

1 Wiz binds active promoters and CTCF-binding

2 sites and is required for normal behaviour in

3 the mouse

4 Luke Isbel¹, Lexie Prokopuk^{1,2}, Haoyu Wu³, Lucia Daxinger^{1,3}, Harald Oey^{1,4}, Alex

5 Spurling¹, Adam J. Lawther⁵, Matthew W. Hale⁵, Emma Whitelaw^{1*}

6 ¹ Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science,

7 Melbourne, Australia

8 ² Current Address: Centre for Genetic Diseases, Hudson Institute of Medical Research,

9 Melbourne, Australia.

10 ³ Department of Human Genetics, Leiden University Medical Centre,

11 Leiden, The Netherlands

12 ⁴ Current Address: Translational Research Institute, University of Queensland Diamantina

13 Institute, Brisbane, Australia

14 ⁵ Department of Psychology and Counselling, School of Psychology and Public Health,

15 La Trobe University, Melbourne, Australia

16 *For Correspondence: E.Whitelaw@latrobe.edu.au (EW)

17 **Abstract**

18 We previously identified *Wiz* in a mouse screen for epigenetic modifiers. Due to its known
19 association with G9a/GLP, *Wiz* is generally considered a transcriptional repressor. Here we
20 provide evidence that it may also function as a transcriptional activator. *Wiz* levels are high in
21 brain but its function and direct targets are unknown. ChIP-seq was performed in adult
22 cerebellum and *Wiz* peaks were found at promoters and transcription factor CTCF binding sites.
23 RNA-seq in *Wiz* mutant mice identified genes differentially regulated in adult cerebellum and
24 embryonic brain. In embryonic brain most decreased in expression and included clustered
25 protocadherin genes. These also decreased in adult cerebellum and showed strong *Wiz* ChIP-
26 seq enrichment. Because a precise pattern of protocadherin gene expression is required for
27 neuronal development, behavioural tests were carried out on mutant mice, revealing an
28 anxiety-like phenotype. This is the first evidence of a role for *Wiz* in neural function.

29 **Introduction**

30 An ENU mutagenesis screen for modifiers of epigenetic reprogramming was carried out in the
31 mouse and the lines produced are termed *MommeDs*, *Modifiers of murine metastable*
32 *epialleles, Dominant* (Daxinger et al., 2013). The screen used a multicopy GFP transgene, under
33 the control of erythroid promoter/enhancer sequences, as a reporter. As this reporter
34 undergoes stochastic silencing among otherwise identical erythroid cells, it is reminiscent of the
35 position effect variegation (PEV) screens used to identify genes with roles in regulating gene
36 expression (Henikoff, 1990, Fodor et al., 2010). Male mice with the reporter were treated with
37 N-ethyl-N-nitrosourea (ENU) and offspring were screened for changes in the percentage of
38 erythrocytes expressing GFP; this was a screen for dominant effects. The mutation underlying
39 the *MommeD30* strain was found to be a single base pair deletion in the gene *Widely*
40 *interspaced zinc finger motifs*, *Wiz*. The deletion, of an adenine, causes a frame-shift that
41 introduces a premature stop codon. *Wiz*^{*MommeD30/MommeD30*} embryos die around mid-gestation
42 and heterozygosity for the mutant allele results in approximately half the normal levels of the
43 protein, confirming that *Wiz*^{*MommeD30*} is a null allele (Daxinger et al., 2013). This is the only
44 mouse strain with a mutant form of this gene to be reported.

45 Little is known about the role of Wiz in any tissue. The protein is predicted to have 11 C2H2
46 type zinc finger domains that are unusually widely spaced. Some of them can bind DNA and
47 some of them can mediate protein-protein interactions with the G9a/GLP (G9a-like) histone
48 methyltransferase complex (Ueda et al., 2006, Bian et al., 2015). The G9a/GLP heterodimer
49 catalyses the methylation of H3K9me1 and H3K9me2 (Jenuwein et al., 1998, Tachibana et al.,
50 2001, Tachibana et al., 2002, Tachibana et al., 2005), which is associated with transcriptional

51 repression (Barski et al., 2007) and is located on chromatin in broad regions, termed LOCKs
52 (large organized chromatin K9-modifications) (Wen et al., 2009). Work carried out in cell lines
53 has shown that Wiz can stabilize the G9a/GLP complex, mediate its localisation to DNA and is
54 required for maintaining normal global levels of H3K9me2, suggesting a role in transcriptional
55 repression (Bian et al., 2015, Simon et al., 2015, Ueda et al., 2006).

56 It has also been shown that G9a can act as a transcriptional activator, independent of its
57 histone methyltransferase activity (Chaturvedi et al., 2009, Lee et al., 2006, Oh et al., 2014,
58 Purcell et al., 2011, Yuan et al., 2007) and it has recently been suggested that WIZ has a role in
59 the activation function of G9a (Simon et al., 2015). *Wiz*^{MommeD30} was identified in the screen
60 because mice heterozygous for the mutant allele showed decreased expression of the GFP
61 transgene. Here, we show at an independent locus known to be sensitive to the dosage of
62 epigenetic modifiers, the *Agouti viable yellow* (*A^{vy}*) allele, that haploinsufficiency for Wiz also
63 results in decreased expression.

64 Wiz is expressed highly in the brain, both in the embryo and in the adult (Matsumoto et al.,
65 1998). As reduced levels of G9a/GLP are known to result in the neurological disease Kleefstra
66 syndrome (Kleefstra et al., 2006, Willemse et al., 2011), it is possible that reduced levels of Wiz
67 might have neurological consequences. Kleefstra syndrome individuals display a complex range
68 of neurological symptoms, including anxiety (Willemse et al., 2011). We decided to study the
69 role of Wiz in the brains of *Wiz*^{MommeD30} mice. We carried out ChIP-seq and RNA-seq in adult
70 cerebellum, a tissue in which Wiz is expressed at high levels (Matsumoto et al., 1998). Our
71 results show that Wiz binds tens of thousands of transcriptional regulatory regions across the
72 genome, some of which decrease in expression in *Wiz*^{MommeD30/+} tissue. ChIP-seq, with an anti-

73 Wiz antibody, has been reported in a human kidney cell line (Bian et al., 2015) and there is little
74 overlap with our findings. Our molecular studies prompted us to assess behaviour in
75 heterozygous mutant mice and we found that *Wiz*^{MommeD30/+} mice display an anxiety-like
76 phenotype.

77 **Results**

78 **Wiz haploinsufficiency increases the probability of silencing the metastable epiallele, *Agouti*
79 **viable yellow.****

80 Haploinsufficiency for Wiz resulted in an increase in the probability of silencing at the GFP
81 transgene reporter, suggesting a role in transcriptional activation at this locus (Daxinger et al.,
82 2013). To test the effect of haploinsufficiency for Wiz at an independent reporter known to be
83 sensitive to dosage of epigenetic modifiers, we used mice carrying the *A^{vy}* allele. This strain has
84 been maintained on the C57BL/6J background in a heterozygous state (*A^{vy}/a*). The coat colour
85 of the mouse can be used as a reporter of transcriptional activity at the *A^{vy}* locus (Duhl et al.,
86 1994). An active locus results in a yellow coat, a silent locus results in a dark brown coat (called
87 pseudoagouti) and a mottled coat arises when some clusters of cells express the locus and
88 some clusters do not. Male *Wiz*^{MommeD30/+} mice (maintained on an FVB/NJ background) were
89 crossed to female *A^{vy}/a* (pseudoagouti) mice and the coat colours of the offspring were scored
90 at weaning (three weeks). FVB/NJ mice carry a wildtype *Agouti* locus, *A/A*, so ~50% of offspring
91 from these crosses will be *A^{vy}/A* and differ only at the *Wiz*^{MommeD30} allele, i.e. *Wiz*^{MommeD30/+};
92 *A^{vy}/A* or *Wiz*^{+/+}; *A^{vy}/A*. Offspring were genotyped for the *A^{vy}* and *Wiz*^{MommeD30} alleles. Mice that
93 did not inherit the *A^{vy}* allele were excluded from analysis. The reciprocal cross was also carried

94 out, i.e. female *Wiz*^{MommeD30/+} mice were crossed to male *A^{v/y}/a* (pseudoagouti) mice. The
95 pedigrees and outcomes for both crosses are shown in Figure 1A and 1B.

96 *Wiz*^{MommeD30/+} offspring from both crosses displayed a shift towards silencing (i.e. more
97 pseudoagouti mice), compared to *Wiz*^{+/+} littermates. The shift was statistically significant in
98 both cases; the shift from *Wiz*^{MommeD30/+} sire cross had a p-value of 0.0002, and from the
99 *Wiz*^{MommeD30/+} dam cross had a p-value of 0.0458 (Figure 1A and 1B). The effect of
100 haploinsufficiency at the *A^{v/y}* locus is similar to the effect at the GFP transgene reporter, i.e.
101 reduced levels of Wiz were associated with increased silencing.

102 **Wiz peaks occur at regulatory regions, many of which bind CTCF**

103 To test the specificity of an anti-Wiz antibody (NBP180586, Novus Biologicals, Littleton, USA) in
104 mouse tissue, Western blotting was carried out using total protein lysates from *Wiz*^{+/+} and
105 *Wiz*^{MommeD30/MommeD30} E12.5 embryonic heads, prior to the death of the *Wiz*^{MommeD30/MommeD30}
106 fetuses (Figure 2 – figure supplement 1). The bands detected in the wildtype sample (at around
107 100 kDa) were absent in the sample from the homozygous mutant, consistent with an antibody
108 that is specific for Wiz. A larger ~160 KDa isoform has previously been reported in adult
109 cerebellum (Matsumoto et al., 1998) and consistent with that report a band of this size is
110 detected with our antibody in protein extracted from this tissue (Figure 2 – figure supplement
111 1). In addition, we used an unbiased tandem affinity purification approach and compared mass
112 spectroscopy data with previously published datasets analysing Wiz binding partners. Co-
113 immunoprecipitation (Co-IP) with the anti-Wiz antibody and an anti-IgG control antibody was
114 carried out on pooled E13.5 brain tissue (n = 1 pooled sample from 5 individuals) and adult
115 cerebellum (n = 2 biological replicates). Four members of the Wiz-Zfp644-EHMT1-EHMT2

116 complex (Ueda et al., 2006, Bian et al., 2015) were in the top five proteins in the dataset,
117 ranked by peptide count in embryonic brain (Figure 2 – figure supplement2).

118 The cerebellums of four adult male *Wiz*^{+/+} mice were pooled for ChIP-seq; two cerebellums per
119 sample and two biological replicates. Cerebellum was chosen for two reasons; it is a relatively
120 simple tissue in the brain with respect to cell types and it is an ENCODE tissue, enabling
121 comparisons to be made with ChIP-seq data obtained for other chromatin proteins. Library
122 preparation and sequencing were carried out by Active Motif (Carlsbad, USA). Over 30 million
123 reads were generated for each of the two Wiz ChIP-seq datasets and the chromatin Input
124 sample dataset. At least 24 million reads could be aligned to the mouse mm10 genome in all
125 cases, after filtering out PCR duplicates.

126 Approximately 40,000 peaks were identified at which the two Wiz-ChIP-seq datasets were
127 significantly enriched compared to the Input control with a p-value of $\leq 1 \times 10^{-20}$ (Supplementary
128 file 1). A representative Wiz peak and ChIP-qPCR validation at the same locus is shown in Figure
129 2A and 2B, respectively. To determine where Wiz was bound with respect to genic regions of
130 transcriptionally active and inactive genes, an analysis of overlap with gene promoters and gene
131 bodies was carried out. Peaks were roughly equally distributed across promoters (defined as 2
132 kb up and downstream of the TSS), intragenic and intergenic regions, despite the fact that
133 promoters comprise only a small fraction of the total mouse genome (Figure 2C). Consistent
134 with this, the deep sequencing read density across the body of all Ensembl genes (normalized
135 for gene length; see methods), as well as 2 kb upstream of the TSS (transcription start site) and
136 downstream of the TTS (transcriptional termination site), showed strong enrichment at the TSS
137 (Figure 2D and 2E).

138 The MEME-ChIP program was used to identify motifs that are common at Wiz-binding locations
139 and several were identified with high statistical significance, i.e. p-values from 1.1×10^{-6902} to
140 2.9×10^{-439} (Figure 3A). One of these, Wiz motif 1, had a significance (MEME-ChIP E-value) E-
141 Value of 2.3×10^{-6903} , an order of magnitude higher than other motifs and aligned almost
142 perfectly with the CTCF-DNA binding motif (Figure 3B). This motif was present in approximately
143 70% of all Wiz binding peaks across the genome.

144 An ENCODE ChIP-seq dataset for CTCF binding in mouse adult cerebellum is publically available.
145 To determine if CTCF peaks overlapped with Wiz peaks, Wiz ChIP-seq read density was
146 compared to ENCODE ChIP-seq data for CTCF in adult cerebellum at Wiz peaks (Figure 3C).
147 Strong overlap was seen. The ENCODE dataset for H3K4me3, a histone modification associated
148 with active transcription, was also mapped to the genome (mm10 build) and ChIP-seq peaks for
149 CTCF, H3K4me3 and Wiz were compared (Figure 3C), as described in Methods. Over half of Wiz
150 ChIP-seq peaks overlapped the promoter mark H3K4me3 and approximately a third of all Wiz
151 ChIP-seq peaks overlapped CTCF ChIP-seq peaks (Figure 3D). Virtually all locations with both
152 CTCF and H3K4me3 peaks also showed a Wiz peak (1356/1359). Interestingly, almost all CTCF
153 peaks overlapped with Wiz peaks.

154 A ChIP-seq dataset was recently published using an anti-Wiz antibody and HEK293T cells (Bian
155 et al., 2015); the authors conclude that Wiz localizes to gene promoters and (using EMSA
156 analysis) can bind to the DNA sequence CATTCCATTCCATT. Shown in their report are
157 representative screen shots of Wiz-ChIP-seq enrichment at the promoter region of seven genes
158 (defined in their study as 5kb up or downstream of the TSS) where there was a “consensus
159 DNA-binding motif confirmed at the locus”. We examined the DNA sequence surrounding each

160 of these seven genes, in all cases the location of the consensus sequence was at least 5Kb from
161 either the promoter where binding was demonstrated or the coordinates of the nearest
162 reported ChIP-seq peak ($n = 11,853$). Instead, the sequence was located downstream in the
163 body of each gene. After remapping this publically available Wiz ChIP-seq dataset (GSE62616),
164 only one of the seven genes was found to have enrichment at the consensus sequence (i.e.
165 reads mapping directly at the CATTCCATTCCATT motif), in the *SENP5* gene. In this case, reads
166 mapped to a sequence located in a repetitive HSATII satellite element and is likely to be a
167 mapping artefact (Figure 3-figure supplement 1). Enrichment was also seen at other copies of
168 the HSATII element across the genome that had the CATTCCATTCCATT sequence (data not
169 shown). We were unable to reproduce the enrichment shown in the report at any of the seven
170 gene promoters (data not shown).

171 **Genome-wide expression changes in embryonic brains and adult cerebellum of *Wiz*^{MommeD30/+}**
172 **mice**

173 Our results have shown that haploinsufficiency for Wiz alters expression at the GFP transgene
174 reporter and the *A^{VY}* locus. We were keen to see if haploinsufficiency for Wiz resulted in
175 changes in the expression of other genes and chose to investigate this in the brain. Firstly, we
176 used the whole brains of E13.5 embryos in an attempt to identify early events and secondly, we
177 used adult cerebellum, to compare expression changes to Wiz ChIP-seq data.

178 Initially, RNA-seq was performed using E13.5 brains from male *Wiz*^{MommeD30/+} and *Wiz*^{+/+}
179 embryos ($n = 3$ biological replicates for each genotype). Reads were aligned to the mouse
180 genome (mm10 build) and differential expression of Ensembl genes between genotypes was

181 calculated using the R package DESeq (Anders and Huber, 2010)(Figure 4A; Supplementary file
182 2).

183 RNA-seq was also carried out using brains from female E13.5 *Wiz*^{MommeD30/+} and *Wiz*^{+/+} embryos
184 and differential expression of genes was calculated as described above (n = 2 biological
185 replicates for each genotype, Figure 4A; Supplementary file 3).

186 Genes with a fold-change in average expression of less than 0.7 or greater than 1.3 (shown in
187 figures on log2 scale) and an adjusted p-value of ≤0.05, were considered significantly
188 differentially expressed. A total of 36 were found to change in males and 44 changed in
189 females. The majority of significantly differentially expressed genes, in both cases, decreased in
190 expression in the mutants (31/36 in males and 43/44 in females). As expected, *Wiz* itself was in
191 this group.

192 Of the combined set of 80 gene expression changes in males and females, 28 genes were
193 common between the sexes and all of these decreased in *Wiz*^{MommeD30/+} samples compared to
194 wildtype samples (Figure 4B and 4C). Genes that were significantly differentially expressed in
195 one sex but not in the other generally trended towards statistical significance in the latter and
196 none of these had a fold-change direction that was different between sexes, i.e. genes
197 statistically down regulated in males trended down in females and vice versa.

198 Among the significantly differentially expressed genes, a relatively high proportion were from
199 two gene clusters, the protocadherin β (Pcdhb) gene cluster on Chromosome 18 and a second
200 cluster consisting of 2610005L07Rik, 6820431F20Rik, AC152164.1, Gm10557, Gm21092,
201 Gm21769, Gm21811, Gm26804 and Gm6483 on Chromosome 8 (Figure 4-figure supplement 1).

202 The latter cluster is poorly annotated due to a nearby repeated region that has made mapping
203 difficult (Boyle and Ward, 1992). Under the annotations listed in the UCSC genome browser
204 (Karolchik et al., 2003), most are annotated as cadherin11 pseudogenes; all have introns and
205 some are predicted to be coding. While it is very likely that Wiz haploinsufficiency affects
206 multiple members of the cadherin-11 pseudogene cluster, it is possible that some expression
207 changes are mapping artefacts. A more complete genome assembly is needed before this can
208 be resolved by RNA-seq or single locus testing (i.e. real time quantitative PCR).

209 To see if the other significantly differentially expressed genes in E13.5 brain were related by
210 location, we examined their locations on each chromosome (Figure 4-figure supplement 1). No
211 other gene clusters were apparent, though it was noticed that several genes were located near
212 telomeres and four of the significantly differentially expressed genes, *Samd11*, *psid-ps1*, *psid-*
213 *ps2* and *Csf2ra*, are the last gene on the chromosome. While it is possible that the locations of
214 these genes and the cadherin-11 pseudogene cluster are incorrect in the current genome build
215 (mm10 build), they are likely to remain adjacent to heterochromatic repeats.

216 Cerebellums were also dissected from adult males ($n = 3$ per genotype) and RNA-seq was
217 carried out on each. A total of 82 genes were significantly differentially expressed; 29 genes
218 decreased and 53 increased in expression in *Wiz*^{MommeD30/+} samples compared to *Wiz*^{+/+} samples
219 (Figure 5A; Supplementary file 4). Of these, some genes ($n = 13$) overlapped with those
220 detected in the E13.5 brain data, including some from the protocadherin β cluster, some from
221 the cadherin 11-like cluster, *Wiz*, *Csf2ra* and *Rps4l* (Figure 5B). These all decreased in expression
222 in *Wiz*^{MommeD30/+} cerebellum compared to *Wiz*^{+/+} cerebellum, as they had done in the embryonic
223 brain.

224 In general, the down regulation of protocadherin β genes was more pronounced in the
225 embryonic brain than in the adult cerebellum, where significance was seen for only a minority
226 of the genes in the clusters (Supplementary file 4).

227 ChIP-seq data were investigated at the protocadherin β cluster (Figure 6A). A genome browser
228 screenshot of the Wiz binding is shown for the protocadherin β cluster and its 5' enhancer
229 (HS5-1bL). Wiz enrichment was detected at all protocadherin β promoters except *Pcdhb1*,
230 which is the only protocadherin β gene that lacks a canonical protocadherin promoter sequence
231 (Guo et al., 2015, Wu et al., 2001). Greatest enrichment was detected at the HS5-1bL enhancer.
232 This enrichment at the enhancer is an order of magnitude higher than that seen at the
233 individual promoters and this might reflect the fact that while single members of the cluster are
234 expressed in individual neurons, the HS5-1bL enhancer is likely to be active in all cell types
235 (Yokota et al., 2011). In all cases tested, enrichment of Wiz was validated using ChIP-qPCR at
236 protocadherin β promoters and the HS5-1bL element (Figure 6B). Conversely, primers located
237 in a gene desert and the *Pcdhb1* promoter, which did not show an expression change in any of
238 the tissues tested, were not enriched for Wiz.

239 Wiz and CTCF ChIP-seq read density were calculated at the transcriptional start sites of the ~82
240 significantly differentially regulated genes in cerebellum. Genes were first separated by
241 whether they decreased or increased in expression in *Wiz*^{MommeD30/+} cerebellum compared to
242 *Wiz*^{+/+} cerebellum. Genes that decreased in expression showed high enrichment for Wiz and some
243 enrichment for CTCF at the TSS (Figure 6C), while genes that increased in expression showed little
244 enrichment for Wiz and no enrichment for CTCF (Figure 6D), suggesting that those that showed
245 increased expression were secondary effects (see Discussion).

246 ***Wiz*^{MommeD30/+} mice have altered behaviour**

247 To characterize the behavioural phenotype of *Wiz*^{MommeD30/+} mice, singly housed males were
248 monitored over a period of 12 days using telemetry devices that record body temperature and
249 locomotor activity. An initial cohort of *Wiz*^{+/+} and *Wiz*^{MommeD30/+} age-matched mice (n = 5 per
250 genotype) was tested and a second cohort (n = 5 *Wiz*^{+/+}, 6 *Wiz*^{MommeD30/+}) was tested two
251 months later to validate results from the first. Data from both cohorts showed similar effects
252 and were combined for presentation. Mice normally display a circadian pattern of activity
253 dependent on the light/dark cycle and so data is shown across the 24-hour period after
254 averaging results at each time point across the 12 days. No shift was seen in the circadian
255 pattern, but decreased activity was seen in the first two hours and the fifth hour after lights off
256 in *Wiz*^{MommeD30/+} mice compared with the *Wiz*^{+/+} mice (Figure 7A; time x genotype interaction,
257 F(23, 391) = 3.923, p < 0.001, ε = 0.350).

258 Temperature profiles were also collected for each of the mice by collapsing three days of
259 monitoring into hourly intervals and taking the average from each 24-hour time point. No
260 difference was seen in *Wiz*^{MommeD30/+} mice compared to *Wiz*^{+/+} mice at any of the time points
261 (Figure 7-figure supplement1).

262 The elevated plus maze (EPM) is a test of anxiety-like behaviour in mice (Walf and Frye, 2007).
263 The same mice used for telemetry monitoring were subjected to an EPM test following the 12-
264 day monitoring period. *Wiz*^{MommeD30/+} showed an approximately 25% reduction in the frequency
265 of entry into the open arms of the EPM apparatus compared with the *Wiz*^{+/+} mice (Figure 7B

266 and C; $t(17) = 2.90$, $p = 0.010$). Importantly, mice were assayed during “lights on” periods
267 (around 2pm - 4pm) when activity was not different between the strains (Figure 7A).

268 To validate the EPM data, two independent cohorts of mice ($Wiz^{MommeD30/+}$, n= 22; $Wiz^{+/+}$, n=21)
269 were subjected to a light-dark box test of anxiety-like behaviour. $Wiz^{MommeD30/+}$ mice showed
270 reduced time spent in the light and reduced rearing activity (Figure 7D and 7E). These findings
271 are consistent with an anxiety-like phenotype in $Wiz^{MommeD30/+}$ mice.

272 **Discussion**

273 ***Wiz* is involved in transcriptional activation and binds many promoters in neural tissue**

274 We previously described a *Wiz* mutation in the *MommeD30* strain (Daxinger et al., 2013). This
275 was the only exonic mutation within the linked interval and in agreement with this being the
276 causative *MommeD30* mutation, RNA-seq data presented here show no other genes within the
277 linked interval were miss-expressed in mutants.

278 While a role for *Wiz* in transcriptional repression is well recognized (Mulligan et al., 2008), the
279 notion that it may be involved in transcriptional activation is less well documented. This study
280 supports the latter idea in a number of ways. Firstly, we show that haploinsufficiency for *Wiz*
281 results in reduced expression at the $A^{\vee Y}$ allele and previous work from our laboratory has shown
282 that half the normal levels of *Wiz* is associated with reduced expression of a multicopy GFP
283 transgene (Daxinger et al., 2013). By focussing on neural tissue, we have gone on to show that
284 haploinsufficiency for *Wiz* results in decreased expression of a number of genes, mainly within gene
285 clusters or adjacent to telomeric repeats in both embryonic brain and adult cerebellum. While any
286 instance of decreased expression could, theoretically, be an indirect effect, our ChIP-seq data suggests

287 otherwise. At most of the sites at which decreased expression is observed, Wiz peaks were detected.
288 Furthermore, analysis of the set of genes with altered expression in adult cerebellum, showed that Wiz
289 was detected at the promoters of those with decreased expression but not at those with increased
290 expression. This suggests that those sites at which reducing the levels of Wiz causes increased
291 expression, i.e. Wiz is acting as a repressor, are indirect effects.

292 **Wiz might bind directly to DNA.**

293 While Wiz has no recognisable catalytic domains, it does have C2H2-type zinc fingers and these
294 have an unusual widely-spaced configuration (Matsumoto et al., 1998). Most proteins that
295 contain multiple zinc fingers have them in a cluster, with short “linker” sequences (usually less
296 than 10 amino acids) separating the fingers (Schuh et al., 1986). Most Wiz zinc fingers are at
297 least 50 amino acids apart. The significance of the large linker configuration remains unclear
298 (Cléard and Spierer, 2001). Other zinc finger-containing proteins with this configuration include
299 Zfp292, Rlf and the *Drosophila* protein Su(Var)3-7 and all of these have been shown to bind
300 DNA *in vitro* using electromobility shift assays (Cléard and Spierer, 2001, Harten et al., 2015,
301 Lipkin et al., 1993). *Rlf*, like *Wiz*, was identified as an *E(Var)* in the *MommeD* screen (Daxinger et
302 al., 2013). Little has been reported about the function of Zfp292.

303 Consistent with previous studies using different Wiz antibodies (Matsumoto et al., 1998, Ueda
304 et al., 2006, Ma et al., 2015, Simon et al., 2015) we identified multiple isoforms of Wiz present
305 in the mouse, three at ~100 kDa and one larger ~160 kDa. The requirement of these different
306 isoforms is unknown and requires careful, isoform-specific study of Wiz. It is also possible that
307 the three ~100 kDa isoforms represent covalent modifications to the Wiz protein. For example,
308 it is known that G9A can methylate non-histone targets, including Wiz (Rathert et al., 2008).

309 **Wiz binding overlaps with CTCF binding in the cerebellum**

310 Our ChIP-seq studies suggest that in adult cerebellum, Wiz binds at promoters and at sites
311 which CTCF, a recognized transcriptional activator, is also known to bind. Because this study has
312 been carried out in a tissue made up of more than one cell type, peak overlap does not
313 necessarily imply that the two proteins bind together. It could be explained by CTCF binding to
314 these locations in one cell type and Wiz binding in another cell type. Competitive binding with
315 CTCF and the closely related BORIS has been demonstrated (Pugacheva et al., 2015). To our
316 knowledge, Wiz has not been pulled down with antibodies to CTCF (Xiao et al., 2011, Yusufzai et
317 al., 2004), though a protein-protein interaction is possible in neural cell types where *Wiz* is
318 expressed highly and co-immunoprecipitation of binding partners remains to be tested.

319 CTCF has multiple roles in organizing the genome; it is enriched at topologically associated
320 domain boundaries and can mediate interactions between enhancer and promoter elements
321 (Bell et al., 1999, Ong and Corces, 2014). Broad regions marked by H3K9me2 often contain
322 small ‘euchromatic islands’ that are devoid of H3K9me2, are DNase1 hypersensitive and have
323 CTCF binding (Wen et al., 2012). A recent study in a number of human cell lines found that
324 reducing WIZ levels resulted in less G9a at CTCF-binding sites across the genome (by carrying
325 out ChIP-seq for G9a) (Simon et al., 2015). Further work is needed to show mechanistically how
326 CTCF, Wiz and G9a establish chromatin at these euchromatic islands.

327 ChIP-seq for WIZ in a human kidney cell line has recently been published (Bian et al., 2015) and
328 no overlap with CTCF-binding sites was reported. An in-house antibody (not commercially
329 available) was used. Approximately 11,000 WIZ-enriched regions were detected and the
330 consensus binding site that emerged, CATTCCATTCCATT, is quite different from the one

331 reported here. This sequence was found in only 4.5% of their peaks, whereas the Wiz consensus
332 reported here (Wiz motif 1) was found in 70% of the peaks identified in our study. Remapping
333 of their Wiz ChIP-seq dataset was carried out, and failed to produce the peaks reported except
334 at a repetitive HSATII satellite element, suggesting a mapping artefact. The differences might
335 also be linked to the different cell types used. Resolution of this discrepancy will require the
336 publication of Wiz ChIP-seq datasets with additional anti-Wiz antibodies and in different
337 tissues.

338 **Wiz regulates protocadherin gene expression**

339 Protocadherin genes are expressed mainly in neurons (Kohmura et al., 1998, Wu and Maniatis,
340 1999) and here we report the involvement of Wiz at protocadherin β genes. CTCF has been
341 shown by others to bind to the promoters and enhancer of this locus, mediating the looping
342 required for the transcription of single protocadherin β genes in single neural cells (Guo et al.,
343 2015). Given that almost all sites in the cerebellum that were bound by CTCF were also bound
344 by Wiz, it is possible that Wiz has a function in chromatin looping between enhancers and
345 promoters. Determining the importance of Wiz in looping will require chromosome
346 conformation capture (3C) analysis in Wiz null cells. Early embryonic lethality precluded testing
347 this in the tissue used in this study.

348 The protocadherin cluster exists as a large tandem array. At this locus, heterochromatinisation
349 facilitates single isoform expression in single cells i.e. the silencing of the majority of
350 protocadherin genes in any single cell is associated with heterochromatin formation (Kawaguchi
351 et al., 2008, Toyoda et al., 2014); the expressed loci must remain active despite being
352 embedded in heterochromatin and it is possible that Wiz has a role in this scenario. In keeping

353 with this hypothesis, the RNA-seq data presented here suggests that other loci sensitive to Wiz
354 haploinsufficiency include genes located adjacent to telomeres, well known heterochromatic
355 regions.

356 **Wiz is required for normal behaviour in the adult mouse**

357 The importance of epigenetic modifications in neural function is emerging; many enzymes that
358 modify histones and DNA have been found to be critical for development and function of the
359 brain. Prior to this study, no link between *Wiz* and behaviour had been shown in either mice or
360 humans. Here we report that *Wiz*^{MommeD30/+} mice showed a more anxious phenotype than
361 wildtype littermates in a number of assays of anxiety. It is possible that heterozygosity for rare
362 *WIZ* mutations in humans could influence neurological disorders that have a complex genetic
363 aetiology.

364 It remains to be seen if the gene expression changes and behavioural phenotype in
365 *Wiz*^{MommeD30/+} mice correlates with altered neuroanatomy. Mice with reduced expression of
366 protocadherin β genes, due to knockout of CTCF (Hirayama et al., 2012) or the downstream
367 enhancer (Yokota et al., 2011), have disorganized formation of the barrel cortex, a region of the
368 brain important for processing information for the tactile somatosensory pathway (Petersen,
369 2007). It remains to be determined if this phenotype in the barrel cortex is recapitulated in
370 *Wiz*^{MommeD30/+} mice.

371 **Methods**

372 **Mouse strains**

373 *Wiz*^{MommeD30} mice were produced in an ENU mutagenesis screen, as previously described
374 (Daxinger et al., 2013). The ENU screen was carried out using a *Line3* strain that is on an FVB/NJ

375 inbred background, as described previously (Blewitt et al., 2005). *Line3* are homozygous for a
376 multicopy GFP transgene with human α -globin promoter and HS-40 elements and the
377 *Wiz*^{MommeD30} mutation was maintained by crossing heterozygous mice to *Line3*. The *A^{VY}* allele
378 arose from a C3H/HeJ colony, was backcrossed to C57BL/6J for at least 20 generations and was
379 maintained on that background in the heterozygous state (*A^{VY/a}*). Genotyping were carried out
380 as previously described for the *Wiz*^{MommeD30} (Daxinger et al., 2013) and *A^{VY}* alleles (Rakyan et al.,
381 2003), primers are listed in Supplementary file 5. All animal work was conducted in accordance
382 with the Australian code for the care and use of animals for scientific purposes, this study was
383 approved by the Animal Ethics Committee of La Trobe University, project numbers 12-74, 12-
384 75, 15-01.

385 **ChIP-seq and ChIP-qPCR for Wiz in cerebellum**

386 Cerebellum tissue from 8 week old wildtype mice were snap-frozen and sent to Active Motif
387 (Carlsbad, USA) for ChIP, library preparation, sequencing and initial data QC. Two cerebellums
388 were pooled per biological replicate. The rabbit polyclonal anti-Wiz (NBP180586, Novus
389 Biologicals, Littleton, USA) was used for ChIP on two biological replicates and an input
390 chromatin sample was made by pooling equal amounts from the two biological replicates.
391 Sequencing was carried out for 75mer read lengths on the NextSeq 500 platform (Illumina, San
392 Diego, USA). At least 30 million reads were generated per library. Read alignment to the mouse
393 mm10 genome build was carried out using the Bowtie2 program with default settings (version
394 2.2.2) (Langmead and Salzberg, 2012) and peak calling was carried out using the MACS2
395 program with the settings “--down-sample --call-summits” (version 2.1.0) (Zhang et al., 2008).
396 Peaks were considered significant if they had a p-value of $\leq 1 \times 10^{-20}$.

397 Heat plots and read sequencing density figures were generated using the seqMiner program
398 (version 1.3.3) (Ye et al., 2011) using default settings. Heat plots were generated by
399 subsampling all datasets to approximately 16 million reads and clustering was carried out using
400 the seqMiner program using default settings. When calculating deep sequencing read density
401 over gene bodies, genes of different lengths were scaled to an equal number and the read
402 density was calculated for each of these by the seqMiner program. Genes were classified as
403 active ($n = 13279$ genes) or silent ($n = 25152$ genes) based on an averaged raw read count of
404 less than 50 from RNA-seq carried out in adult cerebellum tissue (see below), active genes
405 accounted for approximately ~99% of read alignments in the dataset.

406 To analyse the overlap of Wiz ChIP-seq peaks and ENCODE dataset (Shen et al., 2012) peaks, the
407 latter were first downloaded from the UCSC genome browser (Karolchik et al., 2003), datasets
408 were 3' trimmed so that all datasets contained reads of equal length and mapping and peak
409 calling was carried out as described above. ChIP-seq datasets for CTCF and the histone mark
410 H3K4me3 were analysed and in each case this consisted of two biological replicates and an
411 input dataset. Overlap between two peaks was defined as the coordinates of two peaks directly
412 overlapping at one or more bases in the genome. Motif discovery was performed with
413 significant Wiz peaks (region summit +- 500bp) using the MEME-ChIP (version 4.10.0)
414 (Machanick and Bailey, 2011) and MEME suite programs MEME, DREAME, CentriMo and
415 Tomtom with default settings. The publically available Wiz ChIP-seq datasets were downloaded
416 from the GEO website (www.ncbi.nlm.nih.gov/geo/) and mapped to the human genome (hg19
417 build) as described above.

418 ChIP-qPCR was performed with the Anti-Wiz antibody (NBP180586) on three biological
419 replicates by Active Motif and loci were amplified using primers designed by Active Motif at
420 regions of ChIP-seq enrichment indicated in the Figures. Enrichment was calculated by running
421 qPCR reactions alongside known amounts of DNA to generate a standard curve. These are
422 represented as binding events detected per 1000 cells.

423 **Data availability**

424 The data sets supporting the results of this article are available in the NCBI Gene Expression
425 Omnibus under the accession code GSE76909.

426 **Timed matings**

427 Embryos were produced by heterozygous intercrosses ($Wiz^{MommeD30/+}$ x $Wiz^{MommeD30/+}$) and
428 detection of a post coital plug was defined as E0.5 dpc. Dams were euthanized by cervical
429 dislocation and embryos dissected into 1X PBS, tissues were either snap frozen with dry ice for
430 later RNA extraction or lysed overnight for DNA extraction as previously described (Morgan et
431 al., 1999). The sex of embryos was determined by PCR using primer sets annealing either at X or
432 Y chromosomal loci, listed in Supplementary file 5.

433 **RNA-seq**

434 RNA was purified from 8-week old male cerebellum or embryonic brain (E13.5) using Triazol
435 reagent (Life Technologies, Carlsbad, USA), as per manufacture instruction. RNA sequencing
436 was carried out by the Australia Genome Research Facility (AGRF, Parkville, AUS) and at least 20
437 million 100bp single end reads were generated on an Illumina HiSeq platform for each sample,

438 from libraries generated using the Illumina TruSeq RNA Sample Preparation kit (Illumina, San
439 Diego, CA, USA). An initial QC analysis was performed by AGRF. Sequencing reads were mapped
440 to the mouse mm10 genome build using the program Tophat (version 2.0.11) (Trapnell et al.,
441 2009), read counts for gene exons were extracted using the program htseq-count (version
442 0.6.1) (Anders et al., 2015) and differential gene expression was assessed using the R-package
443 DEseq (Anders and Huber, 2010), as described previously (Isbel et al., 2015). Genes were
444 considered significantly differentially regulated if they had a fold-change ($Wiz^{MommelD30/+}/Wiz^{+/+}$)
445 of less than 0.7 or greater than 1.3 (shown in figures on a log2 scale) and an adjusted p-value of
446 ≤ 0.05 .

447 **Western blotting for Wiz in embryonic and adult tissues**

448 Protein lysates from embryos were prepared by homogenising tissue in ten volumes of urea
449 lysis buffer, as described previously (Daxinger et al., 2013). These were quantified using a BCA
450 assay (Thermo Scientific, VIC, Australia), separated on polyacrylamide gels (Bio-Rad, Gladesville,
451 NSW, Australia) and immunoblotted with antibodies directed against Wiz (NBP180586, Novus
452 Biologicals, Littleton, USA) and against γ -Tubulin (T5192, Sigma Aldrich, Castle Hill, AUS). Clarity
453 Western ECL substrate was used for visualisation (Bio-Rad, Gladesville, NSW, Australia).

454 **Affinity purification and mass spectroscopy**

455 **Nuclear extracts were prepared from either pooled embryonic brains (E13.5) or adult**
456 **cerebellum using the Nuclear Complex Co-IP Kit (54001, Active Motif, Carlsbad, CA, USA) as**
457 **described previously (Harten et al., 2015, Isbel et al., 2015). For each sample, 500 µg of**
458 **nuclear lysate was incubated with 2 µg of anti-Wiz antibody (NBP180586, Novus Biologicals,**

459 Littleton, USA) or anti-IgG antibody (sc-2345, Santa Cruz Biotechnology, Dallas, TX, USA)
460 overnight at 4 degrees Celsius. Immunoprecipitations were carried out using bead-conjugated
461 Protein G (Dynabeads 10003D, Invitrogen, Mount Waverley, VIC, Australia) as per the Nuclear
462 Complex Co-IP Kit manufacturer's instructions. Eluted proteins were analysed via mass
463 spectrometry at the La Trobe University Mass Spectrometry Facility and data were analysed
464 as described previously (Harten et al., 2015).

465 **Telemetry monitoring of mice**

466 Age matched male mice, 8-12 weeks old, were individually housed in standard polypropylene
467 cages (14×30×13 cm) with shredded paper as bedding enrichment, ambient temperature was
468 30 ± 1 °C, with a 12 hour light/dark cycle and lights-on at 07:00 hours. Mice received standard
469 rodent chow and water *ad libitum*. After one week of acclimatisation mice were put under
470 general anaesthesia and surgically implanted with a biotelemetry device (Mini-mitter, Bend,
471 USA) into their peritoneal cavities. After one week of recovery, the biotelemetry devices were
472 used to monitor body temperature and locomotor activity. A platform receiver sampled the
473 device's temperature ($\pm 0.1^{\circ}\text{C}$) and location at one minute intervals continuously and this was
474 decoded by a software package VitalView (Starr Life Sciences Corp, Oakmont, USA). Following a
475 12 day period of uninterrupted monitoring, mice were exposed to an elevated plus-maze (EPM)
476 test of anxiety-like behaviour (see below). Immediately following the EPM test, the mice were
477 returned to their home cage for approximately 120 min, then anesthetized using a lethal dose
478 of pentobarbital sodium (Lethabarb, Virbac Pty. Ltd., Milperra, Australia). The telemetry data
479 (activity and body temperature) were collapsed across the 12 days of testing to derive the
480 activity for each hour in the 24 hour cycle. These data were then analysed using a repeated

481 measures analysis of variance (ANOVA) with time as the within subjects factor and genotype
482 and cohort as between subjects factors. A Greenhouse–Geisser correction epsilon (ϵ) was used
483 to correct for potential violation of the sphericity assumption for the within-subject measure.
484 Where significant main or interaction effects were found, post hoc Fisher's LSD tests were used
485 to reveal pairwise differences between groups. As there was no main effect for cohort ($F(1,17)$
486 = 1.313, $p = 0.268$) the data from the two cohorts were combined.

487 **Elevated plus-maze testing**

488 Following the 12 day monitoring period, mice were exposed to the EPM test. Each mouse was
489 placed on the EPM for a duration of 5 min and behaviour was recorded using a digital video
490 camera fixed above and overlooking the apparatus. Position and movement data was analysed
491 using the Ethovision XT behaviour tracking software. The number of entries in the open arms, a
492 common measure of rodent anxiety-like behaviour was analysed using an independent samples
493 t-test.

494 **Light-Dark behaviour testing**

495 Light-dark behaviour testing was carried out on mice aged 2-6 months ($n = 43$) using
496 methodology that was previously described (Balemans et al., 2010). Each animal was housed
497 separately for at least 3 weeks prior to testing. Testing was carried out for a duration of 10
498 minutes and was analysed using an independent samples t-test.

499 **References**

500 ANDERS, S. & HUBER, W. 2010. Differential expression analysis for sequence count data. *Genome Biol*,
501 11, R106 doi:10.1186/gb-2010-11-10-r106.

- 502 ANDERS, S., PYL, P. T. & HUBER, W. 2015. HTSeq--a Python framework to work with high-throughput
503 sequencing data. *Bioinformatics*, 31, 166-9 doi:10.1093/bioinformatics/btu638.
- 504 BALEMANS, M. C. M., HUIBERS, M. M. H., EIKELENBOOM, N. W. D., KUIPERS, A. J., VAN SUMMEREN, R.
505 C. J., PIJPERS, M. M. C. A., TACHIBANA, M., SHINKAI, Y., VAN BOKHOVEN, H. & VAN DER ZEE, C.
506 E. E. M. 2010. Reduced exploration, increased anxiety, and altered social behavior: Autistic-like
507 features of euchromatin histone methyltransferase 1 heterozygous knockout mice. *Behavioural*
508 *Brain Research*, 208, 47-55 doi:DOI 10.1016/j.bbr.2009.11.008.
- 509 BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T. Y., SCHONES, D. E., WANG, Z., WEI, G., CHEPELEV, I. & ZHAO,
510 K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell*, 129, 823-
511 37 doi:10.1016/j.cell.2007.05.009.
- 512 BELL, A. C., WEST, A. G. & FELSENFELD, G. 1999. The Protein CTCF Is Required for the Enhancer Blocking
513 Activity of Vertebrate Insulators. *Cell*, 98, 387-396 doi:[http://dx.doi.org/10.1016/S0092-8674\(00\)81967-4](http://dx.doi.org/10.1016/S0092-8674(00)81967-4).
- 514 BIAN, C., CHEN, Q. & YU, X. 2015. The zinc finger proteins ZNF644 and WIZ regulate the G9a/GLP
515 complex for gene repression. *eLife*, 4 doi:10.7554/eLife.05606.
- 516 BLEWITT, M. E., VICKARYOUS, N. K., HEMLEY, S. J., ASHE, A., BRUXNER, T. J., PREIS, J. I., ARKELL, R. &
517 WHITELAW, E. 2005. An N-ethyl-N-nitrosourea screen for genes involved in variegation in the
518 mouse. *Proc Natl Acad Sci U S A*, 102, 7629-34 doi:10.1073/pnas.0409375102.
- 519 BOYLE, A. L. & WARD, D. C. 1992. Isolation and initial characterization of a large repeat sequence
520 element specific to mouse chromosome 8. *Genomics*, 12, 517-25.
- 521 CHATURVEDI, C.-P., HOSEY, A. M., PALII, C., PEREZ-IRATXETA, C., NAKATANI, Y., RANISH, J. A.,
522 DILWORTH, F. J. & BRAND, M. 2009. Dual role for the methyltransferase G9a in the maintenance
523 of β-globin gene transcription in adult erythroid cells. *P Natl Acad Sci USA*, 106, 18303-18308
524 doi:10.1073/pnas.0906769106.
- 525 CLÉARD, F. & SPIERER, P. 2001. Position-effect variegation in Drosophila: the modifier Su(var)3-7 is a
526 modular DNA-binding protein. *EMBO Reports*, 2, 1095-1100.
- 527 DAXINGER, L., HARTEN, S. K., OEHY, H., EPP, T., ISBEL, L., HUANG, E., WHITELAW, N., APEDAILE, A.,
528 SOROLLA, A., YONG, J., BHARTI, V., SUTTON, J., ASHE, A., PANG, Z., WALLACE, N., GERHARDT, D.
529 J., BLEWITT, M. E., JEDDELOH, J. A. & WHITELAW, E. 2013. An ENU mutagenesis screen identifies
530 novel and known genes involved in epigenetic processes in the mouse. *Genome Biol*, 14, R96
531 doi:10.1186/gb-2013-14-9-r96.
- 532 DUHL, D. M., VRIELING, H., MILLER, K. A., WOLFF, G. L. & BARSH, G. S. 1994. Neomorphic agouti
533 mutations in obese yellow mice. *Nat Genet*, 8, 59-65 doi:10.1038/ng0994-59.
- 534 Fodor, B.D., Shukeir, N., Reuter, G., Jenuwein, T. 2010. Mammalian Su(var) genes in chromatin control.
535 *Annu Rev Cell Dev Biol*, 26, 471-501 doi:10.1146/annurev.cellbio.042308.113225.
- 536 GUO, Y., XU, Q., CANZIO, D., SHOU, J., LI, J., GORKIN, DAVID U., JUNG, I., WU, H., ZHAI, Y., TANG, Y., LU,
537 Y., WU, Y., JIA, Z., LI, W., ZHANG, MICHAEL Q., REN, B., KRAINER, ADRIAN R., MANIATIS, T. & WU,
538 Q. 2015. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter
539 Function. *Cell*, 162, 900-910 doi:10.1016/j.cell.2015.07.038.
- 540 HARTEN, S. K., OEHY, H., BOURKE, L. M., BHARTI, V., ISBEL, L., DAXINGER, L., FAOU, P., ROBERTSON, N.,
541 MATTHEWS, J. M. & WHITELAW, E. 2015. The recently identified modifier of murine metastable
542 epialleles, Rearranged L-Myc Fusion, is involved in maintaining epigenetic marks at CpG island
543 shores and enhancers. *BMC Biol*, 13, 21 doi:10.1186/s12915-015-0128-2.
- 544 HENIKOFF, S. 1990. Position-effect variegation after 60 years. *Trends Genet*, 6, 422-426
545 doi:[http://dx.doi.org/10.1016/0168-9525\(90\)90304-O](http://dx.doi.org/10.1016/0168-9525(90)90304-O).
- 546 HIRAYAMA, T., TARUSAWA, E., YOSHIMURA, Y., GALJART, N. & YAGI, T. 2012. CTCF Is Required for
547 Neural Development and Stochastic Expression of Clustered Pcdh Genes in Neurons. *Cell*
548 *Reports*, 2, 345-357 doi:<http://dx.doi.org/10.1016/j.celrep.2012.06.014>.

- ISBEL, L., SRIVASTAVA, R., OEHY, H., SPURLING, A., DAXINGER, L., PUTHALAKATH, H. & WHITELAW, E. 2015. Trim33 Binds and Silences a Class of Young Endogenous Retroviruses in the Mouse Testis; a Novel Component of the Arms Race between Retrotransposons and the Host Genome. *PLoS Genet*, 11, e1005693 doi:10.1371/journal.pgen.1005693.

JENUWEIN, T., LAIBLE, G., DORN, R. & REUTER, G. 1998. SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci*, 54, 80-93.

KAROLCHIK, D., BAERTSCH, R., DIEKHANS, M., FUREY, T. S., HINRICHES, A., LU, Y. T., ROSKIN, K. M., SCHWARTZ, M., SUGNET, C. W., THOMAS, D. J., WEBER, R. J., HAUSSLER, D. & KENT, W. J. 2003. The UCSC Genome Browser Database. *Nucleic Acids Res*, 31, 51-54 doi:10.1093/nar/gkg129.

KAWAGUCHI, M., TOYAMA, T., KANEKO, R., HIRAYAMA, T., KAWAMURA, Y. & YAGI, T. 2008. Relationship between DNA methylation states and transcription of individual isoforms encoded by the protocadherin-alpha gene cluster. *J Biol Chem*, 283, 12064-75 doi:10.1074/jbc.M709648200.

KLEEFSTRA, T., BRUNNER, H. G., AMIEL, J., OUDAKKER, A. R., NILLESSEN, W. M., MAGEE, A., GENEVIÈVE, D., CORMIER-DAIRE, V., VAN ESCH, H., FRYNS, J.-P., HAMEL, B. C. J., SISTERMANS, E. A., DE VRIES, B. B. A. & VAN BOKHOVEN, H. 2006. Loss-of-Function Mutations in Euchromatin Histone Methyl Transferase 1 (EHMT1) Cause the 9q34 Subtelomeric Deletion Syndrome. *Am J Hum Genet*, 79, 370-377 doi:<http://dx.doi.org/10.1086/505693>.

KOHMURA, N., SENZAKI, K., HAMADA, S., KAI, N., YASUDA, R., WATANABE, M., ISHII, H., YASUDA, M., MISHINA, M. & YAGI, T. 1998. Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. *Neuron*, 20, 1137-51.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*, 9, 357-9 doi:10.1038/nmeth.1923.

LEE, D. Y., NORTHROP, J. P., KUO, M.-H. & STALLCUP, M. R. 2006. Histone H3 Lysine 9 Methyltransferase G9a Is a Transcriptional Coactivator for Nuclear Receptors. *J Biol Chem*, 281, 8476-8485 doi:10.1074/jbc.M511093200.

LIPKIN, S. M., NÄÄR, A. M., KALLA, K. A., SACK, R. A. & ROSENFELD, M. G. 1993. Identification of a novel zinc finger protein binding a conserved element critical for Pit-1-dependent growth hormone gene expression. *Gene Dev*, 7, 1674-1687 doi:10.1101/gad.7.9.1674.

MA, X.-S., CHAO, S.-B., HUANG, X.-J., LIN, F., QIN, L., WANG, X.-G., MENG, T.-G., ZHU, C.-C., SCHATTEN, H., LIU, H.-L., SUN, Q.-Y. 2015. The Dynamics and Regulatory Mechanism of Pronuclear H3k9me2 Asymmetry in Mouse Zygotes. *Scientific Reports*, 5, 17924. Doi:<http://doi.org/10.1038/srep17924>

MACHANICK, P. & BAILEY, T. L. 2011. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics*, 27, 1696-7 doi:10.1093/bioinformatics/btr189.

MATSUMOTO, K., ISHII, N., YOSHIDA, S., SHIOSAKA, S., WANAKA, A. & TOHYAMA, M. 1998. Molecular cloning and distinct developmental expression pattern of spliced forms of a novel zinc finger gene wiz in the mouse cerebellum. *Brain Res Mol Brain Res*, 61, 179-89.

MORGAN, H. D., SUTHERLAND, H. G., MARTIN, D. I. & WHITELAW, E. 1999. Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet*, 23, 314-8 doi:10.1038/15490.

MULLIGAN, P., WESTBROOK, T. F., OTTINGER, M., PAVLOVA, N., CHANG, B., MACIA, E., SHI, Y. J., BARRETINA, J., LIU, J., HOWLEY, P. M., ELLEDGE, S. J. & SHI, Y. 2008. CDYL bridges REST and histone methyltransferases for gene repression and suppression of cellular transformation. *Mol Cell*, 32, 718-26 doi:10.1016/j.molcel.2008.10.025.

OH, S.-T., KIM, K.-B., CHAE, Y.-C., KANG, J.-Y., HAHN, Y. & SEO, S.-B. 2014. H3K9 histone methyltransferase G9a-mediated transcriptional activation of p21. *FEBS Letters*, 588, 685-691 doi:<http://dx.doi.org/10.1016/j.febslet.2014.01.039>.

ONG, C.-T. & CORCES, V. G. 2014. CTCF: an architectural protein bridging genome topology and function. *Nat Rev Genet*, 15, 234-246 doi:10.1038/nrg3663.

- 598 PETERSEN, C. C. H. 2007. The Functional Organization of the Barrel Cortex. *Neuron*, 56, 339-355
599 doi:<http://dx.doi.org/10.1016/j.neuron.2007.09.017>.
- 600 PUGACHEVA, E. M., RIVERO-HINOJOSA, S., ESPINOZA, C. A., MENDEZ-CATALA, C. F., KANG, S., SUZUKI, T.,
601 KOSAKA-SUZUKI, N., ROBINSON, S., NAGARAJAN, V., YE, Z., BOUKABA, A., RASKO, J. E.,
602 STRUNNIKOV, A. V., LOUKINOV, D., REN, B. & LOBANENKOV, V. V. 2015. Comparative analyses of
603 CTCF and BORIS occupancies uncover two distinct classes of CTCF binding genomic regions.
604 *Genome Biol*, 16, 161 doi:[10.1186/s13059-015-0736-8](https://doi.org/10.1186/s13059-015-0736-8).
- 605 PURCELL, D. J., JEONG, K. W., BITTENCOURT, D., GERKE, D. S. & STALLCUP, M. R. 2011. A Distinct
606 Mechanism for Coactivator versus Corepressor Function by Histone Methyltransferase G9a in
607 Transcriptional Regulation. *J Biol Chem*, 286, 41963-41971 doi:[10.1074/jbc.M111.298463](https://doi.org/10.1074/jbc.M111.298463).
- 608 RAKYAN, V. K., CHONG, S., CHAMP, M. E., CUTHBERT, P. C., MORGAN, H. D., LUU, K. V. & WHITELAW, E.
609 2003. Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs
610 after maternal and paternal transmission. *Proc Natl Acad Sci U S A*, 100, 2538-43
611 doi:[10.1073/pnas.0436776100](https://doi.org/10.1073/pnas.0436776100).
- 612 RATHERT, P., DHAYALAN, A., MURAKAMI, M., ZHANG, X., TAMAS, R., JURKOWSKA, R., KOMATSU, Y.,
613 SHINKAI, Y., CHENG, X., JELTSCH, A. 2008. Protein lysine methyltransferase G9a acts on non-histone
614 targets. *Nat Chem Biol*, 4, 344-346 doi:[10.1038/nchembio.88](https://doi.org/10.1038/nchembio.88)
- 615 SCHUH, R., AICHER, W., GAUL, U., CÔTE, S., PREISS, A., MAIER, D., SEIFERT, E., NAUBER, U., SCHRÖDER,
616 C., KEMLER, R. & JÄCKLE, H. 1986. A conserved family of nuclear proteins containing structural elements
617 of the finger protein encoded by Krüppel, a Drosophila segmentation gene. *Cell*, 47, 1025-1032
618 doi:[http://dx.doi.org/10.1016/0092-8674\(86\)90817-2](https://doi.org/10.1016/0092-8674(86)90817-2).
- 619 SHEN, Y., YUE, F., MCCLEARY, D. F., YE, Z., EDSALL, L., KUAN, S., WAGNER, U., DIXON, J., LEE, L.,
620 LOBANENKOV, V. V. & REN, B. 2012. A map of the cis-regulatory sequences in the mouse
621 genome. *Nature*, 488, 116-20 doi:[10.1038/nature11243](https://doi.org/10.1038/nature11243).
- 622 SIMON, J. M., PARKER, J. S., LIU, F., ROTHBART, S. B., AIT-SI-ALI, S., STRAHL, B. D., JIN, J., DAVIS, I. J.,
623 MOSLEY, A. L. & PATTENDEN, S. G. 2015. A Role for Widely Interspaced Zinc Finger (WIZ) in
624 Retention of the G9a Methyltransferase on Chromatin. *J Biol Chem*, 290, 26088-102
625 doi:[10.1074/jbc.M115.654459](https://doi.org/10.1074/jbc.M115.654459).
- 626 TACHIBANA, M., SUGIMOTO, K., FUKUSHIMA, T. & SHINKAI, Y. 2001. Set domain-containing protein,
627 G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and
628 specific selectivity to lysines 9 and 27 of histone H3. *J Biol Chem*, 276, 25309-17
629 doi:[10.1074/jbc.M101914200](https://doi.org/10.1074/jbc.M101914200).
- 630 TACHIBANA, M., SUGIMOTO, K., NOZAKI, M., UEDA, J., OHTA, T., OHKI, M., FUKUDA, M., TAKEDA, N.,
631 NIIDA, H., KATO, H. & SHINKAI, Y. 2002. G9a histone methyltransferase plays a dominant role in
632 euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes*
633 *Dev*, 16, 1779-91 doi:[10.1101/gad.989402](https://doi.org/10.1101/gad.989402).
- 634 TACHIBANA, M., UEDA, J., FUKUDA, M., TAKEDA, N., OHTA, T., IWANARI, H., SAKIHAMA, T., KODAMA, T.,
635 HAMAKUBO, T. & SHINKAI, Y. 2005. Histone methyltransferases G9a and GLP form heteromeric
636 complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev*, 19, 815-26
637 doi:[10.1101/gad.1284005](https://doi.org/10.1101/gad.1284005).
- 638 TOYODA, S., KAWAGUCHI, M., KOBAYASHI, T., TARUSAWA, E., TOYAMA, T., OKANO, M., ODA, M.,
639 NAKAUCHI, H., YOSHIMURA, Y., SANBO, M., HIRABAYASHI, M., HIRAYAMA, T., HIRABAYASHI, T.
640 & YAGI, T. 2014. Developmental epigenetic modification regulates stochastic expression of
641 clustered protocadherin genes, generating single neuron diversity. *Neuron*, 82, 94-108
642 doi:[10.1016/j.neuron.2014.02.005](https://doi.org/10.1016/j.neuron.2014.02.005).
- 643 TRAPNELL, C., PACTER, L. & SALZBERG, S. L. 2009. TopHat: discovering splice junctions with RNA-Seq.
644 *Bioinformatics*, 25, 1105-11 doi:[10.1093/bioinformatics/btp120](https://doi.org/10.1093/bioinformatics/btp120).

- 645 UEDA, J., TACHIBANA, M., IKURA, T. & SHINKAI, Y. 2006. Zinc finger protein Wiz links G9a/GLP histone
646 methyltransferases to the co-repressor molecule CtBP. *J Biol Chem*, 281, 20120-8
647 doi:10.1074/jbc.M603087200.
- 648 WALF, A. A. & FRYE, C. A. 2007. The use of the elevated plus maze as an assay of anxiety-related
649 behavior in rodents. *Nat. Protocols*, 2, 322-328.
- 650 WEN, B., WU, H., LOH, Y.-H., BRIEM, E., DALEY, G. & FEINBERG, A. 2012. Euchromatin islands in large
651 heterochromatin domains are enriched for CTCF binding and differentially DNA-methylated
652 regions. *BMC Genomics*, 13, 566.
- 653 WEN, B., WU, H., SHINKAI, Y., IRIZARRY, R. A. & FEINBERG, A. P. 2009. Large histone H3 lysine 9
654 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet*,
655 41, 246-50 doi:10.1038/ng.297.
- 656 WILLEMSSEN, M. H., VULTO-VAN SILFHOUT, A. T., NILLESEN, W. M., WISSINK-LINDHOUT, W. M., VAN
657 BOKHOVEN, H., PHILIP, N., BERRY-KRAVIS, E. M., KINI, U., VAN RAVENSWAAIJ-ARTS, C. M. A.,
658 DELLE CHIAIE, B., INNES, A. M. M., HOUGE, G., KOSONEN, T., CREMER, K., FANNEMEL, M.,
659 STRAY-PEDERSEN, A., REARDON, W., IGNATIUS, J., LACHLAN, K., MIRCHER, C., HELDERMAN VAN
660 DEN ENDEN, P. T. J. M., MASTEBROEK, M., COHN-HOKKE, P. E., YNTEMA, H. G., DRUNAT, S. &
661 KLEEFSTRA, T. 2011. Update on Kleefstra Syndrome. *Mol Syndromol*, 2, 202-212.
- 662 WU, Q. & MANIATIS, T. 1999. A striking organization of a large family of human neural cadherin-like cell
663 adhesion genes. *Cell*, 97, 779-90.
- 664 WU, Q., ZHANG, T., CHENG, J. F., KIM, Y., GRIMWOOD, J., SCHMUTZ, J., DICKSON, M., NOONAN, J. P.,
665 ZHANG, M. Q., MYERS, R. M. & MANIATIS, T. 2001. Comparative DNA sequence analysis of
666 mouse and human protocadherin gene clusters. *Genome Res*, 11, 389-404
667 doi:10.1101/gr.167301.
- 668 XIAO, T., WALLACE, J. & FELSENFELD, G. 2011. Specific Sites in the C Terminus of CTCF Interact with the
669 SA2 Subunit of the Cohesin Complex and Are Required for Cohesin-Dependent Insulation
670 Activity. *Mol Cell Biol*, 31, 2174-2183 doi:10.1128/mcb.05093-11.
- 671 YE, T., KREBS, A. R., CHOUKRALAH, M. A., KEIME, C., PLEWNIAK, F., DAVIDSON, I. & TORA, L. 2011.
672 seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res*, 39, e35
673 doi:10.1093/nar/gkq1287.
- 674 YOKOTA, S., HIRAYAMA, T., HIRANO, K., KANEKO, R., TOYODA, S., KAWAMURA, Y., HIRABAYASHI, M.,
675 HIRABAYASHI, T. & YAGI, T. 2011. Identification of the Cluster Control Region for the
676 Protocadherin- β Genes Located beyond the Protocadherin- γ Cluster. *J Biol Chem*, 286, 31885-
677 31895 doi:10.1074/jbc.M111.245605.
- 678 YUAN, X., FENG, W., IMHOF, A., GRUMMT, I. & ZHOU, Y. 2007. Activation of RNA Polymerase I
679 Transcription by Cockayne Syndrome Group B Protein and Histone Methyltransferase G9a. *Mol
680 Cell*, 27, 585-595 doi:<http://dx.doi.org/10.1016/j.molcel.2007.06.021>.
- 681 YUSUFZAI, T. M., TAGAMI, H., NAKATANI, Y. & FELSENFELD, G. 2004. CTCF Tethers an Insulator to
682 Subnuclear Sites, Suggesting Shared Insulator Mechanisms across Species. *Mol Cell*, 13, 291-298
683 doi:[http://dx.doi.org/10.1016/S1097-2765\(04\)00029-2](http://dx.doi.org/10.1016/S1097-2765(04)00029-2).
- 684 ZHANG, Y., LIU, T., MEYER, C. A., EECKHOUTE, J., JOHNSON, D. S., BERNSTEIN, B. E., NUSBAUM, C.,
685 MYERS, R. M., BROWN, M., LI, W. & LIU, X. S. 2008. Model-based analysis of ChIP-Seq (MACS).
686 *Genome Biol*, 9, R137 doi:10.1186/gb-2008-9-9-r137.

687

688

689 **Figure 1. Wiz haploinsufficiency increases silencing at the A^{vy} locus.** Pedigree charts (left) are
690 shown for generating F1 mice heterozygous for the A^{vy} allele and either wildtype or
691 heterozygous for $Wiz^{MommeD30}$. $Wiz^{MommeD30/+}$ mice were crossed to pseudoagouti A^{vy}/a mice,
692 numbers of F1 mice in each cohort are indicated. The proportions of coat colours for F1 mice
693 are shown (right) from a cross with either a $Wiz^{MommeD30/+}$ sire (A) or dam (B). Chi-squared tests
694 were carried out to determine significance, * p-value < 0.05, ** p-value < 0.005.

695 **Figure 2. Wiz binds across the genome and at promoter elements.** (A) ChIP-seq was performed
696 for Wiz in adult male cerebellum, a screenshot is shown of a random Wiz enriched site (*
697 indicated) with the two ChIP-seq replicates and input, from the ~40,000 significantly enriched
698 peaks. Encode data for CTCF and H3K4me3 is also included. (B) Enrichment for Wiz in two
699 cerebellum samples (Cb1 and Cb2) is shown by ChIP-qPCR, with primers located in regions not
700 enriched for Wiz in ChIP-seq data and primers flanking the * indicated ChIP-seq peak from (A).
701 Enrichment is represented as binding events detected per 1000 cells and are generated by
702 running samples in parallel with known amounts of genomic DNA. Error bars indicate S.D. from
703 3 technical replicates. (C) The percentage is shown of Wiz peaks that overlap with promoter (up
704 to 2 kb from a TSS) and genic sequence of genes that were classified as active or silenced,
705 according to RNA-seq data mapped to Ensembl genes annotations. (D) Wiz ChIP-seq (and Input)
706 occupancy over all Ensembl gene bodies is shown as deep sequencing read density along the
707 transcription unit, including 2 kb up and downstream of the transcriptional start and stop site.
708 Genes were separated into either active or silenced transcriptional states, as in C.

709 **Figure 2-figure supplement1. Anti-Wiz antibody Western blotting.** (A) Western blotting for Wiz
710 using an NBP180586 anti-Wiz antibody is shown for total protein lysates from E12.5 $Wiz^{+/+}$ and
711 $Wiz^{MommeD30/MommeD30}$ embryo head. Three bands at ~100kD represent Wiz isoforms, consistent
712 with UCSC gene splice variants, in $Wiz^{+/+}$ sample and are absent in $Wiz^{MommeD30/MommeD30}$
713 sample. Western blotting with an anti- γ -Tubulin antibody is shown as a loading control. (B)
714 Anti-Wiz antibody staining in adult cerebellum, showing the short (~100 – 120 KDa) isoforms
715 present in embryo head protein extract, as well as a long (~160 KDa) isoform previously
716 reported to be expressed (Matsumoto et al., 1998).

717 **Figure 2-figure supplement2. Anti-Wiz antibody Co-immunoprecipitation.** (A) Nuclear lysates
718 from either pooled E13.5 embryonic brain or 2 adult cerebellum replicates were subjected to
719 antibody-based affinity purification and mass spectroscopy. The number of peptides are shown
720 for each protein identified in both embryonic (values for 1 sample) and adult brain (values for
721 each of the two samples) anti-Wiz antibody IP samples and absent in anti-IgG antibody controls
722 run for each sample, with greater than 5 peptides in each IP. Protein are ranked by the semi-
723 quantitative number peptides in embryonic brain where Wiz expression is higher. The
724 previously identified Wiz-Zfp644-EHMT1-EHMT2 complex is shaded.

725 **Figure 3. Wiz binding consensus shows a high degree of overlap with that for CTCF.** (A) The six
726 most enriched motifs in Wiz ChIP-seq peaks ($n = 39,781$) and their distribution inside peaks are
727 shown. (B) The most significant of these matches the CTCF-binding site consensus sequence
728 (JASPAR CORE database - MA0139.1). (C) Read density for Wiz and the Encode CTCF and
729 H3K4me3 ChIP-seq datasets were calculated across a 4 kb region centred on the ~40,000 Wiz
730 ChIP-seq peaks. Input sequencing from the Wiz ChIP-seq experiment is also shown. Loci are
731 clustered by similarity of read density and datasets were normalized to the smallest sized
732 library. (D) Venn diagram showing the overlap for Wiz, CTCF and H3K4me3 ChIP-seq peaks. An
733 overlap was defined as at least one base of sequence occupied by a significantly enriched peak
734 ($p \leq 1 \times 10^{-20}$) from two datasets.

735 **Figure 3-figure supplement1. Remapping publically available Wiz ChIP-seq data.** (B) Shown is
736 an UCSC genome browser screenshot of the *SENP5* gene focussing on the consensus binding
737 site (CATTCCATTCCATT motif) reported in Bian et al., 2015. The consensus binding site is
738 indicated. Strong enrichment is seen for the Wiz ChIP-seq data, shown as a continuous read-
739 depth score, at the CATTCCATTCCATT motif, though this occurs over a HSATII satellite element
740 (shown in the repeat masker track) and is likely to be a mapping artefact.

741 **Figure 4. Differential gene expression in male and female *Wiz*^{MommeD30/+} E13.5 brains.** (A)
742 Volcano plots show the log2 fold-change (x-axis) in the average expression of genes in
743 *Wiz*^{MommeD30/+} E13.5 brains compared to *Wiz*^{+/+} E13.5 brains. Shown is data for males (left, $n = 3$
744 per genotype) and females (right, $n = 2$ per genotype). Significance is shown on a $-\log_{10}$ scale,
745 those genes indicated in red are significantly differentially regulated with an adjusted p-value <
746 0.05. (B) The overlap in differentially expressed genes is shown for the male and female
747 embryonic brains, the majority of transcripts differentially expressed represent a common set
748 between the sexes, listed in alphabetical order in (C), and all of these decrease in expression.

749 **Figure 4-figure supplement 1. Location of differentially expressed genes in E13.5**
750 ***Wiz*^{MommeD30/+} brains.** The locations of down regulated genes differentially expressed in E13.5
751 brains of *Wiz*^{MommeD30/+} mice (combined from male and female datasets) are shown on mouse
752 chromosomes. The *Wiz* gene is not shown.

753 **Figure 5. Differential gene expression in *Wiz*^{MommeD30/+} adult cerebellum.** (A) The log2 fold-
754 change (x-axis) in the average expression of genes are shown for *Wiz*^{MommeD30/+} and *Wiz*^{+/+}
755 cerebellum. Three male biological replicates were used per genotype. Significance is shown on
756 a $-\log_{10}$ scale, those genes indicated in red are significantly differentially regulated. (B) A list of
757 significantly differentially regulated genes in *Wiz*^{MommeD30/+} cerebellum that were also
758 deregulated in E13.5 *Wiz*^{MommeD30/+} brains.

759

760 **Figure 6. Wiz binding at the protocadherin β locus in adult cerebellum.** (A) Genome browser
761 screenshots of the protocadherin β locus and the HS5-1bL enhancer (*) indicated) located
762 ~400Kb downstream. Shown are Wiz ChIP-seq data generated in adult cerebellum as the read
763 depth along the locus. Wiz binds promoters of the Pcdhb locus (left) and is enriched more at
764 the downstream HS5-1bL enhancer (right), the relative binding of Wiz is indicated by the
765 altered scale bar. The location of the Pcdhb genes, are indicated. (B) ChIP-qPCR was performed
766 using the anti-Wiz antibody for two negative Wiz-enrichment sites, with primers located in a
767 gene desert and at *Pcdhb1*, which was not enriched in ChIP-seq, and six positive Wiz-
768 enrichment sites, including the promoters of four Pcdhb genes, the promoter of a Pcdhg gene
769 and the downstream HS5-1bL enhancer. Enrichment is represented as binding events detected
770 per 1000 cells and are generated by running samples in parallel with known amounts of
771 genomic DNA. Error bars are the SEM from 3 biological replicates. (C) Shown is Wiz ChIP-seq
772 read density in adult cerebellum as a heatplot for the 29 genes that decreased and (D) the 53
773 genes that increased, in expression in the cerebellum of *Wiz*^{MommeD30/+} mice. Also shown is CTCF
774 ChIP-seq read density. ChIP-seq density was calculated for the 1Kb surrounding the TSS of each
775 gene and genes are ranked from lowest to highest fold change in each case. Datasets were
776 normalized to the size of smallest library.

777 **Figure 7. *Wiz*^{MommeD30/+} mice show decreased activity and increased anxiety-like behaviour.** (A)
778 Line graph illustrating the average 24 hour activity profile from *Wiz*^{+/+} and *Wiz*^{MommeD30/+} mice,
779 collapsed from 12 days of testing. The pattern of activity is broadly similar between the
780 genotypes, though *Wiz*^{MommeD30/+} mice show decreased activity in the first two hours and the
781 fifth hour of the active phase of the light/dark cycle. Grey box indicates the dark phase of the
782 light/dark cycle. (B) Bar graph showing decreased open arm entries among *Wiz*^{MommeD30/+} mice
783 compared with *Wiz*^{+/+} mice, after 5 minute intervals on the EPM test. (C) Heat map of
784 locomotor activity on the EPM from a representative *Wiz*^{+/+} mouse (top) and a representative
785 *Wiz* *Wiz*^{MommeD30/+} mouse (bottom). Broken lines indicate the location of the closed arms of the
786 EPM. (D) Percentage of time *Wiz*^{+/+} (n= 21) and *Wiz*^{MommeD30/+} (n=22) mice spent in the light
787 portion of a light-dark test. (E) The number of vertical counts (rearing) *Wiz*^{+/+} and *Wiz*^{MommeD30/+}
788 mice showed in the light-dark test. Error bars represent \pm SEM, *P-value <0.05, **P-value
789 0.005, ***P-value <0.0005.

790 **Figure 7-figure supplement1. No difference in temperature in *Wiz*^{MommeD30/+} mice.** Telemetry
791 data from across 3 days were collapsed into hourly intervals and the average temperature of
792 *Wiz*^{+/+} (black data points) and *Wiz*^{MommeD30/+} (red data points) mice is shown. Lights on is
793 indicated by light background and lights off is indicated by grey background. Numbers of mice
794 for each cohort are indicated. Error bars are the SEM.

795

796

797 **Supplementary file 1. Wiz genome wide binding sites.** Wiz ChIP-seq binding peaks (mm10
798 genome build) identified in adult wildtype cerebellum (n = 2 Wiz ChIP-seq biological replicates,
799 1 input sample).

800 **Supplementary file 2. Genes significantly differentially expressed in *Wiz*^{MommeD30/+} E13.5 male
801 brain.** The average read count from both *Wiz*^{+/+} and *Wiz*^{MommeD30/+} animals (n = 3 biological
802 replicates per genotype) are shown, with the fold-change (*Wiz*^{MommeD30/+}/*Wiz*^{+/+}) and adjusted
803 significance value for each gene predicted to be differentially expressed. Read counts are
804 normalized for library size.

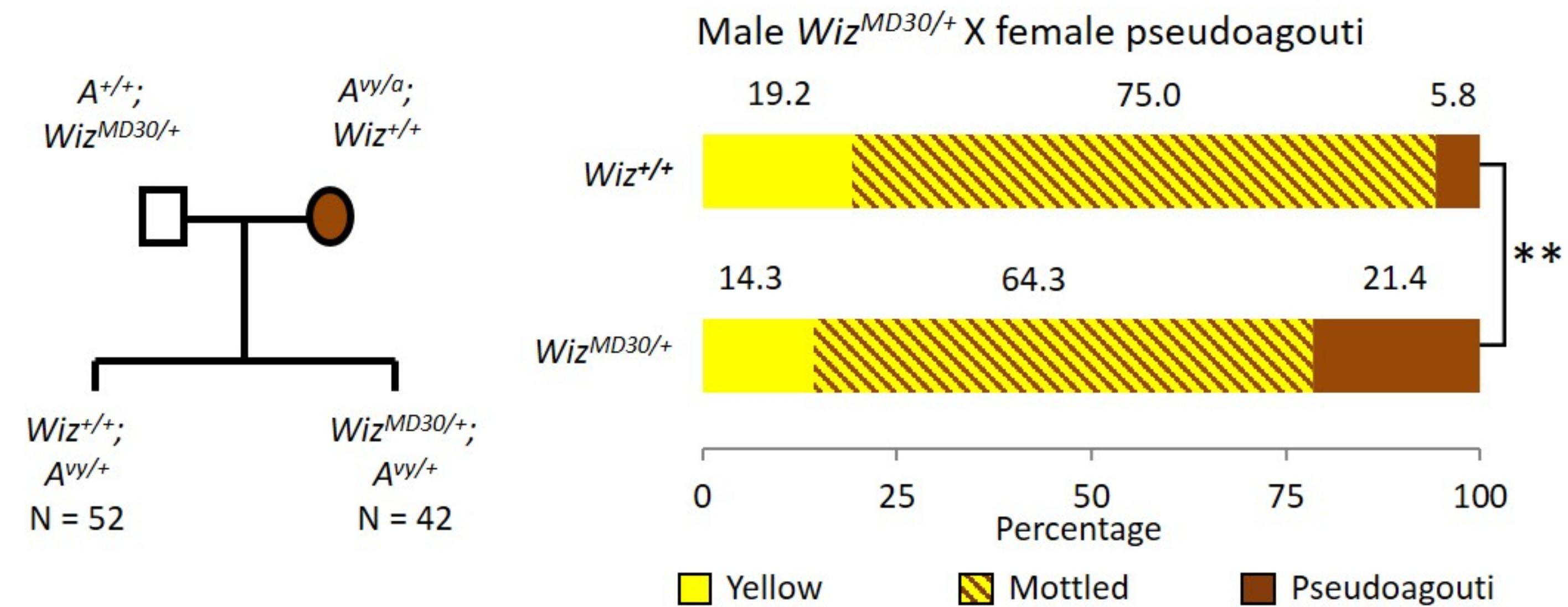
805 **Supplementary file 3. Genes significantly differentially expressed in *Wiz*^{MommeD30/+} E13.5
806 female brain.** The average read count from both *Wiz*^{+/+} and *Wiz*^{MommeD30/+} animals (n = 2
807 biological replicates per genotype) are shown, with the fold-change (*Wiz*^{MommeD30/+}/*Wiz*^{+/+}) and adjusted
808 significance value for each gene predicted to be differentially expressed. Read counts are
809 normalized for library size.

810 **Supplementary file 4. Genes significantly differentially expressed in *Wiz*^{MommeD30/+} cerebellum.**
811 The average read count from both *Wiz*^{+/+} and *Wiz*^{MommeD30/+} animals (n = 3 biological replicates
812 per genotype) are shown, with the fold-change (*Wiz*^{MommeD30/+} / *Wiz*^{+/+}) and adjusted
813 significance value for each gene predicted to be differentially expressed. Cells highlighted
814 orange are protocadherin β cluster genes and cells highlighted green are cadherin 11-like
815 cluster genes. Read counts are normalized for library size.

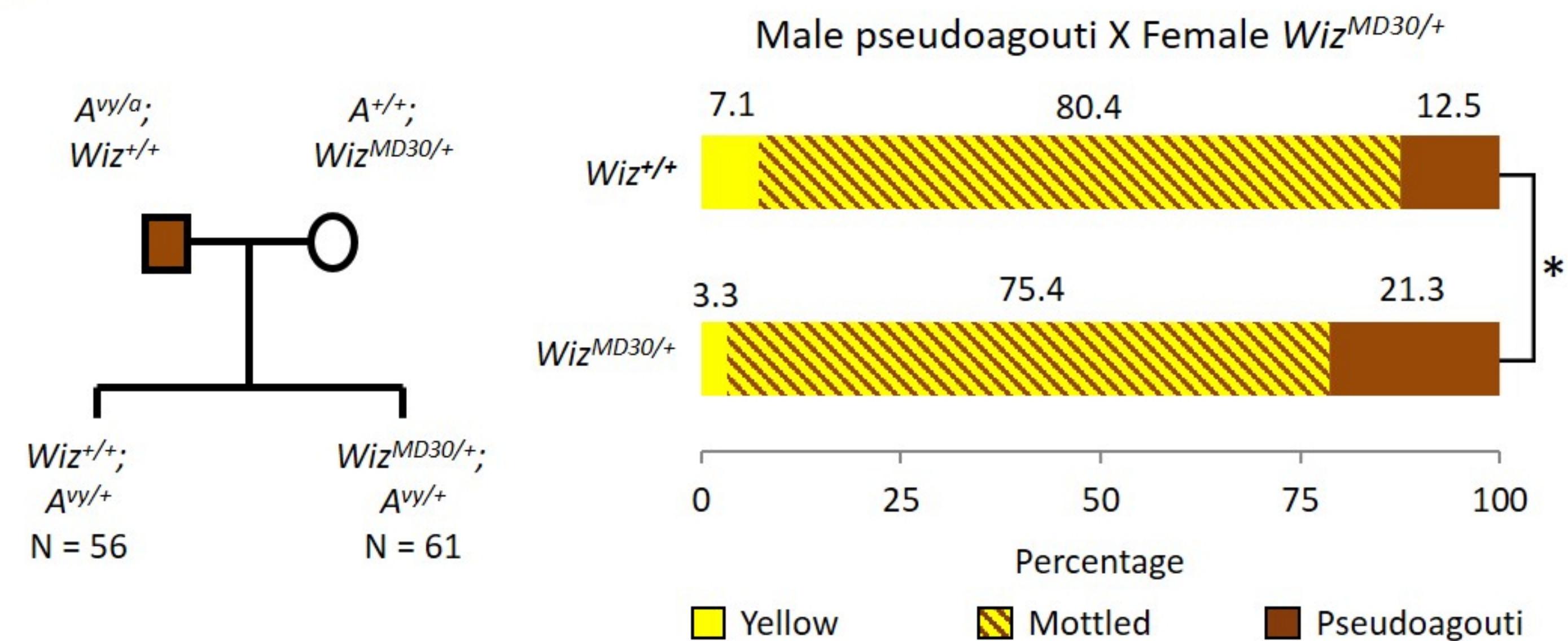
816 **Supplementary file 5. Oligonucleotides used in the study.**

817

A



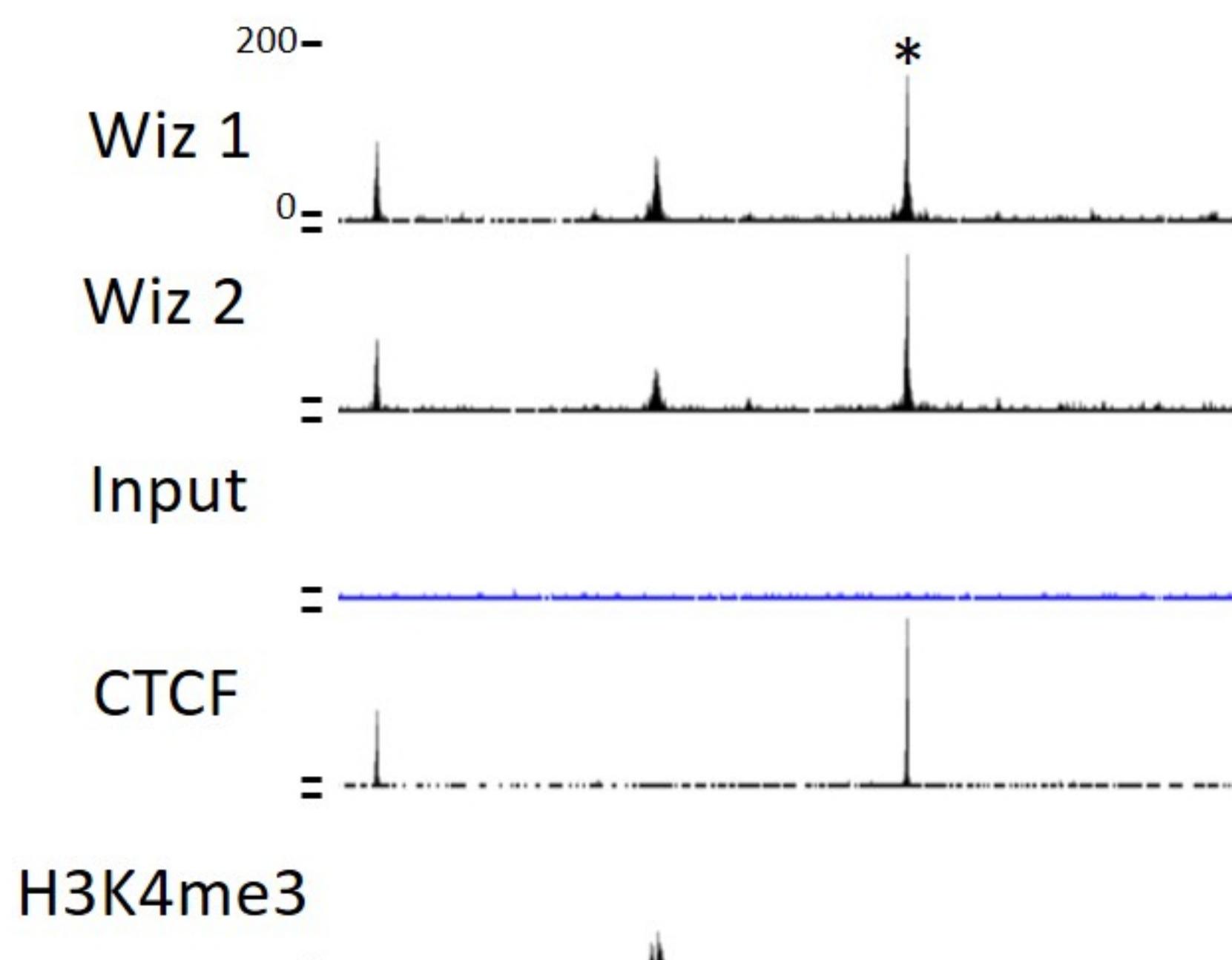
B



A

Genome browser screenshot Wiz ChIP-seq

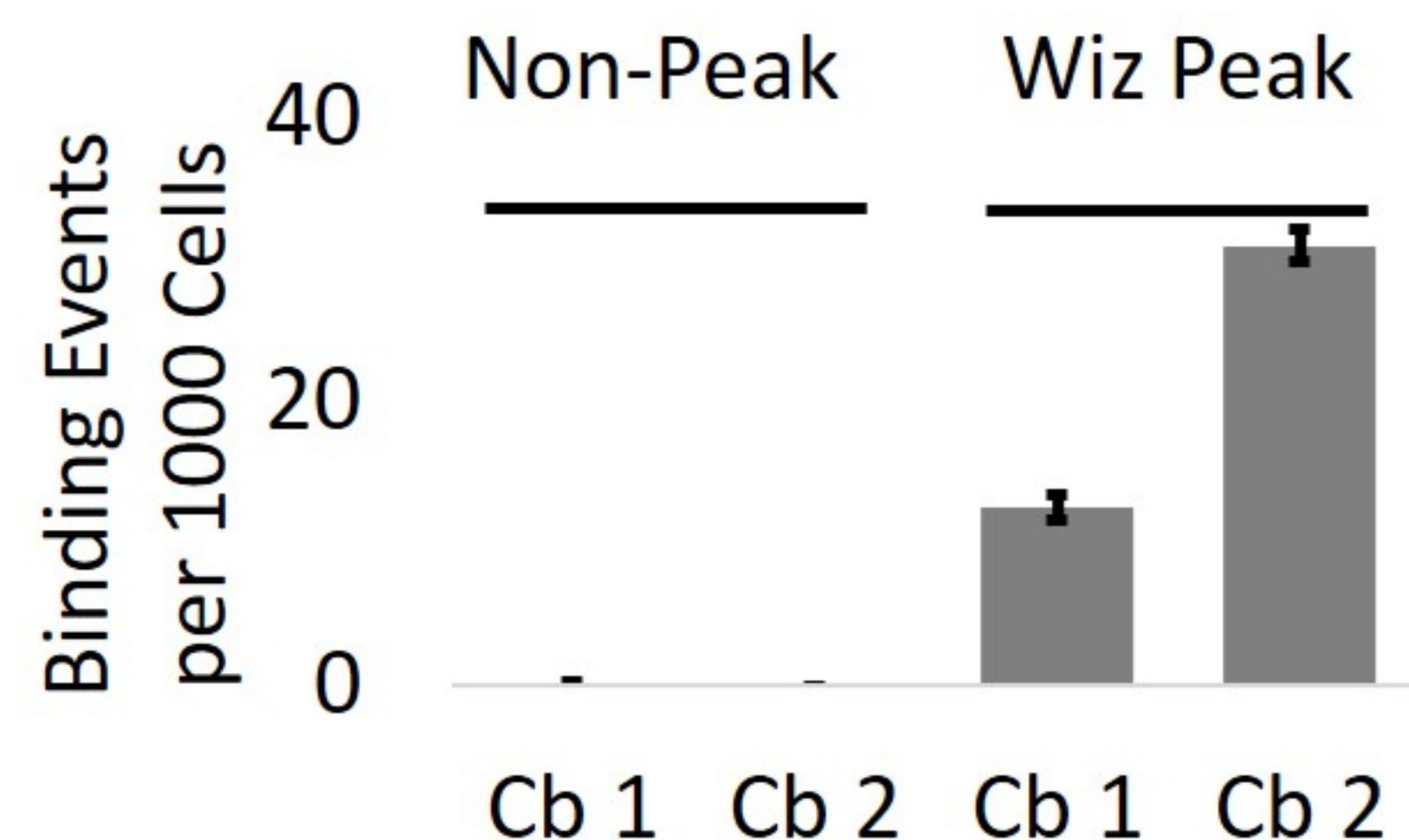
Chr1:39338379-39423607



B

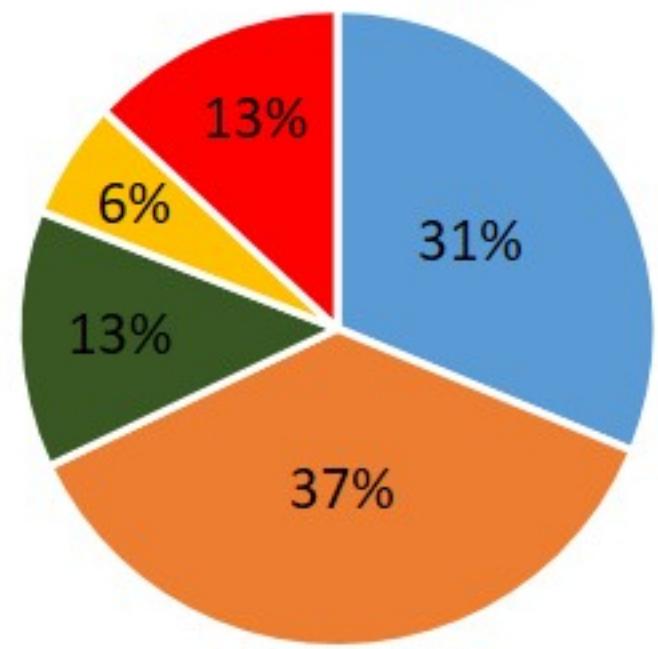
Wiz ChIP-qPCR

Chr1:39338379-39423607



C

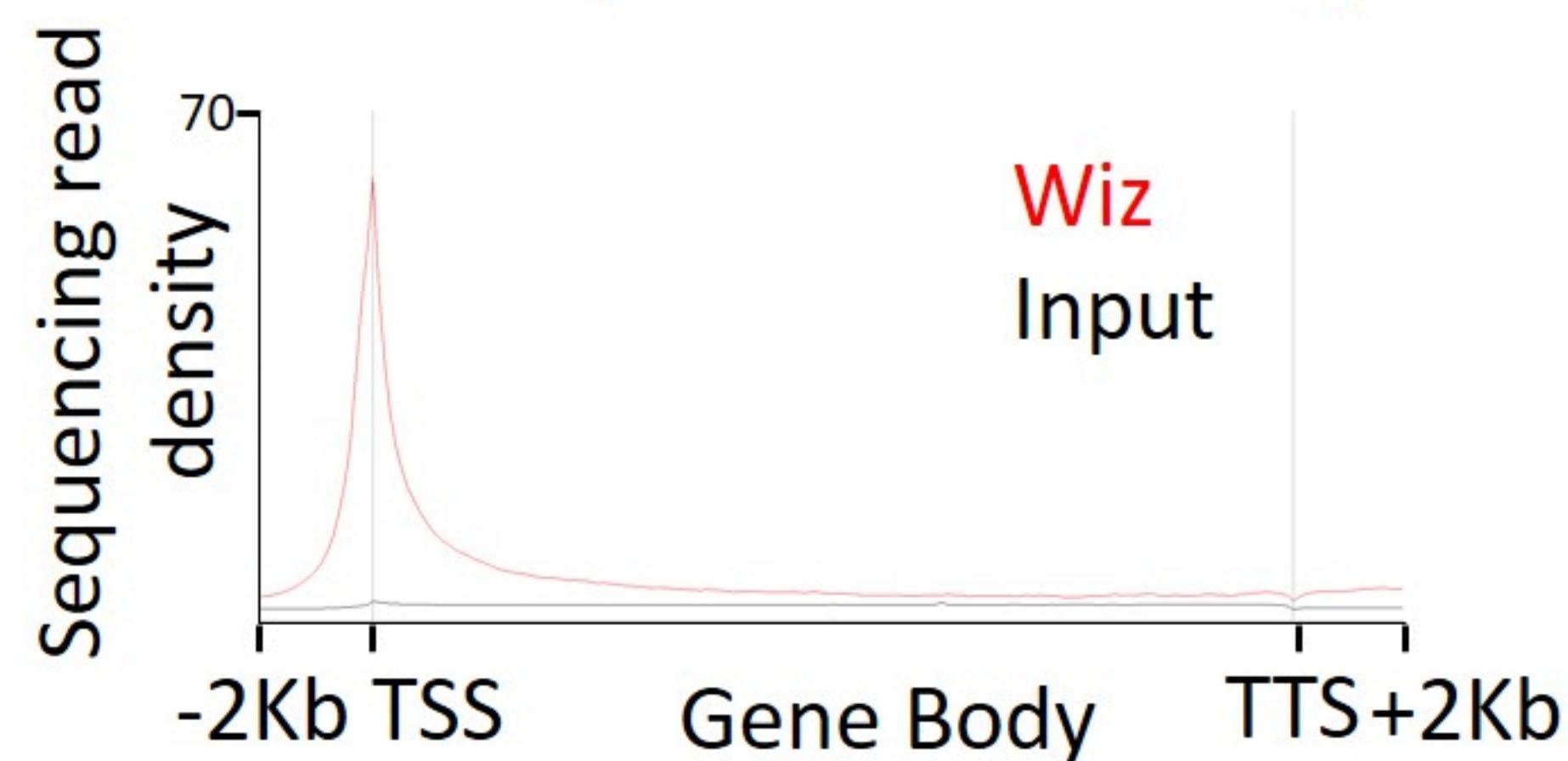
Wiz binding peak locations ($n = 39781$) in genes



- Active gene promoter
- Active gene body
- Silent gene promoter
- Silent gene body
- Intragenic

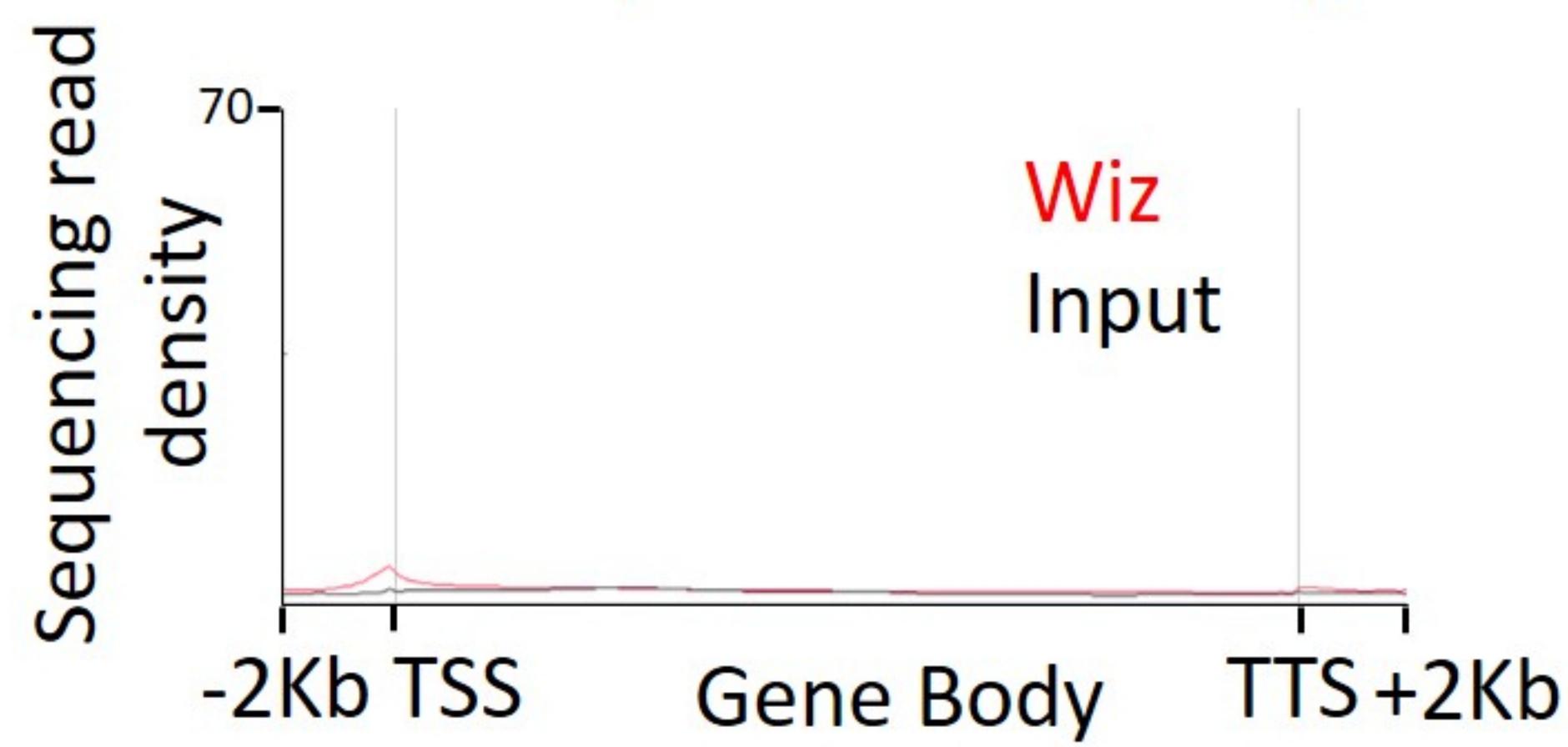
D

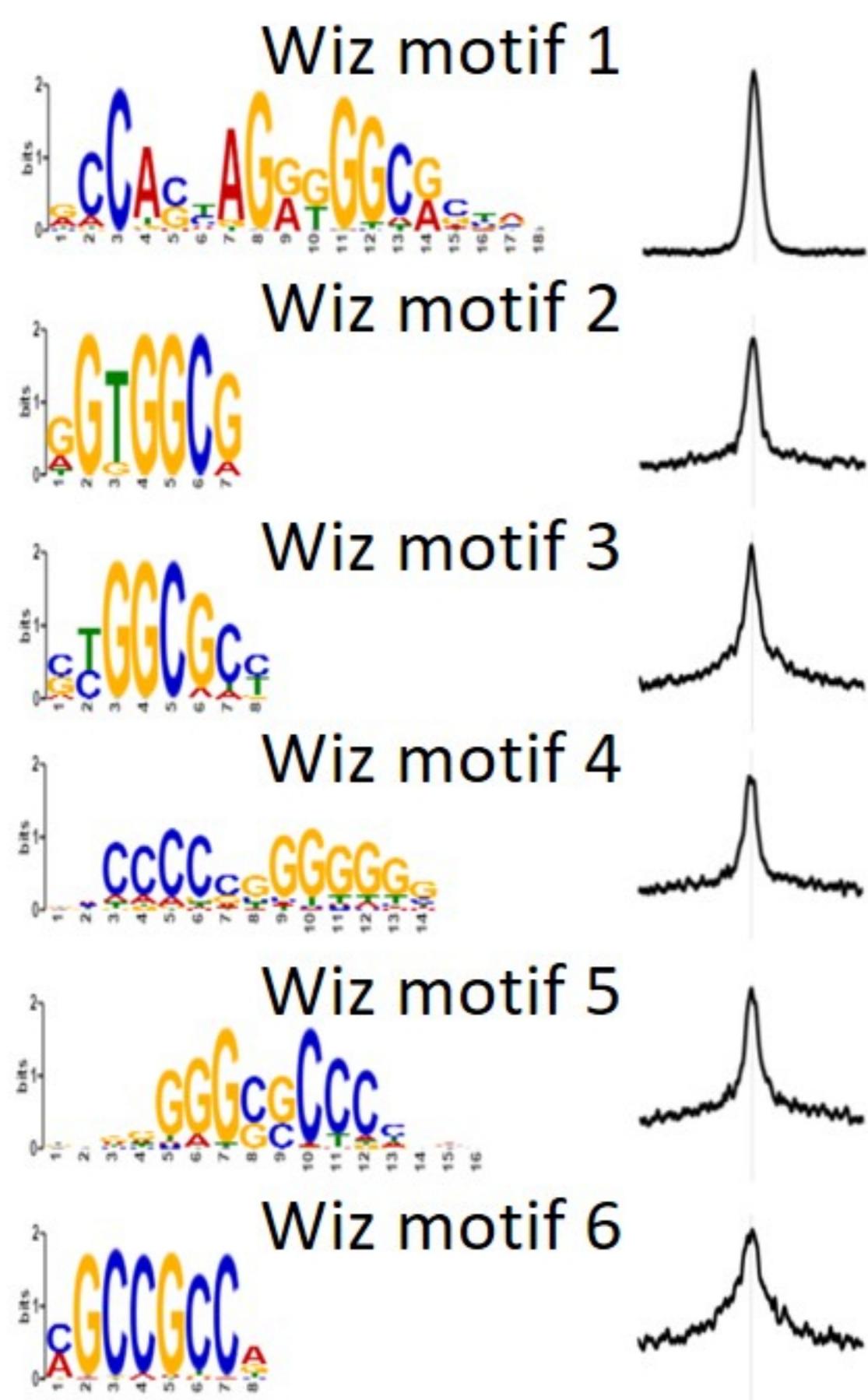
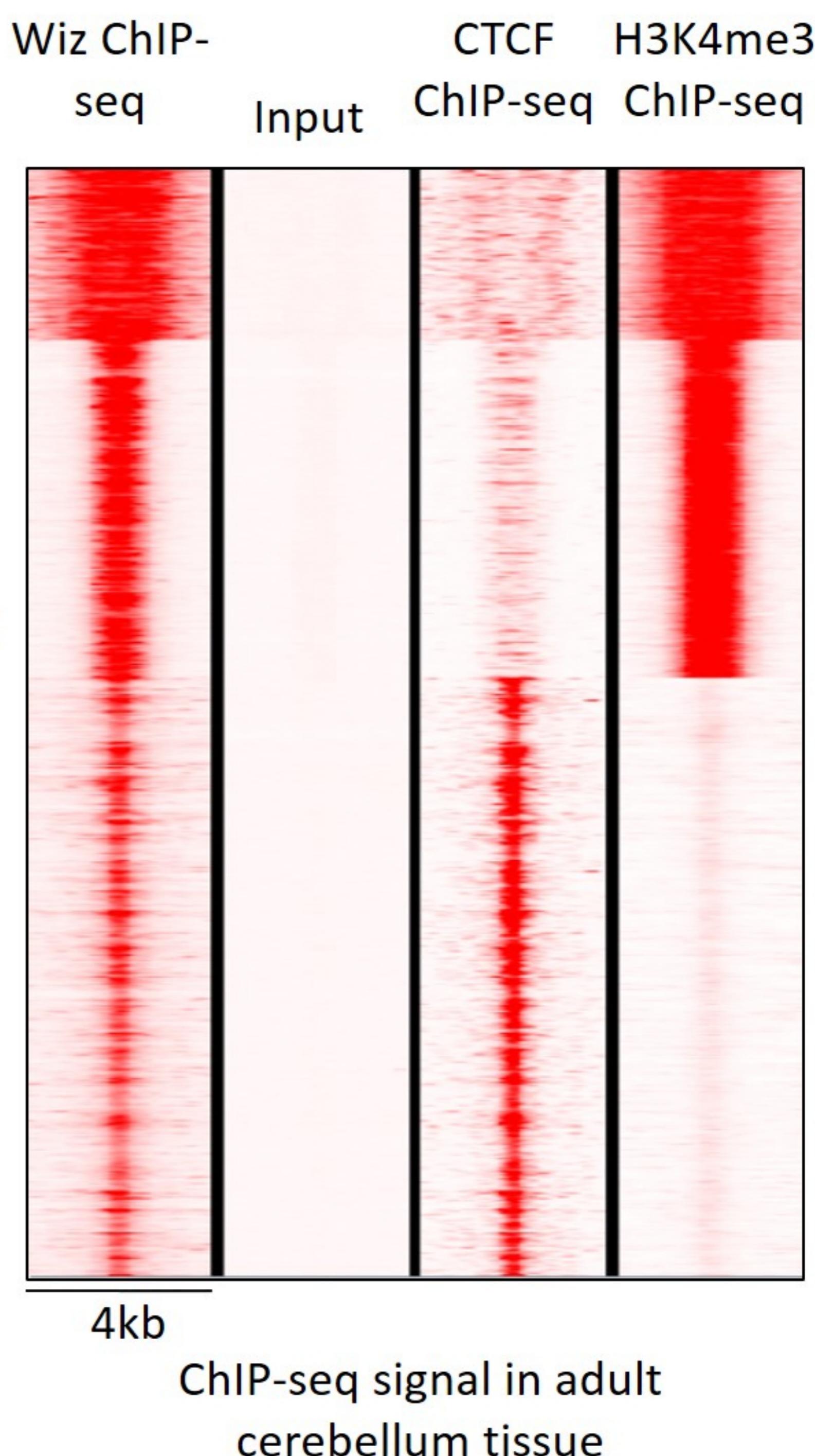
Wiz ChIP-seq reads over active genes



E

Wiz ChIP-seq reads over silent genes

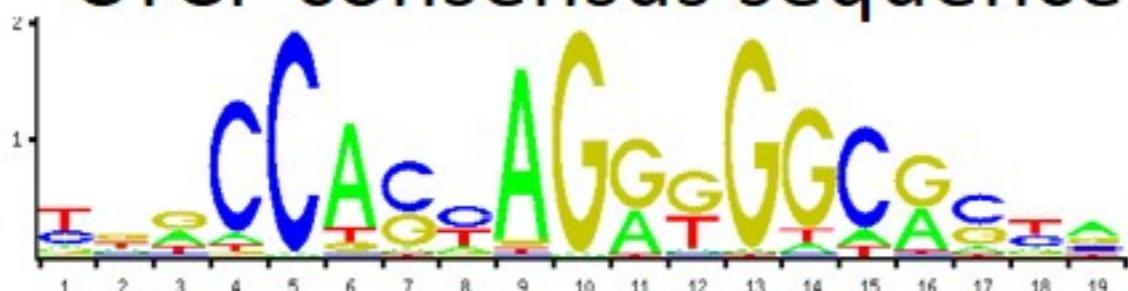
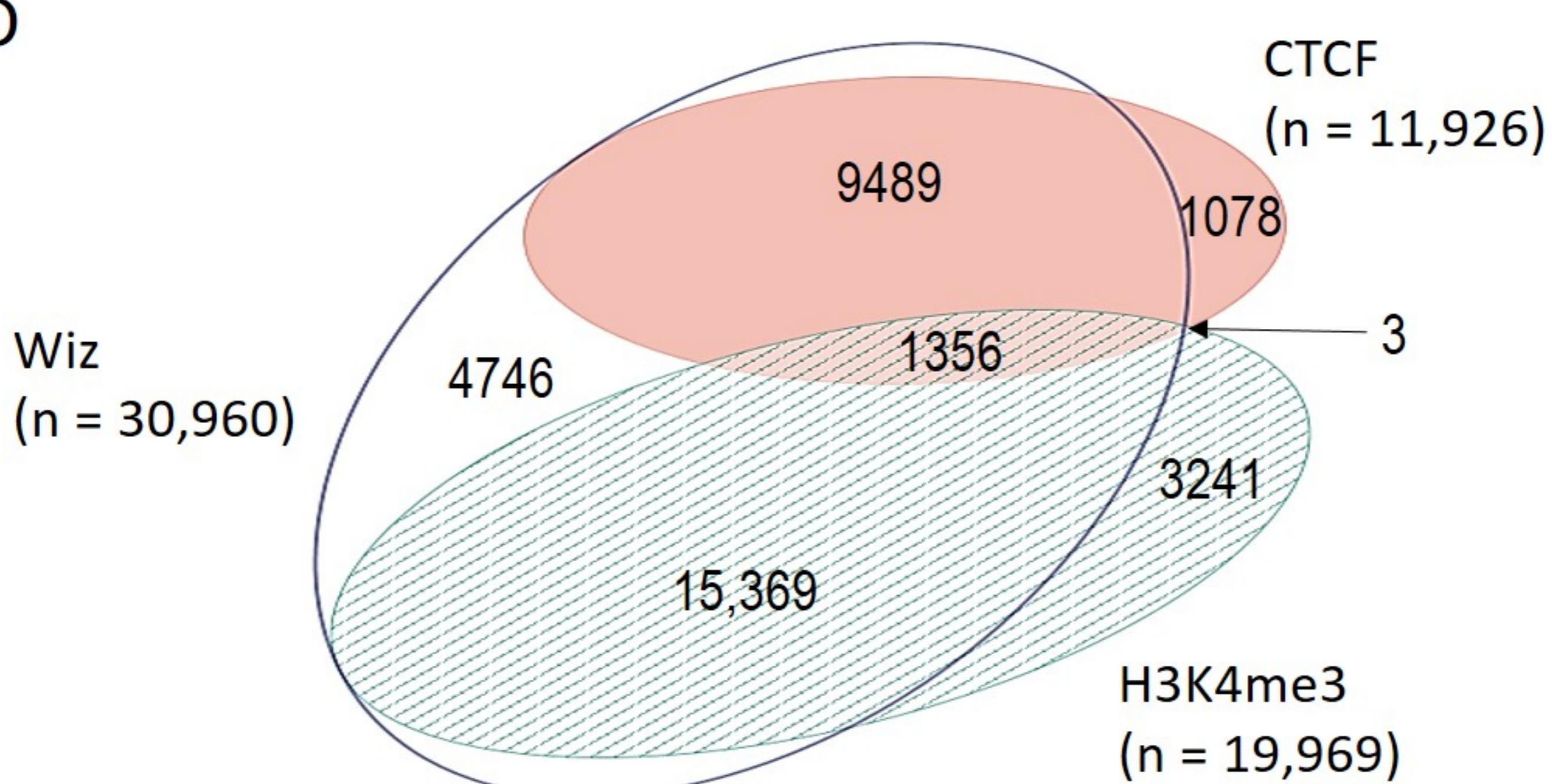


A**C****B**

Wiz motif 1 ~ 70% of Wiz ChIP-seq peaks

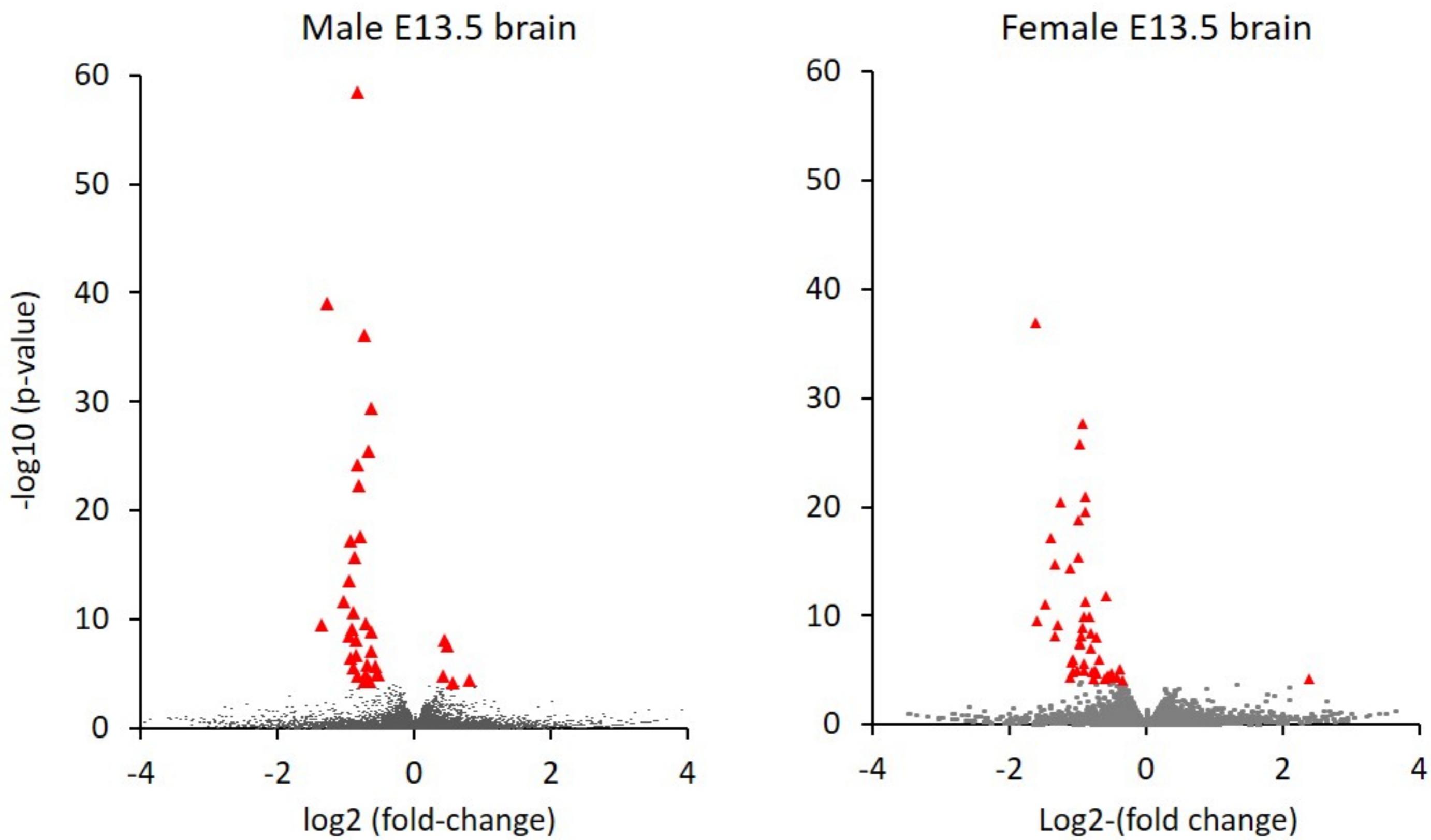


CTCF consensus sequence

**D**

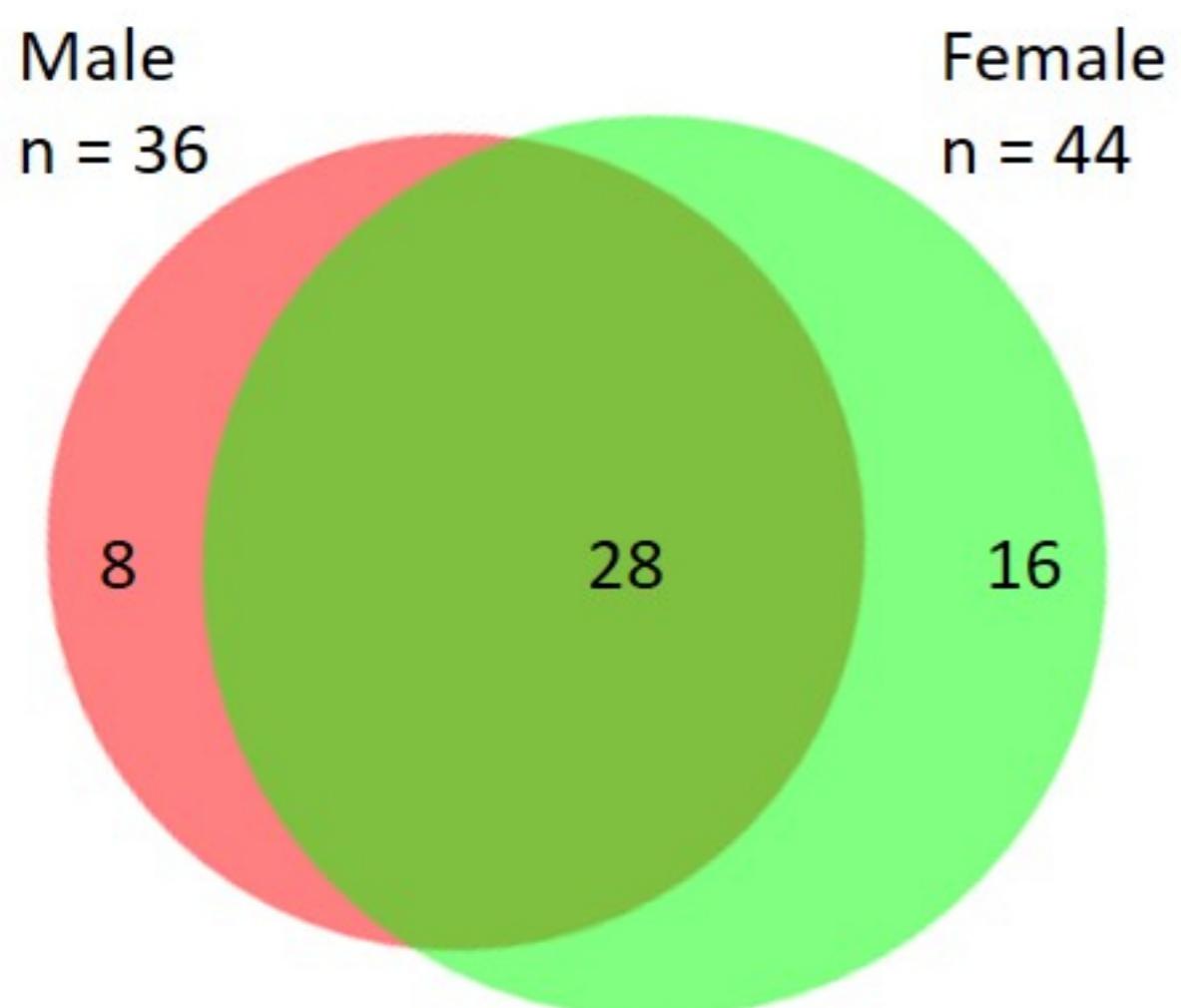
A

Differential expression in *Wiz*^{MD30/+} heterozygous E13.5 brain



B

Overlapping differentially regulated genes in *Wiz*^{MD30/+} E13.5 brain

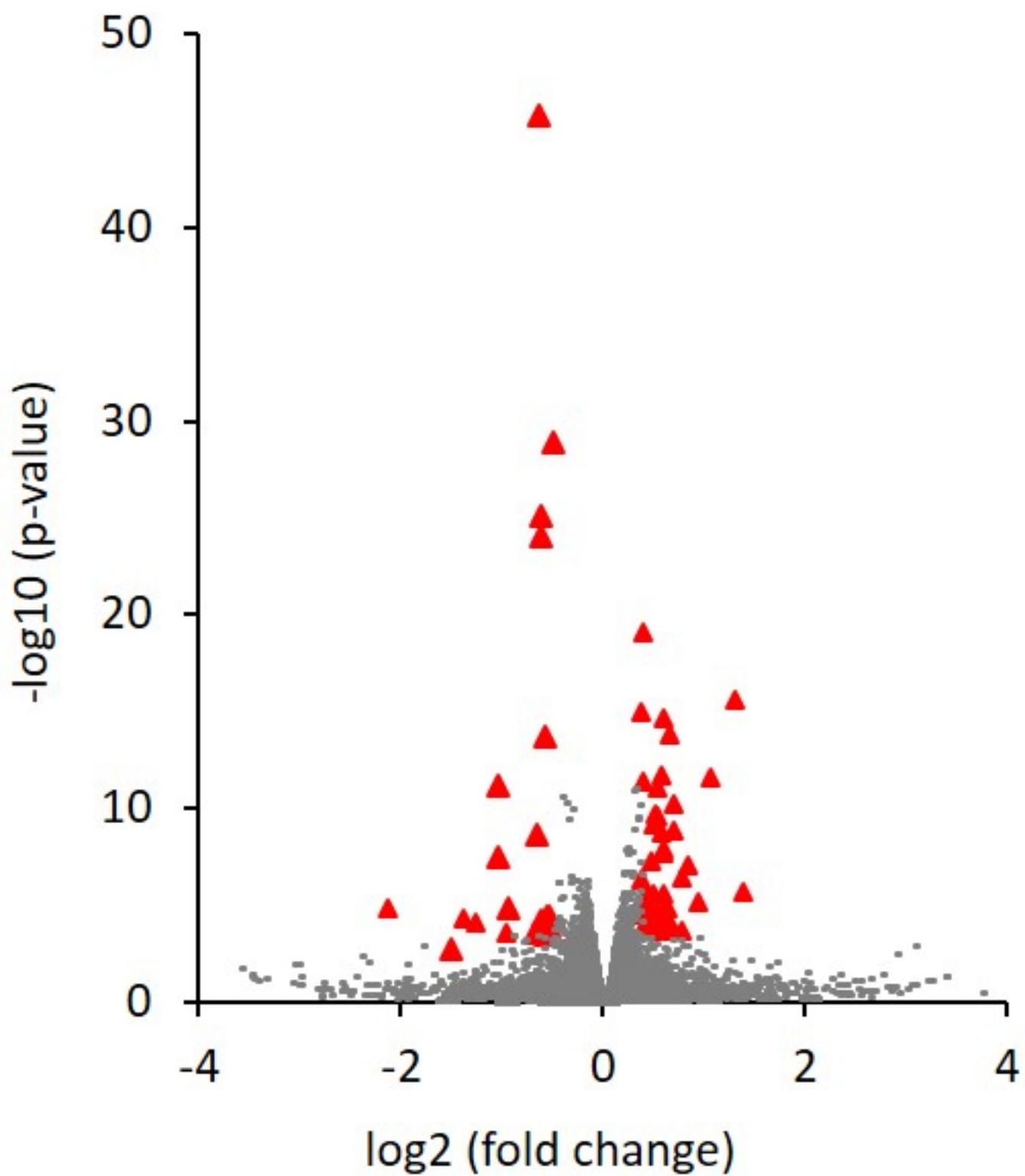


C

Genes downregulated in both male and female <i>Wiz</i> ^{MD30/+} E13.5 brain		
2610005L07Rik	Gm6483	Pcdhb18
3222401L13Rik	Mmp9	Pcdhb19
6820431F20Rik	Pcdhb10	Pcdhb20
AC152164.1	Pcdhb11	Pcdhb22
Csf2ra	Pcdhb12	Pisd-ps1
Gm10557	Pcdhb13	Rps4l
Gm21092	Pcdhb14	Sycp1
Gm21769	Pcdhb15	Wiz
Gm21811	Pcdhb16	
Gm26804	Pcdhb17	

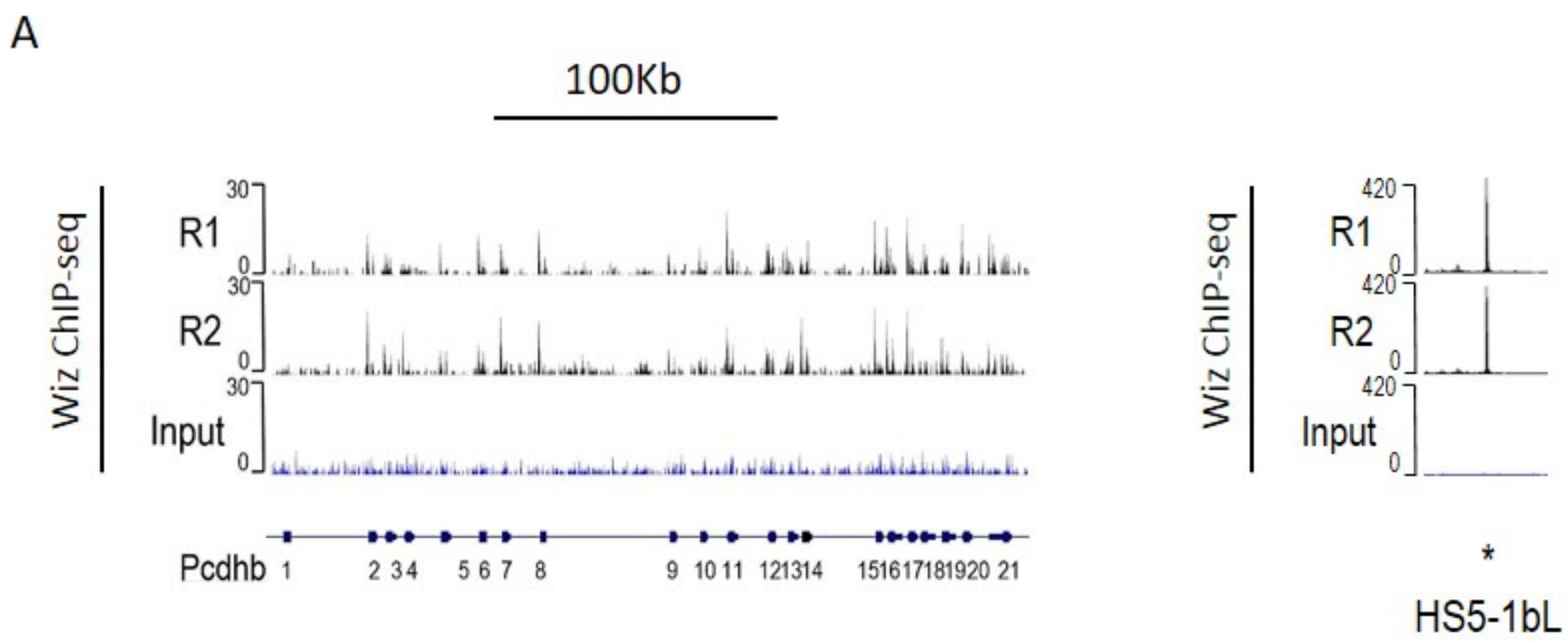
A

Differential expression in $Wiz^{MD30/+}$ adult cerebellum



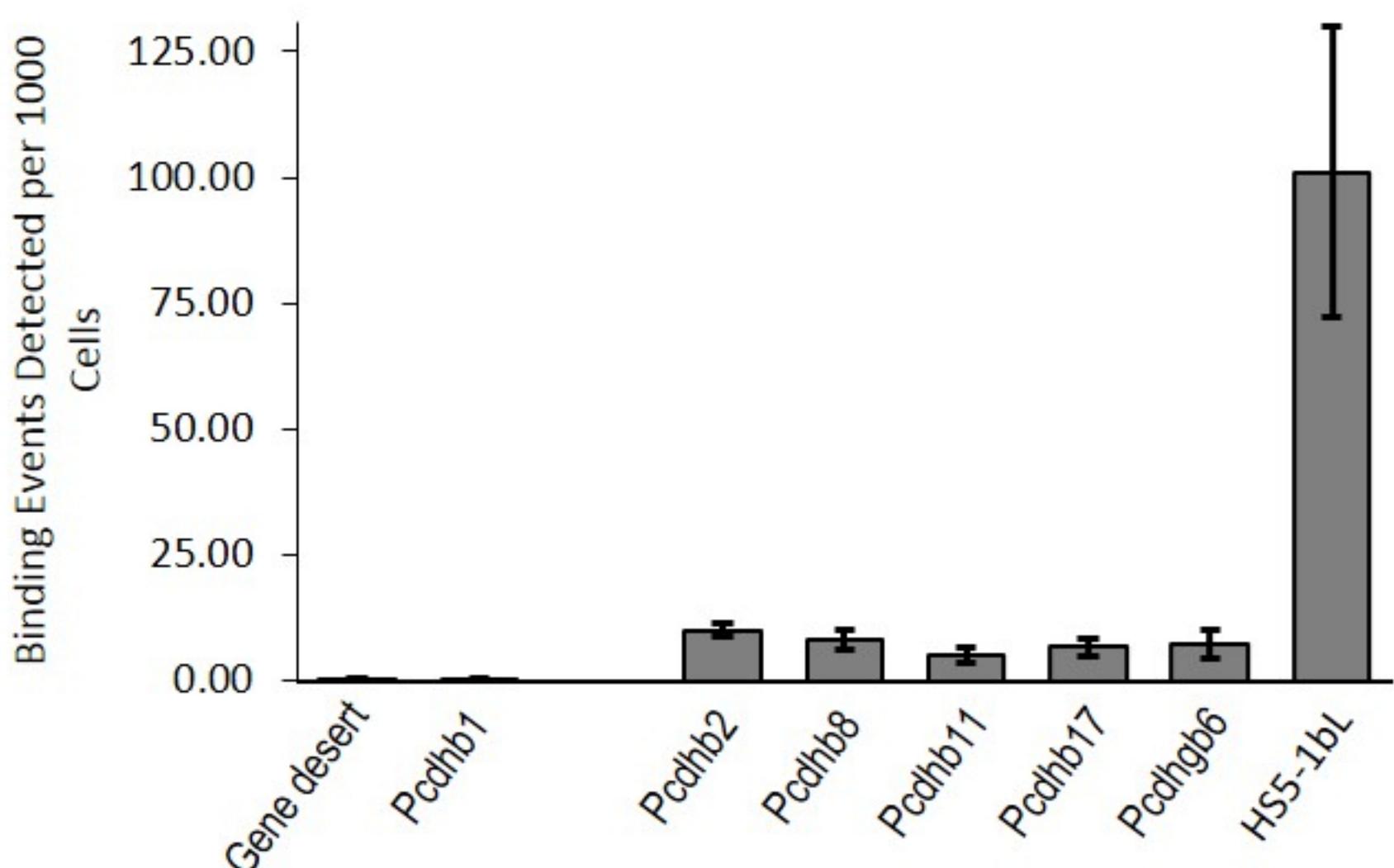
B

Genes downregulated in $Wiz^{MD30/+}$ cerebellum and E13.5 brain
2610005L07Rik
3222401L13Rik
6820431F20Rik
AC152164.1
Csf2ra
Gm21092
Gm21769
Gm21811
Gm6483
Pcdhb14
Rps4l
Wiz

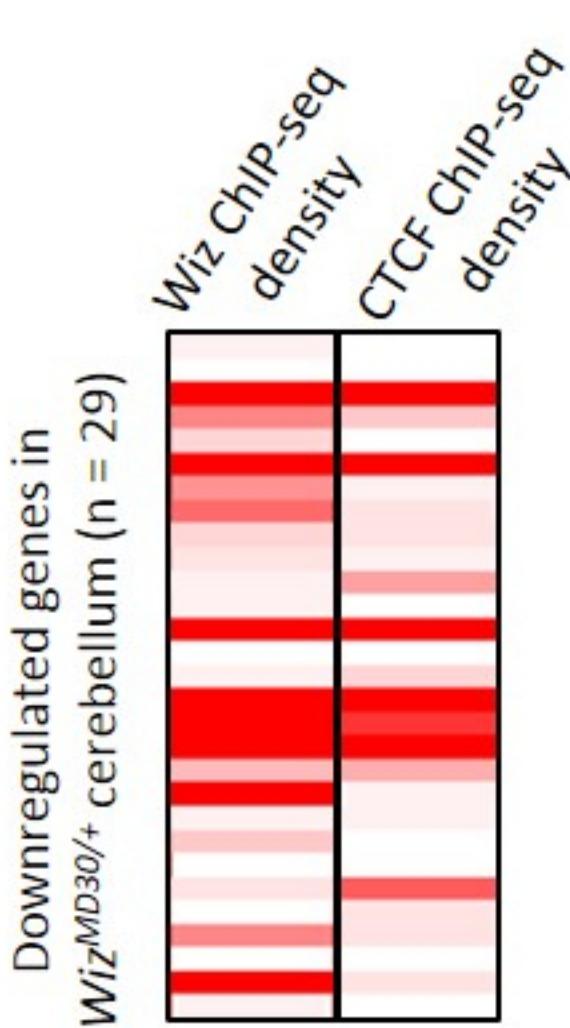


B

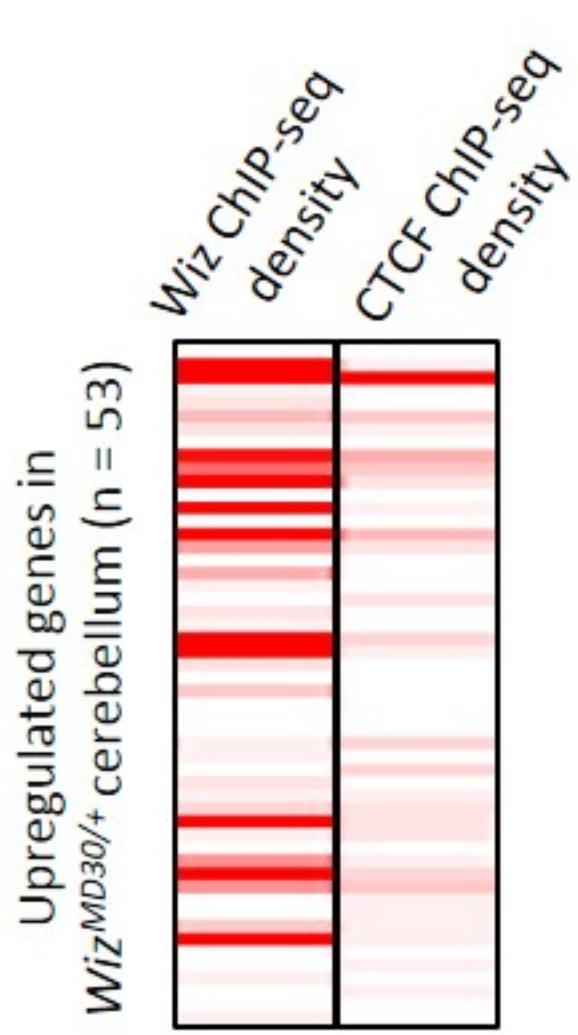
Wiz ChIP-qPCR



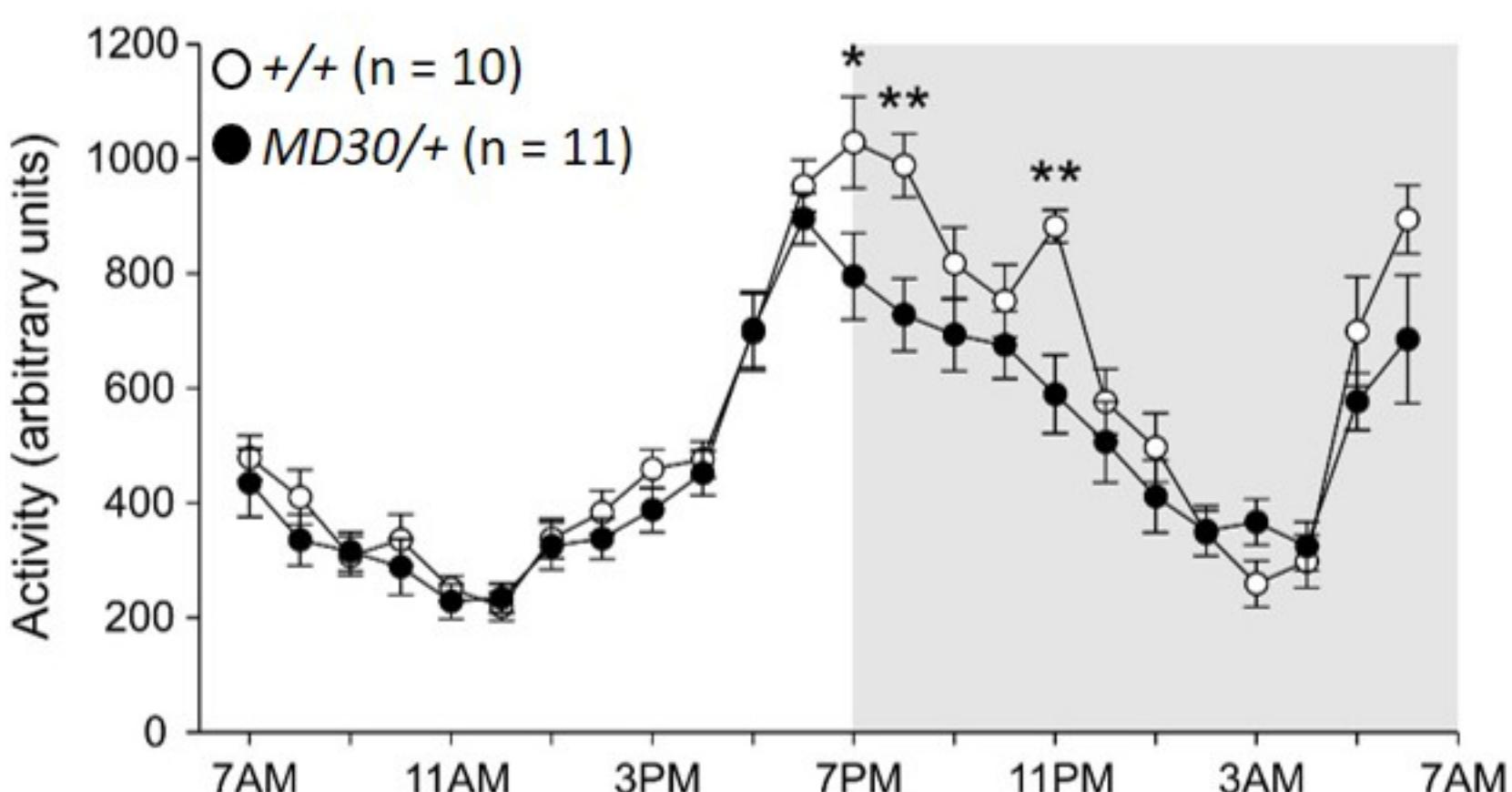
C



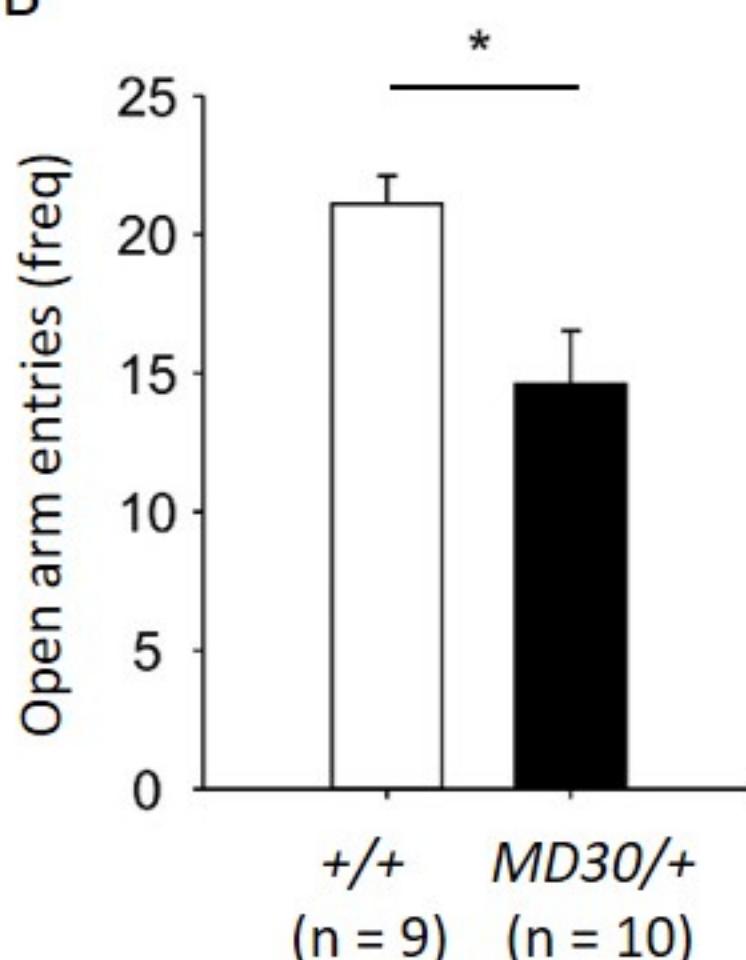
D



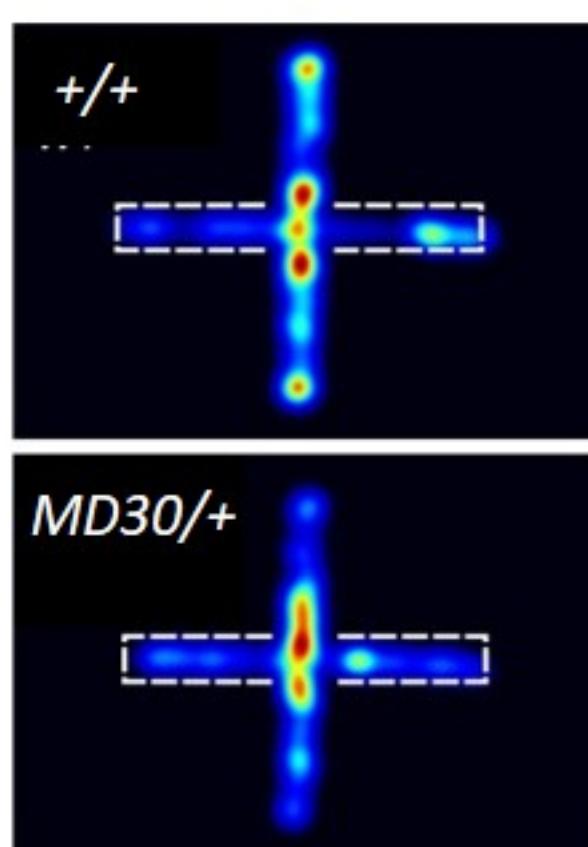
A



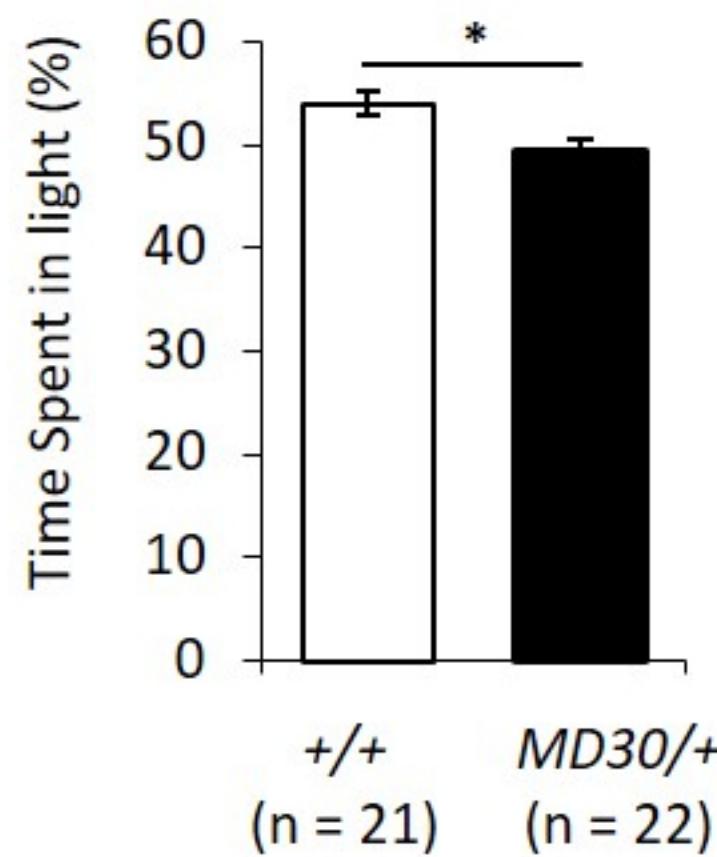
B



C



D



E

