-regulated autism risk genes (e.g. *Kdm5b* and  
202 *Chd2*). M.1 is discussed in more detail below. M.4 represents a network of genes with rising  
203 expression from E12.5 to P0 with much lower expression in adult brain. Down-regulated genes  
204 in M.4 are enriched for GO terms associated with transient neuronal differentiation processes



**Figure 3. Identification of DE genes with correlated expression patterns across brain development reveals perturbation to early and later neurodevelopmental pathways.** **A.** Heatmap representing expression of DE genes across all samples and stages. **B.** Mean expression across developmental stages for the 15 developmental gene expression modules. **C.** Permutation test of overlap between module-specific DE down-regulated genes from M.1 (early neurodevelopment), M.2 (neuronal differentiation), and M.4 (late neuronal/brain genes) (purple boxes in panel D) and autism and *Chd8*-relevant genes sets. M.1 DE down-regulated genes are enriched for early expressed autism-associated networks (Parikshak M2 and M3) and direct binding by *Chd8*, while M.2 and M.4 are enriched for FMRP targets and later autism-associated gene networks (Parikshak M13, M16, M17) but not *Chd8*-bound genes or early developmental networks. **D.** Functional enrichment of GO biological process annotations for stage- and module-specific DE genes.

205 (e.g. axonogenesis) and include autism risk genes involved in post-mitotic migration and  
 206 neuronal maturation (e.g. *Bcl11a*, *Ank2*, and *Ctnnb2*). M.2 is characterized by low expression  
 207 early that gradually increases at each stage including in adult forebrain. Down-regulated genes in

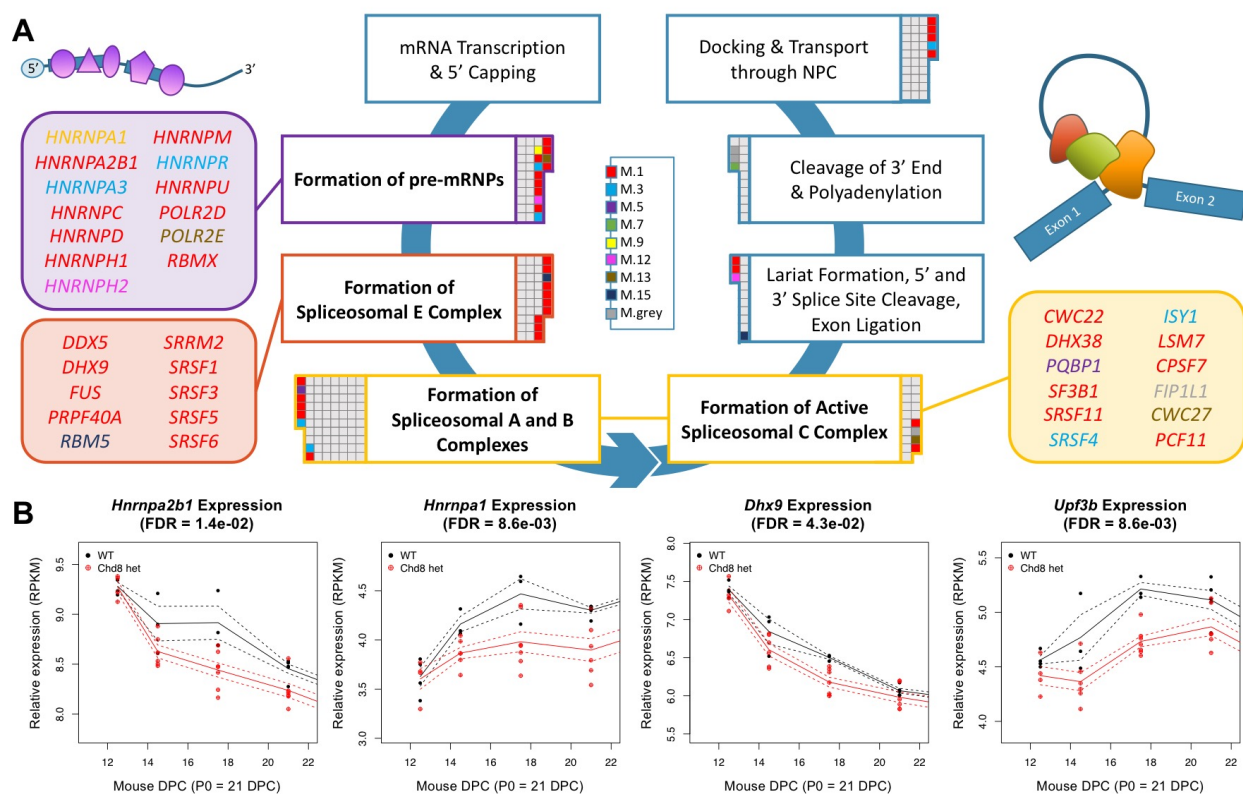
208 M.2 are enriched for GO terms such as synaptic transmission that are hallmarks of more mature  
209 neurons, and include autism risk genes such as *Cers4* and *Grial*. Many synaptic genes are down-  
210 regulated in *Chd8*<sup>+/-</sup> forebrain, with down-regulation strongest in the adult time point but often  
211 present developmentally as well (Figure S5), including high-confidence autism risk genes such  
212 as *Scn2a1* (M.5) and *Cacna1b* (M.2). Segregation of DE genes into modules revealed *Chd8*  
213 haploinsufficiency causes perturbation of autism-relevant gene sets that appeared to be in  
214 separate causal pathways<sup>6</sup>. This examination suggests a developmental hierarchy of autism-  
215 relevant pathology in *Chd8*<sup>+/-</sup> mice involving chromatin remodeling, transcriptional regulation,  
216 and synapse function, the three major pathological pathways implicated via human genetics<sup>6</sup>.

217 In contrast to the spread of high confidence autism risk genes across developmental  
218 modules, we observed strong module-specific enrichment for autism-associated developmental  
219 expression networks identified by Parikshak et al. 2013, as well as for CHD8 binding and  
220 knockdown data, and DE genes whose transcripts are FMRP targets (Table S6). To illustrate this,  
221 we highlighted gene set enrichment for down-regulated DE genes from three modules with  
222 distinct developmental trajectories discussed above, M.1, M.4, and M.2 (Figure 3C). We  
223 observed enrichment in overlap with genes in Parikshak.M3, the earliest expressed network  
224 identified in Parikshak et al. 2013 that was specific to down-regulated genes in M.1, our early  
225 expression module. Genes in Parikshak.M2, a module that is anchored later than Parikshak.M3  
226 but still representative of early brain development, overlap both M.1 and M.4. The three late  
227 modules identified by Parikshak et al. 2013 overlap with down-regulated genes from either M.4  
228 (Parikshak.M13 and Parikshak.M16) or M.2 (Parikshak.M17), but not M.1. While we did not  
229 observe global enrichment among all DE genes for FMRP targets<sup>19</sup>, there was module-specific  
230 enrichment for M.4 DE down-regulated genes among FMRP targets. We observed strong

231 module-specific enrichment of M.1 for Chd8 binding targets identified in embryonic brain and in  
 232 vitro models of neural development<sup>12,14</sup>, suggesting direct regulation of DE genes by Chd8.  
 233 Finally, we found module-specific enrichment of genes down-regulated after CHD8 knockdown  
 234 in hNSC (M.1) and hNPC (M.4), suggesting similar gene regulatory consequences produced by  
 235 *Chd8* haploinsufficiency across systems.

236 ***Chd8*<sup>+/-</sup> mice exhibit down-regulation of genes involved in RNA processing in forebrain across**  
 237 **developmental expression modules**

238 A number of GO terms were strongly enriched for DE genes mapping across modules,

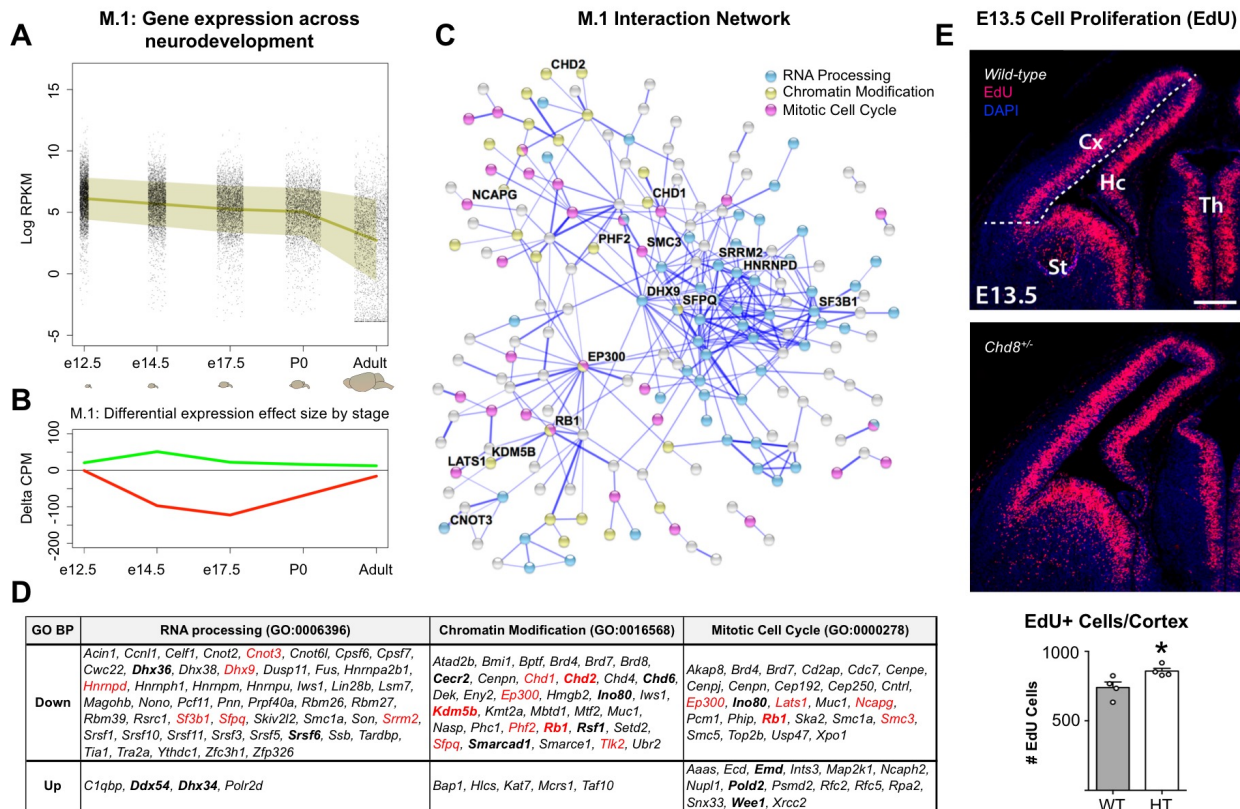


**Figure 4. RNA processing pathways are enriched for differentially expressed genes in *Chd8*<sup>+/-</sup> mice. A.** Condensed model of the RNA processing pathway on Reactome, adapted from the parent pathway *Processing of Capped Intron-Containing pre-mRNA*. Genes annotated to different steps in the RNA processing pathway are denoted as boxes. DE genes are colored by module membership; non-differentially expressed genes are white. Each gene is represented once, at the first step the gene appears in the pathway. **B.** Developmental expression trajectories of example DE down-regulated genes annotated to the formation of pre-mRNPs (*Hnrnpa2b1*, *Hnrnpa1*), formation of the spliceosomal E complex (*Dhx9*), and lariat formation and 5' splice site cleavage (*Upf3b*). DE p-value FDR shown.

239 suggesting perturbation to stage-specific gene sets controlling these processes across  
240 developmental stages (Figure 3D). This signature was particularly strong for genes involved in  
241 RNA processing (Figure 4). DE genes were significantly overrepresented among genes annotated  
242 to RNA processing, e.g. Processing of Capped Intron-Containing pre-mRNA (FDR = 1.59E-09),  
243 mRNA Splicing (FDR = 2.90E-09), and mRNA 3'-end Processing (FDR = 3.29E-04) in the  
244 Reactome database (Figure 4A, Table S4). These genes represented modules with highly  
245 divergent expression trajectories (Figure 4B). For example, among genes annotated to RNA  
246 splicing, expression of *Dhx9* (M.1) decreases across neurodevelopment and has not been  
247 functionally characterized in brain but has been reported in autism-risk networks<sup>17</sup> (Figure 4B),  
248 while *Upf3b* (M.15) expression increases across development and is a neuron-specific factor  
249 required during neuronal differentiation that is implicated in intellectual disability<sup>24,25</sup> (Figure  
250 4B). While RNA processing genes present different overall expression trajectories across brain  
251 development, perturbation to expression generally peaked at E17.5 for these genes, suggesting  
252 that this represents a period where this process is critical in brain development. Unlike synaptic  
253 genes and critical transcription factor genes (e.g. *Tbr1*), the DE genes annotated to have a role in  
254 RNA processing have not been well studied in the context of brain development and represent  
255 candidates for future investigation. RNA processing genes that are DE in *Chd8*<sup>+/-</sup> mice are  
256 reported in Table S8.

257 ***Perturbation to early developmental expression network and associated increase in neuronal***  
258 ***proliferation in Chd8<sup>+/-</sup> mice***

259 M.1 consists of 3,590 genes that show a general trend of decreasing expression levels  
260 across neurodevelopment (Figure 5A). M.1 showed the strongest enrichment for autism risk  
261 early developmental modules<sup>18</sup> and for Chd8 binding targets in embryonic brain and in vitro



**Figure 5. An early neurodevelopmental expression network (M.1) regulated by *Chd8* haploinsufficiency involved in chromatin organization, RNA processing, and cell cycle.** **A.** M.1 gene expression plotted across brain development. Dots represent individual genes, line represents mean expression and shaded area represents  $\pm 1$  SD. **B.** Relative mean differential expression of up-regulated (green) and down-regulated (red) genes in M.1 across brain development. **C.** Protein-protein interaction network of M.1. DE genes are colored by annotation to chromatin organization, RNA processing, and mitotic cell cycle GO Biological Process terms. Labeled genes have been previously identified as autism risk genes. **D.** M.1 *Chd8*-bound DE genes associated with selected GO terms. Red: autism risk genes; black: differential expression in previous *in vitro* *Chd8* knockdown studies. **E.** (Upper) Coronal section of E14.5 stained for EdU (magenta), a marker of proliferation, and DAPI (blue) in WT and *Chd8*<sup>+/-</sup> mice (n = 4 for both genotypes). Scale Bar 200  $\mu$ m. (Lower) Plot (mean  $\pm$  SEM with dots representing individual samples) of EdU positive cells/area. Student's t-test p-value = 0.0338.

262 models. M.1 also had the greatest enrichment of down-regulated genes (p-value = 5.39E-21). 454  
 263 genes in M.1 are DE at FDR < 0.25 (350 down-regulated, 104 up-regulated), accounting for  
 264 ~30% of all DE genes identified in our study. We examined change in expression of DE genes  
 265 within M.1 at each developmental stage, observing that up-regulation peaked at E14.5 compared  
 266 to a fold change peak for down-regulated genes at E17.5 (Figure 5B). Analysis of protein-protein  
 267 interactions (STRING<sup>26</sup>) showed that DE genes in M.1 had more protein-protein interactions  
 268 than expected by chance (observed edges = 972, expected edges = 237, enrichment = 4.10,



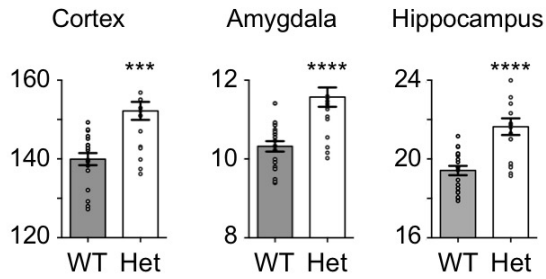
269 STRING p-value < 0.0001). Interacting genes in M.1 were enriched for functions regulating  
270 genome structure, cellular proliferation and differentiation, with enrichment for GO terms (Table  
271 S9) including RNA processing, chromatin organization, and mitotic cell cycle (Figure 5C). M.1  
272 DE genes annotated to these terms, including a number of autism risk genes, were identified as  
273 Chd8 targets. A subset of these genes was also DE after *CHD8 in vitro* knockdown in human  
274 cells (Figure 4D). This analysis suggests that Chd8 binding directly regulates M.1 genes and that  
275 differential expression of M.1 genes in *Chd8*<sup>+/-</sup> embryos may drive changes in chromatin  
276 structure and RNA metabolism linked to early neurodevelopmental pathology associated with  
277 disruption to proliferation and neuronal differentiation.

278 To examine whether these alterations in developmental genes play a functional role in  
279 neuronal development that could lead to megalencephaly, we performed 5-ethynyl-2'-  
280 deoxyuridine (EdU) proliferation assays (Figure 5E). Assessing the number of EdU<sup>+</sup> cells in the  
281 cortical ventricular zone (VZ) after a 1.5 hour pulse at E13.5, a time point of peak neurogenesis,  
282 we found their number significantly increased in the mutant by 15.9% (Student's t-test, p =  
283 0.0338, n = 4 for either genotype). Considering that a number of genes associated with brain  
284 development and cortical structure exhibited DE, we also examined cortical morphology via  
285 analysis of Tbr1 and Ctip2 immunostaining at P0 (Figure S7). We observed no gross alterations  
286 to lamination and found no evidence for focal cortical lesions (n: WT = 8; *Chd8*<sup>+/-</sup> = 10). These  
287 findings experimentally corroborate altered proliferative dynamics in *Chd8*<sup>+/-</sup> mutants, linking  
288 altered neurogenesis and megalencephaly in *Chd8*<sup>+/-</sup> mice.

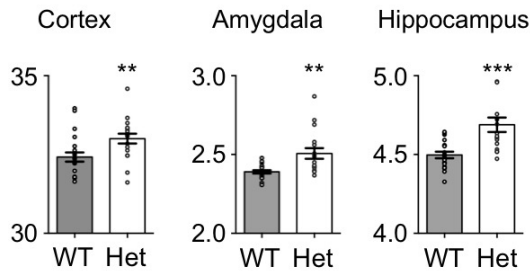
### 289 ***Analysis of Chd8*<sup>+/-</sup> adult brain structure via MRI**

290 To establish whether structural changes persist in the *Chd8*<sup>+/-</sup> mouse brain, structural MRI  
291 was performed to identify changes in regional brain volume and connectivity between adult

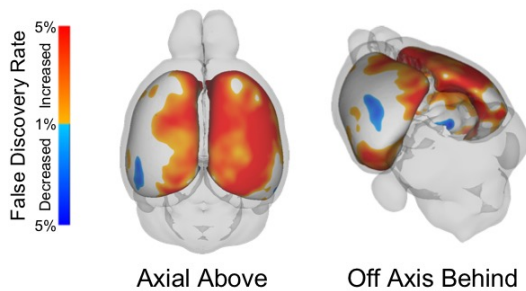
**A. *Chd8*<sup>+/-</sup> Mouse Absolute Volume (mm<sup>3</sup>)**



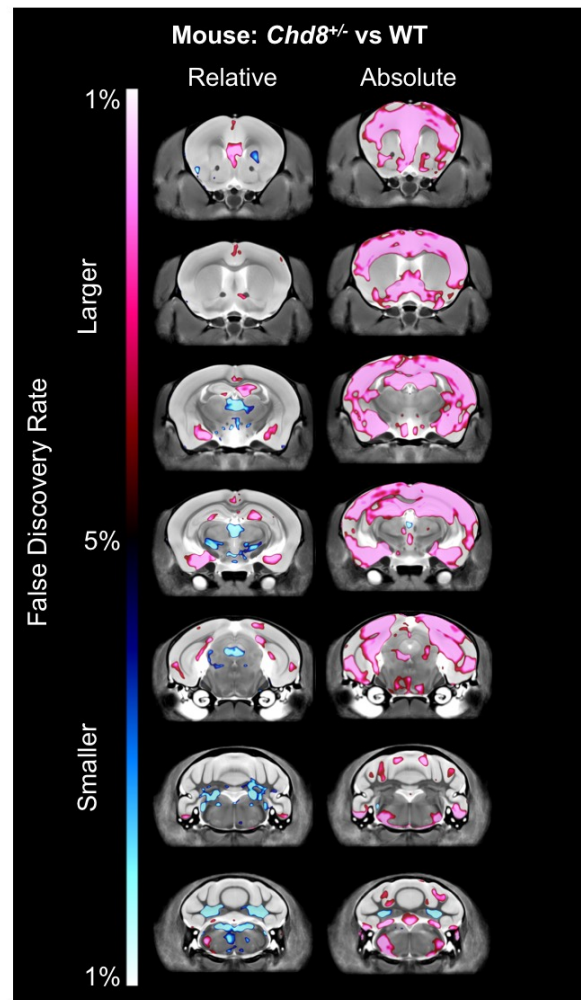
**B. *Chd8*<sup>+/-</sup> Mouse Relative Volume (mm<sup>3</sup>)**



**C. *Chd8*<sup>+/-</sup> Mouse Cortical Thickness**



**D. Brain Volume Differences**



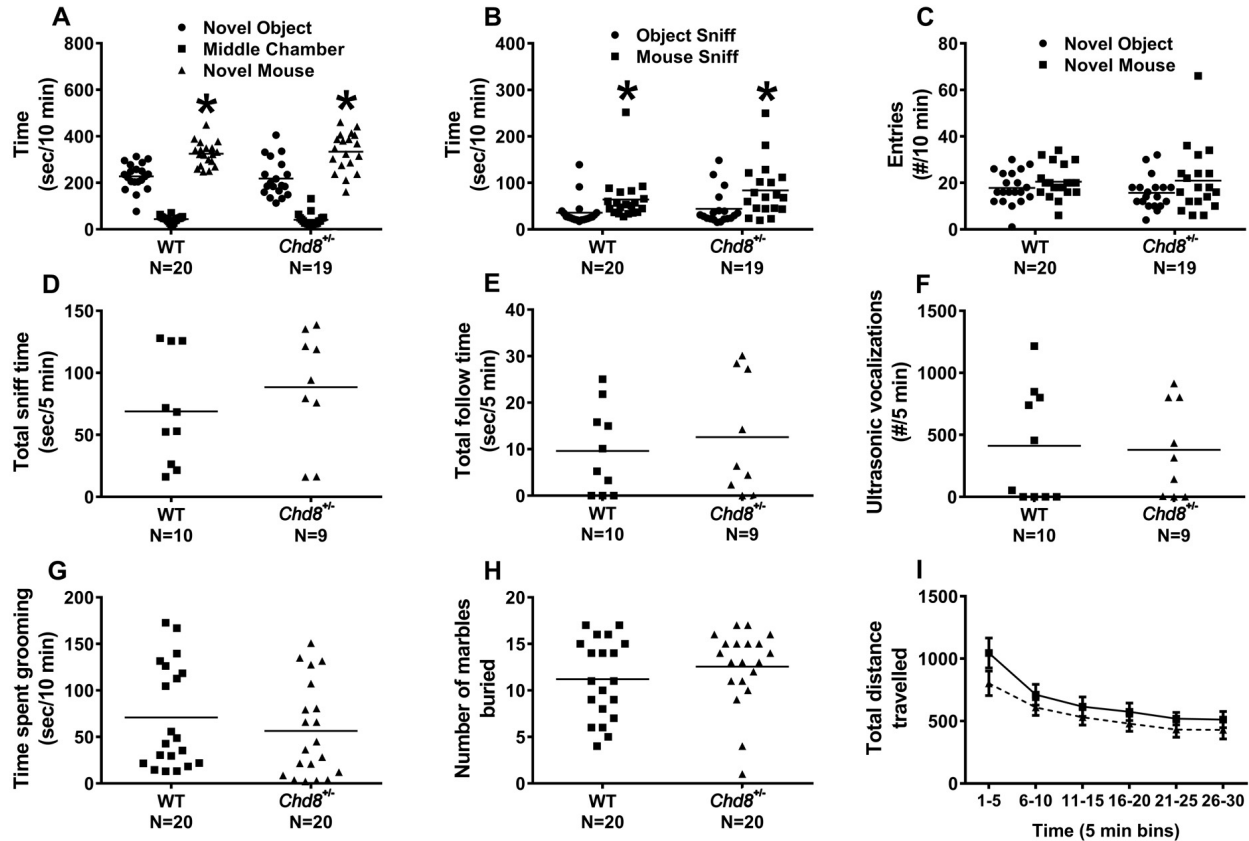
**Figure 6. *Chd8* haploinsufficiency drives macrocephaly and cognitive impairment in mouse and human. A.** MRI revealed significant increases in absolute regional volume of cerebral cortex, amygdala, and hippocampus (\*\*\* $p < 0.0001$ ) in *Chd8*<sup>+/-</sup> mice. N: WT=21, *Chd8*<sup>+/-</sup>=18. See Table S11 for full regional statistical analysis and FDR values. **B.** Increases in volume are still significant for the cortex (\*\* $p = 0.0097$ ), amygdala (\*\* $p = 0.0014$ ) and hippocampus (\*\*\* $p = 0.0004$ ) after correction for absolute brain volume. (Student's *t* test *p*-value, error bars denote SEM). **C.** Increased cortical thickness is present in *Chd8*<sup>+/-</sup> mice. **D.** Voxel-wise differences in volume between *Chd8*<sup>+/-</sup> and WT littermates.

292 *Chd8*<sup>+/-</sup> mice (n=18) and matched wild-type littermates (n=21). No significant differences were  
 293 observed in body weight or other relevant measures of general health in adult *Chd8*<sup>+/-</sup> mice  
 294 (Table S10). Neuroanatomy was assessed and volume was measured as absolute volume (mm<sup>3</sup>)  
 295 and relative volume (% total brain volume). Considering regional differences, the most affected  
 296 region was the cortex, which was increased by 7.5% with a false discovery rate (FDR) of 1%.

297 Similarly, the cerebral white matter and cerebral gray matter were also larger in the *Chd8*<sup>+/-</sup> mice  
298 at 5.4% (FDR of 3%) and 6.1% (FDR of 2%), respectively. When the male and female mice  
299 were examined independently, female *Chd8*<sup>+/-</sup> mice exhibited stronger effect sizes but both sexes  
300 exhibited overall similar trends. In addition to the summary regions, 159 independent brain  
301 regions were assessed with divisions across the cortex, subcortical areas, and cerebellum. Full  
302 results for comparison across individual brain regions are reported in Table S11. *Chd8*<sup>+/-</sup> mice  
303 showed robust increase in absolute volume across cortical regions, hippocampus (+10.3%, FDR  
304 < 1%), and amygdala (+11.0%, FDR < 1%) (Figure 5A-5B). The *Chd8*<sup>+/-</sup> mice also displayed  
305 increased cortical thickness, particularly along the cingulate cortex (Figure 5C). After correction  
306 for total volume, relative volumes were still significantly larger, though cortex failed to surpass  
307 the FDR < 5% cutoff. Deep cerebellar nuclei showed decreased relative volume (-1-3%, FDR <  
308 2%). Voxel-wise differences showed similar trends (Figure 5D). Diffusion Tensor Imaging  
309 (DTI) revealed no differences in fractional anisotropy or mean diffusivity in either the regional  
310 or voxel-wise measurements, indicating that the anatomical connectivity of the white matter in  
311 the *Chd8*<sup>+/-</sup> mice was not significantly different from WT littermates (not shown).

### 312 ***Behavioral phenotyping of adult Chd8*<sup>+/-</sup> mice**

313 Behaviors relevant to ASD were assessed in adult *Chd8*<sup>+/-</sup> mice using two assays of social  
314 behaviors and two assays of repetitive behaviors (Figure 7), as previously described<sup>27</sup>. In the 3-  
315 chambered social approach test<sup>28</sup>, normal sociability was detected in both genotypes (Figure 7A-  
316 7B). Time spent in the chamber with the novel mouse was greater than time spent in the chamber  
317 with the novel object, meeting the definition of sociability in this assay, for both WT and *Chd8*<sup>+/-</sup>  
318 (Figure 7A; WT:  $t(1, 40) = 6.07, p < 0.001$ ; *Chd8*<sup>+/-</sup>:  $t(1, 34) = -3.93, p < 0.001$ ). No sex  
319 differences were detected ( $F(1, 37) = 2.16, p > 0.05$ ). Time spent sniffing the novel mouse was



**Figure 7. Adult *Chd8*<sup>+/-</sup> mice do not differ from WT on ASD-relevant social and repetitive behavior assays.** *Chd8*<sup>+/-</sup> mice and WT littermate controls both met the definition of sociability on the 3-chambered social approach test. (A-C). Each genotype spent significantly more time in the chamber with the novel mouse than in the chamber with the novel object, spent significantly more time sniffing the novel mouse than sniffing the novel object, and showed normal locomotor entries between chamber and normal male-female social interactions with concomitant ultrasonic vocalizations (D-F). *Chd8*<sup>+/-</sup> and WT displayed similar amounts of self-grooming (G), marble burying (H), and open field exploratory locomotion (I).

320 greater than time spent sniffing the novel object in both WT and *Chd8*<sup>+/-</sup> (Figure 7B; WT:  $t(1,$   
 321 40) = 2.47,  $p < 0.02$ ; *Chd8*<sup>+/-</sup>:  $t(1, 34) = -2.33$ ,  $p < 0.03$ ) again without sex effects ( $F(1, 37) =$   
 322 0.107,  $p > 0.05$ ). Number of entries into the side chambers was not affected by genotype in the  
 323 social phase (Figure 7C;  $F(1, 37) = 0.99$ ,  $p > 0.05$ ), nor on the entries into the left or right  
 324 chambers parameter during the previous habituation phase ( $F(1, 37) = 0.069$ ,  $p > 0.05$ ),  
 325 indicating normal general exploratory activity in both genotypes during the social approach  
 326 assay.

327 No abnormalities were observed on social parameters in male *Chd8*<sup>+/-</sup> during the male-  
328 female reciprocal social interaction test, in which male WT and *Chd8*<sup>+/-</sup> subject mice were  
329 individually paired with an unfamiliar estrous B6 female (Figure 7D-7E). WT and *Chd8*<sup>+/-</sup> males  
330 spent similar amounts of time engaged in sniffing (Figure 7D;  $t(1, 17) = 0.35$ ), and following  
331 (Figure 7E;  $t(1, 17) = 0.78$ ,  $p > 0.05$ ) the female, with numerical scores comparable to previous  
332 findings<sup>29,30</sup>. Ultrasonic vocalizations recorded during male-female interaction showed no  
333 genotype difference in the number of emitted calls (Figure 7F;  $t(1, 17) = -0.74$ ,  $p > 0.05$ ).

334 No spontaneous stereotypies or repetitive behaviors were observed. Genotypes did not  
335 differ on time spent in self-grooming behavior (Figure 7G;  $t(1, 38) = -1.29$ ,  $p > 0.05$ ), or in  
336 numbers of marbles buried (Figure 7H;  $t(1, 38) = 0.63$ ,  $p > 0.05$ ). No sex differences were  
337 detected (self-groom:  $t(1, 37) = -1.79$ ,  $p > 0.05$ ; marble burying:  $t(1, 37) = -1.49$ ,  $p > 0.05$ ).  
338 Open field locomotor activity did not differ between genotypes (Figure 7I,  $F(1,27) = 0.28$ ,  
339  $p > 0.05$ ), and both genotypes displayed the expected habituation to the novel environment across  
340 the 30 minute test session (Figure 7I;  $F(5, 90) = 10.62$ ,  $p < 0.001$ ), indicating normal exploratory  
341 behavior.

342 Taken together, these results indicate that heterozygous loss of *Chd8* did not significantly  
343 alter social approach, reciprocal social interaction, or repetitive behaviors in this first cohort of  
344 adult *Chd8*<sup>+/-</sup> mice, as conducted with established assays with face validity to the diagnostic  
345 symptoms of ASD.

## 346 **Discussion**

347 This integrative analysis of a new mouse model demonstrates that *Chd8*  
348 haploinsufficiency results in ASD-relevant neurodevelopmental phenotypes. Our results show  
349 that germline 5 bp and 14 bp deletion mutations in *Chd8* exon 5 result in decreased mRNA,

350 consistent between qPCR and RNA-seq. *Chd8* itself was the most strongly significant DE gene  
351 of our transcriptomic analysis, and was DE at all developmental stages. Further, decrease in  
352 allele-specific expression of the mutant allele was evident in RNA reads. Such reads occur at low  
353 levels relative to the WT allele, suggesting degradation of the mutant frameshift transcript via  
354 nonsense-mediated decay. Finally, the amount of Chd8 protein assayed via western blot was  
355 consistently decreased in *Chd8*<sup>+/-</sup> mouse brain across developmental stages. These results  
356 indicate that the short deletions in *Chd8* exon 5 in our mouse models result in a germline  
357 haploinsufficient state, with concomitant decrease in *Chd8* mRNA and protein observed in  
358 forebrain across all methods and experiments.

359 Our experiments confirm that *Chd8* regulates proliferation and neurogenesis, and suggest  
360 substantial impact across neurodevelopment. They also suggest similar binding targets and  
361 biological function in mouse and human neurodevelopment and a parallel causal role of *CHD8*  
362 haploinsufficiency in megalencephaly. We observed significant overlap between down-regulated  
363 genes in *Chd8*<sup>+/-</sup> mice and autism risk gene sets produced by independent groups using a variety  
364 of approaches<sup>13,16-18</sup>. This overlap suggests that perturbation to neurodevelopment in *Chd8*<sup>+/-</sup>  
365 mice parallels autism-relevant human neurobiology, a finding consistent with our  
366 neuroanatomical and structural MRI results.

367 Our RNA analysis captures subtle changes in transcription across brain development in  
368 *Chd8*<sup>+/-</sup> mice. These changes were consistent across developmental stages for perturbed genes,  
369 were highly relevant to ASD-associated networks, and strongly correlated with biological  
370 pathways and expression modules of interest. Our results parallel *in vitro* findings that suggested  
371 convergence across risk pathways after CHD8 knockdown<sup>12</sup>, providing a developmental  
372 framework revealing disruption of convergent ASD pathways in a genetic mouse model of Chd8

373 haploinsufficiency. Unlike genomic analysis using *in vitro* knockdown studies of *CHD8*<sup>12,14</sup>, our  
374 network analysis using *in vivo* data enabled characterization of the impact of *Chd8*  
375 haploinsufficiency across neurodevelopment.

376 Our results suggest a developmental hierarchy of changes in *Chd8*<sup>+/-</sup> brain development.  
377 For example, M.1, the module with the strongest enrichment for DE genes and genes directly  
378 targeted by CHD8, represents a highly interactive network of genes central to control of  
379 chromatin state and RNA processing, including genes implicated in autism. M.1 may represent a  
380 critical network directly regulated by *Chd8*. The timing disruption of this module supports that it  
381 plays a central role in the neuronal proliferation phenotype observed in E13.5 *Chd8*<sup>+/-</sup> brain and  
382 may be functionally linked to the impact on RNA processing gene expression. The overlap of  
383 RNA processing gene down-regulation in M.1 and in later developmental expression modules  
384 suggests that *Chd8* haploinsufficiency results in general changes in molecules involved in RNA  
385 processing during neuronal differentiation. Considering the large number of DE genes involved  
386 in RNA processing, our results further indicate that disruption to RNA processing is an important  
387 player in neurodevelopmental disorders, in line with other neurodevelopmental disorder genetic  
388 models such as *Fmr1*<sup>31</sup> and *Rbfox1*<sup>32</sup>. Furthermore, though alternative splicing has been well-  
389 established to a critical role in cell differentiation and proliferation in the developing brain<sup>33</sup>, the  
390 roles of other RNA processing pathways such as RNA stability or translocation in brain  
391 development are less studied<sup>34</sup> and could provide potential candidates for future investigation.

392 While dysregulation of genes in M.1 and of RNA processing genes peaks at E17.5, our  
393 analysis suggests changes across neurodevelopment driven by *Chd8* haploinsufficiency. These  
394 changes indicate convergent neuropathology connecting chromatin remodeling, neuronal  
395 differentiation, and synaptic pathways, the principle gene networks identified in autism case

396 sequencing studies. For example, we observed down-regulation of genes involved in RNA  
397 processing (e.g. *Upf3b* and *Hnrnpd*), neuronal differentiation (e.g. *Bcl11a* and *Tbr1*) and in  
398 synapse development and function (e.g. *Scn2a1* and *Cacna1b*), all of which have been implicated  
399 in autism via human genetic studies. The transcription data generated for this study will be a  
400 useful resource for future dissection of pathways involved in the pathogenesis of  
401 neurodevelopmental disorders and in classification of risk genes from genetic studies. Further  
402 studies may capture the neuroanatomical and cellular changes and perturbed signaling pathways  
403 associated with differential expression signatures in *Chd8*<sup>+/-</sup> brain development. Future work will  
404 also be necessary to determine the stage- and cell-specific role of Chd8-binding to establish and  
405 maintain expression patterns of these genes.

406         Structural changes in the brain of adult *Chd8*<sup>+/-</sup> mice parallel other relevant mouse  
407 models. A recent study examined 26 different mouse models related to autism<sup>35</sup>, clustering these  
408 models into 3 distinct groups. Key aspects of Group 1 included larger sizes of cortical structures,  
409 particularly the frontal and parietal lobes, and smaller sizes in the cerebellum, which is in line  
410 with the *Chd8*<sup>+/-</sup> mouse described here. This group of models included *Nrxn1a*, *Shank3*, *En2*, and  
411 *Fmr1*. The *Chd8*<sup>+/-</sup> mouse most resembled the differences found in the *Fmr1* mutant mice.  
412 Further examination may reveal similarities with other mouse models within this group beyond  
413 neuroanatomy (e.g. excitatory deficits in the *Nrxn1a* mouse<sup>36</sup>), as suggested by the widespread  
414 transcriptional changes present in *Chd8*<sup>+/-</sup> neurodevelopment. Increases in cortical  
415 anteroposterior length and developmental neurogenesis appear largely overlapping in *Chd8*<sup>+/-</sup>  
416 mice and *Wdly3* mutants, a recently reported model of megalencephaly in autism<sup>37</sup>.

417         In comparison to the *Chd8*<sup>+/-</sup> mice studied here, heterozygous mouse models of *Pten*,  
418 another gene associated with ASD and macrocephaly, do exhibit core aspects of ASD<sup>38</sup>.



419 However, studies of other mouse models of genes implicated in ASD have not identified  
420 behavioral phenotypes with face value to ASD. Additional studies are necessary to further  
421 examine behavioral phenotypes at earlier developmental stages to test for causal relationship  
422 between structural changes and behavior in *Chd8*<sup>+/-</sup> mice. The presence of genomic and  
423 neuroanatomical phenotypes in *Chd8*<sup>+/-</sup> mice that parallel the clinical signature of human *CHD8*  
424 mutations suggests similar neurodevelopmental pathology between human and mouse.  
425 Intriguingly normal phenotypes on the autism-relevant social and repetitive assays conducted  
426 here highlight future opportunities for comprehensive behavioral phenotyping in a replication  
427 cohort, to evaluate social and repetitive behaviors at juvenile ages and to investigate phenotypes  
428 relevant to other symptom domains described for individuals with *CHD8* mutations, including  
429 cognitive impairments and attentional disorders.

430 Our initial survey of mice heterozygous for mutation to *Chd8* revealed significant  
431 findings across genomic and anatomical axes of neurobiology. These experiments link increased  
432 regional brain volume to perturbations of biological pathways across neurodevelopment,  
433 recapitulating primary neuroanatomical traits observed in *CHD8*<sup>+/-</sup> human individuals. As such,  
434 the results offer insight into the neurodevelopmental pathology associated with mutations to  
435 *CHD8*, a genetic model that appears to be a bellwether for mutations affecting early  
436 transcriptional regulation and chromatin remodeling in autism. In-depth analysis of  
437 developmental neuroanatomy and social and communicative phenotypes as well as associated  
438 attentional and cognitive deficits in these *Chd8*<sup>+/-</sup> mice will be necessary to link observed  
439 changes in brain gene expression and structure with relevant pathology in humans. This study of  
440 the impact of *Chd8* haploinsufficiency *in vivo* in mice demonstrates the power of genomic and

441 systems-level characterization of neurodevelopment in animal models towards resolving major  
442 questions about the genetic and neurodevelopmental origins of autism and intellectual disability.

### 443 **Experimental Procedures**

444 See extended online methods for full description of experimental procedures.

#### 445 ***Generation of Chd8 mutant mice***

446 We used Cas9-mediated mutagenesis of C56BL/6N oocytes to generate two mouse lines  
447 harboring frameshift deletions (5 bp and 14 bp) in mouse *Chd8* exon 5 (gRNA sequence:  
448 GAGGAGGAGGTCGATGTAAC). Guide RNA was synthesized and pooled with Cas9  
449 mRNA<sup>39</sup> and injected into mouse oocytes. We identified F0 pups carrying 5 bp and 14 bp  
450 deletions that overlap the target sequence. Heterozygotes were crossed to WT C57BL/6N mice to  
451 expand the lines and eliminate off-target mutations. We examined Chd8 protein and transcript  
452 levels via western blot (ab114126; Abcam) and qPCR at E14.5, P0, and adult forebrain and  
453 compared cortical length in whole mount P0 brains from *Chd8*<sup>+/-</sup> mice and matched WT  
454 littermates. All mouse studies were approved by the Institutional Animal Care and Use  
455 Committees at the University of California Davis and the Lawrence Berkeley National  
456 Laboratory. Subject mice were housed in a temperature-controlled vivarium maintained on a 12  
457 hour light-dark cycle. Efforts were made to minimize pain and distress and the number of  
458 animals used.

#### 459 ***Developmental neuroanatomy***

460 Litters for neuroanatomy analysis were generated by breeding male *Chd8*<sup>+/-</sup> mice with  
461 WT females. Brains were perfused before isolation, embedding, and sectioning. *P7 Nissl*

462 *staining*: Nissl staining was performed following established protocols and morphological  
463 parameters were measured and compared using Student's t-test. *E13.5 EdU labeling*: Timed-  
464 pregnant females were intraperitoneally injected with 50 mg/kg body weight EdU. After 1.5  
465 hours, females were anesthetized and embryos were perfused, fixed, and sectioned. EdU  
466 detection was performed with the Click-it EdU Alexa 594 imaging kit protocol (Life  
467 Technologies) according to the manufacturer instructions. *P0 lamination*: Slides were incubated  
468 in blocking solution, rinsed in PBS-T, and incubated overnight at 4°C in primary antibody  
469 solution containing anti-Ctip2 (ab18465; Abcam) and anti-Tbr1 (ab31940; Abcam) antibodies.  
470 The slides were rinsed and incubated overnight at 4°C in fluorophore-conjugated secondary  
471 antibodies. Slides were counterstained for 2 hr in DAPI, rinsed, and mounted. All sections over  
472 the entire brain were surveyed for lamination defects and corresponding sections imaged. Within  
473 each genotype, all brains were selected randomly for histological processing without taking  
474 morphological criteria into account. All histology was done blind, by investigators that were  
475 unaware of group allocation. No data points were excluded. All antibodies used for this study  
476 were validated and their use widely reported.

## 477 **Genomics**

478 Bulk forebrain was microdissected from *Chd8*<sup>+/-</sup> and matched WT littermates at E12.5,  
479 E14.5, E17.5, P0, and from adults (>P56). Samples included males and females of each genotype  
480 at each stage. Total RNA was isolated using Ambion RNAqueous and assayed using an Agilent  
481 BioAnalyzer instrument. Stranded mRNA sequencing libraries were prepared using TruSeq  
482 Stranded mRNA kits and 6-12 samples per lane were pooled and sequenced on the Illumina  
483 HiSeq platform using a single end 50 bp (E14.5, E17.5, P0, adult) or paired end 100 bp (E12.5)  
484 strategy. FASTQ files were aligned to the mouse genome (mm9) and counts for mouse genes

485 were calculated. For inclusion in testing, genes were required to have a minimum read count of  
486 at least 10 reads/million in more than two samples. Differential expression was performed with  
487 edgeR<sup>40</sup> using generalized linear models including factor-encoded sex and developmental  
488 stage/sequencing batch as covariates. After normalization, iterative Weighted Gene Correlation  
489 Network Analysis (WGCNA<sup>23</sup>) was used to identify co-expressed gene modules. Fourteen  
490 discrete gene expression networks were identified (numbered by module gene count), and genes  
491 that were not classified into one of these modules were assigned to M.8.grey. Permutation testing  
492 was performed to test for overlap between DE genes and published gene sets. Gene Ontology  
493 Biological Process term enrichment and protein-protein interaction network analysis was  
494 performed using the TopGO Bioconductor package and STRING<sup>26</sup>. Differential expression of  
495 selected targets was verified by qPCR at P0. Primers reported in Table S3. For qPCR analysis, 9  
496 wild-type and 7 *Chd8*<sup>+/-</sup> forebrain samples were used. Samples were excluded if technical  
497 replicates failed. Paired t-test was performed on *Actb* normalized relative gene expression  
498 between WT and *Chd8*<sup>+/-</sup> using  $\Delta\Delta CT$ . To reduce noise, the highest and lowest values from both  
499 groups was discarded. Tbr1 protein level assayed via western blot (ab31940; Abcam) compared  
500 via Student's t-test.

## 501 ***MRI***

502 After perfusion, brains from mice that underwent behavioral screening were scanned  
503 using a multi-channel 7.0 Tesla MRI scanner (Varian Inc., Palo Alto, CA). Diffusion Tensor  
504 Imaging (DTI) was done using a 3D diffusion weighted fast spin echo sequence to create  
505 fractional anisotropy, mean diffusivity, axial diffusivity, and radial diffusivity maps for brains  
506 used in this study. After registration, changes and intensity differences were examined for the  
507 volume or mean diffusion measure for 159 different structures encompassing cortical lobes, large

508 white matter structures, ventricles, cerebellum, brain stem, and olfactory bulbs. Initially seven  
509 summary regions were examined, including the cerebral cortex, olfactory bulbs, cerebral white  
510 matter, cerebral gray matter, ventricles, brainstem, and cerebellum<sup>41</sup>. Multiple comparisons in  
511 this study were controlled for via False Discovery Rate.

## 512 ***Behavioral testing***

513 All procedures were approved by the University of California Davis Institutional Animal  
514 Care and Use Committee, and were conducted in accordance with the National Institutes of  
515 Health Guide for the Care and Use of Laboratory Animals. Efforts were made to minimize pain  
516 and distress and the number of animals used. No previous analyses were performed on animals  
517 used for behavioral testing. Subject mice were housed in a temperature-controlled vivarium  
518 maintained on a 12 hour light-dark cycle. We used mixed genotype home cages with 2-4 animals  
519 per cage and blinded experimenter and video scorer/processor to genotype during testing and  
520 analysis. All tests were conducted during the light cycle. Groups sizes indicated are based on  
521 past experience and power analyses. Effects of genotype and sex evaluated using Multi-factor  
522 ANOVA, as previously published. Significant ANOVAs followed with Tukey's high significant  
523 difference test, or other appropriate post hoc tests including Bonferroni correction tests for  
524 multiple comparisons, to correct for false discovery and to identify specific differences between  
525 groups. Behavioral analysis passed distribution normality tests, was collected using continuous  
526 variables and thus analyzed via parametric analysis, in all assays. For all behavioral analyses,  
527 variances were similar between groups and no data points were excluded. One animal died  
528 during behavioral testing. This happened during Three Chamber Social approach, in the middle  
529 of the behavioral battery. *Chd8*<sup>+/-</sup> male and female mice and WT littermates, ages 2-5 months,  
530 were evaluated in a standard battery of neurobehavioral assays relevant to diagnostic symptoms

531 of autism<sup>27</sup>. In total, 8 male and 10 female *Chd8*<sup>+/-</sup> mice and 11 male and 10 female matched WT  
532 littermates were tested in the following sequence: open field, general health, self-grooming,  
533 marble burying, 3-chambered social approach, and male-female social interactions. Testing was  
534 performed at the UC Davis MIND Institute Intellectual and Developmental Disabilities Research  
535 Center Mouse Behavior Core. Statistical testing was performed using established assay-specific  
536 methods, including Students t-test for single parameter comparisons between genotypes, and  
537 One-Way or Two-Way Repeated Measures Analysis of Variance for comparisons across time  
538 points and/or between sexes.

#### 539 ***Data availability***

540 All relevant data will be available from authors. DOIs for all published gene sets used in  
541 enrichment analysis: Sanders et al. 2015 - 10.1016/j.neuron.2015.09.016; Parikshak et al. 2013 -  
542 10.1016/j.cell.2013.10.031; Cotney et al. 2015 - 10.1038/ncomms7404; Willsey et al. 2015 -  
543 10.1016/j.cell.2013.10.020; Sugathan et al. 2014 - 10.1073/pnas.1405266111; Darnell et al. 2011  
544 - 10.1016/j.cell.2011.06.013; Hormozdiari et al. 2014 - 10.1101/gr.178855.114; Voineagu et al.  
545 2011 - 10.1038/nature10110.

#### 546 ***Code availability***

547 All custom scripts used for data processing and analysis will be available from authors. A  
548 custom sample processing pipeline was used to align raw sequencing samples to mouse genome  
549 mm9 using RNA-seq aligner STAR (version 2.4.2a), features assigned via subreads  
550 featureCounts (version 1.5.0) to UCSC mm9 genes.gtf, and quality check performed on  
551 individual samples using RSeQC (version 2.6.3). Differential expression analysis was done with  
552 a custom pipeline in R Studio using functions from edgeR (version 3.10.5) and limma (version

553 3.24.15). Permutation testing was performed with a custom R script. Iterative co-expression  
554 network analysis was performed with a custom pipeline following the standard WGCNA  
555 (version 3.2.3) workflow and functions. Gene Ontology analysis was performed with a custom  
556 wrapper using standard the TopGO (version 2.20.0) program. See extended methods for  
557 description and parameters.

## 558 **Author Contributions**

559 ALG, JE, JPL, JNC, JLS, KSZ, and ASN designed the experiments. Generation of mouse  
560 model: ASN, DD, AV, LAP, BM, IPF, VA; Mouse behavior: NAC, MCP, MDS, JNC, JLS;  
561 Mouse MRI: JE, JPL; Genomics: ALG, LS-F, IZ, BM, ASN; Neuroanatomy: ALG, TWS, IZ,  
562 KSZ. ALG, LS-F, JE, KSZ, JC, JLS, and ASN drafted the manuscript. All authors contributed to  
563 manuscript revisions.

## 564 **Acknowledgements**

565 Sequencing was performed at the UC Berkeley and UC Davis DNA cores. This work was  
566 supported by institutional funds from the UC Davis Center for Neuroscience and by the UC  
567 Davis MIND Institute Intellectual and Developmental Disabilities Research Center (U54  
568 HD079125). L.S.-F. was supported by the UC Davis Floyd and Mary Schwall Fellowship in  
569 Medical Research. A.V., L.A.P, and D.E.D. were supported by National Institutes of Health  
570 grants R24HL123879, U01DE024427, R01HG003988, U54HG006997, UM1HL098166.  
571 Research conducted at the E.O. Lawrence Berkeley National Laboratory was performed under  
572 Department of Energy Contract DE-AC02-05CH11231, University of California.

## 573 References

- 574 1. Chen, T. & Dent, S. Y. R. Chromatin modifiers and remodellers: regulators of cellular  
575 differentiation. *Nat. Rev. Genet.* **15**, 93–106 (2014).
- 576 2. Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition: unexpected roles  
577 for chromatin. *Nat. Rev. Genet.* **14**, 347–359 (2013).
- 578 3. Sanders, S. J. First glimpses of the neurobiology of autism spectrum disorder. *Curr. Opin. Genet.*  
579 *Dev.* **33**, 80–92 (2015).
- 580 4. McCarthy, S. E. *et al.* De novo mutations in schizophrenia implicate chromatin remodeling and  
581 support a genetic overlap with autism and intellectual disability. *Mol. Psychiatry* **19**, 652–658  
582 (2014).
- 583 5. Vissers, L. E. L. M., Gilissen, C. & Veltman, J. A. Genetic studies in intellectual disability and  
584 related disorders. *Nat. Rev. Genet.* **17**, 9–18 (2016).
- 585 6. De Rubeis, S. *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**,  
586 209–215 (2014).
- 587 7. Iossifov, I. *et al.* The contribution of de novo coding mutations to autism spectrum disorder.  
588 *Nature* **515**, 216–221 (2014).
- 589 8. Bernier, R. *et al.* Disruptive CHD8 mutations define a subtype of autism early in development.  
590 *Cell* **158**, 263–276 (2014).
- 591 9. Damaschke, N. A. *et al.* Frequent disruption of chromodomain helicase DNA-binding protein 8  
592 (CHD8) and functionally associated chromatin regulators in prostate cancer. *Neoplasia* **16**, 1018–  
593 1027 (2014).
- 594 10. Sawada, G. *et al.* CHD8 is an independent prognostic indicator that regulates Wnt/ $\beta$ -catenin  
595 signaling and the cell cycle in gastric cancer. *Oncol. Rep.* **30**, 1137–1142 (2013).
- 596 11. Nishiyama, M. *et al.* Early embryonic death in mice lacking the beta-catenin-binding protein  
597 Duplin. *Mol. Cell. Biol.* **24**, 8386–8394 (2004).
- 598 12. Sugathan, A. *et al.* CHD8 regulates neurodevelopmental pathways associated with autism  
599 spectrum disorder in neural progenitors. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E4468–77 (2014).
- 600 13. Willsey, A. J. *et al.* Coexpression networks implicate human midfetal deep cortical projection  
601 neurons in the pathogenesis of autism. *Cell* **155**, 997–1007 (2013).
- 602 14. Cotney, J. *et al.* The autism-associated chromatin modifier CHD8 regulates other autism risk  
603 genes during human neurodevelopment. *Nat Commun* **6**, 6404 (2015).
- 604 15. Croft, D. *et al.* The Reactome pathway knowledgebase. *Nucleic Acids Res.* **42**, D472–7 (2014).
- 605 16. Sanders, S. J. *et al.* Insights into Autism Spectrum Disorder Genomic Architecture and Biology  
606 from 71 Risk Loci. *Neuron* **87**, 1215–1233 (2015).
- 607 17. Hormozdiari, F., Penn, O., Borenstein, E. & Eichler, E. E. The discovery of integrated gene  
608 networks for autism and related disorders. *Genome Research* **25**, 142–154 (2015).
- 609 18. Parikshak, N. N. *et al.* Integrative functional genomic analyses implicate specific molecular  
610 pathways and circuits in autism. *Cell* **155**, 1008–1021 (2013).
- 611 19. Darnell, J. C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function  
612 and autism. *Cell* **146**, 247–261 (2011).
- 613 20. Voineagu, I. *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular  
614 pathology. *Nature* **474**, 380–384 (2011).
- 615 21. Kobayashi, M. *et al.* Nuclear localization of Duplin, a beta-catenin-binding protein, is essential for  
616 its inhibitory activity on the Wnt signaling pathway. *J. Biol. Chem.* **277**, 5816–5822 (2002).
- 617 22. Nishiyama, M. *et al.* CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment  
618 during early embryogenesis. *Nat. Cell Biol.* **11**, 172–182 (2009).
- 619 23. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis.  
620 *BMC Bioinformatics* **9**, 559 (2008).
- 621 24. Alrahbeni, T. *et al.* Full UPF3B function is critical for neuronal differentiation of neural stem



- 622 cells. *Mol Brain* **8**, 33 (2015).
- 623 25. Laumonnier, F. *et al.* Mutations of the UPF3B gene, which encodes a protein widely expressed in  
624 neurons, are associated with nonspecific mental retardation with or without autism. *Mol.*  
625 *Psychiatry* **15**, 767–776 (2010).
- 626 26. Jensen, L. J. *et al.* STRING 8--a global view on proteins and their functional interactions in 630  
627 organisms. *Nucleic Acids Res.* **37**, D412–6 (2009).
- 628 27. Silverman, J. L., Yang, M., Lord, C. & Crawley, J. N. Behavioural phenotyping assays for mouse  
629 models of autism. *Nat. Rev. Neurosci.* **11**, 490–502 (2010).
- 630 28. Yang, M., Silverman, J. L. & Crawley, J. N. Automated three-chambered social approach task for  
631 mice. *Curr Protoc Neurosci* **Chapter 8**, Unit 8.26–8.26.16 (2011).
- 632 29. Scattoni, M. L., Ricceri, L. & Crawley, J. N. Unusual repertoire of vocalizations in adult BTBR  
633 T+tf/J mice during three types of social encounters. *Genes Brain Behav.* **10**, 44–56 (2011).
- 634 30. Silverman, J. L. *et al.* GABAB Receptor Agonist R-Baclofen Reverses Social Deficits and  
635 Reduces Repetitive Behavior in Two Mouse Models of Autism. *Neuropsychopharmacology* **40**,  
636 2228–2239 (2015).
- 637 31. Hagerman, R., Au, J. & Hagerman, P. FMR1 premutation and full mutation molecular  
638 mechanisms related to autism. *J Neurodev Disord* **3**, 211–224 (2011).
- 639 32. Lee, J.-A. *et al.* Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related  
640 Genes. *Neuron* **89**, 113–128 (2016).
- 641 33. Zhang, X. *et al.* Cell-Type-Specific Alternative Splicing Governs Cell Fate in the Developing  
642 Cerebral Cortex. *Cell* **166**, 1147–1162.e15 (2016).
- 643 34. Pilaz, L.-J. & Silver, D. L. Post-transcriptional regulation in corticogenesis: how RNA-binding  
644 proteins help build the brain. *Wiley Interdiscip Rev RNA* **6**, 501–515 (2015).
- 645 35. Ellegood, J. *et al.* Clustering autism: using neuroanatomical differences in 26 mouse models to  
646 gain insight into the heterogeneity. *Mol. Psychiatry* **20**, 118–125 (2015).
- 647 36. Etherton, M. R., Blaiss, C. A., Powell, C. M. & Südhof, T. C. Mouse neurexin-1alpha deletion  
648 causes correlated electrophysiological and behavioral changes consistent with cognitive  
649 impairments. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 17998–18003 (2009).
- 650 37. Orosco, L. A. *et al.* Loss of Wdfy3 in mice alters cerebral cortical neurogenesis reflecting aspects  
651 of the autism pathology. *Nat Commun* **5**, 4692 (2014).
- 652 38. Clipperton-Allen, A. E. & Page, D. T. Pten haploinsufficient mice show broad brain overgrowth  
653 but selective impairments in autism-relevant behavioral tests. *Hum. Mol. Genet.* **23**, 3490–3505  
654 (2014).
- 655 39. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
- 656 40. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential  
657 expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
- 658 41. Dorr, A. E., Lerch, J. P., Spring, S., Kabani, N. & Henkelman, R. M. High resolution three-  
659 dimensional brain atlas using an average magnetic resonance image of 40 adult C57Bl/6J mice.  
660 *NeuroImage* **42**, 60–69 (2008).
- 661