cis-Regulatory Chromatin Contacts in Neural Cells Reveal Contributions of Genetic Variants to Complex Neurological Disorders

Authors: Michael Song[1,2], Xiaoyu Yang[1], Xingjie Ren[1], Lenka Maliskova[1], Bingkun Li[1], Ian Jones[1], Chao Wang[3], Fadi Jacob[4,5], Kenneth Wu[6], Michela Traglia[7], Tsz Wai Tam[1], Kirsty Jamieson[1], Si-Yao Lu[8], Guo-Li Ming[4,9,10,11], Jun Yao[8], Lauren A. Weiss[1,7], Jesse Dixon[12], Luke M. Judge[6,13], Bruce R Conklin[6,14], Hongjun Song[4,9,10,15], Li Gan[3], Yin Shen[1,2,16]*

Affiliation:

[1]Institute for Human Genetics, University of California, San Francisco, San Francisco, CA, USA
[2]Pharmaceutical Sciences and Pharmacogenomics Graduate Program, University of California, San Francisco, San Francisco, CA, USA
[3]Gladstone Institute for Neurological Disorders, San Francisco, CA, USA
[4]Department of Neuroscience and Mahoney Institute for Neurosciences, Perelman School for Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.
[5]The Solomon H. Snyder Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.
[6]Gladstone Institute for Cardiovascular Disease, San Francisco, CA, USA
[7]Department of Psychiatry, University of California, San Francisco, San Francisco, CA, USA
[8]State Key Laboratory of Membrane Biology, Tsinghua-Peking Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing, China.
[9]Department of Cell and Developmental Biology, University of Pennsylvania, Philadelphia, PA 19104, USA.
[10]Institute for Regenerative Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA.
[11]Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA.
[12]Salk Institute for Biological Studies, La Jolla, CA, USA
[13]Department of Pediatrics, University of California, San Francisco, San Francisco, CA, USA
[14]Departments of Medicine and Ophthalmology, University of California, San Francisco, San Francisco, CA, USA
[15]The Epigenetics Institute, University of Pennsylvania, Philadelphia, PA 19104, USA.
[16]Department of Neurology, University of California, San Francisco, San Francisco, CA, USA

*Corresponding Author: yin.shen@ucsf.edu

Abstract

Mutations in gene regulatory elements have been associated with a wide range of complex neurological disorders. However, due to their inherent cell type-specificity and difficulties in characterizing their regulatory targets, our ability to identify causal genetic variants has remained limited. To address these constraints, we perform integrative analysis of chromatin interactions using promoter capture Hi-C (pcHi-C), open chromatin regions using ATAC-seq, and transcriptomes using RNA-seq in four functionally distinct neural cell types: iPSC-induced excitatory neurons and lower motor neurons, iPSC-derived hippocampal dentate gyrus (DG)-like neurons, and primary astrocytes. We identify hundreds of thousands of long-range *cis* interactions between promoters and distal promoter-interacting regions (PIRs), enabling us to link regulatory elements to their target genes and reveal putative pathways that are dysregulated in disease. We validate several novel PIRs using CRISPR techniques in human excitatory neurons, demonstrating that *CDK5RAP3*, *STRAP*, and *DRD2* are transcriptionally regulated by physically linked enhancers. Finally, we show that physical chromatin interactions mediate genetic interactions in autism spectrum disorder (ASD). Our study illustrates how characterizing the 3D epigenome elucidates novel regulatory relationships in the central nervous system (CNS), shedding light on previously unknown functions for noncoding variants in complex neurological disorders.

Introduction

A large number of genetic variations associated with diverse human traits and diseases are located in putative regulatory regions. Genetic lesions in these regulatory elements can contribute to complex human disease by modulating gene expression and disrupting finely tuned transcriptional networks in development and function. However, deciphering the roles of regulatory variants in disease pathogenesis remains nontrivial due to their lack of annotation in the physiologically relevant cell types. Furthermore, regulatory elements often interact with their cognate genes over long genomic distances, precluding a straightforward mapping of regulatory element connectivity and limiting the functional interpretation of noncoding variants from genome wide association studies (GWAS). Typically, neighboring genes are assigned as risk loci for noncoding variants. However, this nearest gene model is challenged both by experimental and computational evidence[1,2]. For instance, two independent obesity-associated SNPs in the *FTO* gene have been shown not to regulate *FTO*, but *IRX3* in the brain and both *IRX3* and *IRX5* in adipocytes, respectively[3,4]. The *FTO* locus in obesity illustrates the potentially intricate and cell type-specific manner in which noncoding variants contribute to disease. However, such well-annotated cases

are rare, and we still lack systematic mappings of GWAS SNPs to their regulatory targets, especially in the context of complex neurological disorders.

Previous epigenomic annotations of the germinal zone (GZ) and cortical and subcortical plates (CP) in the human brain revealed the importance of 3D chromatin structure in gene regulation and disease[5,6]. However, these studies utilized complex, heterogeneous tissues, limiting their abilities to interpret gene regulation in a cell type-specific manner. Therefore, charting the landscape of epigenomic regulation in well-characterized, physiologically relevant cell types should offer significant advantages for identifying causal variants, deciphering their functions, and enabling novel therapies for previously intractable diseases. Towards this goal, we used wild type human induced pluripotent stem cells (iPSCs) from the WTC11 line[7] to generate three neuronal cell types: excitatory neurons[8], hippocampal dentate gyrus (DG)-like neurons[9], and lower motor neurons[10]. GFAP-positive astrocytes from the gastrulating brains of two individuals were also included for their relevance to human brain development and disease. By performing integrative analysis of promoter-centric, long-range chromatin interactions, open chromatin regions, and transcriptomes (**Fig. 1a**), we provide comprehensive annotations of promoters and distal promoter-interacting regions (PIRs) for each of the neural cell types. We identify putative gene targets for both *in vivo* validated enhancer elements from the VISTA Enhancer Browser[11] and disease-associated variants, enabling the functional validation of PIRs driving diverse processes in cellular identity and disease.

Results

Characterizing the epigenomic landscape of long-range chromatin interactions in human neural cells

To investigate general features of the epigenomic landscape for specific cell types in the human central nervous system (CNS), we focused on isogenic iPSC-induced excitatory neurons, iPSC-derived hippocampal dentate gyrus (DG)-like neurons, and iPSC-induced lower motor neurons, three neuronal subtypes which are currently impractical to isolate from primary tissue. Excitatory neurons were induced from a wild type male iPSC line (WTC11) containing an integrated, isogenic, and inducible neurogenin-2 (Ngn2) cassette (i^3N iPSCs) with doxycycline-inducible *Ngn2* at the AAVS1 safe-harbor locus[8]. The i^3N iPSCs enabled us to obtain homogenous excitatory neurons expressing both the cortical neuron marker CUX1 and the glutamatergic neuron marker VGLUT1[8,12] (**Supplementary Fig. 1a, b**).

109 Hippocampal DG-like neurons expressing the DG granule cell marker PROX1 were differentiated from
110 a WTC11 iPSC line using defined factors as described in previous publications[9,13] (**Supplementary Fig.**
111 **1a, b**). Finally, lower motor neurons were induced from a WTC11 line containing integrated, isogenic,
112 and inducible copies of *NGN2*, *ISL1*, and *LHX3* at the AAVS1 safe-harbor locus (i^3LMN iPSCs)[10]. The
113 cells exhibited homogenous expression of the lower motor neuron markers HB9 and SMI32 in culture
114 (**Supplementary Fig. 1a, b**). In addition, all three neuronal subtypes exhibited high expression of the
115 synaptic genes *SYN1* and *SYN2*, the NMDA receptor genes *GRIN1* and *GRIN2A*, and the AMPA
116 receptor genes *GRIA1* and *GRIA2*, evidencing mature synaptic functions (**Supplementary Fig. 1b**). We
117 included two batches of astrocytes isolated from 19 week gastrulating male fetal brain samples using
118 GFAP as a selection marker (ScienCells). Astrocytes were cultured for two or fewer passages *in vitro*
119 and confirmed for positive expression of GFAP prior to harvesting (**Supplementary Fig. 1a**). Based on
120 the age of the donors and transcriptional signatures for dozens of marker genes distinguishing astrocyte
121 progenitor cells (APCs) and mature astrocytes (e.g. high expression of the APC markers *TOP2A* and
122 *TNC* and low expression of the mature astrocyte markers *AGXT2L1* and *WIF1*)[14], the astrocytes were
123 determined to most likely be APCs (**Supplementary Fig. 1b**).
124
125 We constructed pcHi-C, ATAC-seq, and RNA-seq libraries using two to four biological replicates for
126 each cell type (**Supplementary Table 1**). Specifically, promoter-centric, long-range chromatin
127 interactions were mapped using a set of 280,445 unique RNA probes targeting the promoters of 19,603
128 coding genes in GENCODE 19 (Jung et al., in revision). We first confirm the reproducibility of contact
129 frequency and saturation of inter-replicate correlation for our pcHi-C libraries using HiCRep
130 (**Supplementary Fig. 2c, d)**. Hierarchical clustering of ATAC-seq read density and gene expression
131 similarly group the replicates by cell type (**Supplementary Fig. 2a, b**), evidencing minimal variations
132 during the cell derivation process. Using CHiCAGO[15], we identified significant chromatin interactions
133 with score ≥ 5 at a total of 195,322 unique interacting loci across the four cell types, with 73,890,
134 108,156, 66,978, and 84,087 significant interactions being represented in the excitatory neurons,
135 hippocampal DG-like neurons, lower motor neurons, and astrocytes, respectively (**Supplementary**
136 **Table 2**). Overall, 17,065 or 83.9% of coding gene promoters participate in interactions in at least one
137 cell type (**Supplementary Fig. 1c)**, with 80% of PIRs interacting within a distance of 160 kb (**Fig 1c,**
138 **Supplementary Fig. 3a**). Furthermore, over 97% of interactions reside within topologically associating
139 domains (TADs) from Hi-C datasets in human fetal brain tissue[6] (**Fig. 1b**). Approximately 40% of

140 interactions occur between promoter-containing bins, while the remaining 60% occur between promoter-
141 and non-promoter-containing bins (**Fig. 1d**). The observed numbers of promoter-promoter interactions
142 can potentially be attributed to transcriptional factories of co-regulated genes, widespread co-
143 localization of promoters[16,17], and the ability of many promoters to doubly function as enhancers[18,19].
144 Finally, up to 40% of interacting distal open chromatin peaks are specific to each cell type (**Fig. 1e**),
145 suggesting that PIRs are capable of orchestrating cell type-specific gene regulation. Astrocytes in
146 particular exhibit the largest proportion of cell type-specific open chromatin peaks, likely reflecting
147 basic differences between the neuronal and glial lineages.
148
149 We observe that the majority of promoters interact with more than one PIR (**Fig. 2a**). This observation is
150 consistent with the large number of regulatory elements in the human genome[20] and previous findings
151 that each promoter can be regulated by multiple enhancers[21]. To examine global chromatin signatures at
152 PIRs, we leveraged chromatin states inferred by ChromHMM[22] in matched human brain tissues from the
153 Roadmap Epigenomics Project[23] (dorsolateral prefrontal cortex for excitatory neurons, hippocampus
154 middle for hippocampal DG-like neurons, and normal human astrocytes for fetal astrocytes). We show
155 that PIRs are highly enriched for active chromatin features including open chromatin peaks, enhancers,
156 and transcriptional start sites (TSSs) while simultaneously exhibiting depletion for repressive
157 heterochromatin marks (**Fig. 2b**). PIRs are also enriched for H3K27ac and CTCF binding sites mapped
158 using CUT&RUN[24] in excitatory neurons and lower motor neurons, as well as ChIP-seq in astrocytes
159 from ENCODE (**Supplementary Fig. 3b**). We find that promoters interacting exclusively with
160 enhancer-PIRs exhibit elevated levels of transcription compared to promoters interacting exclusively
161 with repressive-PIRs (two sample t-test, one-sided, $p=9.4\times10^{-63}$) (**Fig. 2c, Supplementary Fig. 3c**).
162 Since multiple enhancers can interact with and regulate the same promoter, we assessed whether
163 interactions with multiple enhancer-PIRs could present evidence for additive effects on transcription. By
164 grouping genes according to the number of interactions their promoters form with enhancer-PIRs in each
165 cell type, we discover a positive correlation between the number of enhancer-PIR interactions and the
166 mean gene expression in each group (linear regression test, $p=2.1\times10^{-3}$) (**Fig. 2d, Supplementary Fig.
167 3d**). Our results demonstrate that chromatin interactions enable the identification of PIRs which are not
168 only enriched for regulatory features, but which can also modulate gene expression.
169
170 **PIRs contribute to cellular identity**

171 We find that chromatin interactions exhibit distinct patterns of cell type specificity, with tens of
172 thousands of interactions that are specific to each cell type (**Fig. 3a, Supplementary Fig. 4a**). These
173 interactions may underlie important functional differences between the cell types, as gene ontology
174 (GO) enrichment analysis for genes interacting with cell type-specific PIRs produced terms associated
175 with neuronal function in the neuronal subtypes and immune function in the astrocytes (**Fig. 3b,**
176 **Supplementary Table 3**). Meanwhile, 58,809 or 30.1% of unique interactions are shared across all four
177 cell types, with neural precursor cell proliferation and neuroblast proliferation ranking among the top
178 GO terms for genes participating in shared interactions. In conjunction with the observed enrichment of
179 active chromatin signatures at PIRs, the cell type-specific nature of PIRs suggests that they harbor
180 lineage-specific regulatory roles. Indeed, numerous promoters of differentially expressed genes form
181 specific contacts with PIRs in the corresponding cell types, including *OPHN1* in hippocampal DG-like
182 neurons, *CHAT* in lower motor neurons, and *TLR4* in astrocytes (**Supplementary Fig. 4b**). OPHN1 is
183 known to stabilize synaptic AMPA receptors and mediate long-term depression in the hippocampus, and
184 its loss of function is associated with X-linked mental retardation[25]. Meanwhile, CHAT is a principal
185 marker for lower motor neuron maturity and function, and TLR4 is a key regulator of immune activation
186 and synaptogenesis in astrocytes[26]. Finally, hierarchical clustering of interaction scores across all
187 significant promoter-PIR interactions demonstrates that cell types can reliably be grouped according to
188 lineage-specific features, with the three neuronal subtypes clustering together more tightly than with the
189 astrocytes (**Fig. 3a**).
190
191 Gene expression is coordinately controlled by transcription factors (TFs) and regulatory elements such
192 as enhancers. Therefore, PIRs identified through chromatin interactions provide a unique perspective to
193 investigate potential mechanisms underlying cell type-specific gene regulation. We use HOMER[27] to
194 evaluate TF motif enrichment at cell type-specific distal open chromatin peaks in PIRs for each cell type
195 (**Fig. 3c, Supplementary Table 4**). We find that the CTCF motif is highly enriched across all cell types,
196 consistent with its role in mediating looping between promoters and regulatory elements within TADs[28-
197 31]. Furthermore, motifs for ASCL1, ISL1, NEUROG2, OLIG2, and ZIC3, TFs linked to neuronal fate
198 commitment, are enriched in various patterns across the neuronal subtypes. Other TFs functioning in
199 brain development include CUX1/CUX2, EBF1/EBF2, HFN6, LHX1/LHX2, NKX6-1, TCF4, TGIF2,
200 and the RFX factors. The TBR1 motif is enriched in hippocampal DG-like neurons, consistent with
201 TBR1's roles in NMDA receptor assembly and maintaining long-term potentiation in the

202 hippocampus[32]. Meanwhile, astrocytes are enriched for motifs in the Fos and Jun families, which
203 contain key regulators for inflammatory and immune pathways. Also enriched in astrocytes are motifs
204 for ATF3 and the RUNX and TEAD families, TFs with established roles in astrocyte differentiation,
205 maturation, and proliferation. Motif enrichment is not always accompanied by expression of the
206 corresponding TFs, potentially reflecting synergistic interactions between different cell types in the
207 CNS. For example, NRF2 is a key regulator of the oxidative stress response, and its activity has been
208 shown to be repressed in neurons while inducing a strong response in astrocytes[33]. Therefore, its shared
209 expression may reflect the neuroprotective role that astrocytes provide for other cell types. Alternatively,
210 TFs do not have to be expressed at high levels to perform their cellular functions due to additional
211 avenues for regulation at the post-transcriptional and protein levels. Overall, our results demonstrate that
212 PIRs contribute to cell fate commitment and are capable both of recapitulating known and revealing
213 novel regulators.
214
215 ***In vivo* validation of interactions linking enhancer elements to their target genes**
216 Regulation of target genes by enhancers is thought to be mediated by physical chromatin looping.
217 Congruent with this concept, interactions detected by pcHi-C can be used to link enhancers with their
218 target genes. The VISTA Enhancer Browser[11] is a database containing experimentally validated human
219 and mouse noncoding sequences with enhancer activity. To date, it contains 2,956 tested elements,
220 1,568 of which exhibit enhancer activity during embryonic development[11]. However, the regulatory
221 targets for these enhancer elements remain largely uncharacterized. To address this knowledge gap, we
222 provide detailed cell type-specific annotations of putative target genes for each enhancer element using
223 our significant promoter-PIR interactions and open chromatin peaks (**Supplementary Table 5**). In total,
224 our interactions recover 589 or 37.6% of positively annotated enhancer elements with human sequences,
225 320 of which were further annotated as neural enhancers according to tissue-specific patterns of LacZ
226 staining in mouse embryos (**Fig. 4a, b**). Of the 589 interacting positive enhancer elements, only 60
227 interact exclusively with their nearest genes (scenario III), whereas 306 interact exclusively with more
228 distal genes (scenario I), identifying 464 novel gene targets (**Fig. 4c**). Meanwhile, 118 elements interact
229 with both their nearest genes and a total of 484 more distal genes (scenario II). The remaining 105
230 elements cannot be resolved at the HindIII fragment level for interactions with their nearest genes
231 (scenario IV), though they interact with 395 additional non-neighboring genes. In total, our chromatin

232 interactions identify 1,343 novel, putative gene targets for positive enhancer elements in the VISTA
233 Enhancer Browser, significantly expanding our knowledge of gene regulatory relationships at these loci.
234
235 **Validation of PIRs detected in human neural cells using CRISPR techniques**
236 We performed validation of two PIRs (regions 1 and 2) physically interacting up to 40 kb away with the
237 promoter of *CDK5RAP3* (**Fig. 4d**). CDK5RAP3 is a known regulator of CDK5, which functions in
238 neuronal development[34] and regulates proliferation in non-neuronal cells[35]. Both PIRs overlap open
239 chromatin peaks as well as enhancer elements with positively annotated forebrain activity in the VISTA
240 Enhancer Browser (mm8766 and mm999 for region 1 and mm1299 for region 2) (**Fig. 4e**). We targeted
241 both regions for CRISPR deletion in i^3N iPSCs, followed by differentiation of the iPSCs into excitatory
242 neurons and quantification of any changes in gene expression by qPCR. Deleting the 2 kb open
243 chromatin peak in region 1 led to significant downregulation of *CDK5RAP3* expression (two sample t-
244 test, two-sided, $p=7.7\times10^{-7}$) (**Supplementary Fig. 4c**). Upon trying to delete the open chromatin peak in
245 region 2, we observed massive cell death of iPSCs immediately following introduction of the Cas9-
246 sgRNA protein complex. We picked 48 individual clones from cells surviving the transfection, but failed
247 to isolate any clones with deletions, suggesting that this locus is essential for maintaining *CDK5RAP3*
248 expression and survival in iPSCs. To circumvent the lethal phenotype for iPSCs associated with region
249 2, we silenced both regions using CRISPR interference (CRISPRi) in excitatory neurons. We also
250 silenced a third region interacting in the other cell types, but not in excitatory neurons (region 3). We
251 show that silencing of regions 1 and 2 but not region 3 leads to significant downregulation of
252 *CDK5RAP3* expression without influencing the expression of nearby genes (two sample t-test, two-
253 sided, $p=3.2\text{x}10^{-8}$ for region 1 and $p=2.2\text{x}10^{-3}$ for region 2) (**Fig. 4f-g**). Interestingly, a neighboring
254 enhancer element annotated with spinal cord activity (mm1576) also participates in interactions with the
255 *CDK5RAP3* promoter in lower motor neurons and astrocytes, but not in the excitatory neurons and
256 hippocampal DG-like neurons (**Fig. 4d, e**). These results demonstrate that chromatin interactions
257 recapitulate cell type-specific patterns of enhancer activity, underscoring the importance of studying
258 epigenomic regulation in the appropriate cell types.
259
260 **Cell type-specific enrichment and regulatory target identification for complex neurological**
261 **disorder-associated variants at PIRs**

262 Previous large-scale epigenomic studies of human tissues and cell lines highlighted the importance of
263 disease-associated variants at distal regulatory regions[23] and the need for high-throughput approaches to
264 prioritize variants for further validation. Therefore, we used our chromatin interactions to annotate
265 complex neurological disorder- and trait-associated variants available from the GWAS Catalog[36]. We
266 downloaded a total of 6,396 unique GWAS SNPs at a significance threshold of 10^{-6} for eleven traits
267 including Alzheimer's disease (AD), attention deficit hyperactivity disorder (ADHD), autism spectrum
268 disorder (ASD), amyotrophic lateral sclerosis (ALS), bipolar disorder (BD), epilepsy (EP),
269 frontotemporal dementia (FTD), mental processing (MP), Parkinson's disease (PD), schizophrenia
270 (SCZ), and unipolar depression (UD). The GWAS SNPs were imputed at a linkage disequilibrium (LD)
271 threshold of 0.8 using HaploReg[37] and filtered to obtain a total of 95,954 unique imputed SNPs across
272 all traits (**Supplementary Table 6**). We find that SNPs are enriched at PIRs in a disease- and cell type-
273 specific manner (**Fig. 5a**), with ASD, MP, and SCZ SNPs enriched at PIRs across all cell types. UD
274 SNPs are exclusively enriched in excitatory neurons and hippocampal DG-like neurons, whereas AD,
275 ADHD, and BD SNPs also exhibit enrichment in lower motor neurons. ALS SNPs are enriched in all the
276 neuronal subtypes but not in astrocytes, consistent with the characterization of ALS as a motor neuron
277 disease and reinforcing evidence for its role in hippocampal degeneration[38]. Notably, PD SNPs are
278 enriched in astrocytes but not in other cell types. The enrichment of PD SNPs at astrocyte-specific PIRs
279 also supports the theory that astrocytes play an initiating role in PD[39], based on evidence that numerous
280 genes implicated in PD possess functions unique to astrocyte biology, as well as the neuroprotective
281 roles that astrocytes serve for dopaminergic neurons in the substantia nigra. Finally, EP and FTD SNPs
282 are not enriched in any of the cell types, indicating their potential functions in alternative cell types,
283 insufficient study power, or mechanisms acting outside of chromatin-mediated gene regulation.
284
285 Up to 70% of GWAS SNPs have at least one SNP in linkage overlapping PIRs in one or more cell types
286 (**Fig. 5b**). As it is common practice in association studies to assign GWAS SNPs to their nearest genes,
287 we evaluated the number of GWAS SNPs with at least one SNP in linkage interacting with their nearest
288 genes. Overall, across all diseases, we find that 248 GWAS SNPs interact exclusively with their nearest
289 genes (scenario III), 1,365 GWAS SNPs interact exclusively with more distal genes (scenario I), and
290 1,243 GWAS SNPs interact with both their nearest genes and more distal genes (scenario II) (**Fig. 5c,**
291 **Supplementary Fig. 5a**). Our interactions identify a total of 16,471 non-neighboring gene targets across
292 all diseases (**Supplementary Table 7**). To prioritize variants potentially disrupting regulatory activity,

293 we focused on SNPs overlapping open chromatin peaks at PIRs. We find that PIRs for these putative
294 regulatory SNPs interact with genes possessing functions that are relevant in the context of their
295 respective disease etiologies. For example, GO enrichment results for genes targeted by AD SNPs
296 include terms associated with amyloid beta formation, interferon beta production, and cranial nerve
297 development (**Supplementary Fig. 6, Supplementary Table 9**). Meanwhile, genes targeted by ASD,
298 BD, SCZ, and UD SNPs are enriched for epigenetic terms including chromatin assembly, nucleosome
299 assembly, and nucleosome organization. For genes targeted by SNPs in the remaining diseases, enriched
300 terms include neuronal processes such as myelin maintenance, neuron projection extension, synapse
301 assembly, synaptic transmission, and nervous system development. A comprehensive annotation of PIRs
302 overlapping putative regulatory SNPs is available in **Supplementary Table 8**. Notably, a previously
303 reported interaction between the *FOXG1* promoter and a PIR containing SCZ SNPs over 700 kb away is
304 recapitulated in our chromatin interactions[6] (**Supplementary Fig. 5b**). In another example, an astrocyte-
305 specific PIR containing AD SNPs targets the promoters of both *CASP2*, a well-known mediator of
306 apoptosis that is also associated with neurodegeneration[40,41], and *FAM131B*, a putative neurokine
307 (**Supplementary Fig. 7a**). Elsewhere, hippocampal DG-like neuron-specific PIRs containing ASD
308 SNPs target the promoter of *BCAS2*, whose knockdown in mice leads to microcephaly-like phenotypes
309 with reduced learning, memory, and DG volume[42] (**Supplementary Fig. 7c**). Finally, the *MSI2*
310 promoter is simultaneously targeted by PIRs containing BD SNPs in hippocampal-DG like neurons,
311 lower motor neurons, and astrocytes, as well as a PIR in astrocytes containing SCZ SNPs
312 (**Supplementary Fig. 7d**). Overall, we show that an approach utilizing epigenomic annotations to
313 jointly prioritize variants and identify their regulatory targets enables the identification of regulatory
314 mechanisms with consequential roles in development and disease.
315

316 **Validation of PIRs containing neurological disorder-associated regulatory variants**

317 We used CRISPR techniques to validate two PIRs containing putative regulatory SNPs targeting the
318 promoters of the *STRAP* and *DRD2* genes. At the *STRAP* locus, PIRs containing MP SNPs in an intron
319 for *PTPRO* interact over 300 kb away with the promoter of *STRAP* (**Fig. 5d**). STRAP influences the
320 cellular distribution of the survival of motor neuron (SMN) complex, which in turn facilitates
321 spliceosome assembly and is associated with spinal muscular atrophy[43]. We derived three independent
322 i^3N iPSC clones containing bi-allelic deletions for this candidate PIR and observed significant
323 downregulation of *STRAP* expression following differentiation of the i^3N iPSCs into excitatory neurons

324 (two sample t-test, two-sided, p=9.4×10^{-13}). Targeting the same region with CRISPRi silencing also
325 consistently downregulated *STRAP* expression in the excitatory neurons (two sample t-test, two-sided,
326 p=1.5x10^{-7}) (**Supplementary Fig. 5c**). Next, we focused on a candidate PIR 20 kb upstream from the
327 promoter of *DRD2*, which encodes the D2 subtype of the dopamine receptor. Previously, rs2514218, a
328 noncoding variant 47 kb upstream from *DRD2*, was found to be associated with anti-psychotic drug
329 response in a cohort of schizophrenia patients[44]. Notably, this variant is in linkage with a cluster of SCZ
330 SNPs overlapping open chromatin peaks within our candidate PIR. *DRD2* is also the gene associated
331 with the Taq1A polymorphism which is linked to reduced dopamine receptor density as well as
332 addiction, anxiety, depression, and social problems in patients[45]. We first demonstrate that mono-allelic
333 deletion of the candidate PIR in three independent clones leads to significant downregulation of *DRD2*
334 expression in excitatory neurons (two sample t-test, two-sided, p=2.7×10^{-7}) (**Fig. 5e**). Next, we confirm
335 that mono-allelic deletion of the candidate PIR leads to allelic imbalance in the expression of *DRD2* by
336 performing TOPO cloning and genotyping cDNA with allele-specific variants (**Supplementary Fig.
337 5d**). *DRD2* is a key gene possessing multifaceted roles in human brain function, and it has been
338 implicated in numerous complex neurological disorders including addiction, bipolar disorder, migraine,
339 and obesity[46]. By prioritizing and validating putative regulatory SNPs for genes such as *DRD2*, our
340 integrative approach enables the development of novel therapeutic and diagnostic strategies targeting
341 specific variants for their roles in otherwise recalcitrant complex neurological disorders.
342
343 **Genetic variants contribute to chromatin interaction bias and alterations in gene expression**
344 Since regulatory variants and other genetic perturbations are thought to disrupt chromatin looping
345 between promoters and PIRs, we were interested to see if we could detect instances of allelic bias across
346 our sets of significant promoter-PIR interactions. We used our chromatin interaction data to perform
347 genome-wide phasing of WTC11 variants using HaploSeq[47] and performed allele-specific mapping at a
348 resolution of 10 kb using HiC-Pro. We identify 41 (0.185%) and 151 (0.703%) of significantly
349 interacting bins to exhibit allelic bias at an FDR cutoff of 5% in excitatory neurons and lower motor
350 neurons, respectively, confirming that genetic diversity can contribute to allelic bias in chromatin
351 interactions (**Fig. 6a, Supplementary Table 10**). In one instance, allelically biased interactions were
352 detected between a PIR containing SNPs for bipolar alcoholism[48] and the promoter of *SYT17* **(Fig. 6b**).
353 *SYT17* encodes a member of a family of membrane-trafficking proteins that mediate synaptic function
354 and regulate calcium-controlled neurotransmitter release[49]. The identification of chromatin interactions

355 with allelic bias at the *SYT17* locus suggests that variants can increase individual risk for bipolar
356 alcoholism by disrupting interactions for *SYT17*, consistent with a model in which regulatory sequences
357 recruit TFs and other activating factors to form physical contacts with their gene targets.
358
359 Physical chromatin interactions have also been theorized to mediate the effects of *cis*-acting regulatory
360 variants such as expression quantitative trait loci (eQTLs) on gene expression. To test this hypothesis,
361 we first show that significant eQTLs filtered at an FDR cutoff of 5% in cortical and hippocampal tissues
362 from GTEx V7[50] are enriched at PIRs for the excitatory neurons and hippocampal DG-like neurons,
363 respectively (one sample z-test, $p<2.2\text{x}10^{-16}$ for both cell types) (**Fig. 6c**). Next, we demonstrate that the
364 mean scores for interactions overlapping significant eQTL-TSS pairs are significantly higher than the
365 scores for interactions overlapping randomly shuffled eQTL-TSS pairs (Kolmogorov-Smirnov test,
366 $p=2.28\text{x}10^{-4}$ for excitatory neurons and $p=1.76\text{x}10^{-6}$ for hippocampal DG-like neurons) (**Fig. 6d**). This
367 indicates that significant promoter-PIR interactions recapitulating regulatory relationships between
368 eQTL-TSS pairs are called with increased levels of confidence. Our results present orthogonal lines of
369 evidence that chromatin interactions can not only be altered by variants in an allele-specific manner, but
370 that variants can also influence gene expression through the formation of chromatin interactions.
371
372 **Enrichment of chromatin interactions at ASD SNP pairs**
373 Genetic epistasis refers to combinations of independent variants exhibiting effects which cannot be
374 predicted from their individual contributions alone, and it has been found to contribute to complex traits,
375 including behavioral ones, in model organisms[51-54]. Epistasis has also been theorized to drive complex
376 traits in humans such as ASD[55]. However, studying epistasis in humans in an unbiased fashion is
377 challenging due to the large number of loci that need to be tested. In addition, most candidate pairs
378 studied for epistasis are based on interactions between proteins encoded at the loci of interest. Therefore,
379 chromatin interactions present a unique opportunity to identify epistatic effects in an unbiased fashion,
380 as they suggest an attractive model in which physical contacts bridge independent variants to mediate
381 genetic interactions in epistasis.
382
383 For each cell type, we tested for epistatic effects between all combinations of SNP pairs occurring at
384 both ends of our significant promoter-PIR interactions. Specifically, we evaluated enrichment in our
385 interactions for independent ASD case-specific SNP pairs over non-overlapping matched pseudo-

386 control-specific SNP pairs (see methods)[56]. Of all the cell types, only the hippocampal DG-like neurons
387 showed enrichment at PLINK's default significance threshold of 10^{-7} (chi-squared test, p=2.7x10^{-16})
388 (**Fig. 6e**). Next, we explored sets of increasingly stringent significance thresholds (**Supplementary**
389 **Table 11**), demonstrating that excitatory neurons also exhibit enrichment at a significance threshold of
390 10^{-20} (chi-squared test, p=0.02). As an additional set of controls, we tested for epistatic effects in
391 nonsignificant promoter-PIR interactions with score < 1, demonstrating that for all cell types, significant
392 interactions exhibit greater enrichments for epistatic effects than are observed in nonsignificant ones.
393 Overall, our results confirm that genetic interactions can indeed be mediated by chromatin interactions,
394 underscoring the importance of the 3D epigenome for elucidating factors underlying individual disease
395 risk.
396
397 **Discussion**
398 There is a distinct lack of 3D epigenomic annotations in cell types that are relevant to disease and
399 development, especially in the field of brain research. Past studies have frequently relied on
400 heterogeneous tissues containing cell types with disparate functions, limiting their abilities to detect and
401 interpret instances of cell type-specific gene regulation. Neurons and glia, for example, represent
402 lineages with divergent functions that coexist in most tissues of the CNS. At the same time, complex
403 diseases often involve multiple dysregulated loci exhibiting cell type-specific patterns of activity. This
404 presents unique challenges for deciphering disease etiology, for example in distinguishing causative
405 mechanisms from secondary, reactive phenotypes across distinct populations of cells. Therefore, the
406 annotation of regulatory relationships in specific, well-characterized cell types should enable the wider
407 community to derive numerous insights into complex disease biology. Chromatin interactions in
408 particular are ideal for mapping promoters to distal regulatory elements, as they provide direct physical
409 evidence of regulatory sequences contacting genes. To date, several studies have characterized
410 chromatin interactions in fetal brain tissues and neural cell types[6,57]. However, these studies used *in situ*
411 Hi-C, which lacks power for calling interactions compared to targeted approaches such as pcHi-C.
412
413 In this study, we leverage pcHi-C, ATAC-seq, and RNA-seq to annotate previously uncharted regulatory
414 relationships between promoters and distal regulatory elements in cell types that are relevant to
415 neurological disease. We show that PIRs in our interactions are not only cell type-specific, but they are
416 also enriched for regulatory chromatin signatures including open chromatin peaks as well as *in vivo*

417 validated enhancers in the VISTA Enhancer Browser. Despite functional similarities between the cell
418 types used in this study, the inspection of cell type-specific distal open chromatin peaks at PIRs reveals
419 subtype-specific binding sites for TFs involved in the specification and maintenance of cellular identity.
420 Furthermore, our interactions enable the identification of novel gene targets for disease-associated
421 variants and the prioritization of variants for validation using CRISPR techniques. We report a large
422 number of putative regulatory variants which may reveal additional aspects of complex disease biology.
423 Finally, the disease- and cell type-specific enrichment of variants at PIRs, combined with the
424 observation that the same PIRs can target different promoters in different cell types, affirms that
425 regulatory variants possess context-dependent functional specificities, underscoring the importance of
426 performing validation in the appropriate cell types.
427
428 The integrative analysis in this study has several limitations including a lack of cell type-specific
429 annotations for various genomic and epigenomic features occurring at PIRs. For example, the analysis of
430 chromatin state enrichment at PIRs utilized data in matched tissues from the Roadmap Epigenomics
431 Project. Although we generated our own maps of H3K27ac and CTCF binding sites, our results would
432 be even more sensitive if chromatin states were inferred in matching cell types. Furthermore, while
433 studying chromatin interactions in healthy cells enables the detection of regulatory interactions in the
434 absence of dysregulation, the characterization of patient-derived iPSCs will also be important in the
435 future to glean specific insights into how the 3D epigenome is altered in disease. Additional follow-up
436 experiments are necessary to determine how the haploinsufficiency of proteins such as STRAP and
437 DRD2 may contribute to phenotypes in disease. Finally, *in vitro* cultured cells can at present only
438 approximate the full set of cellular responses occurring *in vivo*, especially in complex structures such as
439 the brain. Future approaches isolating specific cell populations from tissues, leveraging single cell
440 sequencing, or studying advanced organoid models will be essential for drilling down even deeper into
441 mechanisms driving cellular identity and disease. In conclusion, our study presents a roadmap for the
442 annotation of cell type-specific interactions in the CNS, advancing our ability to elucidate mechanisms
443 by which noncoding variants drive complex neurological disorders. The epigenomic characterization of
444 additional cell types should continue to yield rich insights into the landscape of transcriptional
445 regulation, contributing towards an improved understanding of complex disease biology[58].
446
447 **Methods**

448 **Cell culture**

449 Human excitatory neurons were generated using integrated, isogenic, and inducible neurogenin-2
450 (Ngn2) iPSCs (i^3N iPSCs) with doxycycline-inducible mouse Ngn2 integrated at the AAVS1 safe-
451 harbor locus. The i^3N iPSCs have a well-characterized wild type genetic background (WTC11)[7]. A
452 simplified, two-step pre-differentiation and maturation protocol was used to generate excitatory
453 neurons[8]. Briefly, i^3N iPSCs were incubated with 2 μg/ml doxycycline in pre-differentiation media
454 containing KnockOut DMEM/F12 supplemented with 1x N-2, 1x NEAA, 1 μg/ml mouse laminin, 10
455 ng/ml BDNF, and 10 ng/ml NT3. In addition, 10 μM Rock inhibitor was included in the pre-
456 differentiation media for the first day. Media was changed daily for three days. For maturation, pre-
457 differentiated precursor cells were dissociated and subplated on poly-D-lysine/laminin plates in
458 maturation media containing equal parts DMEM/F12 and Neurobasal-A with 2 μg/ml doxycycline and
459 supplemented with 0.5x B-27, 0.5x N-2, 1x NEAA, 0.5x GlutaMax, 1 μg/ml mouse laminin, 10 ng/ml
460 BDNF, and 10 ng/ml NT3. The doxycycline was omitted from all subsequent media changes. Half of the
461 media was half changed weekly over the first two weeks, then the amount of media was doubled on day
462 21. Thereafter, a third of the media was replaced weekly until harvesting. 7 to 8 week old excitatory
463 neurons were harvested for library preparation.

464

465 Human hippocampal DG-like neurons were generated from dissociated hippocampal organoids
466 (unpublished). Briefly, WTC11 iPSCs were grown on MEF feeder cells and patterned towards a neural
467 ectoderm fate using dual SMAD inhibition as floating embryoid bodies (EBs) in media containing 20%
468 KnockOut Serum Replacement. Next, 4 week old EBs were patterned towards a hippocampal fate using
469 WNT and BMP in media containing 1x N-2. After patterning, organoids were dissociated using a neural
470 tissue dissociation kit (MiltenyiBiotech), plated on PDL- and laminin-coated plates, then cultured for 4
471 weeks in media containing 1x B-27, 10 ng/ml BDNF, 10 ng/ml GDNF, 0.5 mM cAMP, and 200 μM
472 ascorbic acid.

473

474 Human lower motor neurons were differentiated from WTC11 iPSCs with a doxycycline inducible
475 transgene expressing NGN2, ISL1, and LHX3 integrated at the AAVS1 safe-harbor locus (i^3LMN
476 iPSCs) as previously reported[10]. Briefly, i^3LMN iPSCs were maintained on growth factor reduced
477 Matrigel in StemFit media (Nacalai USA). On day 0, $1.5x10^6$ i^3LMN iPSCs were plated on 10-cm
478 dishes, followed 24 hours later by exchange into neural induction media containing doxycycline and

479 compound E. On day 3, the precursor cells were replated onto 12-well plates coated with poly-D-lysine
480 and laminin at a density of $2.5x10^5$ cells per well. From day 3 to day 4, the cells were treated with a
481 pulse of 40 μM BrdU for 24 hours to suppress the proliferation of undifferentiated cells. Media was
482 exchanged on day 4 and every three days thereafter. The cells were harvested 10 days post-
483 differentiation for library preparation.
484
485 Human primary astrocytes (P0) were purchased from ScienCell Research Laboratories (catalog #1800)
486 and cultured using the recommended media (catalog #1801). Briefly, cells were cultured in flasks coated
487 with poly-L-lysine ($2\mu/cm^2$) and passaged once using trypsin and EDTA before harvesting.
488
489 All cells used in the present study were verified as mycoplasma contamination free.
490
491 **Immunofluorescence**
492 Cells were fixed in 4% paraformaldehyde (PFA) for 15 minutes at room temperature, then washed
493 multiple times with PBS containing 0.1% Triton X-100 (PBS-T) before undergoing blocking using a
494 solution of 5% bovine serum albumin (BSA) in PBS-T for 1 hour at room temperature. Primary
495 antibodies against Cux1 (Abcam, ab54583, lot: GR3224721-2), MAP2 (Abcam, ab5392, lot:
496 GR3242762-1), PROX1 (Millipore, MAB5654, lot: 3075604), HB9 (Millipore, ABN174, lot: 3050643),
497 SMI32 (Abcam, ab7795, lot: GR299862-23), and GFAP (Abcam, ab7260, lot: GR3240356-1) were
498 diluted in 5% BSA solution and incubated overnight at 4°C prior to use. Secondary antibodies including
499 Alexa Fluor 568 goat anti-chicken IgG, Alexa Fluor 568 goat anti-mouse IgG, Alexa Fluor 488 donkey
500 anti-rabbit IgG, and Alexa Fluor 488 donkey anti-mouse IgG (Molecular Probes) were diluted in 5%
501 BSA solution and incubated for 1 to 2 hours at room temperature prior to use. Images were acquired
502 using a Leica TCS SP8 confocal microscope with a 40x oil immersion objective lens.
503
504 **Promoter capture Hi-C (pcHi-C)**
505 *In situ* Hi-C libraries for the excitatory neurons, hippocampal DG-like neurons, lower motor neurons,
506 and astrocytes were constructed from 1 to 2 million cells using HindIII as a restriction enzyme as
507 previously described[59]. pcHi-C was performed using biotinylated RNA probes prepared according to an
508 established protocol (Jung et al., under review). Briefly, sets of 120 bp probes with 30 bp overhangs
509 were designed to capture the sequences adjacent to restriction sites flanking each promoter-containing

510 HindIII fragment. Three probes were targeted to each side of each restriction site, such that a total of 12
511 probes targeted each promoter-containing HindIII fragment. In total, promoters (defined as the
512 sequences up to 500 bp upstream and downstream of each transcription start site) for 19,603 of the
513 20,332 protein coding genes in GENCODE 19 were captured using this approach. While noncoding
514 RNA promoters were not explicitly targeted by this design, HindIII fragments containing 3,311 of the
515 14,069 noncoding RNA promoters in GENCODE 19 were also baited by the probes.
516
517 To perform the hybridization, 500 ng of *in situ* Hi-C libraries were first mixed with 2.5 µg human Cot-1
518 DNA (Invitrogen #15279011), 2.5 µg salmon sperm DNA (Invitrogen #15632011), and 0.5 nmol each
519 of the p5 and p7 IDT xGen Universal Blocking Oligos in a volume of 10 µL, then denatured for 5 min at
520 95°C before holding at 65°C. Next, a hybridization buffer mix was prepared by mixing 25 µL 20x SSPE,
521 1 µL 0.5 M EDTA, 10 µL 50x Denhardt's solution, and 13 µL 1% SDS, followed by pre-warming to
522 65°C. Finally, 500 ng of the probe mix was combined with 1 µL 20 U/µL SUPERase-In (Invitrogen
523 #AM2696) in a 6 µL volume, pre-warmed to 65°C, then promptly mixed with the library and
524 hybridization buffer mix. The final solution was transferred to a humidified hybridization chamber and
525 incubated for 24 hours at 65°C. 0.5 mg Dynabeads MyOne Streptavidin T1 magnetic beads (Invitrogen
526 #65601) were used to pull down the captured fragments in a binding buffer consisting of 10 mM Tris-
527 HCl pH 7.5, 1 M NaCl, and 1 mM EDTA. Next, the beads were washed once with 1x SSC and 0.1%
528 SDS for 30 minutes at 25°C, followed by three washes with pre-warmed 0.1X SSC and 0.1% SDS for
529 10 minutes each at 65°C. The final library was eluted in 20 µL nuclease-free water, amplified, then sent
530 for paired-end sequencing on the HiSeq 4000 (50 bp reads), the HiSeq X Ten (150 bp reads), or the
531 NovaSeq 6000 (150 bp reads).
532
533 A detailed description of the capture probe design and experimental procedures can be viewed in the
534 attached manuscript which is under review in Nature.
535
536 **ATAC-seq**
537 ATAC-seq was carried out as previously described using the Nextera DNA Library Prep Kit (Illumina
538 #FC-121-1030)[60]. First, frozen or fresh cells were washed once with ice cold PBS containing 1x
539 protease inhibitor before being exchanged into ice cold nuclei extraction buffer (10 mM Tris-HCl pH
540 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% Igepal CA630, and 1x protease inhibitor) and incubated for 5

541 minutes on ice. Next, 50,000 cells were counted out, exchanged into 1x Buffer TD, then incubated with
542 2.5 µL TDE1 for 30 minutes at 37°C with shaking. The transposed DNA was purified using Qiagen
543 MinElute spin columns, amplified with Nextera primers, then size-selected for fragments between 300
544 and 1000 bp in size using AMPure XP beads. Libraries were sent for single-end sequencing on the
545 HiSeq 4000 (50 bp reads) or paired-end sequencing on the NovaSeq 6000 (150 bp reads). Sequencing
546 reads were mapped to hg19 and processed using the ENCODE pipeline
547 (https://github.com/kundajelab/atac_dnase_pipelines) running the default settings. Only the first read
548 was used, and all sequencing reads were trimmed to 50 bp prior to mapping. Open chromatin peaks
549 called by the pipeline were expanded to a minimum width of 500 bp for all downstream analyses. Peaks
550 overlapping coding gene or noncoding RNA promoters were assigned as promoter open chromatin
551 peaks, while all other peaks were assigned as distal open chromatin peaks. All data processing metrics
552 are reported in **Supplementary Table 1**.
553
554 **RNA-seq**
555 RNA was extracted using the RNeasy Mini Kit (Qiagen #74104). Approximately 500 ng of extracted
556 RNA was used to construct libraries for sequencing using the TruSeq Stranded mRNA Library Prep Kit
557 (Illumina #20020594). Libraries were sent for single-end sequencing on the HiSeq 4000 (50 bp reads) or
558 paired-end sequencing on the NovaSeq 6000 (150 bp reads). Raw sequencing reads were aligned to
559 hg19/GRCh37 using STAR running the standard ENCODE parameters, and transcript quantification
560 was performed in a strand-specific manner using RSEM with the annotation from GENCODE 19. Only
561 the first read was used, and all sequencing reads were trimmed using TrimGalore 0.4.5 running the
562 following options: -q 20 --length 20 --stringency 3 --trim-n. The *edgeR* package in R was used to
563 calculate TMM-normalized RPKM values for each gene based on the expected counts and gene lengths
564 for each replicate reported by RSEM. The mean gene expression across all replicates was used for each
565 cell type. All data processing metrics are reported in **Supplementary Table 1**.
566
567 **CUT&RUN**
568 CUT&RUN libraries for excitatory neurons and lower motor neurons were constructed for 100,000 to
569 250,000 cells with antibodies for H3K27ac and CTCF as previously described[24]. First, cells were lysed
570 in nuclei extraction buffer (20 mM HEPES-KOH pH 7.9, 10 mM KCl, 1 mM $MgCl_2$, 0.1% Triton X-
571 100, 20% glycerol, and 1x protease inhibitor) for 10 minutes on ice. Next, samples were spun down and

572 washed twice with nuclei extraction buffer before being resuspended in 100 µL nuclei extraction buffer.
573 10 µL of Concanavalin A-coated beads previously washed and resuspended in binding buffer (1x PBS, 1
574 mM $CaCl_2$, 1 mM $MgCl_2$, and 1 mM $MnCl_2$) were then added to the samples and incubated with rotation
575 for 15 min at 4°C. Next, samples were washed once each with Buffer 1 (20 mM HEPES-KOH pH 7.9,
576 150 mM NaCl, 2 mM EDTA, 0.5 mM spermidine, 0.1% BSA, and 1x protease inhibitor) and Buffer 2
577 (20 mM HEPES-KOH pH 7.9, 150 mM NaCl, 0.5 mM spermidine, 0.1% BSA, and 1x protease
578 inhibitor) before being resuspended in 50 µL of Buffer 2 containing 0.5 µL antibody (H3K27ac from
579 Active Motif, 39122, lot: 22618011 and CTCF from Millipore 07-729, lot: 305960) and incubating for at
580 least 2 hours with rotation at 4°C. Following the incubation, samples were washed twice with Buffer 2
581 before being incubated in 50 µL of Buffer 2 containing ~700 ng/mL protein A-MNase fusion protein
582 (Batch #6 from the Henikoff Lab) for 1 hour with rotation at 4°C. Samples were washed two more times
583 and resuspended in 100 µL of Buffer 2 before starting the MNase digestion by adding $CaCl_2$ to a
584 concentration of 2 mM (with the samples kept on ice), followed 30 minutes thereafter by the addition of
585 100 µL 2X Stop Buffer (200 mM NaCl, 20 mM EDTA, 4 mM EGTA, 50 ug/mL RNase A, 40 ug/mL
586 glycogen, and 2 pg/mL spike-in DNA) to inactivate the MNase. Samples were incubated for 20 min at
587 37°C and spun down for 5 minutes at 4°C to release DNA fragments that were subsequently extracted
588 from the supernatant using Qiagen MiniElute spin columns. Libraries were prepared using TruSeq
589 adapters and size-selected using SPRIselect beads before being amplified and sent for paired-end
590 sequencing on the NovaSeq 6000 (150 bp reads). Sequencing reads were first trimmed to 50 bp using
591 fastp then mapped to hg19 using bowtie2 running the following options: --local --very-sensitive-local --
592 no-mixed --no-discordant -I 10 -X 700. Picard Tools was used to remove duplicate reads, and MACS2
593 was used to call peaks on merged replicates at an FDR cutoff of 5%.
594
595 **Validation of PIRs using CRISPR deletion**
596 To validate genomic interactions captured by pcHi-C, candidate PIRs were targeted for CRISPR
597 deletion in i^3N iPSCs. At each locus of interest, we designed pairs of sgRNAs to delete the putative
598 regulatory element as localized by open chromatin peaks in the candidate PIR. All sgRNAs were
599 synthesized by Synthego, and Cas9 protein was sourced from QB3-Berkeley. To generate deletion lines,
600 CRISPR/Cas9 nucleofections were performed using the LONZA Human Stem Cell Nucleofector® Kit.
601 For each nucleofection, 500,000 i^3N iPSCs were transfected with Cas9:sgRNA RNP complex
602 (consisting of 12 µg Cas9, 10 µg sgRNA 1, and 10 µg sgRNA 2) using program "A-023" on the LONZA

603 4D-Nucleofector. The cells were then seeded onto Matrigel-coated 6-well plates containing Essential
604 8™ Medium (ThermoFisher #A15169-01) with added Y-27632 for recovery following nucleofection.
605 After 48 hours, the cells were split into new 6-well plates at a concentration of approximately 50 cells
606 per well for picking single colonies. Clones picked from the 6-well plates containing homozygous
607 deletions were confirmed by PCR and induced into excitatory neurons for quantifying the expression of
608 genes targeted by the deleted elements. We used three independent deletion clones for each experiment,
609 and clones with wild type genotypes were used as controls. To perform the quantification, total RNA
610 from the excitatory neurons was extracted using a Qiagen AllPrep DNA/RNA Mini Kit, and cDNA was
611 synthesized using a Bio-RAD iScript™ cDNA Synthesis Kit. qPCR for targeted genes was performed
612 with FastStart Essential DNA Green Master reaction mix (Roche) on the LightCycler® 96 System
613 (Roche). All CRISPR deletion experiments were performed with two independent transfections.
614 Detailed information on all the primers used can be found in **Supplementary Table 12**.
615
616 **Validation of PIRs using CRISPRi**
617 Excitatory neurons induced from i^3N iPSCs were infected with lentivirus carrying dCas9-KRAB-blast
618 (Addgene #89567), and colonies with high expression of dCas9 were picked. The CROP-seq-opti vector
619 (Addgene #106280) was used for sgRNA expression. sgRNAs were designed, cloned, and cotransfected
620 with lentivirus packaging plasmids pMD2.G (Addgene #12259) and psPAX (Addgene #12260) into
621 293T cells using PolyJet (SignaGen Laboratories #SL100688) according to the manufacturer's
622 instructions. Virus-containing media was collected for 72 hours, filtered through a 0.45 μm filter
623 (Millipore #SLHV033RS), and concentrated using an Amicon Ultra centrifugal filter (Millipore
624 #UFC801024). The virus was titrated into the excitatory neurons by qPCR 72 hours after infection. The
625 internal control for qPCR targeted an intronic region (forward primer:
626 TCCTCCGGAGTTATTCTTGGCA, reverse primer: CCCCCCATCTGATCTGTTTCAC). Integration
627 of the WPRE fragment was quantified in comparison with a control cell line containing a known copy
628 number of WPRE. For CRISPRi silencing of putative regulatory elements, excitatory neurons were
629 treated with lentivirus containing sgRNAs (MOI ~3). Cells were collected for mRNA extraction 7 days
630 following transfection, and the expression of target genes was determined by qPCR. All CRISPRi
631 experiments were performed in triplicate, with three technical replicates per experiment. Detailed
632 information on all the primers used can be found in **Supplementary Table 12**.
633

634 **Reproducibility and saturation analysis**

635 We took pcHi-C contact matrices generated at 10 kb resolution using HiC-Pro 2.11.0 with the following
636 settings (MIN_MAPQ=20, MIN_FRAG_SIZE=100, MAX_FRAG_SIZE=5000000,
637 MIN_INSERT_SIZE=100, MAX_INSERT_SIZE=1200, and reporting only bin pairs that are baited on
638 at least one end with our pcHi-C probes, with all other settings set to their default values) and calculated
639 the pairwise stratum adjusted correlation coefficient (SCC) between replicates across all cell types using
640 HiCRep 1.4.0 on chromosome 1 (h=20 and only considering contacts with distances below 5 Mb). SCCs
641 evaluated on the other chromosomes closely resembled the results for chromosome 1 (data not shown).
642 Hierarchical clustering for the pairwise SCC values was performed using the Seaborn *clustermap*
643 function in Python. Pairwise correlation heatmaps and clustering dendrograms for ATAC-seq replicates
644 were generated by counting reads overlapping a set of consensus peaks using the *DiffBind* package in R,
645 with the set of consensus peaks defined as peaks occurring in at least two replicates across all cell types
646 (minOverlap=2). Pairwise distance estimates and clustering dendrograms for RNA-seq replicates were
647 generated using the *DESeq2* package in R. For saturation analysis, we first downsampled all pcHi-C
648 libraries to 5%, 10%, 20%, 40%, 60%, 80%, and 100% of the final sequencing depths used in the study.
649 Next, we computed pairwise SCCs between all pairs of biological replicates using HiCRep at these
650 downsampled sequencing depths.

651

652 **Calling significant promoter-PIR interactions**

653 Paired-end sequencing reads were first trimmed using fastp running the default settings before being
654 mapped, filtered, and deduplicated using HiCUP v0.71 with bowtie2 and filtering for ditags between
655 100 and 1200 bp[61]. In addition, the sequencing depths of all libraries was normalized so that each
656 replicate had the same number of usable reads, or the number of on-target cis pairs interacting over a
657 distance of 10 kb. Significant promoter-PIR interactions were called using CHiCAGO running the
658 default settings, retaining baited fragments that are supported by at least 250 reads (minNPerBaits=250).
659 Promoter-PIR interactions between HindIII fragments with a score (negative log p-value) of 5 or greater
660 in each cell type were determined to be significant. All data processing metrics are reported in
661 **Supplementary Table 1**. In cases where CHiCAGO reported the same interaction twice due to
662 directionality between two bait-containing fragments (i.e. bait A to bait B, bait B to bait A), the two
663 interactions were merged, retaining the more significant score of the two interactions. Interchromosomal
664 interactions were also omitted from the analysis. To call overlaps between our sets of significant

interactions and genomic and epigenomic features including promoters, open chromatin peaks, chromatin states, disease-associated variants, and eQTLs, interacting bins were expanded to a minimum width of 5 kb or retained as the original widths of the HindIII fragments if they exceeded 5 kb. Interactions involving HindIII fragments larger than 100 kb were omitted from our analysis due to low resolution. An interaction was considered to be shared between cell types if both its interacting ends intersected the corresponding ends of an interaction in another cell type. Otherwise, an interaction was classified to be cell type-specific.

Chromatin state analysis

Annotations for the publicly available 15 state ChromHMM model were downloaded from the Roadmap Epigenomics Project for the dorsolateral prefrontal cortex (E073, "Brain Dorsolateral Prefrontal Cortex"), hippocampus (E071, "Brain Hippocampus Middle"), and normal human astrocytes (E125, "NH-A Astrocytes Primary Cells"). The states were available at a resolution of 200 bp and grouped as follows: TssA and TssAFlnk were merged as TSS, TxFlnk, Tx, and TxWk were merged as Tx, EnhG and EnhBiv were merged as other enhancer, and ReprPC and ReprPCWk were merged as ReprPC. All other states (Enh, ZNF/Rpts, Het, TssBiv, and BivFlnk) were used as is. Enrichment analysis was performed for each cell type by counting the number of chromatin states overlapping significant PIRs versus the number of chromatin states overlapping randomly shuffled PIRs with matching distance distributions. A total of 100 sets of randomly shuffled PIRs were sampled in each case.

GO enrichment analysis

Protein coding and noncoding RNA genes from GENCODE 19 participating in significant cell type-specific promoter-PIR interactions were used for cell type-specific GO enrichment analysis. Only genes participating in interactions between promoter-containing and non-promoter-containing bins with a promoter open chromatin peak on one end and a distal open chromatin peak on the other end were used. The promoter open chromatin peaks were used to define the genes with promoters interacting with cell type-specific PIRs. A minimum normalized RPKM of 0.5 was used to filter out genes not significantly expressed in each cell type, and the resulting gene lists were input into Enrichr. Enriched GO terms from the "GO Biological Process 2018" ontology are reported according to their combined scores (calculated by multiplying the log of the p-value by the z-score of the deviation from the expected rank). For disease-specific GO enrichment analysis, target genes across all cell types were combined and input into

696 Enrichr. All promoters in GENCODE 19 were included in this analysis . The top 100 enriched cell type-
697 specific and disease-specific GO terms for each category and their raw p-values are reported in
698 **Supplementary Table 3** and **Supplementary Table 9**.
699
700 **Motif enrichment analysis**
701 We took the sets of all cell type-specific distal open chromatin peaks participating in significant
702 promoter-PIR interactions between promoter-containing and non-promoter-containing bins for each cell
703 type, and used the sequences in 250 bp windows around the peak summits to perform motif enrichment
704 analysis using HOMER running the default settings. The entire genome was used as a background.
705 Significance and expression values for each detected motif and its corresponding TFs are reported in
706 **Supplementary Table 3**. Entries with similar or identical consensus TF motif sequences were grouped
707 for brevity.
708
709 **VISTA enhancer analysis and target gene identification**
710 Human enhancer regions and mouse enhancer regions with orthologous human sequences associated
711 with positive annotations in the VISTA Enhancer Browser were downloaded and analyzed for overlap
712 with our sets of significant promoter-PIR interactions for each cell type. Of the 2,956 total tested
713 elements in their database (January 2019), 1,568 were found to be positive (976 were human elements
714 and 892 were mouse elements with orthologous human sequences). Positive elements (expanded to a
715 minimum width of 5 kb) found to participate in significant interactions are reported in **Supplementary**
716 **Table 5**. For determining whether positive elements interacted with their nearest genes or with more
717 distal genes, we only considered protein coding and noncoding RNA genes in GENCODE 19. To
718 evaluate cases where interactions between positive elements and their nearest genes were unresolvable
719 ("same fragment ambiguity"), we determined if a promoter for the nearest gene overlapped at least one
720 HindIII fragment that the expanded positive element did not also overlap. The following were
721 considered to be neural annotations: neural tube, hindbrain, cranial nerve, midbrain, forebrain,
722 mesenchyme derived from neural crest, dorsal root ganglion, and trigeminal V.
723
724 **SNP enrichment analysis and target gene identification**
725 GWAS SNPs for a total of eleven neurological disorders including Alzheimer's disease (AD), attention
726 deficit hyperactivity disorder (ADHD), amyotrophic lateral sclerosis (ALS), autism spectrum disorder

727 (ASD), bipolar disorder (BD), epilepsy (EP), frontotemporal dementia (FTD), mental processing (MP),
728 Parkinson's disease (PD), and schizophrenia (SCZ), and unipolar depression (UD) were mined from the
729 GWAS Catalog (December 2018) using a p-value threshold of 10^{-6}. The GWAS SNPs were imputed
730 using HaploReg v4.1 at an LD threshold of 0.8 according to the reported study population(s) for each
731 SNP. The imputed SNPs were lifted over to hg19 and filtered for unique SNPs by position. See
732 **Supplementary Table 5** for a detailed summary of the imputation process and the list of studies used.
733 Disease- and cell type-specific enrichment for SNPs was calculated by taking the ratio of the number of
734 SNPs overlapping significant PIRs over the mean number of SNPs over the number of SNPs
735 overlapping randomly shuffled PIRs with matching distance distributions. A total of 100 sets of
736 randomly shuffled PIRs were sampled in each case. To determine whether a GWAS SNP potentially
737 interacted with a target gene, we determined whether it or any of its linked SNPs (expanded to a
738 minimum width of 1 kb) interacted with a promoter for the nearest gene. To evaluate cases where
739 interactions between GWAS SNPs and their nearest genes were unresolvable ("same fragment
740 ambiguity"), we determined if a promoter for the nearest gene overlapped at least one HindIII fragment
741 that a GWAS SNP or any of its linked SNPs did not also overlap. Finally, we derived a list of SNPs for
742 which the SNP was located within 2 kb of the center of an open chromatin peak at a PIR, indicating
743 strengthened evidence for a functional regulatory variant at that locus ("putative regulatory SNPs").
744

745 **Phasing of the WTC11 genome**

746 The raw WTC11 genome sequence can be downloaded from http://genome.ucsc.edu/cgi-
747 bin/hgTracks?db=hg38&hubClear=https://s3-us-west-
748 2.amazonaws.com/downloads.allencell.org/genome-sequence/ucsc_hubs/WTC_genome_hub/hub.txt.
749 Phasing of the WTC11 genome was performed as previously described[47]. Briefly, WTC11 variants were
750 first split by chromosome, and phase-informative pcHi-C reads were extracted using extractHAIRS with
751 the minimum mapping quality set to 10 and the maximum insert size set to 30000000 bp[62]. Phasing was
752 performed using Hapcut with a maximum of 101 iterations. Next, we extracted the maximum variants
753 phased (MVP) haplotype block from the output of Hapcut to use as the seed haplotype. We modified the
754 "neighborhood correction" aspect of phasing by filtering phased variants whose predicted phase would
755 have a marginal probability below 0.99 using an in-house implementation of a hidden Markov model
756 (HMM) as described previously[63,64] with a reference haplotype set from the 1000 Genomes
757 Project. Finally, missing variants were imputed using the same HMM with the reference haplotype set

758 from the 1000 Genomes Project. The chromosome-wide SNP phasing data is available at the Gene
759 Expression Omnibus under the following accessible number: GSE113483.
760
761 **Allelic bias analysis**
762 We utilized the phasing data for the WTC11 genome along with the allele-specific mapping capabilities
763 of HiC-Pro to quantify genome-wide allelic bias between significantly interacting 10 kb bins in
764 excitatory neurons and lower motor neurons. We selected these two cell types because they used
765 homogenous induction of TFs for differentiation, therefore minimizing the noise introduced by
766 conventional differentiation techniques. Briefly, reads were mapped using bowtie to a version of the
767 hg19 reference genome where all sites containing heterozygous phased SNPs were first N-masked. The
768 unfiltered HiC-Pro contact maps were used for this analysis. Next, nucleotides at the masked
769 polymorphic sites were used to assign the reads to either allele, with reads containing conflicting allele
770 assignments or unexpected bases omitted from further analysis. Read pairs with at least one allele-
771 specific mate were used to construct allele-specific Hi-C contact maps at 10 kb resolution, for which
772 interacting bins overlapping with the set of significant promoter-PIR interactions with score ≥ 3 was
773 used to detect bias. Only interacting bins with 10 or more reads across both alleles were kept. A two-
774 sided binomial test was performed to assess allelic bias for each pair of interacting bins, and the
775 resulting p-values were adjusted using the BH method to filter out significantly biased loci at an FDR
776 cutoff of 5%. All allelically biased interactions with p-values < 0.001 are reported in **Supplementary**
777 **Table 10**.
778
779 **eQTL enrichment analysis**
780 1D enrichment of significant eQTLs with an FDR cutoff of 5% from GTEX V7 at significant versus
781 randomly shuffled PIRs in matched tissue types for excitatory neurons (Brain - Cortex, n=136) and
782 hippocampal DG-like neurons (Brain - Hippocampus, n=111) was performed similarly to the chromatin
783 state and SNP enrichment analysis. Overall, we found that 4.7% of significant cortical eQTLs and 6.7%
784 of hippocampal eQTLs interact in excitatory neurons and hippocampal DG-like neurons, respectively.
785 To determine the 2D enrichment of eQTL-TSS pairs in our significant interaction sets, we first filtered
786 out eQTL-TSS pairs that were within 10 kb of each other or on the same HindIII fragment as this would
787 be below the minimum detectable resolution by pcHi-C. Next, we sampled a set of nonsignificant
788 eQTL-TSS pairs with a matching distance distribution as the set of significant eQTL-TSS pairs for each

789 cell type. We also controlled for the number of genes around which the eQTL-TSS pairs were centered.
790 Finally, we compared the distributions of scores for significant interactions supporting the significant
791 and nonsignificant sets of eQTL-TSS pairs by overlapping the eQTL-TSS pairs with our significant
792 interactions.
793
794 **Epistasis analysis**
795 In order to determine whether genetic variation in physically interacting regions might contribute to
796 neurodevelopmental disorders via genetic interactions, we utilized GWAS data for ASD. For epistasis
797 testing, we needed individual-level genotype data, so we used a dataset of 4,109 trios and 4,471,807
798 imputed and genotyped single nucleotide polymorphisms (SNPs), as previously reported[56]. This dataset
799 includes publicly available ASD GWAS data [Autism Genetic Resource Exchange (AGRE), Autism
800 Genome Project (AGP), Simons Simplex Collection (SSC)] in addition to in-house generated data
801 [University of California, San Francisco (UCSF)], harmonized, imputed, and quality controlled (QC+)
802 by us, as previously described[56]. These data comprised trios with one ASD-affected offspring and both
803 parents selected for homogeneous genetic ancestry by multidimensional scaling with PLINK[65]. As our
804 data were family-based and did not include unrelated controls, we used non-transmitted parental alleles,
805 commonly known as pseudo-controls, generated using the *-tucc* option in PLINK. These 4,109 pseudo-
806 controls are perfectly matched to ASD cases for ancestry, thereby serving as a control for any population
807 confounding.
808
809 For each significant promoter-PIR interaction for each cell type, we extracted all independent SNPs in
810 each region from our QC+ imputed GWAS data. We performed a case-only test for pairwise epistasis
811 using the *-fast-epistasis, case-only,* and *set-by-set* options in PLINK with $p \leq 10^{-7}$ as a default threshold
812 in ASD cases and in matched pseudo-controls[56]. Across the interacting loci in the four neural cell types,
813 we performed approximately 19.7 million epistasis tests and 12,637,825 SNP pairs showed $p \leq 10^{-7}$
814 (65% tests performed).
815
816 Because all pairs of regions we tested were on the same chromosome (linked), we expected an excess of
817 false positives (e.g. 65% at $p \leq 10^{-7}$) for the case-only test due to LD or haplotypes containing rare
818 variants. We thus excluded all SNP pairs that showed epistasis with $p \leq 10^{-7}$ in both cases and pseudo-
819 controls to generate case-specific and control-specific epistasis results for comparison, resulting in 0.5%

820 of results with $p \leq 10^{-7}$. We next applied the *–clump* PLINK option to filter SNPs involved in potential
821 epistasis in LD ($r^2 > 0.2$) across pairs (e.g. one promoter SNP putatively interacting with several
822 correlated PIR SNPs), such that only one pair (selected based on epistatic *p*-value) represented the
823 combination of loci. This resulted in reduction of 15 to 19% of the number of pairs considered. The
824 average distance between epistatic SNPs was 104 kb in excitatory neurons, 97 kb in hippocampal DG-
825 like neurons, 87 kb in lower motor neurons, and 90 kb in astrocytes.
826
827 To test for enrichment of epistasis signal specific to ASD, we wanted to compare the number of SNP
828 pairs at $P \leq 10^{-7}$ between the matched cases and pseudo-controls at various signal-to-noise ratios. We
829 divided the set of case-specific epistatic SNP pairs and the set of control-specific epistatic SNP pairs into
830 bins based on *P*-values [$10^{-30} \leq p \leq 10^{-7}$] and a homogeneous number of SNP pairs to distinguish
831 expected signal-to-noise (**Fig. 6e, Supplementary Table 7**). We then compared the number of case-
832 specific and control-specific SNP pairs in each significance category using a proportion test in R. We
833 performed meta-analysis of the case-specific and pseudo-control specific results across the four cell
834 types using the *metafor* package in R. We illustrated the odds ratios of enrichment with the *forestplot*
835 package in R.
836
837 As a negative control, we utilized equivalent promoter-PIR interactions with scores < 1 across the neural
838 cell types. We sampled the same number of non-significant promoter-PIR pairs that we used in the main
839 analysis for each cell type. For each non-significant interaction in each cell type, we filtered out those
840 that overlapped interactions with score > 3 in any of the neural cell types. We also made sure to sample
841 nonsignificant interactions with a similar interaction distance distribution as the significant interactions.
842 We then performed SNP extraction, epistasis testing, and case-control enrichment analysis as described
843 above for the nonsignificant interactions. Before filtering, 15-19% of SNPs showed $p \leq 10^{-7}$, and after
844 filtering 7% met this threshold.
845
846 **Code availability statement**
847 A copy of the custom code used for all the data analysis and figure generation in this study can be
848 viewed and downloaded at the following GitHub repository: https://github.com/stayingsong/brain_pchic
849
850 **Data availability statement**

All datasets used in this study (pcHi-C, ATAC-seq, RNA-seq, CUT&RUN, and chromosome-wide SNP phasing data) are available at the Gene Expression Omnibus under the accession number GSE113483. The reviewer access token is mjmrcsuaddkfhut.

Data can also be visualized on the WashU Epigenome Browser at the following link: http://epigenomegateway.wustl.edu/legacy/?genome=hg19&session=zEdB7v5de4&statusId=33592151

Tracks include ATAC-seq signal, RNA-seq plus/minus strand signal, CTCF CUT&RUN signal, and promoter-PIR interactions with score ≥ 5 for each cell type. HindIII fragments, positive Vista elements, GENCODE 19 genes, and SNPs for each disease are also shown.

Acknowledgements

We thank Anthony Schmitt and Bing Ren for sharing pcHi-C probes and the pcHi-C protocol. Genomic analysis of the WTC11 line in this study was made possible by the whole genome sequencing data generated by the Allen Institute for Cell Science. We thank the Institute and its founder Paul G. Allen for making this work possible. We also thank G. Hon and S. Henikoff for providing reagents and Y. Guo, N. Ahituv, R. D. Hawkins, M. McManus, and B. Ren for providing critical feedback on the manuscript. This work was made possible in part by NIH-NEI EY002162 (Core Grant for Vision Research and the Research to Prevent Blindness Unrestricted Grant). This work was supported by the US National Institute of Health (NIH) grants R01AG057497 to YS, LG, and HS, R01EY027789 and UM1HG009402 to YS, the UCSF Weill Institute for Neuroscience Innovation Award, the Hillblom Foundation, and the American Federation for Aging Research New Investigator Award in Alzheimer's Disease to YS. R01EY028249, R01HL130533, R01-HL13535801 to BRC as well as P01NS097206 and U19MH106434 to HS. R01MH105128, R35NS097370, and U19AI131130 to GLM. MS is supported by T32GM007175. FJ is supported by an NIH training grant T32GM73009.

Author contribution

MS and YS designed the study. MS, XY, XR, LM, IJ, KJ and TWT performed the experiments. MS, BL, and IJ performed data analysis. LW supervised the epistasis analysis by MT. JD contributed to genomic phasing using HaploSeq. SL, JY, KW, BRC, FJ, GLM, HS, LD, CW, and LG contributed by providing cell samples. MS and YS prepared the manuscript with assistance from all authors.

882

883 **Competing interests**

884 The authors declare no competing financial interests.

885

886 **Figure legends**

887 **Figure 1. Genome-wide mapping of physical chromatin interactions in functionally distinct neural**
888 **cell types.**

889 (**a**) Schematic of the study design for generating four functionally distinct cell types in the CNS and
890 performing integrative analysis of chromatin interactions using promoter capture Hi-C, open chromatin
891 regions using ATAC-seq, and transcriptomes using RNA-seq. For pcHi-C, we used 3, 2, 3, and 4
892 biological replicates respectively for the excitatory neurons, hippocampal DG-like neurons, lower motor
893 neurons, and astrocytes. For ATAC-seq, we used 2, 2, 3, and 4 biological replicates respectively for the
894 cell types. For RNA-seq, we used 2, 2, 2, and 4 biological replicates respectively for the cell types. (**b**)
895 Proportions of interactions occurring within TADs for each cell type. (**c**) Histogram and empirical CDF
896 plots of interaction distances for each cell type. (**d**) Proportions of interactions between promoter-
897 containing bins (blue) and promoter- and non-promoter-containing bins (purple) for each cell type. (**f**)
898 Proportions of cell type-specific (blue) and shared (grey) distal open chromatin peaks at PIRs for each
899 cell type.

900

901 **Figure 2. Integrative analysis of chromatin interactions, epigenomic features, and gene expression.**

902 (**a**) Histograms of the numbers of PIRs interacting with each promoter in each cell type. The means are
903 indicated with red lines. Only promoters interacting with at least one PIR are included (15,316
904 promoters in excitatory neurons, 19,546 promoters in hippocampal DG-like neurons, 14,990 promoters
905 in lower motor neurons, and 15,397 promoters in astrocytes, out of a total of 34,401 protein coding and
906 noncoding RNA promoters in GENCODE 19). (**b)** Bar plots showing counts of epigenomic chromatin
907 states (inferred at a 200 bp resolution using the ChromHMM core 15 state model in matched tissues)
908 overlapping significant (solid bars) versus randomly shuffled (striped bars) PIRs for each cell type. Error
909 bars represent the standard deviation over 100 sampled sets of randomly shuffled PIRs. No matching
910 tissue data was available for the lower motor neurons so they were omitted from the analysis. (**c**)
911 Comparative gene expression analysis across all cell types for expressed genes (normalized RPKM >
912 0.5) whose promoters interact exclusively with either enhancer-PIRs (n=6836) or repressive-PIRs

913 (n=2612). Distributions of gene expression values are shown for each group. (**d**) Boxplots showing
914 distributions of gene expression values across all cell types for expressed genes (normalized RPKM >
915 0.5) grouped according to the numbers of interactions their promoters form with enhancer-PIRs. Linear
916 regression was performed on the mean gene expression values for each bin. Only bins containing at least
917 10 genes were included in the analysis.
918
919 **Figure 3. Cell type-specific PIRs and TF motif enrichment analysis.**
920 (**a**) Classification of unique promoter-PIR interactions with interaction score ≥ 5 in at least one cell type
921 into specificity categories based on their scores in each cell type. The numbers of promoter-PIR
922 interactions in each specificity category are summarized in **Supplementary Fig. 3a**. Cell types are also
923 hierarchically clustered based on their interaction scores over all interacting loci. (**b**) Top enriched GO
924 terms from the "GO Biological Process 2018" ontology in Enrichr for genes whose promoters
925 participate in cell type-specific interactions with distal open chromatin peaks in each cell type (groups 1-
926 4). Also shown are top enriched GO terms for genes participating in shared interactions across all cell
927 types (group 15). Enriched GO terms are ranked by their combined scores (calculated by multiplying the
928 log of the p-value by the z-score of the deviation from the expected rank). An expanded list of enriched
929 GO terms and their raw p-values is available in **Supplementary Table 3**. (**c**) Enrichment of consensus
930 TF motif sequences at open chromatin peaks in cell type-specific PIRs using HOMER, organized by
931 motifs (rows) and cell types (columns). The color of each dot represents the degree of enrichment
932 (negative log p-value) for each motif in each cell type, while the size of each dot represents gene
933 expression (normalized RPKM) for the corresponding TFs for each motif. Entries with similar or
934 identical consensus TF motif sequences are grouped for brevity.
935
936 **Figure 4. Validation of PIRs in human neural cells.**
937 (**a**) *In vivo* validated enhancers with neural annotations overlap a significantly higher proportion of open
938 chromatin peaks in the neural cells compared to enhancers with non-neural annotations (chi-squared test,
939 $p<2.2x10^{-16}$). (**b**) Pie chart showing counts of *in vivo* validated enhancers with human sequences
940 participating in chromatin interactions (589 out of 1568 total elements). Counts of interacting positive
941 enhancer elements with neural and non-neural annotations are also shown. (**c**) Counts of interacting
942 positive enhancer elements interacting exclusively with their nearest genes (blue), interacting
943 exclusively with more distal genes (pink), or interacting with both their nearest genes and more distal

944 genes (orange). Positive enhancer elements that could not be resolved for interactions with their nearest
945 genes are also shown (grey). The number of regulatory targets interacting with positive enhancer
946 elements in each category is shown on the right. (**d**) Promoter-PIR interactions at the *CDK5RAP3* locus.
947 Open chromatin peaks in PIRs up to 40 kb downstream of *CDKRAP3* (regions 1, 2, and 3, yellow
948 highlight) interact with the promoter of *CDK5RAP3* in a cell type-specific manner. Notably, only
949 regions 1 and 2 participate in interactions with the promoter of *CDK5RAP3* in excitatory neurons. In
950 addition, both *in vivo* validated enhancers (pink) and CTCF binding sites in excitatory neurons (dark
951 blue) are shown to be localized to all three candidate regulatory regions. All interactions fall within a CP
952 TAD (chr17:45,920,000-47,480,000). (**e**) LacZ staining in mouse embryos shows tissue-specific patterns
953 of enhancer activity. (**f**) CRISPRi silencing of region 1 using two independent sgRNAs results in
954 significant downregulation of *CDK5RAP3* expression in excitatory neurons (two sample t-test, two-
955 sided, p=3.2x10^{-8}). No significant downregulation was detected for the neighboring genes *MRPL10*,
956 *PNPO*, and *NFE2L1*. Each CRISPRi experiment was performed in triplicate, with three technical
957 replicates per experiment. (**g**) CRISPRi silencing of region 2, but not region 3, results in significant
958 downregulation of *CDK5RAP3* expression in excitatory neurons (two sample t-test, two-sided,
959 p=2.2x10^{-3}).
960
961 **Figure 5. Genetic analysis of promoter-PIR interactions with complex neurological disorder-**
962 **associated variants.**
963 (**a**) Enrichment analysis for complex neurological disorder-associated SNPs in Alzheimer's disease
964 (AD), attention deficit hyperactivity disorder (ADHD), amyotrophic lateral sclerosis (ALS), autism
965 spectrum disorder (ASD), bipolar disorder (BP), epilepsy (EP), frontotemporal dementia (FTD), mental
966 processing (MP), Parkinson's disease (PD), schizophrenia (SCZ), and unipolar depression (UD). The
967 color and size of each dot respectively represent the enrichment p-value and raw fold enrichment
968 (calculated as the number of SNPs overlapping significant PIRs divided by the mean number of SNPs
969 overlapping randomly shuffled PIRs across 100 sampled sets) for each cell type and disease pairing. (**b**)
970 Proportions and total counts of GWAS SNPs with at least one SNP in linkage participating in chromatin
971 interactions. Cell type-specific SNPs for excitatory neurons (blue), hippocampal DG-like neurons
972 (orange), lower motor neurons (yellow), and astrocytes (green) are highlighted. (**c**) Counts of GWAS
973 SNPs across all diseases with at least one SNP in linkage interacting exclusively with their nearest genes
974 (scenario III, blue), interacting exclusively with more distal genes (scenario I, pink), or interacting with

975 both their nearest genes and more distal genes (scenario II, orange). GWAS SNPs that could not be
976 resolved for interactions with their nearest genes are also shown (scenario IV, grey). Counts of
977 regulatory targets interacting with GWAS SNPs in each scenario are shown on the right. (**d**) PIRs
978 containing MP SNPS (yellow highlight) in an intron for *PTPRO* interact with the promoter of *STRAP*
979 over 300 kb away. All interactions fall within a CP TAD (chr12:14,960,000-16,040,000). Homozygous
980 deletion of this PIR in three independent clones results in significant downregulation of *STRAP*
981 expression in excitatory neurons (two sample t-test, two-sided, p=9.4x10^{-13}). Error bars represent the
982 SEM. (**e**) A PIR containing SCZ SNPs interacts with the *DRD2* promoter 20 kb upstream of the PIR. All
983 interactions fall within a CP TAD (chr11:113,200,000-114,160,000). Mono-allelic deletion of this PIR
984 in three independent clones results in significant downregulation of *DRD2* expression in excitatory
985 neurons (two sample t-test, two-sided, p=2.7x10^{-7}). Error bars represent the SEM.
986
987 **Figure 6. Genetics variants contribute to chromatin interaction bias and alterations in gene**
988 **expression.**
989 (**a**) Quantile-quantile plots showing the proportions of interacting 10 kb bins exhibiting significant
990 allelic bias at an FDR of 5% in excitatory neurons and lower motor neurons. (**b**) An example of an
991 interaction exhibiting significant allelic bias in excitatory neurons (binomial test, two-sided, p=5.4x10^{-4})
992 and lower motor neurons (binomial test, two-sided, p=4.2x10^{-7}). The interaction occurs between a PIR
993 containing SNPs for bipolar alcoholism at an open chromatin peak (green highlight) and the promoter of
994 *SYT17* (orange highlight). Heterozygous WTC11 variants at the PIR are shown, along with bar graphs of
995 detected read counts for each allele in our chromatin interactions. (**c**) Enrichment of significant eQTLs
996 from GTEX V7 at significant versus randomly shuffled PIRs in matched tissue types for excitatory
997 neurons and hippocampal DG-like neurons (one sample z-test, p<2.2x10^{-16} for both cell types). Error
998 bars show the standard deviation over 100 sampled sets of randomly shuffled PIRs. (**d**) Distributions of
999 interaction scores for chromatin interactions overlapping significant eQTL-TSS pairs versus randomly
000 sampled nonsignificant eQTL-TSS pairs in excitatory neurons and hippocampal DG-like neurons.
001 Interaction scores are significantly enriched for significant eQTL-TSS pairs (Kolmogorov-Smirnov test,
002 p=2.3x10^{-4} for excitatory neurons and p=1.8x10^{-6} for hippocampal DG-like neurons). (**e**) Forest plot for
003 independent ASD case-specific and non-overlapping matched pseudo-control-specific SNP pairs for
004 each cell type. The x-axis shows the odds ratio (OR) estimated between the numbers of case- and
005 control-specific SNP pairs at a significance threshold of 10^{-7} in each cell type. The area inside the

006 squares is proportional to the number of observations for each comparison. Hippocampal DG-like
007 neuron promoter-PIR pairs showed case-specific enrichments of ASD SNP pairs (chi-squared test,
008 $p<2.7x10^{-16}$). The confidence intervals for each OR estimation are shown in blue, and the red line
009 represents a baseline OR of 1.
010
011 **Supplementary Figure 1. Characterization of the cell types used in the study.**
012 (**a**) Immunofluorescence staining of key markers in excitatory neurons, hippocampal DG-like neurons,
013 lower motor neurons, and astrocytes. Excitatory neurons were positively stained for CUX1, an upper
014 cortical layer marker, and MAP2, a neuronal marker which is specifically expressed in dendrites. The
015 yield of excitatory neurons is calculated as the number of CUX1 and MAP2 double positive cells
016 divided by the total number of live cells. Hippocampal DG-like neurons were positively stained for
017 PROX1, a transcription factor specifying granule cell identity in the DG. The yield of mature
018 hippocampal DG-like neurons is calculated as the number of PROX1 and MAP2 double positive cells
019 divided by the total number of live cells. Lower motor neurons were positively stained for HB9, a motor
020 neuron marker, and the pan-neuronal neurofilament marker SMI32. The yield of mature lower motor
021 neurons is calculated as the number of HB9 and SMI32 double positive cells divided by the total number
022 of live cells. Finally, astrocytes were positively stained for GFAP. The yield of GFAP-positive
023 astrocytes is calculated as the number of GFAP positive cells divided by the total number of live cells.
024 The number of staining experiments and the total number of cells is indicated, and error bars represent
025 the SEM. (**b**) Heatmaps displaying the expression of key marker genes for the neural cell types.
026 Astrocytes used in this study exhibit an expression profile consistent with APC identity. (**c**) Counts of
027 protein coding (dark blue) and noncoding RNA (light blue) genes with promoters interacting in each cell
028 type.
029
030 **Supplementary Figure 2. Correlation between pcHi-C, ATAC-seq, and RNA-seq replicates.**
031 (a) Gene expression values for each RNA-seq replicate were hierarchically clustered according to
032 sample distances using DESeq2. **(b**) Heatmap with pairwise correlations and hierarchical clustering of
033 read densities at the set of unified open chromatin peaks for the ATAC-seq replicates. (**c**) Heatmap with
034 pairwise correlations based on the stratum-adjusted correlation coefficient (SCC) from HiCRep
035 (evaluated at a resolution of 10 kb) for the pcHi-C replicates. (**d**) Saturation of the SCC between
036 biological replicates for the pcHi-C libraries as a function of total sequencing depth.

037

038 **Supplementary Figure 3. Integrative analysis of chromatin interactions in individual cell types.**
039 (**a**). Histograms of interaction distances for each cell type. The mean interaction distances for each cell
040 type are indicated with red lines. (**b**) Bar plots showing counts of H3K27ac and CTCF binding sites
041 overlapping significant (solid bars) versus randomly shuffled (striped bars) PIRs for excitatory neurons,
042 lower motor neurons, and astrocytes. Error bars represent the standard deviation over 100 sampled sets
043 of randomly shuffled PIRs. (**c**) Comparative gene expression analysis in individual cell types for
044 expressed genes (normalized RPKM > 0.5) whose promoters interact exclusively with either enhancer-
045 PIRs (n=6836) or repressive-PIRs (n=2612). Distributions of gene expression values are shown for each
046 group. (**d**) Boxplots showing distributions of gene expression values in individual cell types for
047 expressed genes (normalized RPKM > 0.5) grouped according to the numbers of interactions their
048 promoters form with enhancer-PIRs. Linear regression was performed on the mean gene expression
049 values for each bin. Only bins containing at least 10 genes were included in the analysis.
050

051 **Supplementary Figure 4. Cell type-specific aspects of chromatin interactions.**
052 (**a**) Venn diagram displaying counts of unique promoter-PIR interactions across excitatory neurons,
053 hippocampal DG-like neurons, lower motor neurons, and astrocytes for each specificity pattern (groups
054 1-15 in **Fig. 3a**). (**b**) Examples of interactions between cell type-specific PIRs (yellow highlight) and the
055 promoters for *OPHN1*, *CHAT*, and *TLR4* (orange highlight). Open chromatin peaks and gene expression
056 are also displayed for each of the cell types. (**c**) Significant downregulation of *CDK5RAP5* expression
057 was observed across three independent clones with homozygous deletions for the candidate PIR in
058 excitatory neurons (two sample t-test, two-sided, p=7.7x10^{-7}). Error bars represent the SEM.
059

060 **Supplementary Figure 5. Using chromatin interactions to elucidate the functions of GWAS**
061 **variants.**
062 (**a**) Counts of GWAS SNPs for each disease with at least one SNP in linkage interacting exclusively
063 with their nearest genes (scenario III), interacting exclusively with more distal genes (scenario I), or
064 interacting with both their nearest genes and more distal genes (scenario II). GWAS SNPs that could not
065 be resolved for interactions with their nearest genes are also tabulated (scenario IV), along with counts
066 of regulatory targets interacting with GWAS SNPs in each scenario. (**b**) Significant promoter-PIR
067 interactions in hippocampal DG-like neurons and astrocytes recapitulate a previously reported

068 interaction between the *FOXG1* promoter and a distal open chromatin peak containing rs1191551, a
069 schizophrenia-associated variant[6]. (**c**) CRISPRi silencing of a candidate PIR for *STRAP* using two
070 independent sgRNAs results in significant downregulation of *STRAP* expression in excitatory neurons
071 (two sample t-test, two-sided, p=1.5x10^{-7}). No significant downregulation was detected for the
072 neighboring genes *MGST1* and *WBP1*, though the expression of *PTPRO* was affected. Each CRISPRi
073 experiment was performed in triplicate, with three technical replicates per experiment. (**d**) Schematic of
074 detected genotypes in the *DRD2* gene and its candidate PIR in mono-allelic deletion clones. Genotyping
075 and RT-PCR sequencing for WTC11 variants in the *DRD2* gene reveal allele-specific imbalances in
076 *DRD2* expression, consistent with the deletion of the candidate PIR in one of the alleles. Results for two
077 wild type control clones and three mono-allelic deletion clones are shown for comparison.
078
079 **Supplementary Figure 6. Top enriched GO terms for genes targeted by complex neurological**
080 **disorder-associated variants.**
081 Top enriched GO terms from Enrichr for genes whose promoters are targeted by variants for each
082 disease. EP and FTD are omitted due to their low numbers of reported variants and target genes
083 identified by our significant promoter-PIR interactions. An expanded list of enriched GO terms is
084 available in **Supplementary Table 9**. In general, we observe enrichment of terms associated with
085 epigenetic, neuronal, and disease-specific processes across the diseases.
086
087 **Supplementary Figure 7. Examples of putative regulatory SNPs at cell type-specific PIRs.**
088 In all examples, interacting PIRs are highlighted in yellow and the targeted promoters are highlighted in
089 orange. (**a**) A PIR containing AD SNPs interacts with the promoters of *FAM131B* and *CASP2* in
090 astrocytes, but interacts instead with the *ZYX* promoter in hippocampal DG-like neurons and lower
091 motor neurons. (**b**) PIRs containing MP SNPs in an intron for *PTPRO* interacts with the *STRAP*
092 promoter in all four cell types. (**c**) A PIR containing SCZ SNPs interacts with the *TRIM33* promoter in
093 astrocytes. Two other PIRs containing SCZ SNPs interact with the promoters of *TRIM33* and *BCAS2* in
094 hippocampal DG-like neurons. (**d**) A PIR containing BD SNPs interacts with the *MSI2* promoter in
095 hippocampal DG-like neurons, lower motor neurons, and astrocytes while also interacting with the
096 *AKAP1* promoter in lower motor neurons and astrocytes. Meanwhile, another group of PIRs containing
097 SCZ SNPs interact with the *MSI2* promoter exclusively in astrocytes.
098

099 **Tables**

100 **Supplementary Table 1.** pcHi-C, ATAC-seq, and RNA-seq data processing metrics.

101 pcHi-C data processing metrics include output from the HiCUP mapping pipeline (columns C through
102 O), the total number of reads in the CHiCAGO input file (columns P through Q), and the total number of
103 processed interactions with score ≥ 5 from the CHiCAGO pipeline (column R). ATAC-seq data
104 processing and QC metrics from the ENCODE pipeline are reported (columns C through K). The
105 number of peaks called in individual as well as across all replicates is shown (column L). RNA-seq data
106 processing metrics from STAR are reported (columns C through G).

107

108 **Supplementary Table 2.** Processed significant interactions called by CHiCAGO.

109 For each cell type, the left and right interacting BED intervals are shown in columns A through C and
110 columns E through G, respectively. The number of supporting reads, interaction score, and specificity
111 string for each interaction are shown in columns I through K. The number of overlaps with promoters,
112 promoter open chromatin peaks, and distal open chromatin peaks, as well as the interacting gene IDs are
113 shown in columns L through W. Overlaps for the left and right interacting BED intervals are shown
114 separately, and "promoter" refers to protein coding and noncoding RNA transcripts in GENCODE 19
115 while "promoter other" refers to all other transcripts in GENCODE 19. The number of overlaps with
116 positive Vista elements and SNPs for each disease and their associated IDs are shown separately for the
117 left and right interacting BED intervals in columns X through BS.

118

119 **Supplementary Table 3.** GO enrichment results for genes interacting with cell type-specific PIRs.

120 GO enrichment results from the "Biological Process 2018" ontology in Enrichr are shown for genes
121 interacting with PIRs specific to each of the cell types, as well as for genes interacting with PIRs shared
122 across all the cell types ("shared terms"). In each tab, the top 100 GO terms and their associated p-
123 values, Z-scores, combined scores, and genes are shown.

124

125 **Supplementary Table 4.** Motif enrichment results at cell type-specific PIRs.

126 For each cell type, the complete set of known motif results detected by HOMER are reported. This
127 includes the motif name, consensus sequence, p-value, # of target sequences with motif, and # of
128 background sequences with motif.

129

130 **Supplementary Table 5.** Putative target genes for *in vivo* validated enhancer elements.
131 Information about each positive Vista element including its position, ID, and annotation are shown in
132 columns A through F. Information about its nearest gene, as well as genes whose promoters fall on the
133 same HindIII fragment as the element are shown in columns G through I (protein coding and noncoding
134 RNA transcripts in GENCODE 19 only). Column J represents all other transcripts from GENCODE 19
135 whose promoters fall on the same HindIII fragment as the element. Column K reports whether or not
136 interactions are resolvable between the Vista element and its nearest gene in column G. Columns L
137 through W contain information on whether the Vista element overlaps open chromatin peaks,
138 participates in interactions, or targets genes on the other ends of interactions for each cell type.
139
140 **Supplementary Table 6.** GWAS Catalog SNP imputation summary.
141 The first tab contains a summary of the SNP imputation process for each disease in the study. This
142 includes the number of GWAS SNPs downloaded from the GWAS Catalog, the number of GWAS SNPs
143 passing the significance cutoff of 10^{-6}, the numbers of GWAS SNPs associated with each study
144 population, and the numbers of imputed SNPs for each study population. The remaining tabs contain
145 lists of all the studies included for each disease along with their associated information.
146
147 **Supplementary Table 7.** Putative target genes for neurological disorder-associated SNPs.
148 Two tabs are included for each disease in Supplementary Table 7 ("GWAS SNPs" and "all SNPs"). The
149 first tab ("GWAS SNPS") contains the results for all GWAS SNPs downloaded from the GWAS
150 Catalog. Information about each GWAS SNP including its position, rsid, allele information, query SNP
151 status, and whether or not it overlaps any exons are shown in columns A through I. Information about its
152 nearest gene, as well as genes whose promoters fall on the same HindIII fragment as the element are
153 shown in columns J through L (protein coding and noncoding RNA transcripts in GENCODE 19 only).
154 Column M represents all other transcripts from GENCODE 19 whose promoters fall on the same
155 HindIII fragment as the element. Column N reports whether or not interactions are resolvable between
156 the GWAS SNP (or any of its linked SNPs) and its nearest gene in column G. The total number of
157 linked SNPs for each GWAS SNP is shown in column O. Columns P through AE contain information
158 on whether the GWAS SNP itself or any of its linked SNPs participate in interactions or target genes on
159 the other ends of interactions for each cell type. The second tab ("all SNPs") contains similar
160 information for all imputed SNPs. Columns A through I contain information about each imputed SNP,

161 columns J through N contain information about its nearest and same fragment gene(s), and columns O
162 through AD contain information on whether the imputed SNP overlaps with promoter or distal open
163 chromatin peaks, participates in interactions, or targets genes on the other ends of interactions for each
164 cell type.
165
166 **Supplementary Table 8.** Putative target genes for imputed SNPs overlapping open chromatin peaks.
167 Supplementary Table 8 contains subsets of imputed SNPs from the "all SNPs" tabs in Supplementary
168 Table 7, for which the imputed SNPs overlap with promoter or distal open chromatin peaks in at least
169 one cell type.
170
171 **Supplementary Table 9.** GO enrichment results for disease-specific target genes.
172 GO enrichment results from the "Biological Process 2018" ontology in Enrichr are shown for genes
173 interacting with PIRs containing variants for each disease. In each tab, the top 100 GO terms and their
174 associated p-values, Z-scores, combined scores, and genes are shown.
175
176 **Supplementary Table 10.** Interactions exhibiting significant allelic bias.
177 A list of allelically biased interactions with a p-value cutoff of 10^{-3} are shown for the excitatory neurons
178 and lower motor neurons. Supplementary Table 10 follows the format of Supplementary Table 2 with a
179 few exceptions. The number of reads supporting interactions in each allele is reported in column I. The
180 negative log p-value is reported in column J.
181
182 **Supplementary Table 11.** ASD epistatic SNP pairs by cell type.
183 The total numbers of epistatic SNP pairs in ASD cases and matched pseudo-controls with the following
184 P-value thresholds: $p \leq 1.0\text{x}10^{-20}$, $1.0\text{x}10^{-20} \leq p < 1.0\text{x}10^{-17}$, $1.0\text{x}10^{-17} \leq p < 1.0\text{x}10^{-14}$, $1.0\text{x}10^{-14} \leq p <$
185 $1.0\text{x}10^{-13}$, $1.0\text{x}10^{-13} \leq p < 1.0\text{x}10^{-12}$, $1.0\text{x}10^{-12} \leq p < 1.0\text{x}10^{-11}$, $1.0\text{x}10^{-11} \leq p < 1.0\text{x}10^{-10}$, $1.0\text{x}10^{-10} \leq p <$
186 $1.0\text{x}10^{-9}$, $1.0\text{x}10^{-9} \leq p < 1.0\text{x}10^{-8}$, and $1.0\text{x}10^{-8} \leq p \leq 1.0\text{x}10^{-7}$ are shown for the entire dataset including
187 for excitatory neurons, hippocampal DG-like neurons, lower motor neurons, and astrocytes. Results for
188 nonsignificant interactions in the corresponding cell types are also shown.
189
190 **Supplementary Table 12.** sgRNA and primer sequences.
191 A full list of sgRNAs and primers for the CRISPR deletion and CRISPRi experiments are reported here.

192 **References**

193 1 Mumbach, M. R. *et al.* Enhancer connectome in primary human cells identifies target genes
194 of disease-associated DNA elements. *Nat Genet* **49**, 1602-1612, doi:10.1038/ng.3963
195 (2017).
196 2 Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex
197 trait gene targets. *Nat Genet* **48**, 481-487, doi:10.1038/ng.3538 (2016).
198 3 Smemo, S. *et al.* Obesity-associated variants within FTO form long-range functional
199 connections with IRX3. *Nature* **507**, 371-375, doi:10.1038/nature13138 (2014).
200 4 Claussnitzer, M. *et al.* FTO Obesity Variant Circuitry and Adipocyte Browning in Humans. *N*
201 *Engl J Med* **373**, 895-907, doi:10.1056/NEJMoa1502214 (2015).
202 5 de la Torre-Ubieta, L. *et al.* The Dynamic Landscape of Open Chromatin during Human
203 Cortical Neurogenesis. *Cell* **172**, 289-304 e218, doi:10.1016/j.cell.2017.12.014 (2018).
204 6 Won, H. *et al.* Chromosome conformation elucidates regulatory relationships in developing
205 human brain. *Nature* **538**, 523-527, doi:10.1038/nature19847 (2016).
206 7 Miyaoka, Y. *et al.* Isolation of single-base genome-edited human iPS cells without antibiotic
207 selection. *Nat Methods* **11**, 291-293, doi:10.1038/nmeth.2840 (2014).
208 8 Wang, C. *et al.* Scalable Production of iPSC-Derived Human Neurons to Identify Tau-
209 Lowering Compounds by High-Content Screening. *Stem Cell Reports* **9**, 1221-1233,
210 doi:10.1016/j.stemcr.2017.08.019 (2017).
211 9 Mertens, J. *et al.* Differential responses to lithium in hyperexcitable neurons from patients
212 with bipolar disorder. *Nature* **527**, 95-99, doi:10.1038/nature15526 (2015).
213 10 Fernandopulle, M. S. *et al.* Transcription Factor-Mediated Differentiation of Human iPSCs
214 into Neurons. *Curr Protoc Cell Biol* **79**, e51, doi:10.1002/cpcb.51 (2018).
215 11 Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser--a
216 database of tissue-specific human enhancers. *Nucleic Acids Res* **35**, D88-92,
217 doi:10.1093/nar/gkl822 (2007).
218 12 Zhang, Y. *et al.* Rapid single-step induction of functional neurons from human pluripotent
219 stem cells. *Neuron* **78**, 785-798, doi:10.1016/j.neuron.2013.05.029 (2013).
220 13 Yu, D. X. *et al.* Modeling hippocampal neurogenesis using human pluripotent stem cells.
221 *Stem Cell Reports* **2**, 295-310, doi:10.1016/j.stemcr.2014.01.009 (2014).
222 14 Zhang, Y. *et al.* Purification and Characterization of Progenitor and Mature Human
223 Astrocytes Reveals Transcriptional and Functional Differences with Mouse. *Neuron* **89**, 37-
224 53, doi:10.1016/j.neuron.2015.11.013 (2016).
225 15 Cairns, J. *et al.* CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data.
226 *Genome Biol* **17**, 127, doi:10.1186/s13059-016-0992-2 (2016).
227 16 Jackson, D. A., Hassan, A. B., Errington, R. J. & Cook, P. R. Visualization of focal sites of
228 transcription within human nuclei. *EMBO J* **12**, 1059-1065 (1993).
229 17 Zhang, Y. *et al.* Chromatin connectivity maps reveal dynamic promoter-enhancer long-
230 range associations. *Nature* **504**, 306-310, doi:10.1038/nature12716 (2013).
231 18 Diao, Y. *et al.* A tiling-deletion-based genetic screen for cis-regulatory element identification
232 in mammalian cells. *Nat Methods* **14**, 629-635, doi:10.1038/nmeth.4264 (2017).
233 19 Engreitz, J. M. *et al.* Local regulation of gene expression by lncRNA promoters, transcription
234 and splicing. *Nature* **539**, 452-455, doi:10.1038/nature20149 (2016).
235 20 Consortium, E. P. An integrated encyclopedia of DNA elements in the human genome.
236 *Nature* **489**, 57-74, doi:10.1038/nature11247 (2012).

21 Shen, Y. *et al.* A map of the cis-regulatory sequences in the mouse genome. *Nature* **488**, 116-120, doi:10.1038/nature11243 (2012).
22 Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and characterization. *Nat Methods* **9**, 215-216, doi:10.1038/nmeth.1906 (2012).
23 Roadmap Epigenomics, C. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-330, doi:10.1038/nature14248 (2015).
24 Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**, doi:10.7554/eLife.21856 (2017).
25 Nadif Kasri, N., Nakano-Kobayashi, A. & Van Aelst, L. Rapid synthesis of the X-linked mental retardation protein OPHN1 mediates mGluR-dependent LTD through interaction with the endocytic machinery. *Neuron* **72**, 300-315, doi:10.1016/j.neuron.2011.09.001 (2011).
26 Shen, Y. *et al.* Postnatal activation of TLR4 in astrocytes promotes excitatory synaptogenesis in hippocampal neurons. *J Cell Biol* **215**, 719-734, doi:10.1083/jcb.201605046 (2016).
27 Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* **38**, 576-589, doi:10.1016/j.molcel.2010.05.004 (2010).
28 Ren, G. *et al.* CTCF-Mediated Enhancer-Promoter Interaction Is a Critical Regulator of Cell-to-Cell Variation of Gene Expression. *Mol Cell* **67**, 1049-1058 e1046, doi:10.1016/j.molcel.2017.08.026 (2017).
29 Guo, Y. *et al.* CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* **162**, 900-910, doi:10.1016/j.cell.2015.07.038 (2015).
30 Handoko, L. *et al.* CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* **43**, 630-638, doi:10.1038/ng.857 (2011).
31 Hou, C., Dale, R. & Dean, A. Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc Natl Acad Sci U S A* **107**, 3651-3656, doi:10.1073/pnas.0912087107 (2010).
32 Chung, W. C., Huang, T. N. & Hsueh, Y. P. Targeted deletion of CASK-interacting nucleosome assembly protein causes higher locomotor and exploratory activities. *Neurosignals* **19**, 128-141, doi:10.1159/000327819 (2011).
33 Liddell, J. R. Are Astrocytes the Predominant Cell Type for Activation of Nrf2 in Aging and Neurodegeneration? *Antioxidants (Basel)* **6**, doi:10.3390/antiox6030065 (2017).
34 Yin, X., Warner, D. R., Roberts, E. A., Pisano, M. M. & Greene, R. M. Novel interaction between nuclear co-activator CBP and the CDK5 activator binding protein - C53. *Int J Mol Med* **16**, 251-256 (2005).
35 Xie, Y. H. *et al.* Cloning and characterization of human IC53-2, a novel CDK5 activator binding protein. *Cell Res* **13**, 83-91, doi:10.1038/sj.cr.7290153 (2003).
36 Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* **47**, D1005-D1012, doi:10.1093/nar/gky1120 (2019).
37 Ward, L. D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* **40**, D930-934, doi:10.1093/nar/gkr917 (2012).
38 Abdulla, S. *et al.* Hippocampal degeneration in patients with amyotrophic lateral sclerosis. *Neurobiol Aging* **35**, 2639-2645, doi:10.1016/j.neurobiolaging.2014.05.035 (2014).

39 Booth, H. D. E., Hirst, W. D. & Wade-Martins, R. The Role of Astrocyte Dysfunction in Parkinson's Disease Pathogenesis. *Trends Neurosci* **40**, 358-370, doi:10.1016/j.tins.2017.04.001 (2017).
40 Tiwari, M., Lopez-Cruzan, M., Morgan, W. W. & Herman, B. Loss of caspase-2-dependent apoptosis induces autophagy after mitochondrial oxidative stress in primary cultures of young adult cortical neurons. *J Biol Chem* **286**, 8493-8506, doi:10.1074/jbc.M110.163824 (2011).
41 Zhao, X. *et al.* Caspase-2 cleavage of tau reversibly impairs memory. *Nat Med* **22**, 1268-1276, doi:10.1038/nm.4199 (2016).
42 Huang, C. W. *et al.* Conditional Knockout of Breast Carcinoma Amplified Sequence 2 (BCAS2) in Mouse Forebrain Causes Dendritic Malformation via beta-catenin. *Sci Rep* **6**, 34927, doi:10.1038/srep34927 (2016).
43 Burghes, A. H. & Beattie, C. E. Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? *Nat Rev Neurosci* **10**, 597-609, doi:10.1038/nrn2670 (2009).
44 Zhang, J. P. *et al.* Association of a Schizophrenia Risk Variant at the DRD2 Locus With Antipsychotic Treatment Response in First-Episode Psychosis. *Schizophr Bull* **41**, 1248-1255, doi:10.1093/schbul/sbv116 (2015).
45 Eisenstein, S. A. *et al.* Prediction of striatal D2 receptor binding by DRD2/ANKK1 TaqIA allele status. *Synapse* **70**, 418-431, doi:10.1002/syn.21916 (2016).
46 Pearlson, G. D. *et al.* In vivo D2 dopamine receptor density in psychotic and nonpsychotic patients with bipolar disorder. *Arch Gen Psychiatry* **52**, 471-477 (1995).
47 Selvaraj, S., J, R. D., Bansal, V. & Ren, B. Whole-genome haplotype reconstruction using proximity-ligation and shotgun sequencing. *Nat Biotechnol* **31**, 1111-1118, doi:10.1038/nbt.2728 (2013).
48 Lydall, G. J. *et al.* Confirmation of prior evidence of genetic susceptibility to alcoholism in a genome-wide association study of comorbid alcoholism and bipolar disorder. *Psychiatr Genet* **21**, 294-306, doi:10.1097/YPG.0b013e32834915c2 (2011).
49 Sudhof, T. C. Calcium control of neurotransmitter release. *Cold Spring Harb Perspect Biol* **4**, a011353, doi:10.1101/cshperspect.a011353 (2012).
50 Consortium, G. T. *et al.* Genetic effects on gene expression across human tissues. *Nature* **550**, 204-213, doi:10.1038/nature24277 (2017).
51 Shao, H. *et al.* Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. *Proc Natl Acad Sci U S A* **105**, 19910-19914, doi:10.1073/pnas.0810388105 (2008).
52 Gale, G. D. *et al.* A genome-wide panel of congenic mice reveals widespread epistasis of behavior quantitative trait loci. *Mol Psychiatry* **14**, 631-645, doi:10.1038/mp.2008.4 (2009).
53 Huang, W. *et al.* Epistasis dominates the genetic architecture of Drosophila quantitative traits. *Proc Natl Acad Sci U S A* **109**, 15553-15559, doi:10.1073/pnas.1213423109 (2012).
54 Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T. L. & Kruglyak, L. Finding the sources of missing heritability in a yeast cross. *Nature* **494**, 234-237, doi:10.1038/nature11867 (2013).
55 Blair, D. R. *et al.* A nondegenerate code of deleterious variants in Mendelian loci contributes to complex disease risk. *Cell* **155**, 70-80, doi:10.1016/j.cell.2013.08.030 (2013).

327 56 Mitra, I. *et al.* Reverse Pathway Genetic Approach Identifies Epistasis in Autism Spectrum
328 Disorders. *PLoS Genet* **13**, e1006516, doi:10.1371/journal.pgen.1006516 (2017).
329 57 Rajarajan, P. *et al.* Neuron-specific signatures in the chromosomal connectome associated
330 with schizophrenia risk. *Science* **362**, doi:10.1126/science.aat4311 (2018).
331 58 Straws in a haystack. *Nat Genet* **50**, 631, doi:10.1038/s41588-018-0125-9 (2018).
332 59 Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of
333 chromatin looping. *Cell* **159**, 1665-1680, doi:10.1016/j.cell.2014.11.021 (2014).
334 60 Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of
335 native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-
336 binding proteins and nucleosome position. *Nat Methods* **10**, 1213-1218,
337 doi:10.1038/nmeth.2688 (2013).
338 61 Wingett, S. *et al.* HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res* **4**, 1310,
339 doi:10.12688/f1000research.7334.1 (2015).
340 62 Bansal, V. & Bafna, V. HapCUT: an efficient and accurate algorithm for the haplotype
341 assembly problem. *Bioinformatics* **24**, i153-159, doi:10.1093/bioinformatics/btn298
342 (2008).
343 63 Li, N. & Stephens, M. Modeling linkage disequilibrium and identifying recombination
344 hotspots using single-nucleotide polymorphism data. *Genetics* **165**, 2213-2233 (2003).
345 64 Delaneau, O., Zagury, J. F. & Marchini, J. Improved whole-chromosome phasing for disease
346 and population genetic studies. *Nat Methods* **10**, 5-6, doi:10.1038/nmeth.2307 (2013).
347 65 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based
348 linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).

Figure 1

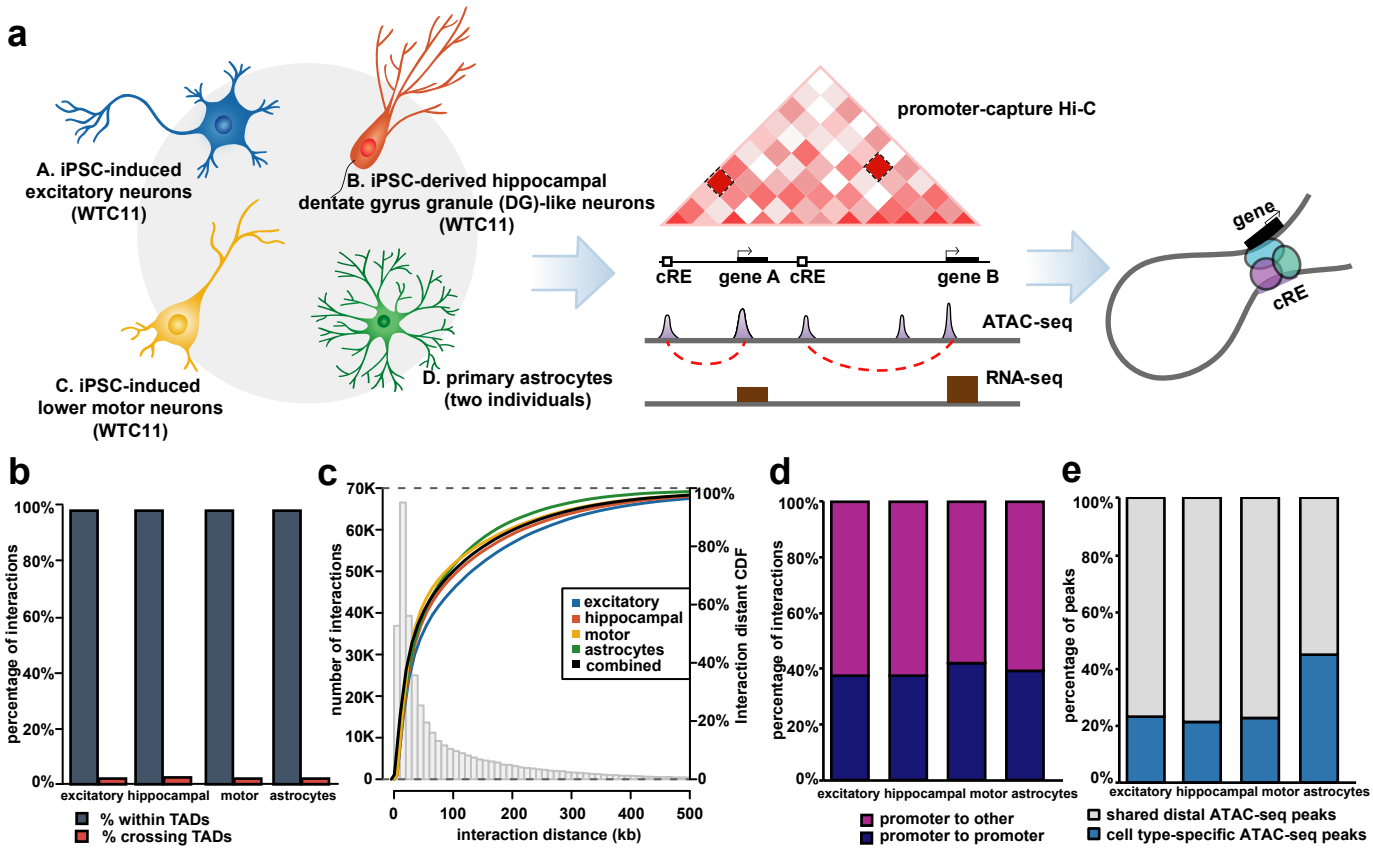


Figure 1. Genome-wide mapping of physical chromatin interactions in functionally distinct neural cell types.
(**a**) Schematic of the study design for generating four functionally distinct neural cell types in the CNS and performing integrative analysis of chromatin interactions using promoter capture Hi-C, open chromatin regions using ATAC-seq, and transcriptomes using RNA-seq. For pcHi-C, we used 3, 2, 3, and 4 biological replicates respectively for the excitatory neurons, hippocampal DG-like neurons, lower motor neurons, and astrocytes. For ATAC-seq, we used 2, 2, 3, and 4 biological replicates respectively for the cell types. For RNA-seq, we used 2, 2, 2, and 4 biological replicates respectively for the cell types.
(**b**) Proportions of interactions occuring within TADs for each cell type.
(**c**) Histogram and empirical CDF plots of interaction distances for each cell type.
(**d**) Proportions of interactions between promoter-containing bins (blue) and promoter- and non-promoter-containing bins (purple) for each cell type.
(**e**) Proportions of cell type-specific (blue) and shared (grey) distal open chromatin peaks at PIRs for each cell type.

Figure 2

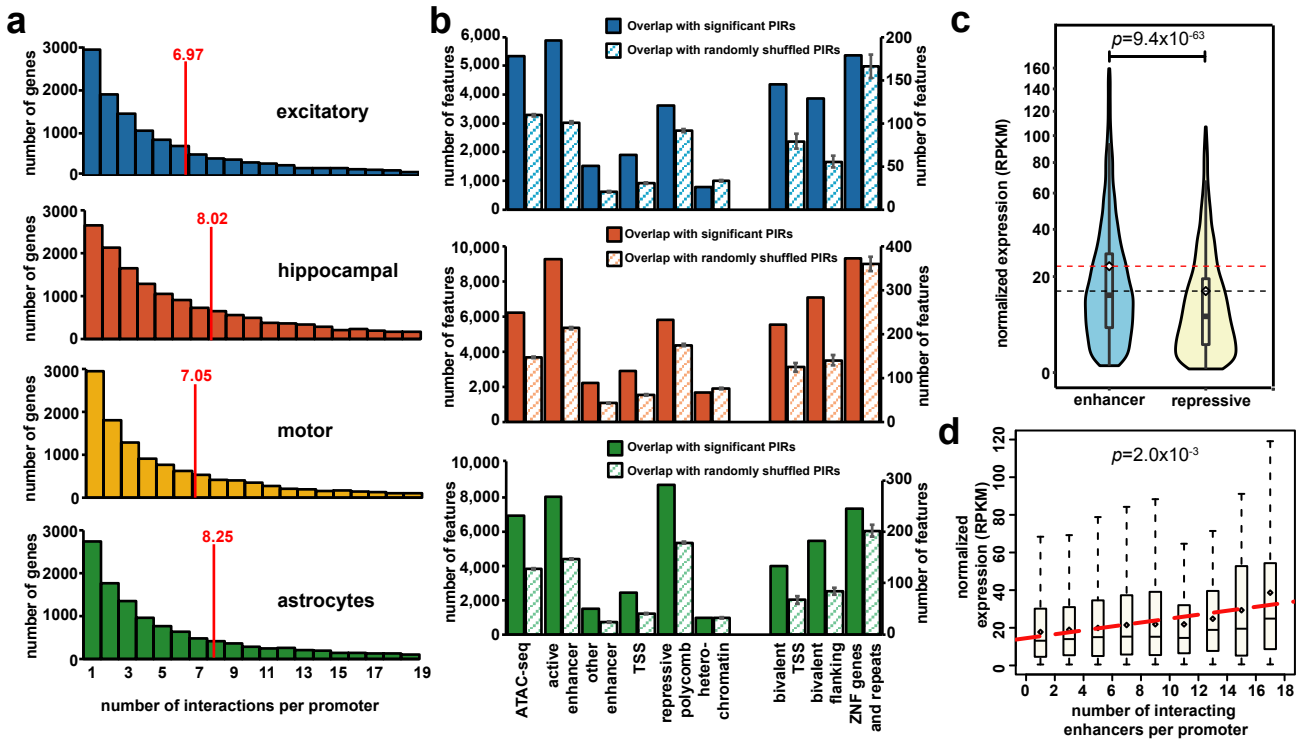


Figure 2. Integrative analysis of chromatin interactions, epigenomic features, and gene expression.
(**a**) Histograms of the numbers of PIRs interacting with each promoter in each cell type. The means are indicated with red lines. Only promoters interacting with at least one PIR are included (15,316 promoters in excitatory neurons, 19,546 promoters in hippocampal DG-like neurons, 14,990 promoters in lower motor neurons, and 15,397 promoters in astrocytes, out of a total of 34,401 protein coding and noncoding RNA promoters in GENCODE 19).
(**b**) Bar plots showing counts of epigenomic chromatin states (inferred at a 200 bp resolution using the ChromHMM core 15 state model in matched tissues) overlapping significant (solid bars) versus randomly shuffled (striped bars) PIRs for each cell type. Error bars show the standard deviation over 100 sampled sets of randomly shuffled PIRs. No matching tissue data was available for the lower motor neurons so they were omitted from the analysis.
(**c**) Comparative gene expression analysis across all cell types for expressed genes (normalized RPKM > 0.5) whose promoters interact exclusively with either enhancer-PIRs (n=6,836) or repressive-PIRs (n=2,612). Distributions of gene expression values are shown for each group.
(**d**) Boxplots showing distributions of gene expression values across all cell types for expressed genes (normalized RPKM > 0.5) grouped according to the numbers of interactions their promoters form with enhancer-PIRs. Linear regression was performed on the mean gene expression values for each bin. Only bins containing at least 10 genes were included in the analysis.

Figure 3

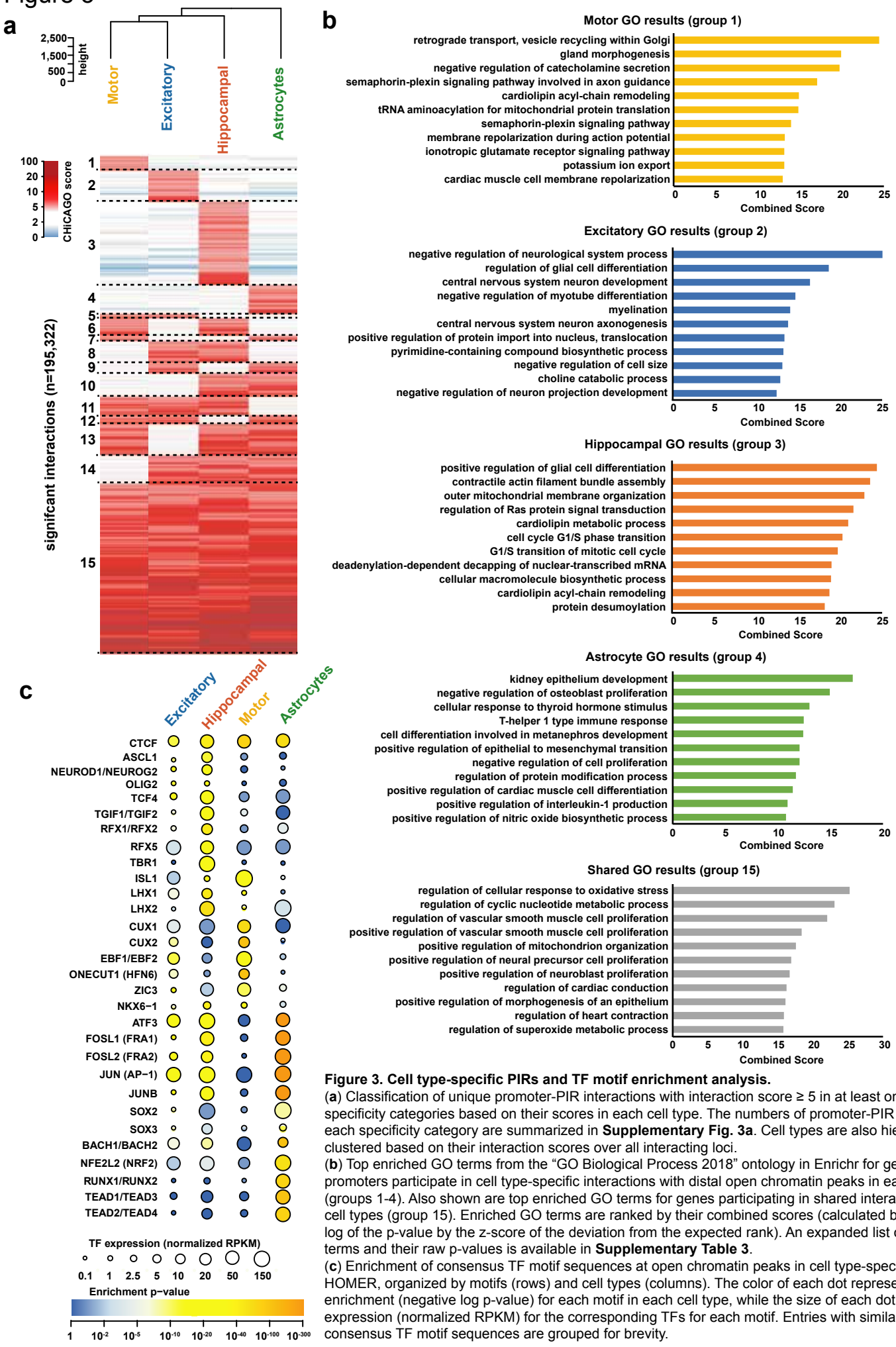


Figure 3. Cell type-specific PIRs and TF motif enrichment analysis.
(**a**) Classification of unique promoter-PIR interactions with interaction score ≥ 5 in at least one cell type into specificity categories based on their scores in each cell type. The numbers of promoter-PIR interactions in each specificity category are summarized in **Supplementary Fig. 3a**. Cell types are also hierarchically clustered based on their interaction scores over all interacting loci.
(**b**) Top enriched GO terms from the "GO Biological Process 2018" ontology in Enrichr for genes whose promoters participate in cell type-specific interactions with distal open chromatin peaks in each cell type (groups 1-4). Also shown are top enriched GO terms for genes participating in shared interactions across all cell types (group 15). Enriched GO terms are ranked by their combined scores (calculated by multiplying the log of the p-value by the z-score of the deviation from the expected rank). An expanded list of enriched GO terms and their raw p-values is available in **Supplementary Table 3**.
(**c**) Enrichment of consensus TF motif sequences at open chromatin peaks in cell type-specific PIRs using HOMER, organized by motifs (rows) and cell types (columns). The color of each dot represents the degree of enrichment (negative log p-value) for each motif in each cell type, while the size of each dot represents gene expression (normalized RPKM) for the corresponding TFs for each motif. Entries with similar or identical consensus TF motif sequences are grouped for brevity.

Figure 4

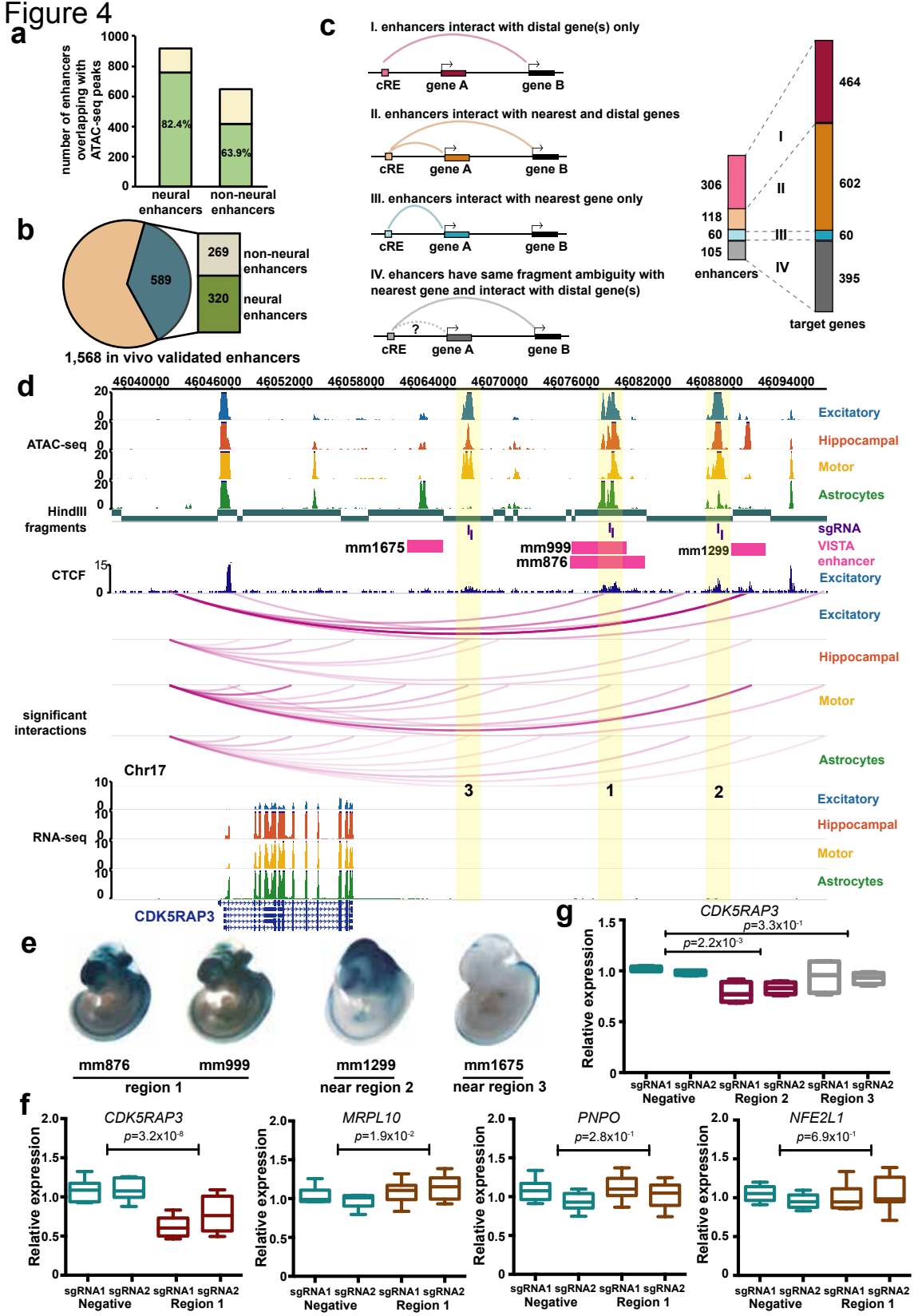


Figure 4. Validation of PIRs in human neural cells.

(**a**) *In vivo* validated enhancers with neural annotations overlap a significantly higher proportion of open chromatin peaks in the neural cells compared to enhancers with non-neural annotations (Pearson's chi-squared test, $p<2.2\times10^{-16}$).

(**b**) Pie chart showing counts of in vivo validated enhancers with human sequences participating in chromatin interactions (589 out of 1,568 total elements). Counts of interacting positive enhancer elements with neural and non-neural annotations are also shown.

(**c**) Counts of interacting positive enhancer elements interacting exclusively with their nearest genes (blue), interacting exclusively with more distal genes (pink), or interacting with both their nearest genes and more distal genes (orange). Positive enhancer elements that could not be resolved for interactions with their nearest genes are also shown (grey). The number of regulatory targets interacting with positive enhancer elements in each category is shown on the right.

(**d**) Promoter-PIR interactions at the *CDK5RAP3* locus. Open chromatin peaks in PIRs up to 40 kb downstream of *CDKRAP3* (regions 1, 2, and 3, yellow highlight) interact with the promoter of *CDK5RAP3* in a cell type-specific manner. Notably, only regions 1 and 2 participate in interactions with the promoter of *CDK5RAP3* in excitatory neurons. In addition, both in vivo validated enhancers (pink) and CTCF binding sites in cortical excitatory neurons (dark blue) are shown to be localized or near all three candidate regulatory regions. All interactions fall within a CP TAD (chr17:45,920,000-47,480,000).

(**e**) LacZ staining in mouse embryos shows tissue-specific patterns of enhancer activity.

(**f**) CRISPRi silencing of region 1 using two independent sgRNAs results in significant down-regulation of *CDK5RAP3* expression in excitatory neurons (two sample t-test, two-sided, $p=3.2\times10^{-8}$). No significant down-regulation was detected for the neighboring genes *MRPL10*, *PNPO*, and *NFE2L1*. Each CRISPRi experiment was conducted in triplicate, with three technical replicates per experiment.

(**g**) CRISPRi silencing of region 2, but not region 3, results in significant down-regulation of *CDK5RAP3* expression in excitatory neurons (two sample t-test, two-sided, $p=2.2\times10^{-3}$).

Figure 5

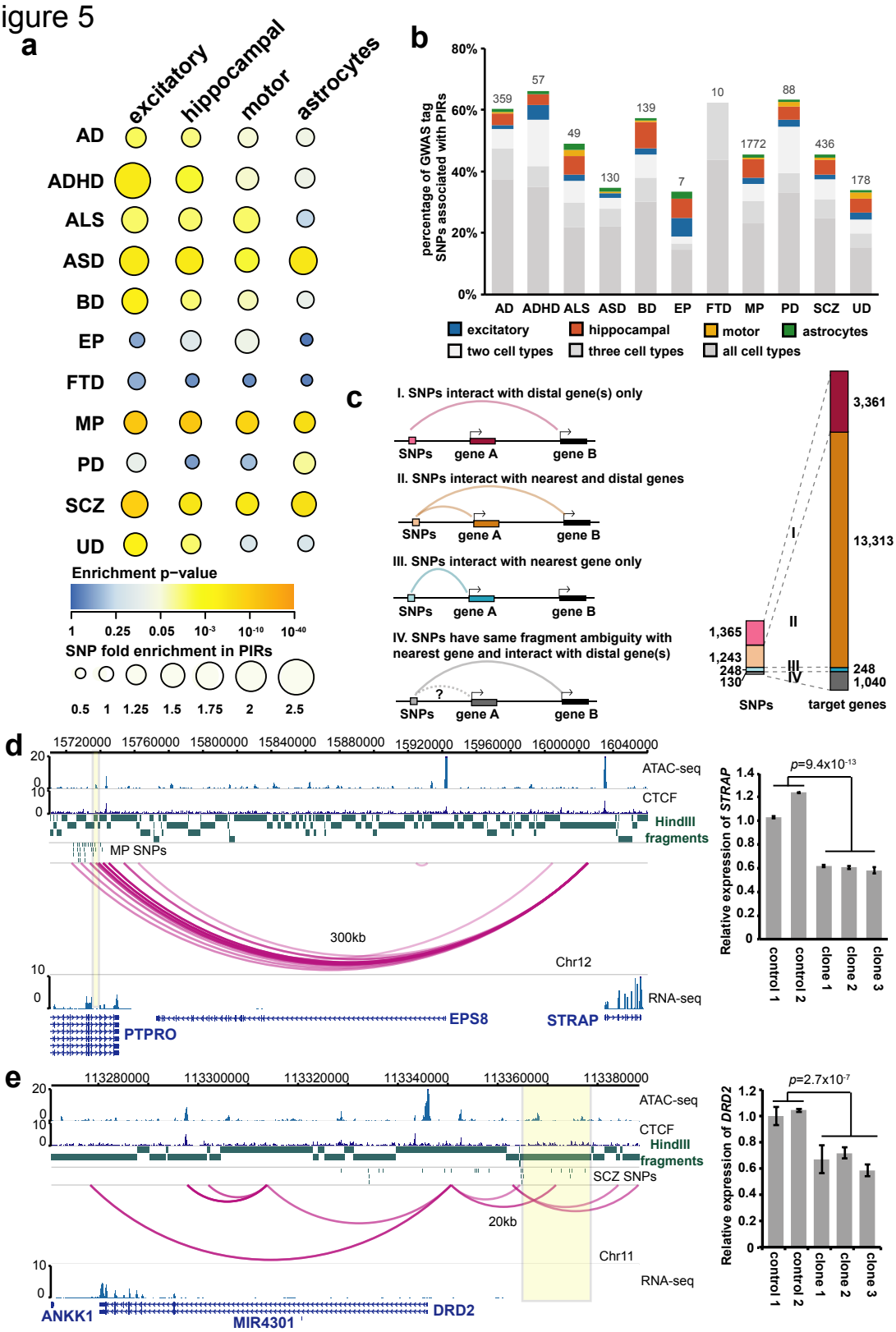


Figure 5. Genetic analysis of promoter-PIR interactions with complex neurological disorder-associated variants.
(**a**) Enrichment analysis for complex neurological disorder-associated SNPs in Alzheimer's disease (AD), attention deficit hyperactivity disorder (ADHD), autism spectrum disorder (ASD), amyotrophic lateral sclerosis (ALS), bipolar disorder (BP), epilepsy (EP), frontotemporal dementia (FTD), mental processing (MP), Parkinson's disease (PD), schizophrenia (SCZ), and unipolar depression (UD). The color and size of each dot respectively represent the enrichment *p*-value and raw fold enrichment (calculated as the number of SNPs overlapping significant PIRs divided by the mean number of SNPs overlapping randomly shuffled PIRs across 100 sampled sets) for each cell type and disease pairing.
(**b**) Proportions and total counts of GWAS SNPs with at least one SNP in linkage participating in chromatin interactions. Cell type-specific SNPs for excitatory neurons (blue), hippocampal DG-like neurons (orange), lower motor neurons (yellow), and astrocytes (green) are highlighted.
(**c**) Counts of GWAS SNPs across all diseases with at least one SNP in linkage interacting exclusively with their nearest genes (scenario III, blue), interacting exclusively with more distal genes (scenario I, pink), or interacting with both their nearest genes and more distal genes (scenario II, orange). GWAS SNPs that could not be resolved for interactions with their nearest genes are also shown (scenario IV, grey). Counts of regulatory targets interacting with GWAS SNPs in each scenario are shown on the right.
(**d**) PIRs containing MP SNPS at an open chromatin peak (yellow highlight) in an intron for *PTPRO* interacts with the promoter of *STRAP* over 300 kb away. All interactions fall within a CP TAD (chr12:14,960,000-16,040,000). Homozygous deletion of this PIR in three independent clones results in significant downregulation of *STRAP* expression in excitatory neurons (two sample t-test, two-sided, p=9.4x10^{-13}). Error bars represent the SEM.
(**e**) A PIR containing SCZ SNPs interacts with the *DRD2* promoter 20 kb upstream of the PIR. All interactions fall within a CP TAD (chr11:113,200,000-114,160,000). Mono-allelic deletion of this PIR in three independent clones results in significant downregulation of *DRD2* expression in excitatory neurons (two sample t-test, two-sided, p=2.7x10^{-7}). Error bars represent the SEM.

Figure 6

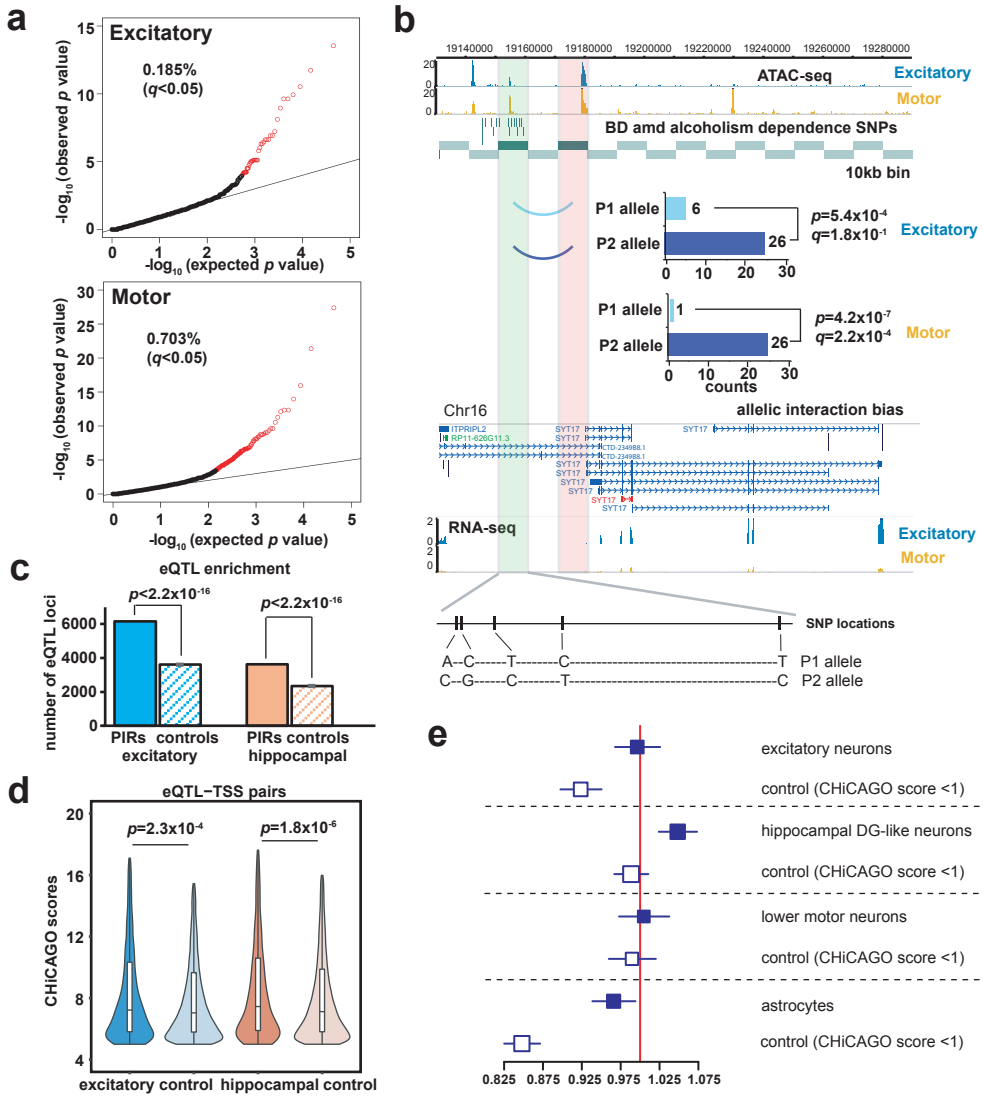
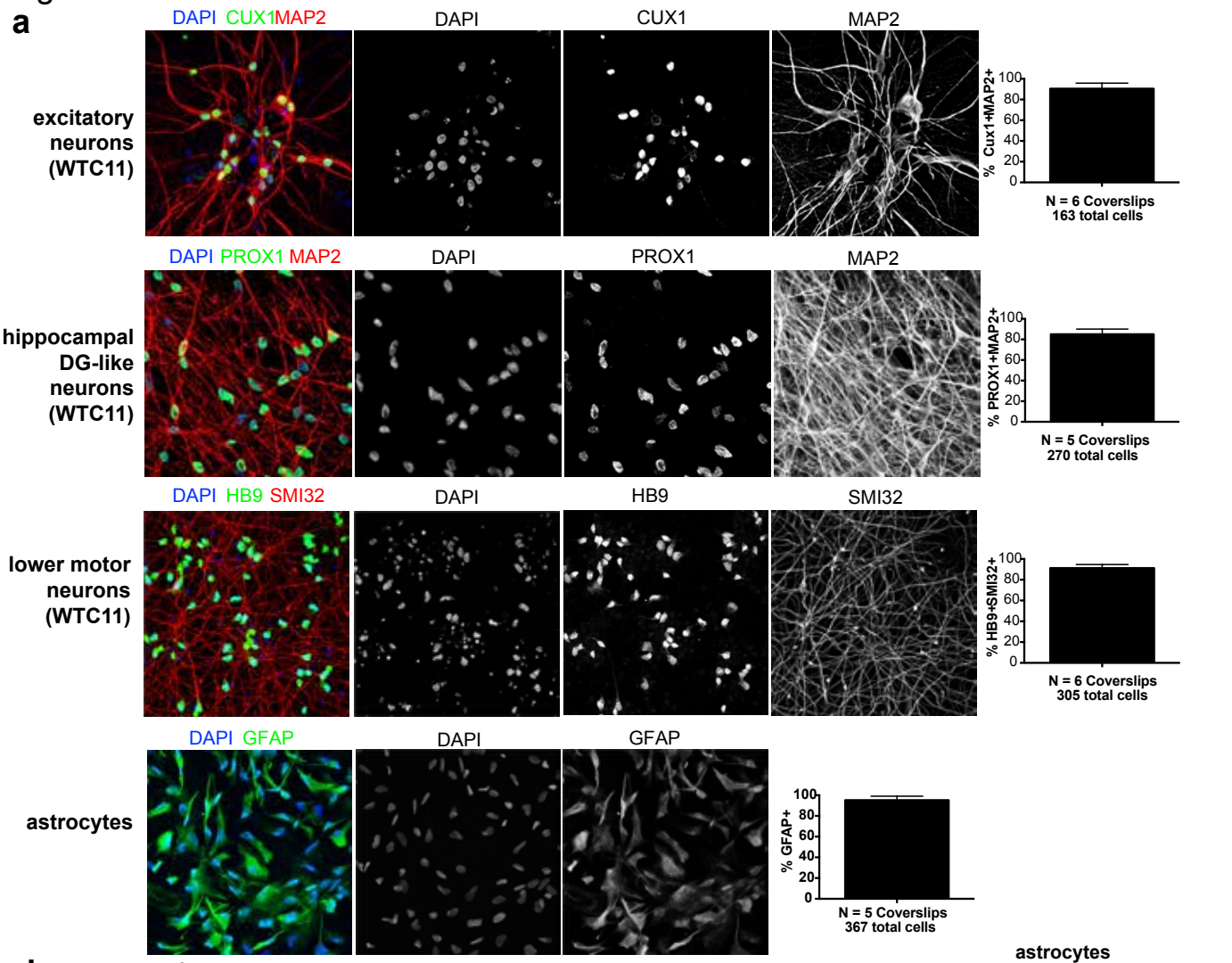
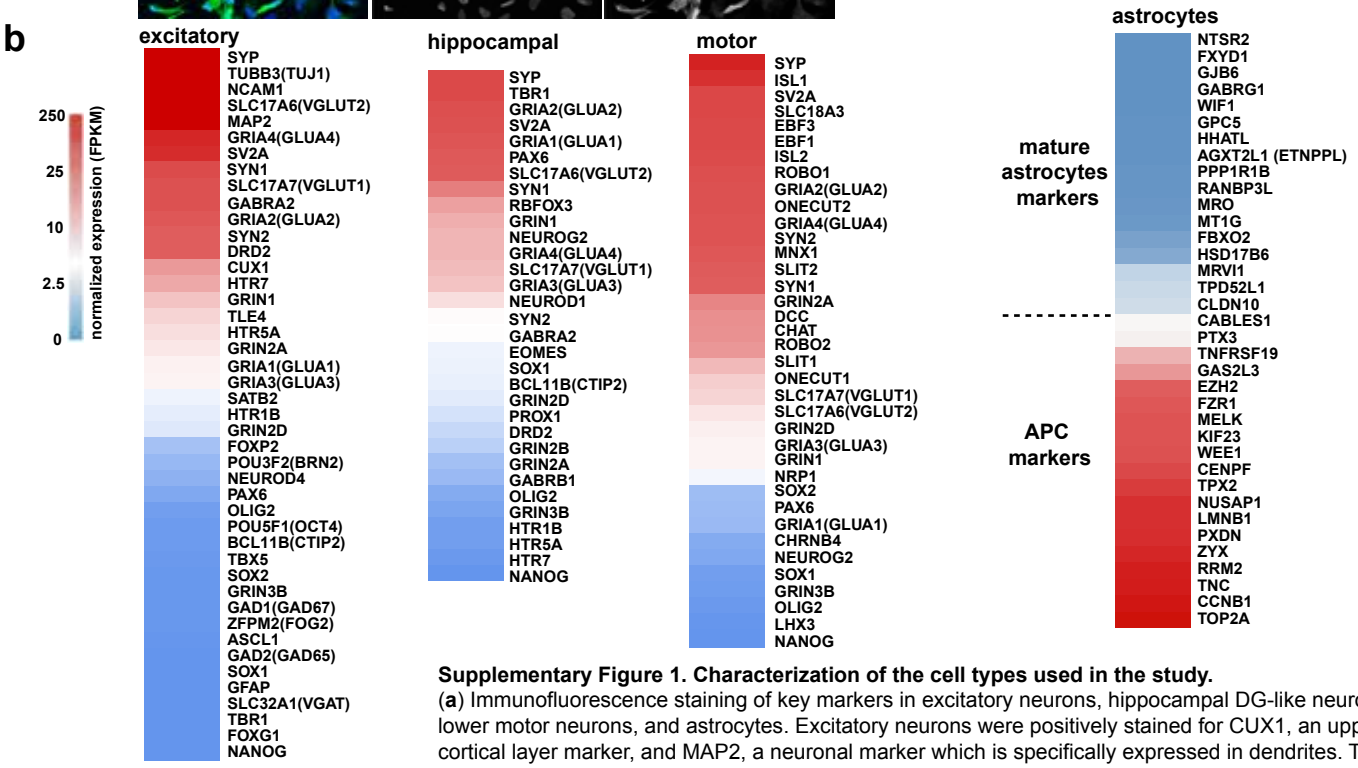
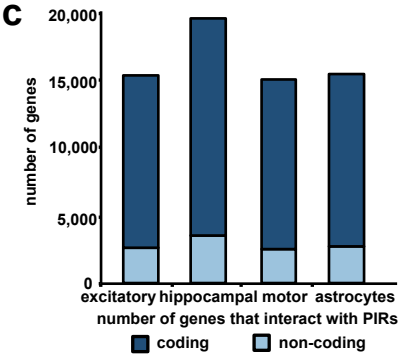


Figure 6. Genetics variants contribute to chromatin interaction bias and alterations in gene expression.
(**a**) Quantile-quantile plots showing the proportions of interacting 10 kb bins exhibiting significant allelic bias at an FDR of 5% in excitatory neurons and lower motor neurons.
(**b**) An example of an interaction exhibiting significant allelic bias in excitatory neurons (binomial test, two-sided, p=5.4x10^{-4}) and lower motor neurons (binomial test, two-sided, p=4.2x10^{-7}). The interaction occurs between a PIR containing SNPs for bipolar alcoholism at an open chromatin peak (green highlight) and the promoter of *SYT17* (orange highlight). Heterozygous WTC11 variants at the PIR are shown, along with bar graphs of detected read counts for each allele in our chromatin interactions.
(**c**) Enrichment of significant eQTLs from GTEX V7 at significant versus randomly shuffled PIRs in matched tissue types for excitatory neurons and hippocampal DG-like neurons (one sample z-test, p<2.2x10^{-16} for both cell types). Error bars show the standard deviation over 100 sampled sets of randomly shuffled PIRs.
(**d**) Distributions of interaction scores for chromatin interactions overlapping significant eQTL-TSS pairs versus randomly sampled nonsignificant eQTL-TSS pairs in excitatory neurons and hippocampal DG-like neurons. Interaction scores are significantly enriched for significant eQTL-TSS pairs (Kolmogorov-Smirnov test, p=2.3x10^{-4} for excitatory neurons and p=1.8x10^{-6} for hippocampal DG-like neurons).
(**e**) Forest plot for independent ASD case-specific and non-overlapping matched pseudo-control-specific SNP pairs for each cell type. The x-axis shows the odds ratio (OR) estimated between the numbers of case- and control-specific SNP pairs at a significance threshold of 10^{-7} in each cell type. The area inside the squares is proportional to the number of observations for each comparison. Hippocampal DG-like neuron promoter-PIR pairs showed case-specific enrichments of ASD SNP pairs (chi-squared test, p<2.7x10^{-16}). The confidence intervals for each OR estimation are shown in blue, and the red line represents a baseline OR of 1.

Figure S1

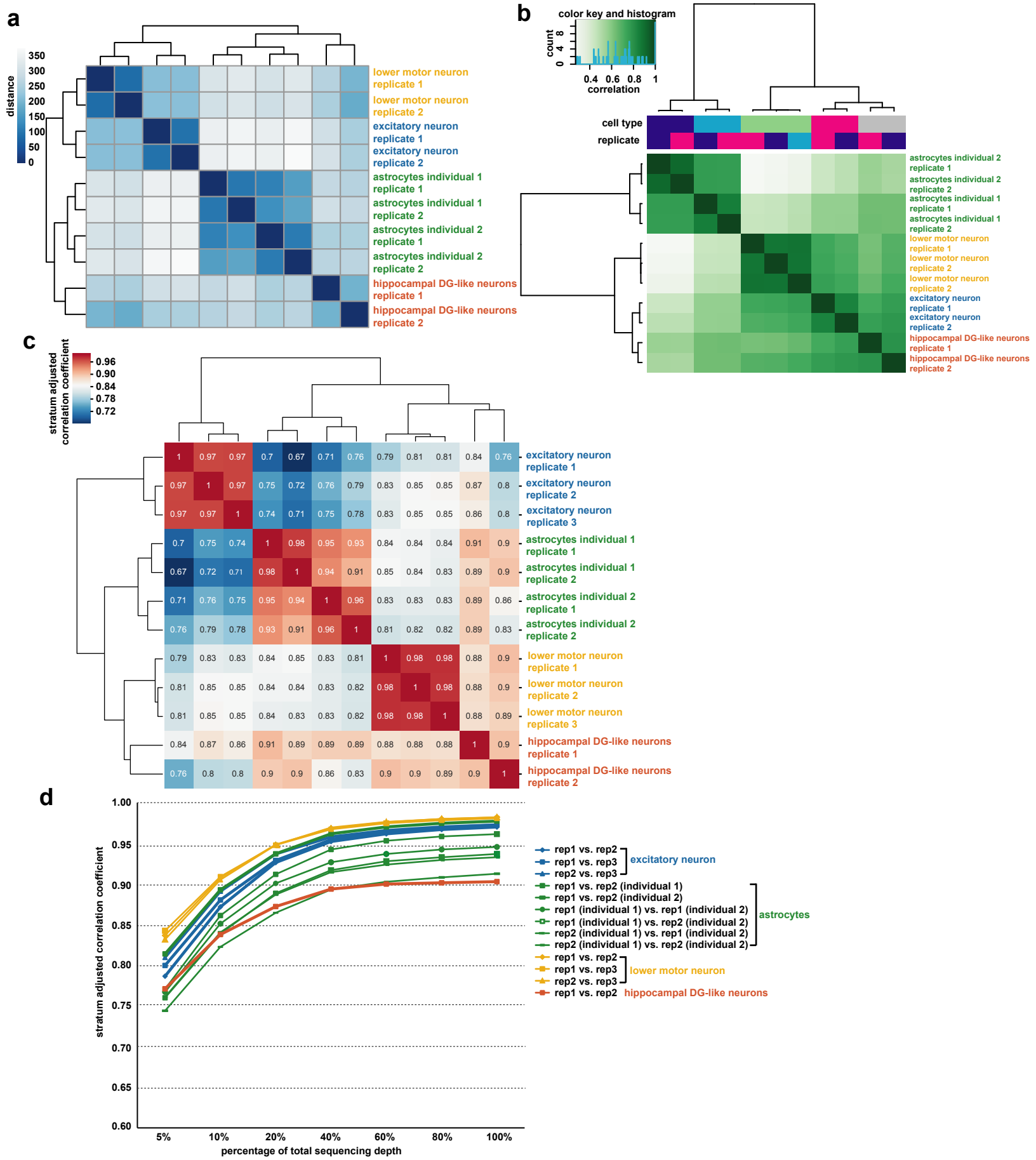






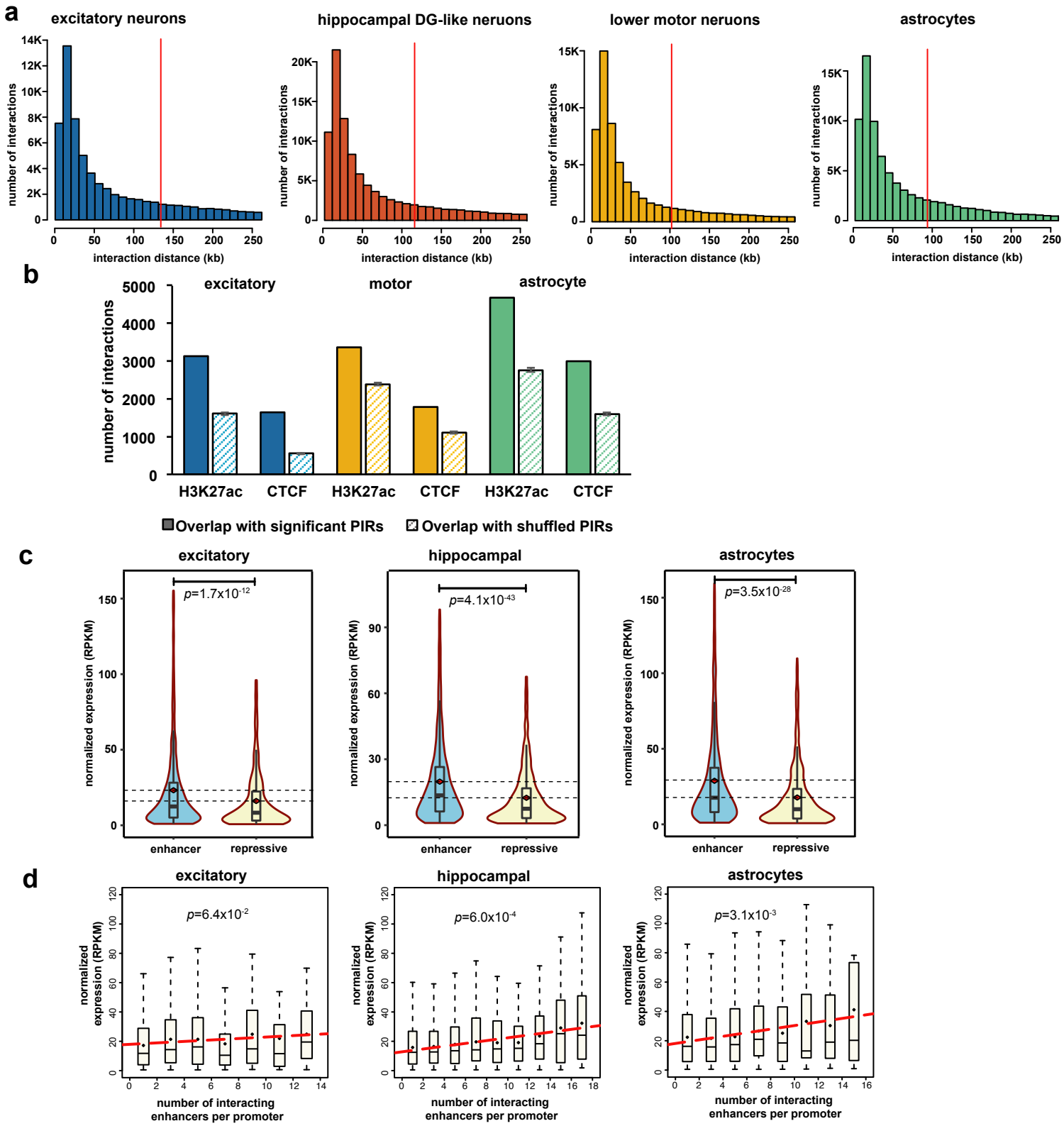
Supplementary Figure 1. Characterization of the cell types used in the study.
(**a**) Immunofluorescence staining of key markers in excitatory neurons, hippocampal DG-like neurons, lower motor neurons, and astrocytes. Excitatory neurons were positively stained for CUX1, an upper cortical layer marker, and MAP2, a neuronal marker which is specifically expressed in dendrites. The yield of excitatory neurons is calculated as the number of CUX1 and MAP2 double positive cells divided by the total number of live cells. Hippocampal DG-like neurons were positively stained for PROX1, a transcription factor specifying granule cell identity in the DG. The yield of mature hippocampal DG-like neurons is calculated as the number of PROX1 and MAP2 double positive cells divided by the total number of live cells. Lower motor neurons were positively stained for HB9, a motor neuron marker, and the pan-neuronal neurofilament marker SMI32. The yield of mature lower motor neurons is calculated as the number of HB9 and SMI32 double positive cells divided by the total number of live cells. Finally, astrocytes were positively stained for GFAP. The yield of GFAP-positive astrocytes is calculated as the number of GFAP positive cells divided by the total number of live cells. The number of staining experiments and the total number of cells is indicated, and error bars represent the SEM.
(**b**) Heatmaps displaying the expression of key marker genes for the neural cell types. Astrocytes used in this study exhibit an expression profile consistent with APC identity.
(**c**) Counts of protein coding (dark blue) and noncoding RNA (light blue) genes with promoters interacting in each cell type.

Figure S2



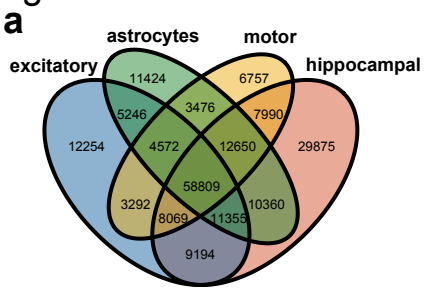
Supplementary Figure 2. Correlation between pcHi-C, ATAC-seq, and RNA-seq replicates.
(**a**) Gene expression values for each RNA-seq replicate were hierarchically clustered according to sample distances using DESeq2.
(**b**) Heatmap with pairwise correlations and hierarchical clustering of read densities at the set of unified open chromatin peaks for the ATAC-seq replicates.
(**c**) Heatmap with pairwise correlations based on the stratum-adjusted correlation coefficient (SCC) from HiC-Rep (evaluated at a resolution of 10 kb) for the pcHi-C replicates.
(**d**) Saturation of the SCC between biological replicates for the pcHi-C libraries as a function of total sequencing depth.

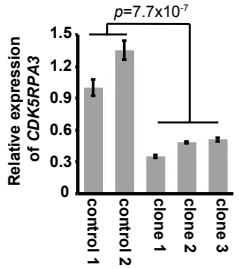
Figure S3

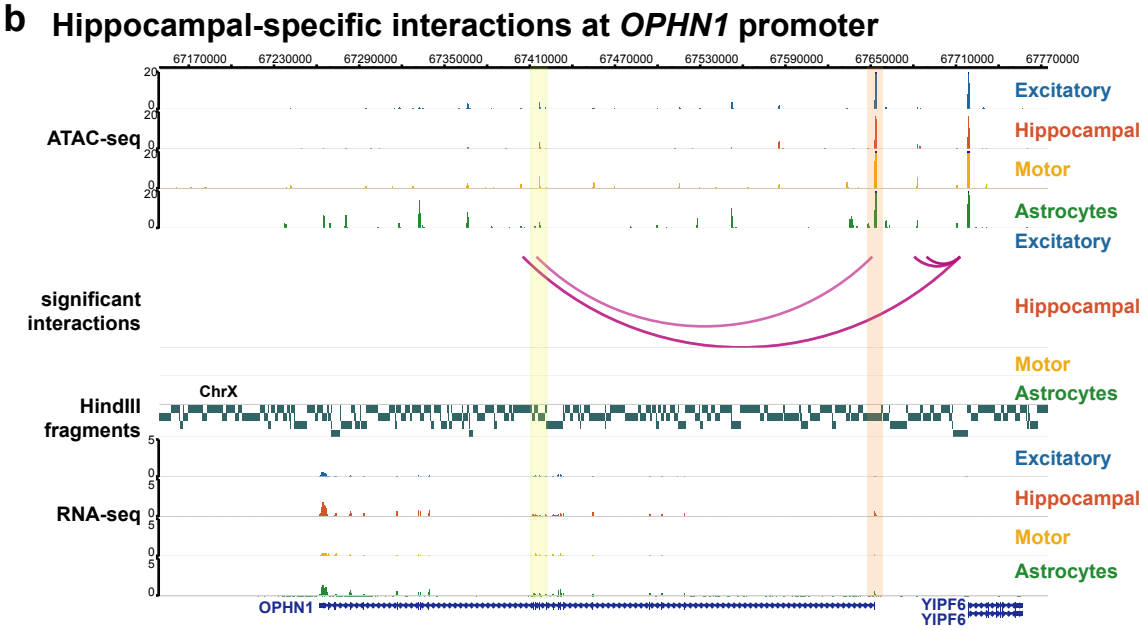


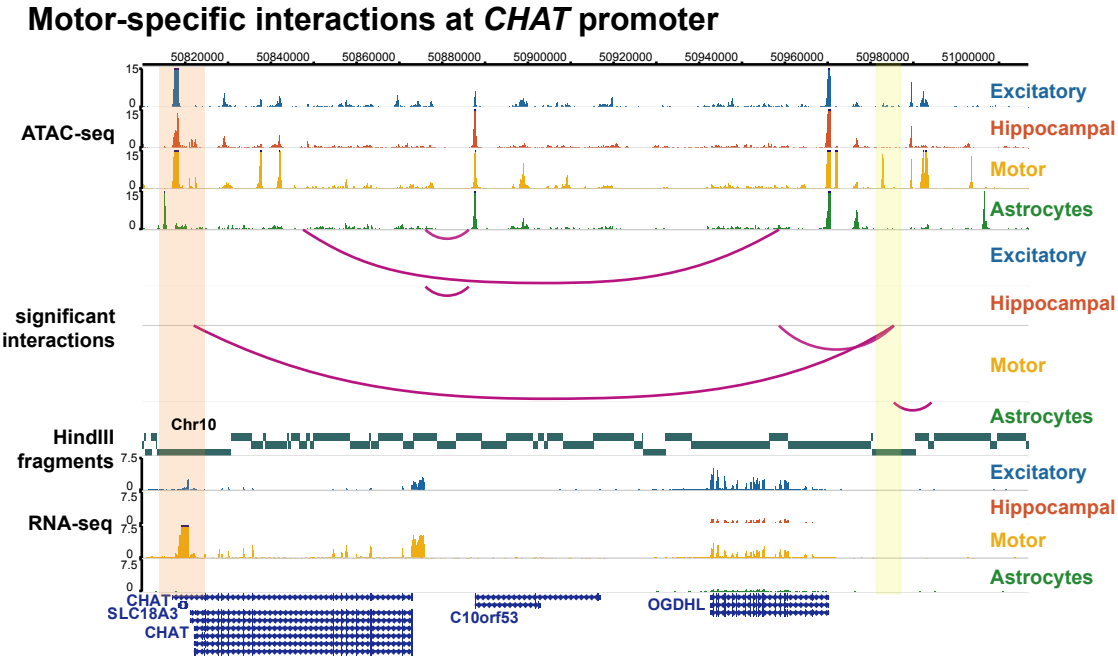
Supplementary Figure 3. Integrative analysis of chromatin interactions in individual cell types.
(**a**). Histograms of interaction distances for each cell type. The mean interaction distances for each cell type are indicated with red lines.
(**b**) Bar plots showing counts of H3K27ac and CTCF binding sites overlapping significant (solid bars) versus randomly shuffled (striped bars) PIRs for excitatory neurons, lower motor neurons, and astrocytes. Error bars represent the standard deviation over 100 sampled sets of randomly shuffled PIRs.
(**c**) Comparative gene expression analysis in individual cell types for expressed genes (normalized RPKM > 0.5) whose promoters interact exclusively with either enhancer-PIRs (n=6836) or repressive-PIRs (n=2612). Distributions of gene expression values are shown for each group.
(**d**) Boxplots showing distributions of gene expression values in individual cell types for expressed genes (normalized RPKM > 0.5) grouped according to the numbers of interactions their promoters form with enhancer-PIRs. Linear regression was performed on the mean gene expression values for each bin. Only bins containing at least 10 genes were included in the analysis.

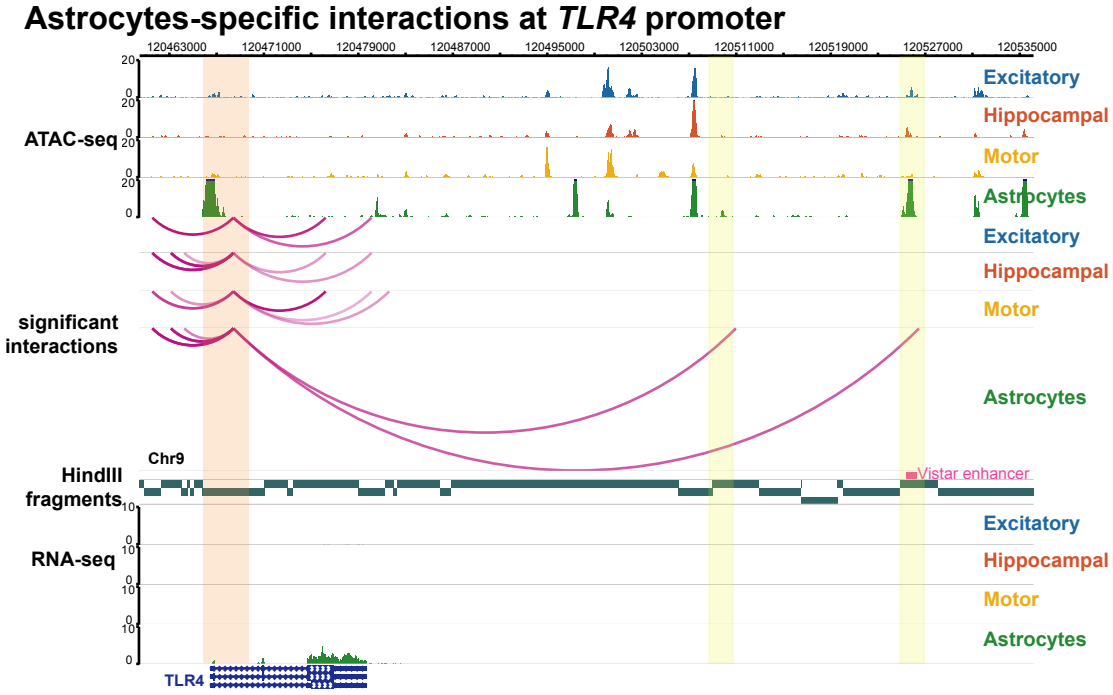
Figure S4







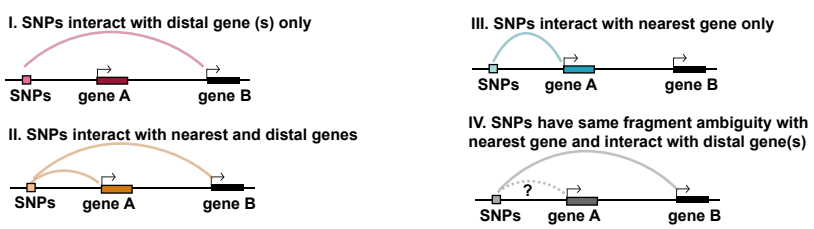




Supplementary Figure 4. Cell type-specific aspects of chromatin interactions.
(**a**) Venn diagram displaying counts of unique promoter-PIR interactions across excitatory neurons, hippocampal DG-like neurons, lower motor neurons, and astrocytes for each specificity pattern (groups 1-15 in **Fig. 3a**).
(**b**) Examples of interactions between cell type-specific PIRs (yellow highlight) and the promoters for *OPHN1*, *CHAT*, and *TLR4* (orange highlight). Open chromatin peaks and gene expression are also displayed for each of the cell types.
(**c**) Significant downregulation of *CDK5RAP5* expression was observed across three independent clones with homozygous deletions for the candidate PIR in excitatory neurons (two sample t-test, two-sided, $p=7.7\times10^{-7}$). Error bars represent the SEM.

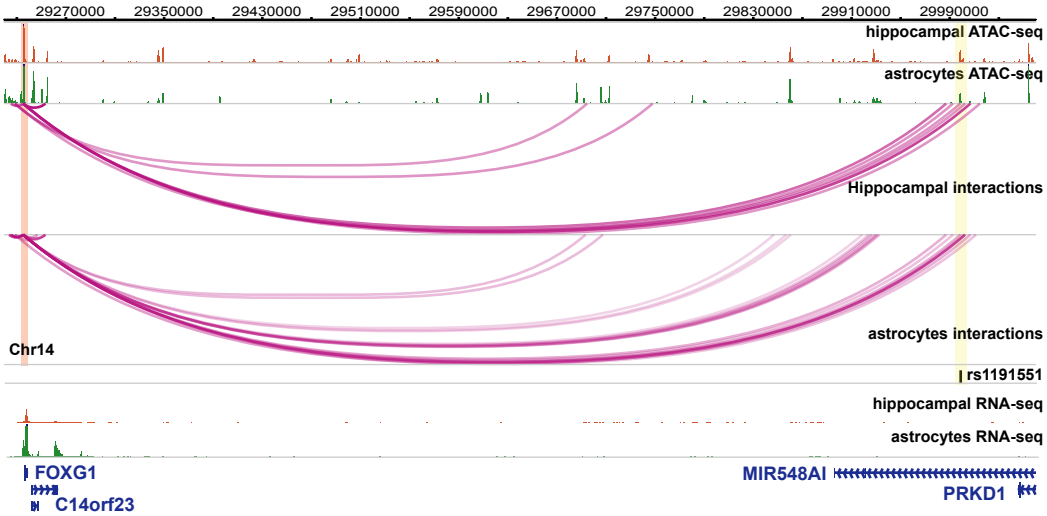
Figure S5

a

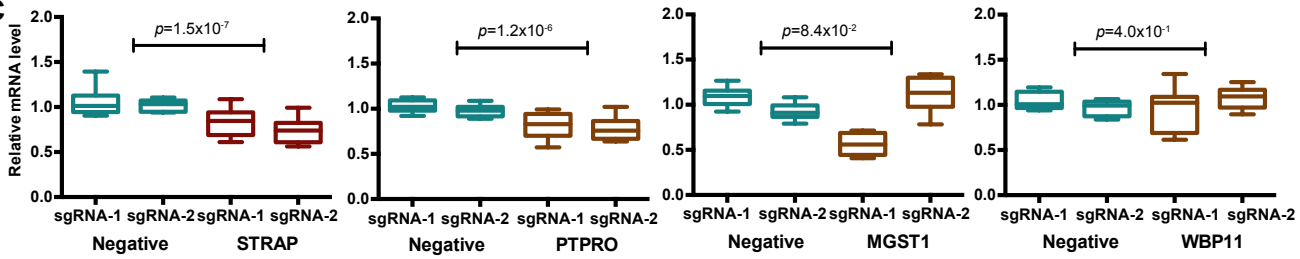


	I		II		III		IV	
	SNPs	genes	SNPs	genes	SNPs	genes	SNPs	genes
AD	100	338	179	1,545	19	19	61	529
ADHD	25	14	25	174	6	6	1	0
ALS	30	75	14	110	4	4	1	15
ASD	55	200	60	637	9	9	6	89
BD	67	84	59	713	11	11	2	9
EP	9	20	3	31	4	4	0	0
FTD	8	64	2	14	0	0	0	0
MP	874	2,082	695	7,820	152	152	51	371
PD	31	46	46	431	9	9	2	3
SCZ	190	435	203	2,444	33	33	10	107
UD	96	278	61	650	16	16	5	14

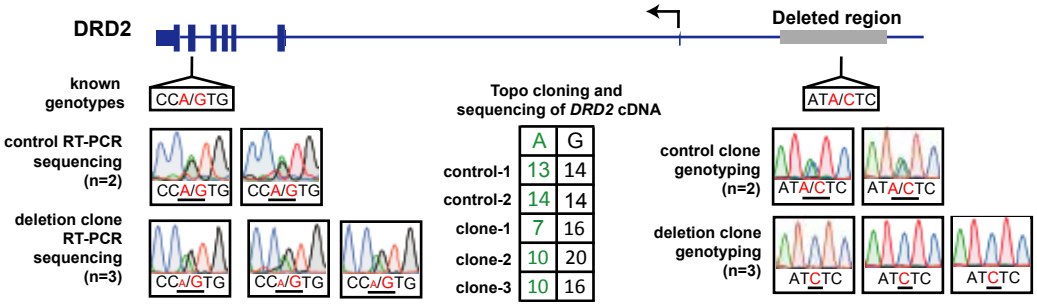
b



c

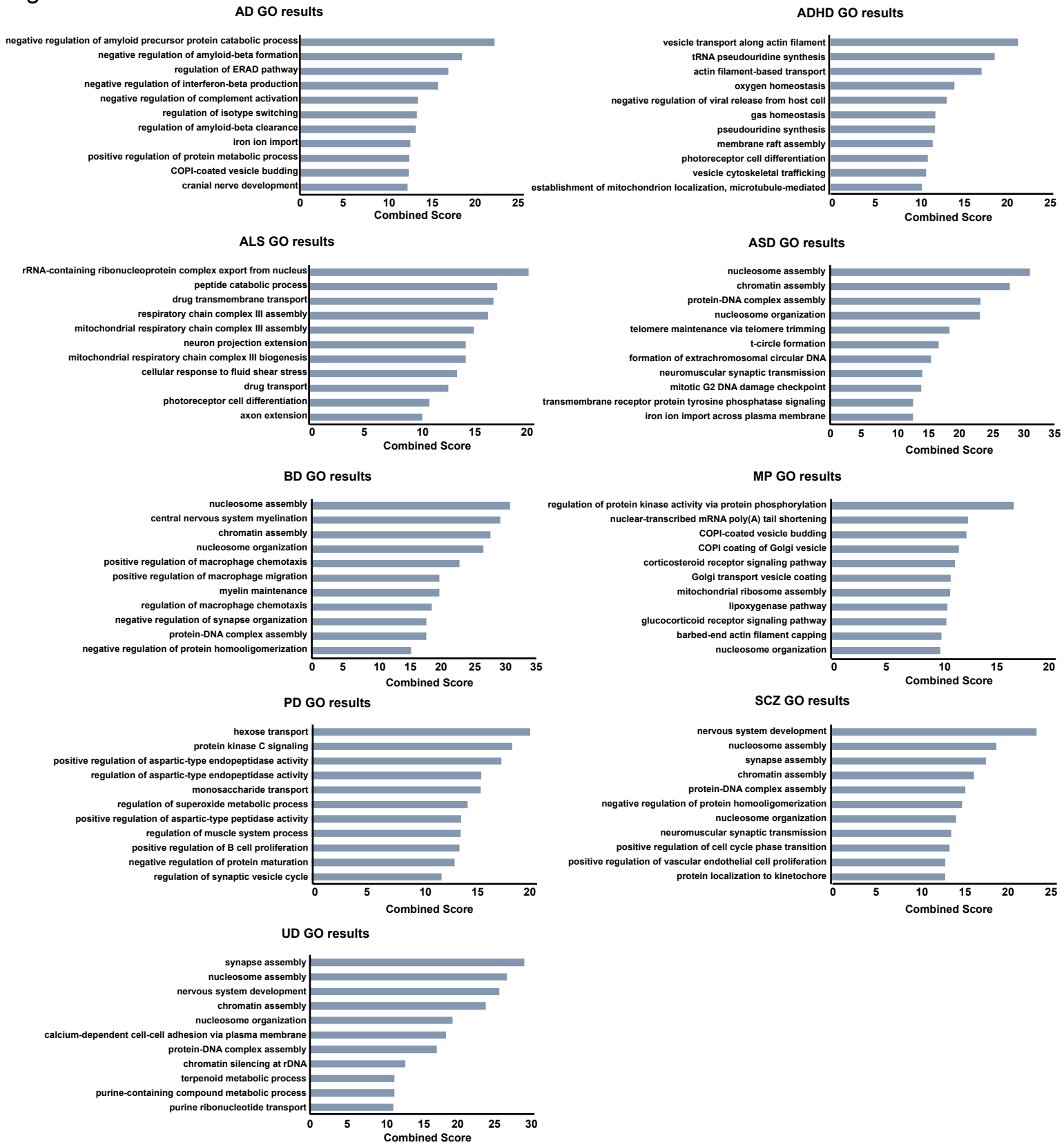


d



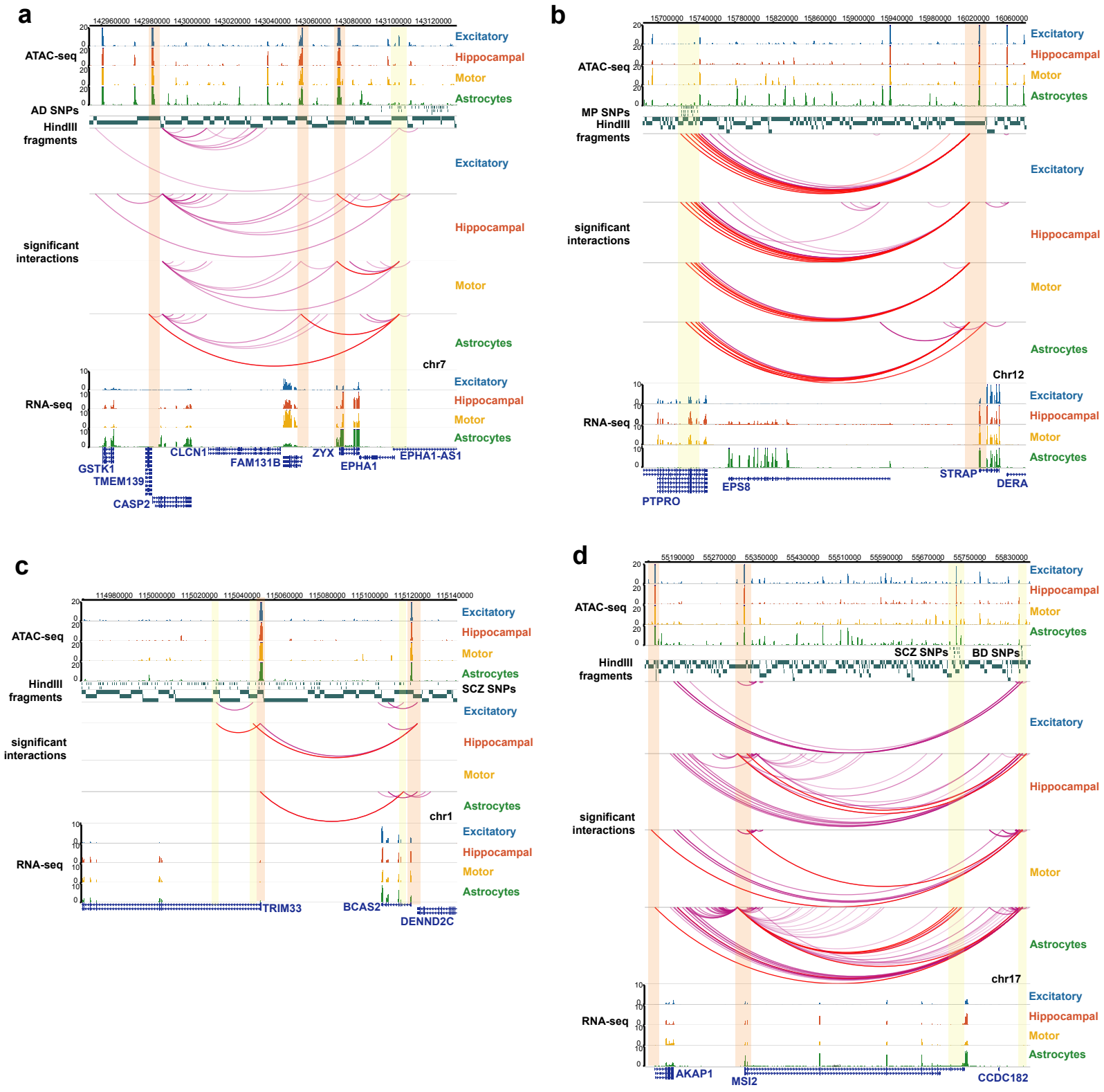
Supplementary Figure 5. Using chromatin interactions to elucidate the functions of GWAS variants.
(**a**) Counts of GWAS SNPs for each disease with at least one SNP in linkage interacting exclusively with their nearest genes (scenario III), interacting exclusively with more distal genes (scenario I), or interacting with both their nearest genes and more distal genes (scenario II). GWAS SNPs that could not be resolved for interactions with their nearest genes are also tabulated (scenario IV), along with counts of regulatory targets interacting with GWAS SNPs in each scenario.
(**b**) Significant promoter-PIR interactions in hippocampal DG-like neurons and astrocytes recapitulate a previously reported interaction between the *FOXG1* promoter and a distal open chromatin peak containing rs1191551, a schizophrenia-associated variant6.
(**c**) CRISPRi silencing of a candidate PIR for *STRAP* using two independent sgRNAs results in significant downregulation of *STRAP* expression in excitatory neurons (two sample t-test, two-sided, $p=1.5\times10^{-7}$). No significant downregulation was detected for the neighboring genes *MGST1* and *WBP1*, though the expression of *PTPRO* was affected. Each CRISPRi experiment was performed in triplicate, with three technical replicates per experiment.
(**d**) Schematic of detected genotypes in the *DRD2* gene and its candidate PIR in mono-allelic deletion clones. Genotyping and RT-PCR sequencing for WTC11 variants in the *DRD2* gene reveal allele-specific imbalances in *DRD2* expression, consistent with the deletion of the candidate PIR in one of the alleles. Results for two WT control clones and three mono-allelic deletion clones are shown for comparison.

Figure S6



Supplementary Figure 6. Top enriched GO terms for genes targeted by complex neurological disorder-associated variants.
Top enriched GO terms from Enrichr for genes whose promoters are targeted by variants for each disease. EP and FTD are omitted due to their low numbers of reported variants and target genes identified by our significant promoter-PIR interactions. An expanded list of enriched GO terms is available in **Supplementary Table 9**. In general, we observe enrichment of terms associated with epigenetic, neuronal, and disease-specific processes across the diseases.

Figure S7



Supplementary Figure 7. Examples of putative regulatory SNPs at cell type-specific PIRs.
In all examples, interacting PIRs are highlighted in yellow and the targeted promoters are highlighted in orange.
(**a**) A PIR containing AD SNPs interacts with the promoters of *FAM131B* and *CASP2* in astrocytes, but interacts instead with the *ZYX* promoter in hippocampal DG-like neurons and lower motor neurons.
(b) PIRs containing MP SNPs in an intron for *PTPRO* interacts with the *STRAP* promoter in all four cell types.
(**c**) A PIR containing SCZ SNPs interacts with the *TRIM33* promoter in astrocytes. Two other PIRs containing SCZ SNPs interact with the promoters of *TRIM33* and *BCAS2* in hippocampal DG-like neurons.
(**d**) A PIR containing BD SNPs interacts with the *MSI2* promoter in hippocampal DG-like neurons, lower motor neurons, and astrocytes while also interacting with the *AKAP1* promoter in lower motor neurons and astrocytes. Meanwhile, another group of PIRs containing SCZ SNPs interact with the *MSI2* promoter exclusively in astrocytes.