

1 Multi-omics profiling of DNA methylation and gene expression alterations in 2 human cocaine use disorder

3 Eric Zillich¹, Hanna Belschner¹, Diana Avetyan¹, Diego Andrade-Brito^{1,2,3}, José Jaime
4 Martínez-Magaña^{2,3}, Josef Frank¹, Naguib Mechawar^{4,5}, Gustavo Turecki^{4,5}, Judit Cabana-
5 Domínguez⁶, Noèlia Fernández-Castillo⁷, Bru Cormand⁷, Janitza L. Montalvo-Ortiz^{2,3,8}, Markus
6 M. Nöthen⁹, Anita C. Hansson¹⁰, Marcella Rietschel¹, Rainer Spanagel^{10,12}, Stephanie H.
7 Witt^{1,11,12*}, Lea Zillich^{1,12,13*}

8 ¹ Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical
9 Faculty Mannheim, Heidelberg University, Mannheim, Germany

10 ² Department of Psychiatry, Yale University School of Medicine, New Haven, CT, USA

11

12 ³ VA CT Healthcare Center, West Haven, CT, USA

13 ⁴ McGill Group for Suicide Studies, Douglas Mental Health University Institute, Montreal, QC, Canada

14 ⁵ Department of Psychiatry, McGill University, Montreal, QC, Canada

15 ⁶ Psychiatric Genetics Unit, Group of Psychiatry, Mental Health and Addiction, Vall d'Hebron Research
16 Institute (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona,
17 Spain. Biomedical Network Research Centre on Mental Health (CIBERSAM), Madrid, Spain.

18 ⁷ Department de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona,
19 and Institut de Biomedicina de la Universitat de Barcelona (IBUB); Barcelona, Catalonia, Spain; Centro
20 de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III
21 (ISCIII); Madrid, Spain; Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat,
22 Barcelona, Catalonia, Spain

23 ⁸ US Department of Veterans Affairs National Center of Posttraumatic Stress Disorder, Clinical
24 Neurosciences Division, West Haven, CT, USA

25 ⁹ Institute of Human Genetics, University of Bonn, School of Medicine and University Hospital Bonn,
26 Bonn, Germany

27 ¹⁰ Institute of Psychopharmacology, Central Institute of Mental Health, Medical Faculty Mannheim,
28 Heidelberg University, Mannheim, Germany

29 ¹¹ Center for Innovative Psychiatric and Psychotherapeutic Research, Biobank, Central Institute of
30 Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

31 ¹² German Center for Mental Health (DZPG), partner site Mannheim/Heidelberg/Ulm

32 ¹³ HITBR Hector Institute for Translational Brain Research gGmbH, Mannheim, Germany

33

34 *shared last-authorship

35 corresponding author: stephanie.witt@zi-mannheim.de

36

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

37

38 **Abstract**

39 Structural and functional changes of the brain are assumed to contribute to excessive cocaine intake,
40 craving, and relapse in cocaine use disorder (CUD). Epigenetic and transcriptional changes were
41 hypothesized as a molecular basis for CUD-associated brain alterations. Here we performed a multi-
42 omics study of CUD by integrating epigenome-wide methylomic (N=42) and transcriptomic (N=25) data
43 from the same individuals using postmortem brain tissue of Brodmann Area 9 (BA9). Of the N=1,057
44 differentially expressed genes ($p < 0.05$), one gene, *ZFAND2A*, was significantly upregulated in CUD at
45 transcriptome-wide significance ($q < 0.05$). Differential alternative splicing (AS) analysis revealed N=98
46 alternatively spliced transcripts enriched in axon and dendrite extension pathways. Strong convergent
47 overlap in CUD-associated expression deregulation was found between our BA9 cohort and
48 independent replication datasets. Epigenomic, transcriptomic, and AS changes in BA9 converged at two
49 genes, *ZBTB4* and *INPP5E*. In pathway analyses, synaptic signaling, neuron morphogenesis, and fatty
50 acid metabolism emerged as the most prominently deregulated biological processes. Drug repositioning
51 analysis revealed glucocorticoid receptor targeting drugs as most potent in reversing the CUD
52 expression profile. Our study highlights the value of multi-omics approaches for an in-depth molecular
53 characterization and provides insights into the relationship between CUD-associated epigenomic and
54 transcriptomic signatures in the human prefrontal cortex.

55

56 **Introduction**

57 Cocaine use disorder (CUD) is a globally prevalent substance use disorder (SUD) with around 4.2 million
58 people worldwide being diagnosed with CUD¹. Individuals suffering from CUD present with compulsive
59 cocaine use patterns, strong cocaine craving, and high rates of relapse even after prolonged time of
60 abstinence². Currently, there is no FDA-approved pharmacotherapy for CUD and treatment is mainly
61 focused on symptom reduction³. Neurobiological alterations in the brain are assumed to contribute to
62 the observed clinical symptoms in CUD⁴. This is supported by neuroimaging studies that have shown
63 profound structural and functional alterations in the brain in individuals with CUD^{5,6}. In addition to striatal
64 brain regions involved in reward processing⁷, frontal cortical areas that are neuroanatomically connected
65 with limbic structures, are implicated in addiction due to their importance for inhibitory control^{5,6,8}.

66

67 Dynamic changes in epigenetics and gene expression were hypothesized as a molecular basis of CUD-
68 associated brain changes^{9,10}. So far, the majority of studies investigating brain tissue focused on rodent
69 models of cocaine addiction, identifying specific genomic loci to be differentially methylated in brain
70 regions such as the prefrontal cortex (PFC)¹¹ and nucleus accumbens (NAc)¹². Gene expression levels
71 are tightly regulated by epigenetic mechanisms and DNA methylation (DNAm) changes especially in
72 gene promoter regions were shown to alter transcript abundance¹³. In line with this, differential gene
73 expression in rodent models of cocaine addiction was reported in different brain regions where
74 transcription factors of the immediate early gene (IEG) family such as *Egr1*, *Nr4a1*, and *Fos* were found
75 to be differentially expressed¹⁴⁻¹⁷. At the transcriptome-wide scale, differentially expressed genes were
76 consistently enriched in biological processes related to neurotransmission and ion channel activity, but
77 also metabolic alterations related to lipid metabolism and ATP homeostasis were found¹⁴.

78
79 Few studies have been performed investigating genome-wide DNAm or transcriptomic changes in CUD
80 in human postmortem brain tissue. Two epigenome-wide studies using reduced representation bisulfite
81 sequencing (RRBS) in a cohort of N=25 individuals with CUD and N=25 control individuals identified
82 N=145 and N=173 CUD-associated differentially methylated regions (DMRs) in the nucleus accumbens
83 (NAc)¹⁸ and in the caudate nucleus (CN)¹⁹, respectively. Investigating the same brain regions in a
84 different cohort (N=25 CUD cases, N=20 controls), another study characterized transcriptome-wide
85 gene expression changes and reported on the upregulation of synaptic transmembrane transporter
86 genes while immune processes were downregulated²⁰. The largest study in the human PFC
87 investigating CUD-associated transcriptomic changes (N=19 CUD, N=17 controls) identified N=883
88 nominally significant ($p < 0.05$) differentially expressed genes (DEGs) in neuronal nuclei from the
89 Brodmann Area 46 subregion²¹. CUD-associated co-expression networks were enriched for GTPase
90 signaling and neurotransmitter secretion. Regarding epigenomic alterations in the PFC, we were
91 previously able to identify 20 CUD-associated DMRs in Brodmann Area 9, a subregion of the PFC, and
92 further detected that co-methylation networks in CUD were enriched for synaptic signaling processes²².
93 Although epigenetics represents an important regulatory mechanism for transcription, the co-regulation
94 of DNAm and gene expression in the same brain samples has not yet been investigated in CUD, limiting
95 the comparability of results between epigenetic and gene expression studies.

96

97 In addition to epigenetics and transcription, alterations of alternative splicing might contribute to the
98 neurobiological changes in the CUD brain, as shown in other SUDs. Previous studies using postmortem
99 human brain tissue from individuals with alcohol use disorder (AUD)²³⁻²⁵ and opioid use disorder (OUD)²⁶
100 detected differential alternative splicing in transcripts of genes implicated in neuropsychiatric disorders,
101 such as *BIN1*, *FLOT1*, and *ELOVL7* suggesting RNA splicing alterations to be a further molecular
102 mechanism in the neurobiology of SUDs. While a recent study using a cocaine self-administration model
103 in mice showed widespread changes in alternative splicing in multiple brain regions²⁷, no systematic
104 evaluation of splicing alterations in human CUD was performed so far.

105
106 In the present study, we aimed to characterize the molecular underpinnings of CUD in the human
107 prefrontal cortex by applying a multi-omics analysis approach. We investigated differentially expressed
108 genes in postmortem brain tissue from deceased CUD cases compared to well-matched controls and
109 integrated them with the results of our epigenome-wide DNAm analysis¹⁸ from the same individuals of
110 the BA9 subregion of the human PFC. Further, we characterized differential alternative splicing in BA9.
111 We then performed replication analysis of CUD-associated DEGs in two other independent RNA-seq
112 datasets of human dlPFC. Gene expression data, including alternative splicing results, and DNA
113 methylation data were then integrated and put into a biological context. Finally, we addressed the urgent
114 need for novel therapeutic approaches, by performing a drug repositioning analysis based on the CUD-
115 associated transcriptional profile in BA9.

116
117 Collectively, our multi-omics study design represents an integrated analysis of DNAm and gene
118 expression data together with alternative transcript splicing that highlights the role of synaptic and
119 metabolic alterations in CUD and the glucocorticoid receptor as a pharmacological candidate target.

120

121 **Results**

122 **Individuals with and without CUD do not differ in sociodemographic characteristics and cell type** 123 **composition**

124 We first assessed the phenotypic similarities between CUD cases and controls. No significant
125 differences were observed regarding the pH value of the brain, postmortem interval (PMI), RNA integrity
126 number (RIN), and occurrence of comorbid depressive and alcohol use disorders (Supplementary Table
127 S1). We further investigated the variance partition of potential covariates in the RNA-seq dataset

128 (Supplementary Fig. S1a) and found age, PMI, brain pH, and RIN to be associated with gene expression
129 levels, and hence included them as covariates in further analyses. To explore whether major cell type
130 composition could affect analysis results, we performed a cell type deconvolution analysis using
131 CIBERSORT based on human PFC major cell type marker gene signatures (Supplementary Fig. S1b,
132 Supplementary Table S2a). No significant differences in the distribution of major cell types such as
133 astrocytes, oligodendrocytes, microglia, neurons and others were detected between samples from
134 individuals with and without CUD as all 95% high-density intervals from the Bayesian estimation
135 contained 0 (Supplementary Table S2b).

136

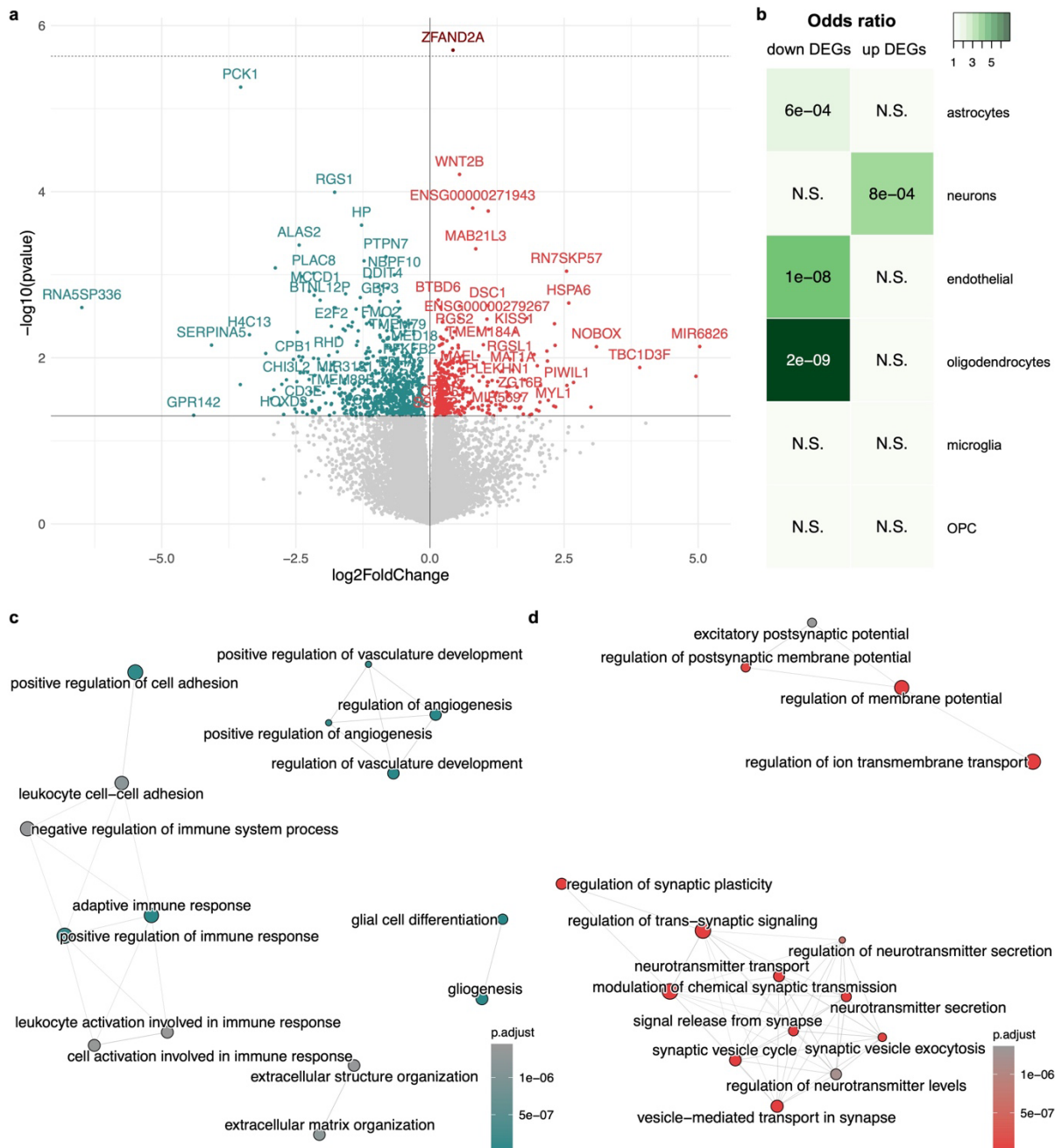
137 **Transcriptome-wide differential gene expression patterns in CUD are related to synaptic** 138 **signaling, ion transport, and inflammatory processes**

139 The transcriptome-wide analysis of differential expression in BA9 revealed a total of N=1,057 DEGs
140 associated with CUD ($p < 0.05$). Of these, N=378 were upregulated and N=679 were downregulated (Fig.
141 1a, Supplementary Table S3). After adjustment for multiple testing, *ZFAND2A* (*Zinc Finger AN1-Type*
142 *Containing 2A*, $\log_2FC = 0.43$, $p = 1.98e-06$, $q = 0.04$), remained significantly upregulated in individuals with
143 CUD (5% FDR). We observed a genomic inflation factor of $\lambda = 1.29$ (Supplementary Fig. S1c) and results
144 were stable in a sensitivity analysis without individuals with AUD or MDD (Supplementary Fig. S1d). To
145 evaluate whether BA9 DEGs are significantly enriched within cell-type specific genes of the human PFC,
146 we performed an overlap analysis, using the same set of major brain cell type marker genes as in the
147 cell type composition analysis. Upregulated DEGs were significantly enriched for neuron marker genes
148 exclusively, whereas downregulated DEGs were significantly enriched in markers of non-neuronal cell
149 types such as astrocytes, endothelial cells, and oligodendrocytes (Fig. 1b).

150

151 Next, we were interested in the biological functions related to the identified CUD-associated DEGs. After
152 adjusting for multiple testing, we detected N=276 statistically significant GO terms for positive GSEA
153 normalized enrichment scores (NES, Supplementary Table S4) and N=782 significant GO terms for
154 negative NES (Supplementary Table S5). Among significantly enriched pathways, the largest positive
155 NES was detected for “vesicle-mediated transport in synapse” (NES=2.63, $q = 5.31e-15$), whereas
156 “superoxide metabolic process” (NES=-2.43, $q = 2.08e-06$) was the top finding with negative NES. To
157 identify functional modules of pathways consisting of multiple GO terms related to similar biological
158 functions, we created an enrichment map (emap) visualization based on the significant findings from

159 GSEA. For GO terms with negative NES, we detected one large cluster related to inflammatory and
 160 immune signaling and several smaller clusters consisting of pathways involved in angiogenesis,
 161 extracellular matrix (ECM) organization, and gliogenesis (Fig. 1c). Two major clusters emerged for
 162 pathways with positive NES. The first was related to neurotransmission and synaptic signaling whereas
 163 the second cluster consisted of GO terms involved in transmembrane transporter activity (Fig. 1d).



164
 165 **Fig. 1 – Differential expression analysis in CUD suggests synaptic signaling and immunological alterations in Brodmann**
 166 **Area 9**

167 **a** Volcano plot of the differential expression (DE) analysis showing the N=378 upregulated (red) and N=679 downregulated genes
 168 (blue) at nominal significance ($p < 0.05$). Solid black line indicates nominal significance ($p < 0.05$), dashed gray line indicates
 169 transcriptome-wide significance (FDR $q < 0.05$). **b** Results of the overlap analysis for upregulated (up) and downregulated (down)

170 DEGs among cell-type specific marker genes. Green color depicts the odds ratio (OR) of overlap, p-values inside the panels
171 indicates significance of overlap based on Fisher-Test. Gene-set enrichment analysis (GSEA) was performed for the DEGs in
172 BA9 ranked by the Wald test statistic from DESeq2. Statistically significant results ($q < 0.05$) from GSEA separated by **c** negative
173 and **d** positive normalized enrichment scores (NES) are shown in an enrichment map visualization. N.S. = not significant, OPC =
174 oligodendrocyte progenitor cell.

175

176 **Network analysis highlights fatty-acid metabolism and morphogenesis processes in CUD**

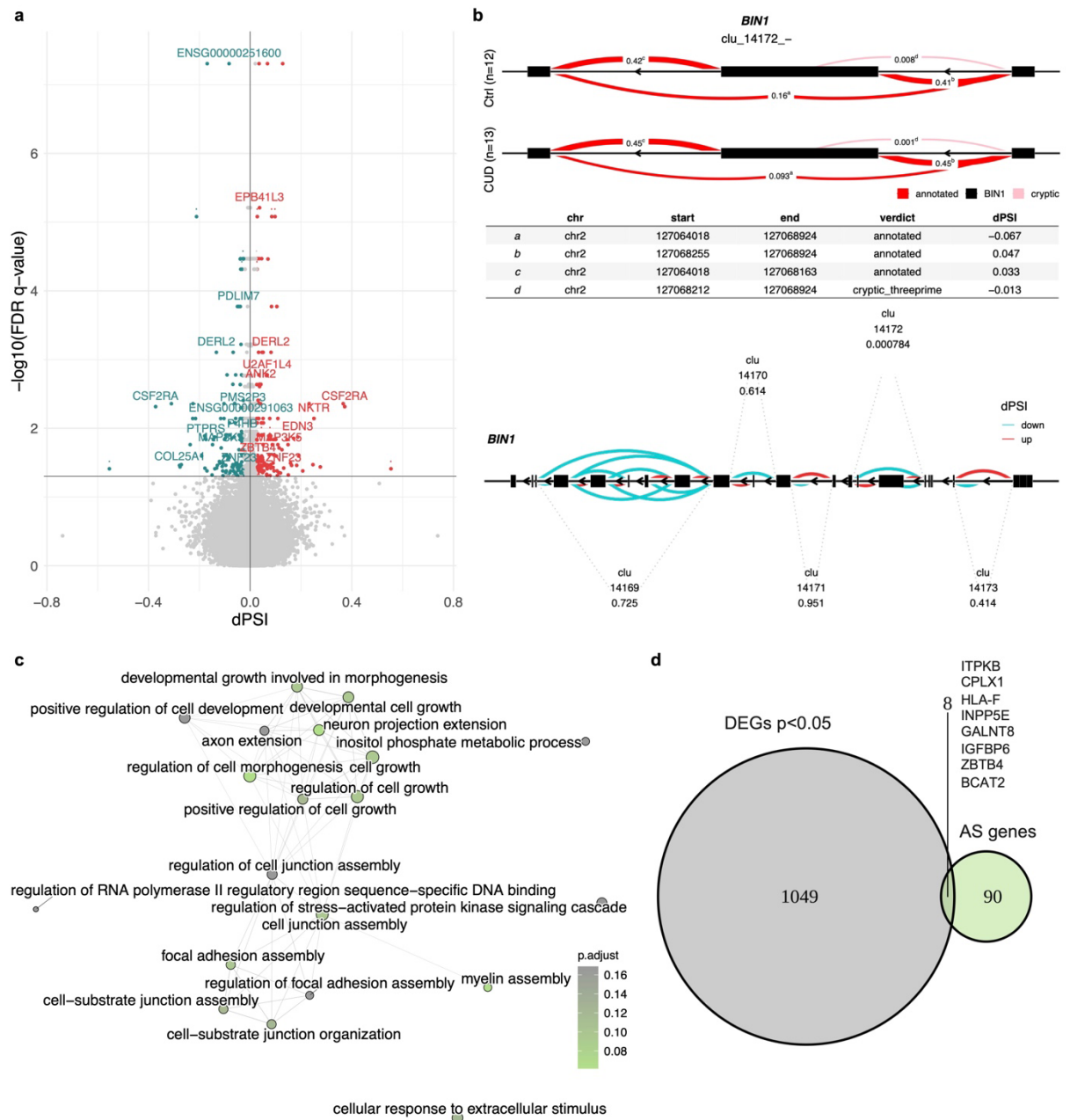
177 We next performed weighted gene co-expression network analysis (WGCNA) to investigate gene co-
178 expression patterns in CUD and detected a total of $N=27$ co-expression modules (Supplementary Fig.
179 S2a+2b). Co-expression module yellow was significantly correlated with CUD ($r=-0.47$, $p=0.02$) while
180 no significant association with other known covariates was observed (Supplementary Fig. S2a). Module
181 yellow consisted of $N=2,517$ genes and module membership was highly correlated with gene
182 significance for CUD ($r=0.61$, $p < 1e-200$, Supplementary Fig. S2c, Table S6). GO enrichment analysis
183 for module yellow genes revealed $N=519$ statistically significant GO terms after multiple testing
184 correction (Supplementary Table S7). Strongest associations were detected for “small molecule
185 catabolic process” ($q=2.22e-11$), and more specifically, “carboxylic acid catabolic process” ($q=1.59e-09$,
186 Supplementary Fig. S2d). After clustering the significant terms, a prominent GO term cluster related to
187 fatty acid metabolism was detected, while another cluster was related to organ developmental and
188 morphogenesis processes. To further characterize WGCNA expression module yellow, we generated a
189 protein-protein interaction (PPI) network based on module hub genes and identified APOE ($N=9$ edges),
190 ERBB2 ($N=8$ edges), ALDH7A1 ($N=7$ edges), PPARA ($N=7$ edges), and TLR4 ($N=7$ edges) as the most
191 strongly connected nodes in the PPI network (Supplementary Fig. S3, Supplementary Table S6).

192

193 **Genes with alternative splicing events in CUD are involved in cell junction formation and the** 194 **morphogenesis of axons and dendrites**

195 To investigate alternative splicing in CUD and its potential relevance for contributing to altered
196 neurobiological functions in the brain, we performed a differential alternative splicing analysis using
197 LeafCutter²⁸. After multiple testing correction, we identified a total of $N=108$ differentially spliced intron
198 clusters in BA9 ($FDR < 0.05$, Fig. 2a, Supplementary Tables S8-S10). These clusters were distributed
199 among $N=98$ genes that we further denote as alternatively spliced (AS) genes. One of the top findings
200 in our AS analysis of CUD was *BIN1* (*Bridging Integrator 1*, $q=7.8e-04$, Fig. 2b) previously identified as
201 a conserved AS genes in other substance use disorders. We next investigated the biological pathways

202 enriched for alternative splicing events based on our list of AS genes. Statistically significant enrichment
203 after multiple testing correction was detected for N=15 GO terms (Supplementary Table S10). Strongest
204 enrichment was found for GO terms “cell junction assembly” and “neuron projection extension” (both
205 $q=3.62e-03$). In the emap visualization of enriched GO terms with a more lenient threshold of 25% FDR
206 ($q<0.25$), we detected a well-connected cluster containing GO terms related to cellular growth and cell-
207 cell junction development, while also more brain-specific processes such as myelination and the
208 extension of axons and dendrites were found (Fig. 2c). While differential alternative splicing itself
209 contributes to altered biological functions by inducing different abundances of transcript isoforms, this
210 effect might be potentiated by differential gene expression. We thus investigated the overlap of AS and
211 DEGs in CUD and identified 8 genes that were differentially spliced and differentially expressed in BA9:
212 *ITPKB*, *CPLX1*, *HLA-F*, *INPP5E*, *GALNT8*, *IGFBP6*, *ZBTB4*, and *BCAT2* (Fig. 2d).
213



214

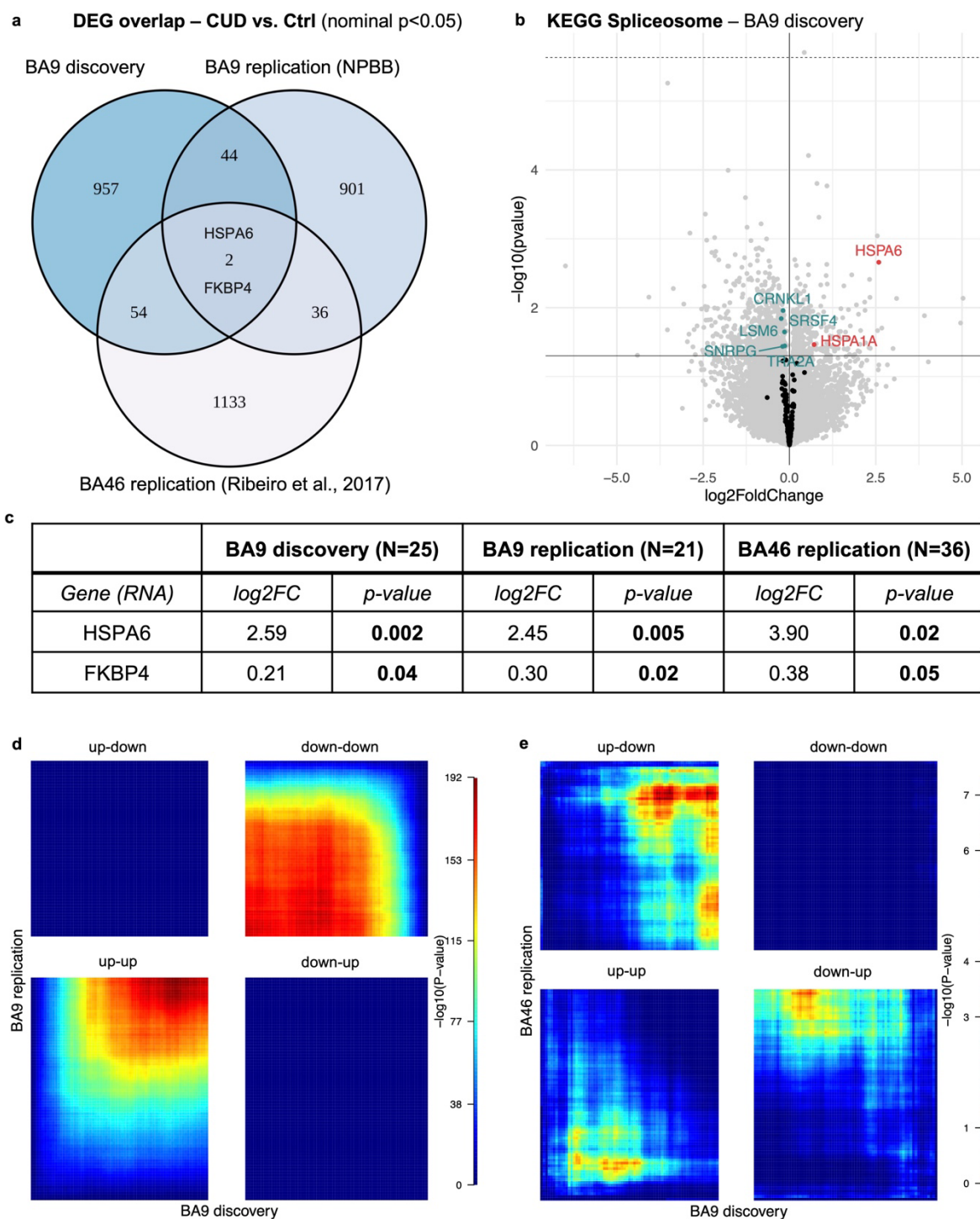
215 **Fig. 2 – Differential alternatively spliced genes in CUD are related to neuron morphogenetic processes**

216 **a** Volcano plot of the differential alternative splicing (AS) results in Brodmann Area 9 (BA9). Statistically significant intron clusters
 217 (N=108) identified by LeafCutter ($q < 0.05$) were annotated by gene name while dots represent individual introns of an intron cluster.
 218 Introns highlighted in red ($dPSI > 0$) are more abundant in CUD while introns highlighted in blue ($dPSI < 0$) are less abundant in
 219 CUD. **b** Results of the differential AS analysis at the cluster and gene level for one of the top findings, an intron cluster (clu_14172_-
 220) in the Bridging Integrator 1 (*BIN1*) gene. Upper panel: visualization of *BIN1* exons and introns with percent spliced in (PSI)
 221 measures related to the significant cluster clu_14172_-. The table indicates delta percent spliced in (dPSI) values from the CUD
 222 vs. Ctrl comparison. Lower panel: gene-level summary of all intron clusters detected in *BIN1*. FDR q-values are shown below
 223 cluster names. **c** GO enrichment analysis for the N=98 AS genes harboring differentially statistically significant ($q < 0.05$) intron
 224 clusters in CUD. **d** Overlap of findings from differential expression (DE) analysis (N=1057 DEGs at $p < 0.05$) and differential AS
 225 analysis.

226 **Replication analysis in independent cohorts reveals *FKBP4* and *HSPA6* as conserved DEGs in**
227 **CUD**

228 To evaluate the potential replication of CUD-associated DEGs in other RNA-seq datasets of human
229 PFC, we performed an overlap analysis of nominally significant DEGs ($p < 0.05$) across studies. CUD-
230 associated differential expression testing was performed in two independent replication datasets, the
231 first originating from BA9 (BA9 replication, bulk RNA-seq) and the second from BA46 (BA46 replication,
232 neuron-specific RNA-seq). Two genes, *HSPA6* and *FKBP4*, were shared upregulated DEGs at nominal
233 significance and showed comparable effect sizes (\log_2FC) in CUD across all three PFC datasets (Fig.
234 3a+c). As *HSPA6* is a spliceosome-associated gene with conserved differential expression across
235 datasets, we performed a look-up of genes related to the KEGG Spliceosome pathway (hsa03040) in
236 DE results from our discovery cohort (Fig. 3b). Here, we aimed to address the hypothesis of
237 spliceosomal differential gene expression as a potential mechanism for splicing alterations in CUD²⁹.
238 *HSPA6* was the spliceosome-associated gene showing strongest CUD-associated expression changes
239 in BA9 ($\log_2FC = 2.59$, $p = 0.002$). We detected six additional spliceosome-associated genes that were
240 among nominally significant DEGs: *HSPA1A* ($\log_2FC = 0.71$, $p = 0.034$), *CRNKL1* ($\log_2FC = -0.19$,
241 $p = 0.011$), *LSM6* ($\log_2FC = -0.24$, $p = 0.014$), *SRSF4* ($\log_2FC = -0.14$, $p = 0.022$), *SNRPG* ($\log_2FC = -0.20$,
242 $p = 0.037$), and *TRA2A* ($\log_2FC = -0.14$, $p = 0.036$).

243
244 Using rank-rank hypergeometric overlap (RRHO) visualization for a more unbiased evaluation of
245 convergent and divergent gene expression patterns across studies, we found strong convergent overlap
246 between the BA9 discovery and BA9 replication datasets indicating similar patterns of CUD-associated
247 expression deregulation (Fig. 3d). In the comparison with neuron-specific expression data from BA46,
248 we found prominent divergent gene expression patterns between datasets, while convergent expression
249 patterns were enriched in the shared upregulated genes across studies (Fig. 3e).



250 **Fig. 3 – Replication analysis of CUD associated transcriptomic alterations in independent datasets**

251 **a** Overlap of nominally significant ($p < 0.05$) differentially expressed genes across datasets reveals two shared DEGs, HSPA6 and
 252 FKBP4. The replication datasets were based on N=21 BA9 samples from the National PTSD Brain Bank (NPBB) and neuronal-
 253 specific transcriptomic data of N=36 BA46 samples available under GEO accession number GSE99349 (BA46 replication). **b**
 254 HSPA6 is the strongest spliceosome-associated DEG in BA9. **c** Results of the look-up analysis for shared DEGs HSPA6 and
 255 FKBP4 - \log_2 -fold change and p -value: association p -value from the DESeq2-based differential expression results. Significant
 256 associations are highlighted in bold. Rank-rank hypergeometric overlap (RRHO) visualization for **d** the BA9 replication dataset
 257 and **e** the neuronal-specific BA46 dataset indicating convergent and divergent expression patterns across studies using full

258 differential expression statistics as the input datasets. Color scale represents $-\log_{10}(p)$ of the hypergeometric testing procedure in
259 RRHO. Convergent expression across datasets: up-up and down-down, divergent: up-down and down-up.

260

261 **Drug repositioning analysis highlights glucocorticoid receptor targeting drugs to reverse the**

262 **CUD gene expression profile**

263 To evaluate the potential use of small molecule drugs to revert the gene expression pattern of CUD, we
264 performed drug repositioning analysis based on the L1000 assay as implemented in CMap
265 (Supplementary Fig. S4a), using the top 150 up- and downregulated genes as input (Supplementary
266 Table S11). Among the results with negative normalized connectivity score (NCS), i.e. perturbagens
267 that revert the DE profile in CUD, the most significant finding for small molecule drugs after multiple
268 testing correction was the glucocorticoid receptor agonist medrysone (NCS=-1.78, $q=2.2e-16$,
269 Supplementary Fig. S4b). Glucocorticoid receptor agonists were the only FDR-significant perturbagen
270 class overrepresented among CMap GSEA results (Supplementary Fig. S4c). When we further
271 investigated connectivity scores for all glucocorticoid receptor targeting drugs including agonists and
272 antagonists in CMap, we found exclusively significant negative connectivity scores (Supplementary Fig.
273 S4d) suggesting glucocorticoid receptor targeting molecules as potential pharmacological drugs to
274 revert the CUD expression changes in BA9. In line with this finding, the biological pathway “response to
275 glucocorticoid” (NES=-1.54, $q=0.019$, Supplementary Table S5) was among the FDR-significant
276 pathways with negative NES in the GSEA analysis of DEGs from BA9.

277

278 **Findings of the integrated analysis of DNA methylation and gene expression data converge at** 279 **the gene and pathway levels**

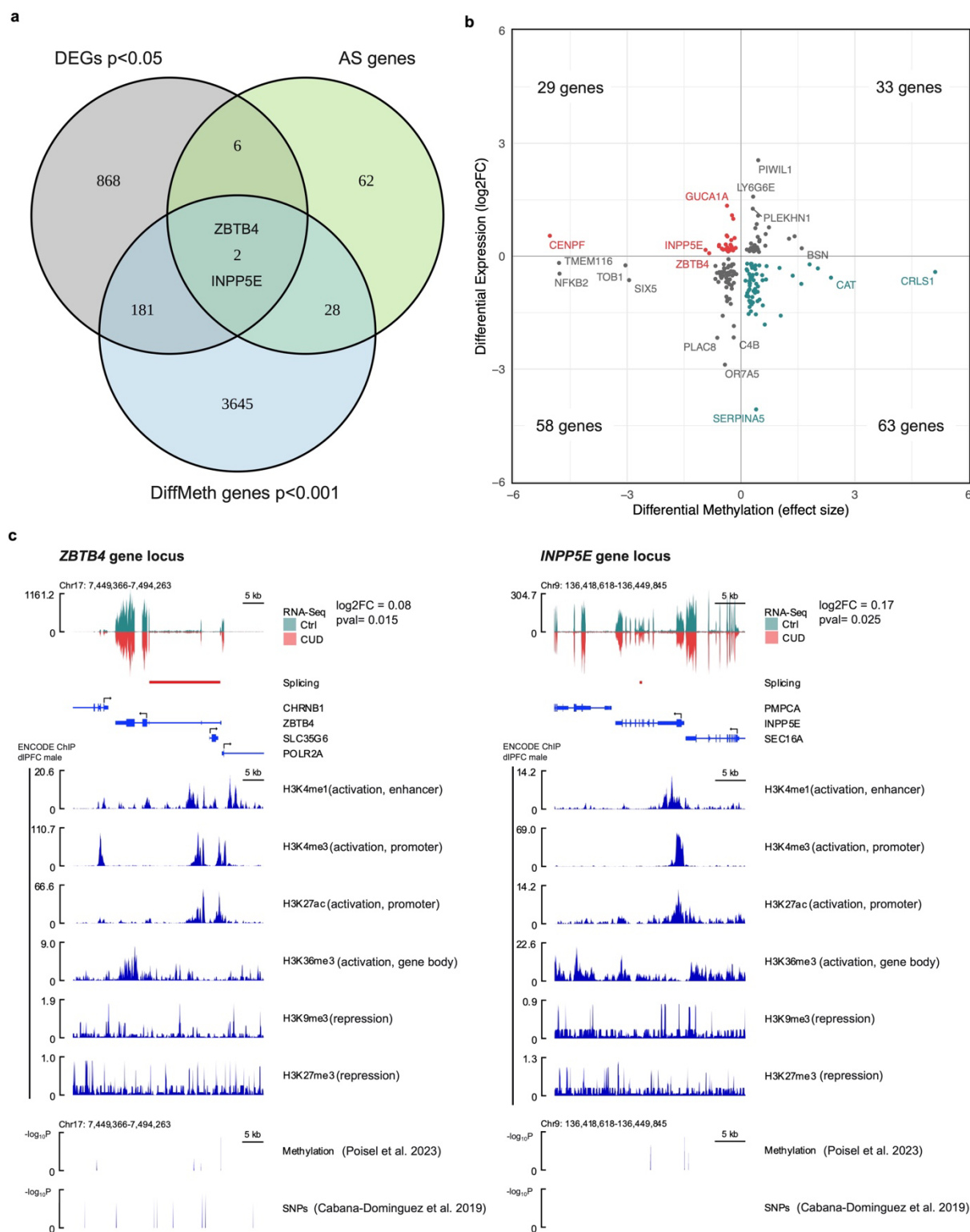
280 As DNA methylation data was available and previously analyzed for the same cohort in BA9, we next
281 aimed to integrate findings from the epigenome-wide and transcriptome-wide studies on the gene-level,
282 applied multi-omics factor analysis, and performed an integrative functional GO-term enrichment
283 analysis across all -omics layers.

284

285 *Gene-level integration of epigenomic, transcriptomic and splicing alterations highlights ZBTB4 and* 286 *INPP5E in CUD*

287 Of the overlapping genes between the differential methylation, expression, and alternative splicing
288 analyses, two genes were consistently altered across all the investigated molecular views in BA9:
289 *ZBTB4* (*Zinc Finger And BTB Domain Containing 4*) and *INPP5E* (*Inositol polyphosphate-5-*

290 *phosphatase E*) (Fig. 4a, Supplementary Table S12). Both genes were characterized by a
291 hypomethylated CpG site and increased transcript levels in CUD (Fig. 4b). For *ZBTB4*, the strongest
292 association for a CpG site was found for cg03443505 (chr17:7387573, $\beta=-0.84$, $p=1.01e-05$). *ZBTB4*
293 was upregulated with a log2FC of 0.08 ($p=0.015$) and it contained the differentially spliced intron cluster
294 *chr17:clu_10246_-* ($q=0.028$). The strongest association for CpG differential methylation in the *INPP5E*
295 gene was found for cg18558462 (chr9:139334381, $\beta=-0.93$, $p=8.55e-03$). It was differentially expressed
296 with log2FC of 0.17 ($p=0.025$) and intron cluster *chr9:clu_25078_-* was differentially alternatively spliced
297 ($q=0.015$). In the replication datasets, we detected conserved transcript upregulation of *ZBTB4* (BA9
298 replication, log2FC=0.12, pval=0.21; BA46, log2FC=0.12, pval=0.09) and *INPP5E* (BA9 replication,
299 log2FC=0.11, pval=0.47; BA46, log2FC=0.08, pval=0.49), however not statistically significant. To
300 deeper characterize the *ZBTB4* and *INPP5E* gene loci in BA9 and specifically in the context of CUD, we
301 performed an integrative gene locus visualization approach by combining GWAS, EWAS, alternative
302 splicing, and RNA-seq results for CUD with ENCODE ChIP-seq reference data from human dorsolateral
303 prefrontal cortex. ChIP-seq data confirmed the presence of activating chromatin marks at promoter
304 (H3K4me4, H3K27ac) and gene body regions (H3K36me3) at *ZBTB4* and *INPP5E* gene loci in the
305 human dlPFC. In addition, for *ZBTB4*, multiple nominally significant associations for SNPs and CpG
306 sites were detected that were most prominent within intronic and intergenic regions, while no SNP but
307 CUD-associated CpG sites were identified in the *INPP5E* gene locus (Fig. 4c). In line with this, when
308 we quantified the association of genetic variants with CUD at the gene level using a gene-based
309 association analysis in MAGMA, we detected stronger association for *ZBTB4* with CUD ($Z=1.74$; $p=0.04$)
310 compared to *INPP5E* ($Z=-1.14$; $p=0.87$).



311
 312 **Fig. 4 – Convergence of DNA methylation, alternative splicing, and gene expression alterations in CUD at the *ZBTB4* and**
 313 ***INPP5E* gene loci**
 314 **a** Overlap of differential expression (DE), differential DNA methylation (DiffMeth), and differential alternative splicing analyses
 315 suggest two genes, *ZBTB4* and *INPP5E*, where alterations are consistently associated with CUD. **b** Relationship between DE and
 316 DiffMeth genes in Brodmann Area 9 based on log2FC (y-axis) from DE analysis and effect size β from linear regression in the
 317 EWAS of CUD (x-axis). For both genes, *ZBTB4* and *INPP5E* (highlighted in red), hypomethylation of the strongest significant

318 CUD-associated CpG site and increased transcript levels are observed. c Integrated visualization of functional genomic datasets
319 for *ZBTB4* and *INPP5E* gene loci. CUD-associated genomic variants (SNPs $p < 0.05$ from³⁰), CUD-associated CpG sites ($p < 0.05$
320 from²²), RNA-seq data and intron clusters ($q < 0.05$) from the present study were visualized together with ENCODE ChIP-seq data
321 for different chromatin marks in human dorsolateral prefrontal cortex.

322

323 *Multi-Omics Factor Analysis confirms cell junction, synaptic signaling, and neurogenesis as important*
324 *biological processes in CUD*

325 The integration of DNA methylation and gene expression data as described above is based on results
326 of the EWAS, DE, and AS analyses which is one possible way of integrating multiple omics datasets. In
327 addition, multi-omics analysis tools such as MOFA enable an integrated analysis of omics datasets in a
328 single statistical framework. Using MOFA on our DNA methylation and gene expression data from BA9,
329 we identified one factor representation of the multi-omics dataset (factor 9) that was significantly
330 correlated with CUD ($r = -0.48$, $p = 0.02$) and age ($r = 0.47$, $p = 0.02$, Supplementary Fig. S5a+S5b). Factor
331 9 displayed significantly smaller factor values in CUD cases compared to individuals without CUD in a
332 Wilcoxon test ($p = 0.02$, Supplementary Fig. S5c). When we extracted the CpG sites with the strongest
333 weights on factor 9, cg23859635 annotated to *MTA3* was the CpG site with the strongest positive weight
334 on factor 9 ($w = 0.31$), while cg24621354 in the gene *TES* displayed the strongest negative weight ($w = -$
335 0.33 , Supplementary Fig. S5d, Supplementary Table S13). In the gene expression dataset, the small
336 GTPase *RAB6A* had the strongest positive weight ($w = 0.07$), while *HIVEP2* had the strongest negative
337 weight ($w = -0.05$) on factor 9 (Supplementary Fig. S5e, Table S13). Results of a GSEA on negative
338 expression weights on factor 9 revealed FDR-significant ($q < 0.05$) enrichment for synaptic signaling, cell
339 junction organization, and neurogenesis pathways, confirming the results from the previous analyses.
340 In contrast, GSEA on positive expression weights revealed enrichment for cellular respiration and small
341 molecule metabolic processes (Supplementary Fig. S5f, Supplementary Table S14). When we used
342 missMethyl to investigate the biological pathways for DNA methylation features with strong weights on
343 factor 9, we detected enrichment for similar biological pathways as in the analysis of expression
344 features. While none of the enrichment results remained FDR-significant after multiple testing correction,
345 strongest enrichment for CpG sites with negative weights on factor 9 was detected for intracellular
346 calcium concentration regulation and synaptic vesicle related processes (Supplementary Fig. S5g,
347 Supplementary Table S15). The pathways showing the strongest enrichment for the positive weight
348 CpG sites were related to monocarboxylic acid and specifically, lactate transmembrane transporter
349 activity, and ER stress pathways (Supplementary Fig. S5h, Supplementary Table S15).

350 *Integrative functional analysis reveals functional modules related to neurotransmission, cell*
351 *differentiation, cell junction organization, and fatty acid metabolism*

352 In an integrative functional analysis approach, we used all available information from our study on DNA
353 methylation, gene expression, and alternative splicing alterations in CUD to identify potential
354 convergence of association results at the pathway level in BA9. We thus performed a GO enrichment
355 analysis based on 10 curated lists with CUD-associated genes derived from EWAS, DE analysis,
356 alternative splicing analysis, WGCNA modules based on DNA methylation and expression data, and
357 MOFA (Supplementary Table S16). In the enrichment map for GO terms, we identified several functional
358 modules where the same biological pathway was detected for multiple gene lists at FDR-adjusted
359 statistical significance ($q < 0.05$) indicating convergence of the results from different analysis approaches
360 (Supplementary Fig. S6). The two largest functional modules (FM) contained pathways involved in
361 neurotransmission and synaptic signaling (FM1), while FM2 was enriched for neuron and glial cell
362 differentiation, growth, and morphogenesis processes. Two further prominent modules were related to
363 synapse and cell junction organization (FM3) and fatty acid metabolism (FM4).

364

365 **Discussion**

366 By applying a multi-omics data integration approach on DNA methylation and gene expression data
367 from postmortem human brain tissue we aimed for a deeper understanding of the neurobiology of CUD
368 in the human prefrontal cortex. At the gene level, our differential expression analysis suggests two
369 candidates, *FKBP4* and *HSPA6*, which were replicated as nominally significant findings in two
370 independent cohorts. In addition, our multi-omics analyses highlight *ZBTB4* and *INPP5E*, that were
371 consistently altered across omics analyses in BA9 and displayed consistent upregulation patterns in
372 independent replication datasets. At the pathway level, we found converging evidence for CUD-
373 associated DNAm and transcriptional alterations that were related to neurotransmission, fatty-acid
374 metabolism, and changes in neuronal morphology.

375

376 Analysis of the transcriptome in BA9 revealed *ZFAND2A* as the DEG showing the strongest association
377 with CUD. *ZFAND2A* is a canonical heat shock gene in humans encoding a zinc-finger containing
378 protein that is involved in the regulation of proteasomal protein degradation^{31,32}. It was further identified
379 as a DEG in a study on transcriptomic signatures of Alzheimer's disease (AD)³³. Another AD-related
380 finding emerged in co-expression network analysis. *APOE* showed the strongest connectivity in the PPI

381 network for module yellow genes and has been intensively characterized due to its association with age
382 of onset in AD³⁴. While SUDs and neurodegenerative disorders such as AD depict different
383 neuropsychiatric disorders based on the current understanding of disease mechanisms, CUD and AD
384 share brain atrophic changes as a clinical symptom³⁵ and our results suggest that there might be shared
385 molecular mechanisms involved.

386
387 Previous studies have identified differential alternative splicing in alcohol use disorder²⁵ and opioid use
388 disorder (OUD)²⁶ in the human brain, however, RNA splicing alterations have not been characterized in
389 human CUD so far. In the differential alternative splicing analysis, we found N=98 statistically significant
390 genes containing AS intron clusters. Interestingly, among our top findings, we found *Bridging Integrator*
391 *1 (BIN1)* for which differential alternative splicing in the brain has been described in OUD. *BIN1* was the
392 only differential AS gene in OUD that was conserved across all investigated brain regions; dlPFC, NAc,
393 and midbrain²⁶. Further, AS events in *Bin1* were identified in the mouse brain in a study on splicing
394 alterations associated with cocaine self-administration²⁷. As dendrite and axon morphogenesis
395 processes were among the enrichment results for AS genes in BA9, we hypothesize that AS is directly
396 related to neuroplastic changes in the CUD brain. Mechanistically, AS processes change the abundance
397 of transcript isoforms with different biological functions that might contribute to the neuroadaptations in
398 CUD. We explored the mechanism of spliceosomal gene expression alterations as a potential
399 contributor to differential AS events in CUD. Exposure to cocaine was previously hypothesized to alter
400 spliceosomal gene expression²⁹ and our results suggest spliceosomal genes such as *HSPA6* and
401 *HSP1A1* as DEGs in BA9. As spliceosomal gene alterations were also detected in the replication
402 analysis with *HSPA6* as a shared upregulated DE gene across studies, AS events might be an important
403 mechanism in CUD contributing to neurobiological changes in the PFC.

404
405 In the last step of the RNA-sequencing analysis, we aimed to address the urgent need for novel
406 pharmacotherapeutic approaches for the treatment of CUD by performing a drug repositioning analysis.
407 We detected glucocorticoid receptor-targeting drugs having consistently negative connections with the
408 CUD expression profile in BA9. In addition, *FKBP4*, an important regulator of glucocorticoid receptor
409 signaling was identified as a conserved upregulated DEG in CUD based on three independent dlPFC
410 datasets. *FKBP4* has a key role in the nuclear translocation of the glucocorticoid receptor, as it replaces
411 *FKBP5* upon cortisol binding to the receptor complex leading to its nuclear translocation³⁶.

412 Pharmacological targeting of glucocorticoid receptor signaling was tested in rodent models of cocaine
413 addiction³⁷⁻³⁹. Reduced behavioral response to cocaine was observed when glucocorticoid receptor
414 antagonists such as mifepristone were applied³⁷. In contrast, corticosterone was shown to promote
415 cocaine intake in rats^{38,39}. In the drug repositioning analysis, results for glucocorticoid receptor agonists
416 were more prominent compared to antagonists which appears to be in conflict with previous literature.
417 However, glucocorticoid receptor antagonists such as mifepristone also displayed significant negative
418 connectivity scores with the BA9 expression signature supporting previous findings. Further, synthetic
419 glucocorticoid receptor agonists such as dexamethasone were shown to impair cocaine self-
420 administration in rats³⁸ indicating a more complex relationship between the endogenous glucocorticoid
421 system and exogenously applied glucocorticoid receptor targeting drugs. We thus suggest that
422 glucocorticoid receptor targeting drugs should be further investigated for their potential use as a
423 pharmacotherapy in CUD.

424
425 Using multi-omics data integration, we identified two genes, *ZBTB4* and *INPP5E*, for which CUD-
426 associated alterations were consistently detected across DNAm, gene expression, and alternative
427 splicing analyses. Both genes contained a hypomethylated CpG site, stronger transcript expression was
428 found in individuals with CUD, and significant differentially spliced intron clusters were identified. Despite
429 being strongly expressed in the brain and most prominently in neurons⁴⁰, the role of *ZBTB4* in
430 neuropsychiatric disorders remains poorly understood. However, due to the DNA binding capacity and
431 its role as a transcriptional repressor, *ZBTB4* deregulation in CUD could lead to downstream expression
432 changes of its target genes. Further, protein-protein interaction data suggests interaction of *ZBTB4* with
433 the transcription factor PRDM5 as well as with the AP2M1 and AP2A1 subunits of the adapter protein 2
434 (AP-2) complex that is involved in endocytosis of neurotransmitter receptors in neurons^{41,42}. The second
435 finding at the gene level, *INPP5E*, encodes a phosphatidylinositol-phosphatase specific to cilia and
436 *INPP5E* mutations were found in Joubert syndrome which is characterized by cerebellar and cerebral
437 malformation⁴³. A possible link to CUD provide neuronal primary cilia, known as key signaling hubs on
438 somata enriched for G-protein-coupled receptors (GPCRs)⁴⁴. As *INPP5E* is required for proper
439 trafficking of GPCRs along ciliary microtubules⁴⁵, deregulation of *INPP5E* might lead to aberrant ciliary
440 signaling that has recently gained attention in the addiction field: cell type-specific ablation of neuronal
441 primary cilia in mice was shown to affect body weight as well as locomotor response to psychostimulants

442 such as cocaine⁴⁶ and amphetamine⁴⁷. In humans, further studies on INPP5E are required to evaluate
443 its role in SUDs.

444

445 Evaluating the convergence of results at the pathway level revealed widespread molecular alterations
446 in synaptic signaling represented by functional module FM1 in the GO enrichment analysis. Our findings
447 are well in line with previous literature that reported on cocaine-associated DNAm and expression
448 changes in genes involved in neurotransmission^{14-16,18,48}. The observed overrepresentation of neuronal
449 marker genes in the upregulated DEGs together with the non-neuronal marker gene enrichment in the
450 downregulated DEGs further suggests a particular importance of CUD-associated expression changes
451 in altering neurotransmission. In a study on CUD-associated gene expression changes in neuronal
452 nuclei of the human dlPFC²¹, the authors found a WGCNA co-expression module that was significantly
453 associated with CUD and was enriched for GTPase signaling and neurotransmitter transport that well
454 matches our results in BA9. Neuronal function thus appears to be strongly influenced by altered
455 epigenetic and transcriptional programs in the CUD brain.

456

457 Functional modules FM2 and FM3 were related to pathways involved in neuron, synapse, and axon
458 morphogenetic processes. This is supported by literature from animal models of CUD, where alterations
459 in dendritic branches and spine density were observed in the PFC of cocaine self-administering rats⁴⁹.
460 Even a single cocaine exposure was sufficient to reduce dendritic spine density in neurons⁵⁰. In
461 summary, brain morphological changes depict an interesting link between molecular and behavioral
462 aspects of addiction as neuroplastic changes are the basis of neurocircuit alterations in the SUD brain
463 that are related to compulsive drug-seeking and relapse⁵¹.

464

465 Another converging finding were metabolic changes related to fatty acid metabolism (FM4). This finding
466 was especially prominent in the CUD-associated WGCNA module yellow where we found a functional
467 module of pathways related to fatty acid metabolism. Further, results from MOFA suggested gene
468 expression changes related to the electron transport chain as another key metabolic pathway alteration.
469 This is supported by findings from animal models of cocaine addiction where a downregulation of
470 glycolysis and oxidative phosphorylation were observed in the brain⁵² while fatty acid metabolism genes
471 were upregulated⁵³. It has to be noted that metabolic changes in CUD are most likely not brain-specific
472 but also appear on a systemic level as individuals with CUD were found to have reduced body fat in

473 comparison to a healthy control group⁵⁴. To follow up on this finding, future studies should evaluate if
474 interfering with fatty acid metabolism could depict a therapeutic strategy in CUD as a ketogenic diet has
475 been shown to alter the behavioral response to cocaine in rats⁵⁵.

476
477 There are some limitations that apply to our multi-omics study of CUD. First, depicting an inherent
478 limitation of analyses in human postmortem brain tissue, our cross-sectional analysis design can only
479 reflect the endpoint of CUD limiting the identification of dynamic changes in DNAm and gene expression
480 during the disease course. Second, considering the sample size and the few DEGs at transcriptome-
481 wide significance, it remains unclear whether the findings are generalizable to the general population
482 highlighting the need for studies in larger and more diverse cohorts. Third, while the homogeneity of our
483 sample consisting of only males from EA ancestry is a strength in statistical analysis, sex-specific and
484 ancestry-related molecular signatures of CUD remain an open question. At least in the analysis of the
485 more diverse replication cohorts we were able to show comparable CUD-associated gene expression
486 patterns when compared to our discovery cohort.

487
488 In summary, our study identifies novel associations with CUD at the gene level, confirms these on the
489 multi-omics level, and suggests differential alternative splicing as an important molecular hallmark of CUD
490 in the human prefrontal cortex. At the same time, our study supports previous findings of synaptic
491 signaling alterations that have been robustly detected when investigating the neurobiological effects of
492 cocaine. We highlight drugs targeting glucocorticoid receptor signaling to be further tested as a
493 treatment for CUD.

494 **Author contributions**

495 Conceptualization, E.Z., L.Z., M.R., R.S., and S.H.W.; Methodology, E.Z., H.B., J.F., and L.Z.;

496 Resources, G.T., N.M., A.C.H., J.C-D., N.F-C., B.C., and J.M.O.; Data curation, E.Z., D.A., H.B., D.A-

497 B., J.J.M-M., and L.Z. ; Data Analysis, E.Z., H.B., D.A., D.A-B., L.Z., and J.M.M-M., Investigation, E.Z.,

498 H.B., A.C.H., M.R., R.S., S.H.W, and L.Z.; Writing – Original Draft, E.Z., H.B.; Writing – Reviewing &

499 Editing, E.Z., H.B., D.A., D.A-B, J.M-M., J.F., N.M., G.T., J.C-D., N.F-C., B.C., J.L.M-M., A.C.H., M.R.,

500 R.S., S.H.W., and L.Z.; Supervision, J.F., M.R., and R.S., J.M.O., L.Z., and S.H.W; Project

501 Administration, R.S., A.C.H., M.R. and S.H.W., Funding Acquisition, R.S., A.C.H., M.R., and S.H.W.

502 **Competing interests**

503 The authors declare that there are no competing interests.

504

505 **Data availability statement**

506 Raw methylation data is deposited in the European Genome Phenome Archive (EGA) under accession

507 number EGAS00001006828 (<https://ega-archive.org/studies/EGAS00001006826>). RNA-sequencing

508 data has been uploaded to the EGA will be listed with an accession number by the time of publication.

509

510 **Code availability statement**

511 All original code used for data analysis and figure preparation will be available in a GitHub repository

512 (https://github.com/lzillich/BA9_multi_omics) by the time of publication.

513

514 **Ethics statement**

515 Postmortem human brain tissue was obtained from the Douglas Bell Canada Brain Bank where tissue

516 sampling is performed based on their established ethical guidelines. Our multi-omics study was

517 approved by the Ethics Committee II of the University of Heidelberg, Medical Faculty Mannheim,

518 Germany, under the register number 2021-681.

519

520

521

522 **Funding**

523 Funding supporting this study was provided by the German Federal Ministry of Education and Research
524 (BMBF) within the e:Med research program SysMedSUDs: “A systems-medicine approach toward
525 distinct and shared resilience and pathological mechanisms of substance use disorders” (01ZX01909
526 to R.S., M.R., A.C.H., and S.H.W.). In addition, by the Deutsche Forschungsgemeinschaft (DFG)
527 through the collaborative research centre TRR265: “Losing and Regaining Control over Drug Intake”⁵⁶
528 (Project ID 402170461 to S.H.W., R.S., A.C.H. and M.R.), the Hetzler Foundation for Addiction Research
529 (to A.C.H.), the ERA-NET program: Psi-Alc (01EZ1908), Spanish ‘Ministerio de Ciencia, Innovación y
530 Universidades’ (PID2021-1277760B-I100, to B.C.), ‘Generalitat de Catalunya/AGAUR’ (2021-SGR-
531 01093, to B.C.), ICREA Academia 2021, ‘Fundació La Marató de TV3’ (202218-31, to B.C.) and from
532 ‘Ministerio de Sanidad, Servicios Sociales e Igualdad/Plan Nacional Sobre Drogas’ (PNSD-2020I042,
533 to N.F.-C.). The project has been carried out using the Mannheim (CIMH) infrastructure of the German
534 Center for Mental Health (DZPG).
535

536 **Methods**

537 **Postmortem human brain tissue**

538 The sample of human postmortem brain tissue of BA9 was obtained from the Douglas Bell Canada
539 Brain Bank (DBCBB). Inclusion criteria were age > 18 and a diagnosis of cocaine dependence based
540 on DSM-IV. Throughout this study, we will nevertheless use the more recent terminology from DSM-5
541 i.e., cocaine use disorder. Individuals were excluded from the study if they were diagnosed with severe
542 neurodevelopmental or psychiatric disorders other than depressive disorders or had received additional
543 diagnoses of substance use disorders other than alcohol use disorder. All included subjects were male
544 and of European American descent. Demographic information for the cohort of N=42 BA9 tissue donors
545 is described in²² and for the subset of N=25 individuals with RNA-seq data in Table S1.

546

547 **DNA methylation data generation**

548 DNA extraction was performed as described in²². In brief, DNA was extracted from the full set of N=42
549 BA9 samples using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). The epigenome-wide
550 DNAm profile was determined using the Illumina MethylationEPIC BeadChip v1 (850k) (Illumina, San
551 Diego, CA, USA). During sample processing and analysis of DNAm levels, randomization was applied
552 based on CUD status and known comorbidities such as AUD and depressive disorders.

553

554 **Generation of gene expression data**

555 Using the miRNAeasy Mini extraction Kit (Qiagen, Hilden, Germany), total RNA was extracted from the
556 N=42 BA9 samples using ~5mg of frozen tissue from each individual. The RNA integrity number (RIN)
557 was measured using a TapeStation 4200 (Agilent, Santa Clara, CA, USA) resulting in a total of N=25
558 samples remaining for RNA sequencing (RIN > 5.5). Following ribosomal RNA (rRNA) depletion,
559 libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit (New England
560 Biolabs, Ipswich, MA, USA) followed by sequencing with an average of 60 million read pairs (2x100bp)
561 per sample. RNA sequencing was performed using an Illumina NovaSeq 6000 device.

562

563 **Statistical analyses**

564 All statistical analyses in the R programming environment were performed using R version 4.2.1. If not
565 otherwise stated, adjustment for multiple testing was performed using the Benjamini-Hochberg (FDR)

566 procedure⁵⁷. An analysis workflow for the multi-omics study of DNA methylation and gene expression in
567 CUD is shown in Supplementary Fig. S7.

568

569 **DNA methylation analysis**

570 Methylation data was analyzed as part of the Poisel, et al.²² study where a detailed description of the
571 analysis pipeline can be found in the methods section. In brief, DNA methylation levels were
572 preprocessed using an in-house quality control (QC) pipeline based on CPACOR⁵⁸. The neuronal cell
573 fraction was estimated based on the Houseman algorithm⁵⁹ using a dIPFC reference dataset⁶⁰. Quantile-
574 normalized beta values were derived from raw-intensities, followed by logit-transformation to M values
575 of methylation. An epigenome-wide association study (EWAS) was performed using a linear regression
576 model while adjusting for covariates that have a known effect on DNA methylation such as age,
577 postmortem interval (PMI), pH of the brain tissue, neuronal cell fraction, comorbid depressive and/or
578 alcohol use disorder, and technical factors. Downstream analyses based on the results of the EWAS
579 included the identification of differentially methylated regions (DMRs), a gene ontology enrichment
580 analysis using CUD-associated CpG sites ($p_{\text{assoc}} < 0.001$), and a network analysis in WGCNA to evaluate
581 CUD-associated co-methylation modules.

582

583 **Gene expression analysis**

584 Sequencing quality metrics were inspected using FastQC v.0.12.1 confirming all 25 fastq files to be used
585 in further analysis. Reads were mapped to the GRCh38 genome primary assembly using STAR
586 v.2.7.10b⁶¹. Quantification of features was performed using the featureCounts implementation in the R
587 package Rsubread v.2.12.3⁶² with the genome annotation gtf-file v.43 from GENCODE
588 (<https://www.genencodegenes.org>). The raw count matrix was imported in DESeq2 v.1.38.3⁶³ and
589 differential expression (DE) testing was performed while adjusting for the covariates age, PMI, brain pH
590 and RIN in the DESeq2 experimental design formula. The distribution of resulting p-values was
591 assessed in a quantile-quantile plot to evaluate potential genomic inflation (Fig. S1A). Fold-change cut-
592 offs for DEGs were an absolute log₂ fold change of larger than 0.07, corresponding to a 5% change in
593 transcript abundance. Statistical significance cut-offs were $p < 0.05$ for nominal significance and $q < 0.05$
594 for a 5% FDR-adjusted significant association with CUD. All covariates included in the DESeq2 model
595 are known to influence the gene expression profile and were confirmed in a variance partition analysis
596 in our dataset using the R package variancePartition v.1.28.7 (Supplementary Fig. S1a). As comorbid

597 MDD and AUD explained only minimal variance in the expression data and only 25 of the 42 samples
598 were available in the expression analysis, MDD and AUD were not included as covariates in the
599 statistical model. A sensitivity analysis was performed including MDD and AUD as covariates
600 (Supplementary Fig. S1b) confirming a strong correlation between the log₂ fold-changes of the
601 nominally significant results.

602

603 **Cell-type deconvolution analysis**

604 Based on reference signatures of gene expression derived from single-cell studies, the distribution of
605 cell types in bulk expression data can be inferred using cell-type deconvolution algorithms such as
606 CIBERSORT⁶⁴. We used a curated set of cell type-specific marker genes of the human prefrontal cortex
607 based on a study from Yu and He⁶⁵ where a gene was required to have a 10-fold stronger expression
608 in a specific cell type compared to all other cell types to be considered a marker gene. DESeq2-
609 normalized counts of the BA9 expression dataset were used and cell type deconvolution was performed
610 using the CIBERSORT R script v1.04. To test for significant differences in cell type distribution in
611 samples from individuals with and without CUD, we performed a Bayesian estimation of the difference
612 in means and evaluated the 95% high-density interval. The Bayesian testing was based on BEST⁶⁶ as
613 implemented in the R package BayesianFirstAid v.0.1. Further, an overlap analysis of DEGs in cell type
614 markers was performed in GeneOverlap v.1.34.0⁶⁷ using the 10-fold marker gene list from⁶⁵ in a Fisher
615 test.

616

617 **Functional enrichment analysis**

618 To characterize altered biological functions related to the observed gene expression differences, we
619 performed a gene set enrichment analysis (GSEA) for Gene Ontology (GO) terms using the gseGO
620 function from the R package clusterProfiler v.4.6.2⁶⁸. The DESeq2 Wald statistic defined as the log₂FC
621 divided by its standard error was used for ranking of the results. A significance threshold of $q < 0.05$ (5%
622 FDR) was considered statistically significant. Results of the GSEA were visualized using the emaplot
623 function in enrichplot v.1.18.3.

624

625 **Weighted gene co-expression network analysis (WGCNA)**

626 To identify CUD-associated co-expression patterns, we constructed co-expression modules using
627 network analysis in WGCNA (R package v.1.72.1)⁶⁹ and related them to CUD and other phenotypic

628 variables available in the DBCBB cohort. Using the input matrix of normalized and variance stabilization
629 transformed (vst) gene counts from DESeq2, a soft power threshold of 9 was estimated to achieve the
630 criterion of scale free topology ($R^2 > 0.85$). For the construction of networks, we used the parameters
631 $\text{minModuleSize}=10$, $\text{mergeCutHeight}=0.25$, and $\text{maxBlockSize}=36,000$. The Pearson correlation of the
632 module eigengene derived from each of the resulting $n=27$ co-expression modules with the phenotypes
633 of interest including CUD was calculated to identify significant associations of the modules with
634 phenotypes (Fig. S3A). Downstream analyses of modules significantly associated with CUD included a
635 GO enrichment analysis using the genes assigned to the modules using the full genome as the
636 background. Next, module genes were ranked by the product of gene significance*module membership
637 to identify hub genes. The top 10% of module hub genes were further investigated by constructing
638 protein-protein interaction (PPI) networks. For this, Cytoscape v.3.9.1⁷⁰ with stringApp v.1.7.0⁷¹ was
639 used. A detailed description of the PPI visualization settings in Cytoscape is found in²².

640

641 **Replication analysis of differential expression results**

642 Replication analysis of CUD-associated DEGs was performed in two independent datasets where RNA-
643 seq data from postmortem human brain tissue of the prefrontal cortex from individuals with and without
644 CUD was available. As the first replication dataset, BA9 bulk RNA-sequencing data from $N=7$ individuals
645 with CUD and $N=14$ control individuals originating from the National PTSD Brain Bank (NPBB)⁷² was
646 used. Phenotypic information for the BA9 replication cohort is shown in Table S17. RNA-seq data
647 sequenced and pre-processed as described in⁷³ was analyzed for CUD-associated differential gene
648 expression in DESeq2 using donor age, sex, PMI, and RIN as covariates. The second replication cohort
649 was based on a neuronal-specific RNA-sequencing dataset (GEO accession number: GSE99349) as
650 described in²¹. In this study, neuronal nuclei were isolated from postmortem human brain tissue of the
651 Brodmann Area 46 subregion of the dlPFC that is laterally adjacent to BA9. Here, bulk RNA-seq data
652 was generated from $N=19$ individuals with CUD and $N=17$ without CUD from a male mixed ancestry
653 cohort originating from the University of Miami Brain Bank (MBB). Raw sequencing data from the
654 replication cohort was downloaded from GEO and processed using the same analysis pipeline as in the
655 BA9 discovery sample: 1) mapping using STAR, 2) quantification using featureCounts, and 3) DE
656 analysis in DESeq2. For the replication analysis in MBB data, we used the same statistical model as in
657 the discovery analysis with differential expression testing for CUD while adjusting for donor age, RIN,
658 pH, and PMI. To explore the results, we first performed an overlap analysis of nominally significant CUD-

659 associated DEGs ($p < 0.05$) identified in the three datasets. Second, a targeted look-up of effect sizes
660 (\log_2FC) and association p -values was performed for overlapping DEGs across datasets and for the top
661 findings from the BA9 discovery sample, ZBTB4 and INPP5E. As an additional replication approach, we
662 performed rank-rank hypergeometric overlap (RRHO) using the R package RRHO2 v.1.0⁷⁴ to evaluate
663 convergent and divergent expression patterns at the transcriptome-wide scale between studies. RRHO
664 scores were generated based on full differential expression statistics from discovery and replication
665 datasets followed by the evaluation of overlapping signatures between studies using the hypergeometric
666 testing procedure as implemented in RRHO2.

667

668 **Signature-based drug repositioning analysis**

669 With the top 150 upregulated and downregulated genes ranked by the DESeq2 test statistic from the
670 differential expression analysis, the maximum input size in the Connectivity Map (CMap) query tool
671 (<https://clue.io/query>, software version 1.1.1.43) was used to evaluate the connectivity of expression
672 signatures (Table S8). CMap query uses the L1000 assay from the NIH LINCS project
673 (<https://lincsproject.org/>) as a drug-gene expression relationship database. In L1000, expression
674 changes for a representative set of 978 landmark transcripts are measured in response to treatment
675 with a perturbagen such as a pharmaceutical drug⁷⁵. In addition to the connectivity scores for individual
676 perturbagens, CMap also provides information on perturbagen classes and a GSEA output for pathways
677 and drug targets. Normalized connectivity scores and FDR-adjusted p -values for perturbagens and
678 GSEA results were obtained from the CMap query tool and visualized as waterfall plots in R using
679 ggplot2 v.3.4.2.

680

681 **Differential splicing analysis**

682 Alternative splicing was evaluated using the annotation-free quantification approach of RNA splicing in
683 LeafCutter v.0.2.9²⁸. First, raw sequencing data were aligned to the GRCh38 reference genome using
684 STAR with an adapted 2-pass mapping procedure. For this, the first mapping step was performed using
685 a regular gtf-file derived genome index. The resulting splice junctions (SJ_out.tab-files) from the N=25
686 samples were combined and filtered so that non-canonical junctions, junctions that were supported by
687 less or equal than 2 uniquely mapping reads, annotated junctions already covered by the gtf-file, and
688 duplicated junctions were removed. Using the filtered splice junction output, a modified genome index
689 was derived using STAR in genomeGenerate mode. This extended genome index containing

690 information on gene annotation and splice junctions was used in the second mapping step resulting in
691 the final bam-file output after mapping. Generation of junc-files, intron clustering, and differential intron
692 excision analysis was performed as outlined by the authors of leafCutter
693 (<https://davidaknowles.github.io/leafcutter/>) while including age, PMI, pH, and RIN as covariates into the
694 Dirichlet-Multinomial generalized linear model. Default settings were used in the leafcutter_ds.R script
695 i.e. maximum cluster size =10, minimum samples per intron = 5, minimum samples per group = 3, and
696 a minimum coverage of 20 reads. The differential intron excision analysis results in an estimate for the
697 change in the percent spliced in measure (Δ PSI) for each intron in a cluster and an FDR-adjusted p-
698 value for the cluster in which the differential splicing events were detected. Differential splicing events
699 in clusters with $|\Delta$ PSI| > 0.025 and an FDR-adjusted q-value<0.05 were considered statistically
700 significant²⁶. Visualizations for the differentially spliced clusters and genes were created using the leafviz
701 extension in leafCutter. GO enrichment analysis for genes containing differentially alternatively spliced
702 intron clusters was performed using the enrichGO function with GO “BP” ontology terms in
703 clusterProfiler.

704

705 **Integrative gene locus analysis**

706 Integrated visualization of functional genomics data was performed using SparK v.2.6.2⁷⁶. Summary
707 statistics from a meta-analysis GWAS of cocaine dependence (CD) in an EA population (N=6,378)³⁰
708 was used to cover SNPs that are associated with CD. The EWAS summary statistics from ²² were used
709 as the DNAm dataset. To prioritize the association results for visualization, SNPs and CpG sites with
710 nominal significant association p-value ($p < 0.05$) were filtered from the GWAS and EWAS results. ChIP-
711 seq datasets for different activating and repressing chromatin marks were downloaded from ENCODE⁷⁷
712 as deposited in the Human Reference Epigenome Matrix for dorsolateral PFC in males: ENCF241REN
713 (H3K4me1), ENCF752EVS (H3K4me3), ENCF866IWY (H3K27ac), ENCF149DDW (H3K36me3),
714 ENCF784SSN (H3K9me3), ENCF167ASN (H3K27me3). BigWig files were converted to BedGraph
715 using the UCSC bigWigToBedGraph tool. Bam-files from the RNA-seq analysis were indexed using
716 samtools v.1.5⁷⁸ and then converted to BedGraph using the bamCoverage function from deeptools
717 v.3.5.3⁷⁹. For the genetic dataset, we performed an additional gene-based association analysis using
718 Multi-marker Analysis of GenoMic Annotation (MAGMA)⁸⁰. Here, we aimed to quantify the combined
719 association of all SNPs annotated to a gene of interest with CUD as the phenotype.

720

721 **Multi-omics factor analysis**

722 Multi-omics factor analysis (MOFA)⁸¹ was used to jointly analyze the DNAm and gene expression
723 datasets in BA9 aiming for the identification of CUD-associated factors. The factor analysis framework
724 enables an improved characterization of gene and pathway alterations across different omics datasets
725 by investigating the contribution of each omics view such as DNAm or gene expression to a learned
726 factor. Downstream analyses such as GSEA enable the analysis of biological functions that are
727 associated with a factor based on the factor loading of features such as genes that contribute to the
728 biological pathway. As the DNAm input dataset for MOFA (R package v.1.3.1), we used methylation M-
729 values from the 20,000 most variant promoter CpG sites (TSS200 and TSS1500 annotations) under the
730 assumption of their prominent role in regulating transcription levels of nearby genes. Methylation data
731 was extracted for the individuals that also had expression data available (N=25). For the expression
732 dataset, we used normalized and variance stabilization transformed counts from the 20,000 most variant
733 genes to obtain an equal number of features in each view. The MOFA model was trained on the matched
734 DNAm and expression data from N=25 individuals using default model options with a total of 10 factors
735 and the training options `convergence_mode = "slow"`, `seed = 42`, and `maxiter=10,000`. Association of
736 factors with phenotypes was evaluated using the `correlate_factors_with_covariates` function. GSEA was
737 performed on negative and positive weights individually using the `run_enrichment` function based on the
738 `c5.go.bp.v2023.1.Hs.symbols.gmt` gene set reference file from MSigDB⁸². Functional characterization
739 of DNAm weights was performed by subsetting the top 2.5% of CpG sites from both sides of the weight
740 distribution on factor 9 resulting in N=500 CpG sites with strongest positive and negative weights on
741 factor 9, respectively. Next, GO enrichment analysis was performed in `missMethyl v.1.33.1` using the
742 full set of N=20,000 CpG sites as background.

743

744 **GO enrichment analysis of CUD-associated gene sets**

745 Convergence of CUD association signals at the pathway level was evaluated by pathway enrichment
746 analysis for GO terms using the `enrichGO` function on the GO "BP" ontology in the `compareCluster`
747 functionality of `clusterProfiler`. A total of 10 gene lists were included in the input dataset: 1) CUD-
748 associated CpG sites (N=394, $p>0.001$) from the EWAS of CUD²², genes in the CUD-associated
749 WGCNA methylation modules 2) blue (N=9,201), 3) steelblue (N=390), 4) brown (N=5,268), 5) brown4
750 (N=205), 6) nominally significant DEGs (N=1,057, $p<0.05$), 7) genes in the CUD-associated WGCNA
751 expression module yellow (N=2,517), 8) AS genes (N=98, $q<0.05$), 9) MOFA methylation weights factor

752 9 (N=983 genes based on the 2.5 and 97.5 percentiles of the weight distribution for CpG sites), and 10)
753 MOFA methylation weights factor 9 (N=1,000 genes based on the 2.5 and 97.5 percentiles of the weight
754 distribution for genes). Pathways remaining statistically significant after FDR correction ($q < 0.05$) were
755 displayed in an enrichment map with a pie plot visualization scheme for GO terms that were repeatedly
756 identified for the different gene lists.

757

758 References

- 759 1 Degenhardt, L. *et al.* The global burden of disease attributable to alcohol and drug use in 195
760 countries and territories, 1990–2016: a systematic analysis for the Global Burden of
761 Disease Study 2016. *The Lancet Psychiatry* **5**, 987-1012 (2018).
762 [https://doi.org:10.1016/S2215-0366\(18\)30337-7](https://doi.org/10.1016/S2215-0366(18)30337-7)
- 763 2 American Psychiatric Association & American Psychiatric Association. DSM-5 Task Force.
764 (2013). Diagnostic and statistical manual of mental disorders : DSM-5. Arlington, V. A. P. A.
- 765 3 Kampman, K. M. The treatment of cocaine use disorder. *Sci Adv* **5**, eaax1532 (2019).
766 [https://doi.org:10.1126/sciadv.aax1532](https://doi.org/10.1126/sciadv.aax1532)
- 767 4 Nestler, E. J. The neurobiology of cocaine addiction. *Sci Pract Perspect* **3**, 4-10 (2005).
768 [https://doi.org:10.1151/spp05314](https://doi.org/10.1151/spp05314)
- 769 5 Ersche, K. D. *et al.* Abnormal structure of frontostriatal brain systems is associated with
770 aspects of impulsivity and compulsivity in cocaine dependence. *Brain* **134**, 2013-2024 (2011).
771 [https://doi.org:10.1093/brain/awr138](https://doi.org/10.1093/brain/awr138)
- 772 6 Goldstein, R. Z. & Volkow, N. D. Drug addiction and its underlying neurobiological basis:
773 neuroimaging evidence for the involvement of the frontal cortex. *Am J Psychiatry* **159**, 1642-
774 1652 (2002). [https://doi.org:10.1176/appi.ajp.159.10.1642](https://doi.org/10.1176/appi.ajp.159.10.1642)
- 775 7 Cox, J. & Witten, I. B. Striatal circuits for reward learning and decision-making. *Nat Rev*
776 *Neurosci* **20**, 482-494 (2019). [https://doi.org:10.1038/s41583-019-0189-2](https://doi.org/10.1038/s41583-019-0189-2)
- 777 8 Koob, G. F. & Volkow, N. D. Neurobiology of addiction: a neurocircuitry analysis. *Lancet*
778 *Psychiatry* **3**, 760-773 (2016). [https://doi.org:10.1016/s2215-0366\(16\)00104-8](https://doi.org/10.1016/s2215-0366(16)00104-8)
- 779 9 Robison, A. J. & Nestler, E. J. Transcriptional and epigenetic mechanisms of addiction. *Nature*
780 *Reviews Neuroscience* **12**, 623-637 (2011). [https://doi.org:10.1038/nrn3111](https://doi.org/10.1038/nrn3111)
- 781 10 Fernández-Castillo, N., Cabana-Domínguez, J., Corominas, R. & Cormand, B. Molecular
782 genetics of cocaine use disorders in humans. *Molecular Psychiatry* **27**, 624-639 (2022).
783 [https://doi.org:10.1038/s41380-021-01256-1](https://doi.org/10.1038/s41380-021-01256-1)
- 784 11 Baker-Andresen, D. *et al.* Persistent variations in neuronal DNA methylation following cocaine
785 self-administration and protracted abstinence in mice. *Neuroepigenetics* **4**, 1-11 (2015).
786 [https://doi.org:10.1016/j.nepig.2015.10.001](https://doi.org/10.1016/j.nepig.2015.10.001)
- 787 12 Anier, K., Malinovskaja, K., Aonurm-Helm, A., Zharkovsky, A. & Kalda, A. DNA methylation
788 regulates cocaine-induced behavioral sensitization in mice. *Neuropsychopharmacology* **35**,
789 2450-2461 (2010). [https://doi.org:10.1038/npp.2010.128](https://doi.org/10.1038/npp.2010.128)
- 790 13 Gibney, E. R. & Nolan, C. M. Epigenetics and gene expression. *Heredity* **105**, 4-13 (2010).
791 [https://doi.org:10.1038/hdy.2010.54](https://doi.org/10.1038/hdy.2010.54)
- 792 14 Campbell, R. R. *et al.* Cocaine induces paradigm-specific changes to the transcriptome within
793 the ventral tegmental area. *Neuropsychopharmacology* **46**, 1768-1779 (2021).
794 [https://doi.org:10.1038/s41386-021-01031-4](https://doi.org/10.1038/s41386-021-01031-4)
- 795 15 Savell, K. E. *et al.* A dopamine-induced gene expression signature regulates neuronal function
796 and cocaine response. *Science Advances* **6**, eaba4221
797 [https://doi.org:10.1126/sciadv.aba4221](https://doi.org/10.1126/sciadv.aba4221)
- 798 16 Li, M. *et al.* Dynamic Expression Changes in the Transcriptome of the Prefrontal Cortex after
799 Repeated Exposure to Cocaine in Mice. *Front Pharmacol* **8**, 142 (2017).
800 [https://doi.org:10.3389/fphar.2017.00142](https://doi.org/10.3389/fphar.2017.00142)
- 801 17 Teague, C. D. & Nestler, E. J. Key transcription factors mediating cocaine-induced plasticity in
802 the nucleus accumbens. *Molecular Psychiatry* **27**, 687-709 (2022).
803 [https://doi.org:10.1038/s41380-021-01163-5](https://doi.org/10.1038/s41380-021-01163-5)
- 804 18 Vaillancourt, K. *et al.* Methylation of the tyrosine hydroxylase gene is dysregulated by cocaine
805 dependence in the human striatum. *iScience* **24**, 103169 (2021).
806 [https://doi.org:10.1016/j.isci.2021.103169](https://doi.org/10.1016/j.isci.2021.103169)

- 807 19 Vaillancourt, K. *et al.* Cocaine-related DNA methylation in caudate neurons alters 3D
808 chromatin structure of the IRXA gene cluster. *Mol Psychiatry* **26**, 3134-3151 (2021).
809 <https://doi.org/10.1038/s41380-020-00909-x>
- 810 20 Mews, P. *et al.* Convergent abnormalities in striatal gene networks in human cocaine use
811 disorder and mouse cocaine administration models. *Science Advances* **9**, eadd8946 (2023).
812 <https://doi.org/doi:10.1126/sciadv.add8946>
- 813 21 Ribeiro, E. A. *et al.* Gene Network Dysregulation in Dorsolateral Prefrontal Cortex Neurons of
814 Humans with Cocaine Use Disorder. *Sci Rep* **7**, 5412 (2017). [https://doi.org/10.1038/s41598-](https://doi.org/10.1038/s41598-017-05720-3)
815 [017-05720-3](https://doi.org/10.1038/s41598-017-05720-3)
- 816 22 Poisel, E. *et al.* DNA methylation in cocaine use disorder—An epigenome-wide approach in the
817 human prefrontal cortex. *Frontiers in Psychiatry* **14** (2023).
818 <https://doi.org/10.3389/fpsy.2023.1075250>
- 819 23 Van Booven, D. *et al.* Alcohol use disorder causes global changes in splicing in the human
820 brain. *Transl Psychiatry* **11**, 2 (2021). <https://doi.org/10.1038/s41398-020-01163-z>
- 821 24 Li, R. *et al.* RNA alternative splicing impacts the risk for alcohol use disorder. *Molecular*
822 *Psychiatry* (2023). <https://doi.org/10.1038/s41380-023-02111-1>
- 823 25 Huggett, S. B., Ikeda, A. S., Yuan, Q., Benca-Bachman, C. E. & Palmer, R. H. C. Genome-
824 and transcriptome-wide splicing associations with alcohol use disorder. *Sci Rep* **13**, 3950
825 (2023). <https://doi.org/10.1038/s41598-023-30926-z>
- 826 26 Huggett, S. B., Ikeda, A. S., McGeary, J. E., Kaun, K. R. & Palmer, R. H. C. Opioid Use
827 Disorder and Alternative mRNA Splicing in Reward Circuitry. *Genes (Basel)* **13** (2022).
828 <https://doi.org/10.3390/genes13061045>
- 829 27 Xu, S.-J. *et al.* Chromatin-mediated alternative splicing regulates cocaine-reward behavior.
830 *Neuron* **109**, 2943-2966.e2948 (2021). <https://doi.org/10.1016/j.neuron.2021.08.008>
- 831 28 Li, Y. I. *et al.* Annotation-free quantification of RNA splicing using LeafCutter. *Nature Genetics*
832 **50**, 151-158 (2018). <https://doi.org/10.1038/s41588-017-0004-9>
- 833 29 Huggett, S. B. & Stallings, M. C. Cocaine'omics: Genome-wide and transcriptome-wide
834 analyses provide biological insight into cocaine use and dependence. *Addict Biol* **25**, e12719
835 (2020). <https://doi.org/10.1111/adb.12719>
- 836 30 Cabana-Domínguez, J., Shivalikanjli, A., Fernández-Castillo, N. & Cormand, B. Genome-wide
837 association meta-analysis of cocaine dependence: Shared genetics with comorbid conditions.
838 *Prog Neuropsychopharmacol Biol Psychiatry* **94**, 109667 (2019).
839 <https://doi.org/10.1016/j.pnpbp.2019.109667>
- 840 31 Rossi, A. *et al.* The Proteasome Inhibitor Bortezomib Is a Potent Inducer of Zinc Finger AN1-
841 type Domain 2a Gene Expression: ROLE OF HEAT SHOCK FACTOR 1 (HSF1)-HEAT
842 SHOCK FACTOR 2 (HSF2) HETEROCOMPLEXES*. *Journal of Biological Chemistry* **289**,
843 12705-12715 (2014). <https://doi.org/https://doi.org/10.1074/jbc.M113.513242>
- 844 32 Lee, D., Takayama, S. & Goldberg, A. L. ZFAND5/ZNF216 is an activator of the 26S
845 proteasome that stimulates overall protein degradation. *Proceedings of the National Academy*
846 *of Sciences* **115**, E9550-E9559 (2018). <https://doi.org/10.1073/pnas.1809934115>
- 847 33 Shippy, D. C. & Ulland, T. K. Exploring the zinc-related transcriptional landscape in
848 Alzheimer's disease. *IBRO Neurosci Rep* **13**, 31-37 (2022).
849 <https://doi.org/10.1016/j.ibneur.2022.06.002>
- 850 34 Yamazaki, Y., Zhao, N., Caulfield, T. R., Liu, C.-C. & Bu, G. Apolipoprotein E and Alzheimer
851 disease: pathobiology and targeting strategies. *Nature Reviews Neurology* **15**, 501-518
852 (2019). <https://doi.org/10.1038/s41582-019-0228-7>
- 853 35 Ersche, K. D., Jones, P. S., Williams, G. B., Robbins, T. W. & Bullmore, E. T. Cocaine
854 dependence: a fast-track for brain ageing? *Mol Psychiatry* **18**, 134-135 (2013).
855 <https://doi.org/10.1038/mp.2012.31>
- 856 36 Davies, T. H., Ning, Y. M. & Sánchez, E. R. A new first step in activation of steroid receptors:
857 hormone-induced switching of FKBP51 and FKBP52 immunophilins. *J Biol Chem* **277**, 4597-
858 4600 (2002). <https://doi.org/10.1074/jbc.C100531200>
- 859 37 Deroche-Gamonet, V. *et al.* The glucocorticoid receptor as a potential target to reduce cocaine
860 abuse. *J Neurosci* **23**, 4785-4790 (2003). <https://doi.org/10.1523/jneurosci.23-11-04785.2003>
- 861 38 John, R. M., David, S. & Nick, E. G. Corticosterone Facilitates the Acquisition of Cocaine Self-
862 Administration in Rats: Opposite Effects of the Type II Glucocorticoid Receptor Agonist
863 Dexamethasone. *Journal of Pharmacology and Experimental Therapeutics* **287**, 72 (1998).
- 864 39 Evan, N. G. *et al.* Corticosterone Acts in the Nucleus Accumbens to Enhance Dopamine
865 Signaling and Potentiate Reinstatement of Cocaine Seeking. *The Journal of Neuroscience* **33**,
866 11800 (2013). <https://doi.org/10.1523/JNEUROSCI.1969-13.2013>
- 867 40 Filion, G. J. *et al.* A family of human zinc finger proteins that bind methylated DNA and repress
868 transcription. *Mol Cell Biol* **26**, 169-181 (2006). <https://doi.org/10.1128/mcb.26.1.169-181.2006>

- 869 41 Guardia, C. M., De Pace, R., Mattera, R. & Bonifacino, J. S. Neuronal functions of adaptor
870 complexes involved in protein sorting. *Curr Opin Neurobiol* **51**, 103-110 (2018).
871 <https://doi.org/10.1016/j.conb.2018.02.021>
- 872 42 Szklarczyk, D. *et al.* The STRING database in 2023: protein-protein association networks and
873 functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res* **51**,
874 D638-d646 (2023). <https://doi.org/10.1093/nar/gkac1000>
- 875 43 Hardee, I. *et al.* Defective ciliogenesis in INPP5E-related Joubert syndrome. *Am J Med Genet*
876 *A* **173**, 3231-3237 (2017). <https://doi.org/10.1002/ajmg.a.38376>
- 877 44 DeMars, K. M., Ross, M. R., Starr, A. & McIntyre, J. C. Neuronal primary cilia integrate
878 peripheral signals with metabolic drives. *Front Physiol* **14**, 1150232 (2023).
879 <https://doi.org/10.3389/fphys.2023.1150232>
- 880 45 Garcia-Gonzalo, F. R. *et al.* Phosphoinositides Regulate Ciliary Protein Trafficking to Modulate
881 Hedgehog Signaling. *Dev Cell* **34**, 400-409 (2015).
882 <https://doi.org/10.1016/j.devcel.2015.08.001>
- 883 46 Everett, T. *et al.* Cilia loss on distinct neuron populations differentially alters cocaine-induced
884 locomotion and reward. *J Psychopharmacol* **38**, 200-212 (2024).
885 <https://doi.org/10.1177/02698811231219058>
- 886 47 Ramos, C. *et al.* Neuron-specific cilia loss differentially alters locomotor responses to
887 amphetamine in mice. *J Neurosci Res* **99**, 827-842 (2021). <https://doi.org/10.1002/jnr.24755>
- 888 48 Gao, P., Limpens, J. H. W., Spijker, S., Vanderschuren, L. J. M. J. & Voorn, P. Stable
889 immediate early gene expression patterns in medial prefrontal cortex and striatum after long-
890 term cocaine self-administration. *Addiction Biology* **22**, 354-368 (2017).
891 [https://doi.org:https://doi.org/10.1111/adb.12330](https://doi.org/https://doi.org/10.1111/adb.12330)
- 892 49 Robinson, T. E., Gorny, G., Mitton, E. & Kolb, B. Cocaine self-administration alters the
893 morphology of dendrites and dendritic spines in the nucleus accumbens and neocortex.
894 *Synapse* **39**, 257-266 (2001). [https://doi.org/10.1002/1098-2396\(20010301\)39:3<257::Aid-](https://doi.org/10.1002/1098-2396(20010301)39:3<257::Aid-syn1007>3.0.Co;2-1)
895 [syn1007>3.0.Co;2-1](https://doi.org/10.1002/1098-2396(20010301)39:3<257::Aid-syn1007>3.0.Co;2-1)
- 896 50 Caffino, L., Messa, G. & Fumagalli, F. A single cocaine administration alters dendritic spine
897 morphology and impairs glutamate receptor synaptic retention in the medial prefrontal cortex
898 of adolescent rats. *Neuropharmacology* **140**, 209-216 (2018).
899 [https://doi.org:https://doi.org/10.1016/j.neuropharm.2018.08.006](https://doi.org/https://doi.org/10.1016/j.neuropharm.2018.08.006)
- 900 51 Koob, G. F. & Volkow, N. D. Neurocircuitry of Addiction. *Neuropsychopharmacology* **35**, 217-
901 238 (2010). <https://doi.org/10.1038/npp.2009.110>
- 902 52 Zhou, Z., Yuan, Q., Mash, D. C. & Goldman, D. Substance-specific and shared transcription
903 and epigenetic changes in the human hippocampus chronically exposed to cocaine and
904 alcohol. *Proceedings of the National Academy of Sciences* **108**, 6626-6631 (2011).
905 <https://doi.org/10.1073/pnas.1018514108>
- 906 53 Pati, S., Angel, P., Drake, R. R., Wagner, J. J. & Cummings, B. S. Lipidomic changes in the
907 rat hippocampus following cocaine conditioning, extinction, and reinstatement of drug-seeking.
908 *Brain Behav* **9**, e01451 (2019). <https://doi.org/10.1002/brb3.1451>
- 909 54 Ersche, K. D., Stochl, J., Woodward, J. M. & Fletcher, P. C. The skinny on cocaine: insights
910 into eating behavior and body weight in cocaine-dependent men. *Appetite* **71**, 75-80 (2013).
911 <https://doi.org/10.1016/j.appet.2013.07.011>
- 912 55 Martinez, L. A., Lees, M. E., Ruskin, D. N. & Masino, S. A. A ketogenic diet diminishes
913 behavioral responses to cocaine in young adult male and female rats. *Neuropharmacology*
914 **149**, 27-34 (2019). <https://doi.org/10.1016/j.neuropharm.2019.02.001>
- 915 56 Heinz, A. *et al.* Addiction Research Consortium: Losing and regaining control over drug intake
916 (ReCoDe)-From trajectories to mechanisms and interventions. *Addict Biol* **25**, e12866 (2020).
917 <https://doi.org/10.1111/adb.12866>
- 918 57 Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful
919 Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B*
920 *(Methodological)* **57**, 289-300 (1995).
- 921 58 Lehne, B. *et al.* A coherent approach for analysis of the Illumina HumanMethylation450
922 BeadChip improves data quality and performance in epigenome-wide association studies.
923 *Genome Biol* **16**, 37 (2015). <https://doi.org/10.1186/s13059-015-0600-x>
- 924 59 Houseman, E. A. *et al.* DNA methylation arrays as surrogate measures of cell mixture
925 distribution. *BMC Bioinformatics* **13**, 86 (2012). <https://doi.org/10.1186/1471-2105-13-86>
- 926 60 Jaffe, A. E. & Kaminsky, Z. FlowSorted. *Blood*. 450k: *Illumina HumanMethylation data on*
927 *sorted blood cell populations*. *R package version 1* (2017).
- 928 61 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
929 <https://doi.org/10.1093/bioinformatics/bts635>

- 930 62 Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better
931 for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research* **47**, e47-
932 e47 (2019). <https://doi.org/10.1093/nar/gkz114>
- 933 63 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
934 RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014). <https://doi.org/10.1186/s13059-014-0550-8>
- 935 64 Newman, A. M. *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat*
936 *Methods* **12**, 453-457 (2015). <https://doi.org/10.1038/nmeth.3337>
- 937 65 Yu, Q. & He, Z. Comprehensive investigation of temporal and autism-associated cell type
938 composition-dependent and independent gene expression changes in human brains. *Scientific*
939 *Reports* **7**, 4121 (2017). <https://doi.org/10.1038/s41598-017-04356-7>
- 940 66 Kruschke, J. K. Bayesian estimation supersedes the t test. *J Exp Psychol Gen* **142**, 573-603
941 (2013). <https://doi.org/10.1037/a0029146>
- 942 67 Shen L, S. I. GeneOverlap: Test and visualize gene overlaps. R package version 1.36.0.
943 (2023).
- 944 68 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing
945 biological themes among gene clusters. *Omics* **16**, 284-287 (2012).
946 <https://doi.org/10.1089/omi.2011.0118>
- 947 69 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network
948 analysis. *BMC Bioinformatics* **9**, 559 (2008). <https://doi.org/10.1186/1471-2105-9-559>
- 949 70 Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular
950 interaction networks. *Genome Res* **13**, 2498-2504 (2003). <https://doi.org/10.1101/gr.1239303>
- 951 71 Doncheva, N. T., Morris, J. H., Gorodkin, J. & Jensen, L. J. Cytoscape StringApp: Network
952 Analysis and Visualization of Proteomics Data. *J Proteome Res* **18**, 623-632 (2019).
953 <https://doi.org/10.1021/acs.jproteome.8b00702>
- 954 72 Friedman, M. J. *et al.* VA's National PTSD Brain Bank: a National Resource for Research.
955 *Curr Psychiatry Rep* **19**, 73 (2017). <https://doi.org/10.1007/s11920-017-0822-6>
- 956 73 Girgenti, M. J. *et al.* Transcriptomic organization of the human brain in post-traumatic stress
957 disorder. *Nat Neurosci* **24**, 24-33 (2021). <https://doi.org/10.1038/s41593-020-00748-7>
- 958 74 Cahill, K. M., Huo, Z., Tseng, G. C., Logan, R. W. & Seney, M. L. Improved identification of
959 concordant and discordant gene expression signatures using an updated rank-rank
960 hypergeometric overlap approach. *Sci Rep* **8**, 9588 (2018). <https://doi.org/10.1038/s41598-018-27903-2>
- 961 75 Subramanian, A. *et al.* A Next Generation Connectivity Map: L1000 Platform and the First
962 1,000,000 Profiles. *Cell* **171**, 1437-1452.e1417 (2017).
963 <https://doi.org/10.1016/j.cell.2017.10.049>
- 964 76 Kurtenbach, S. & Harbour, J. W. SparK: A Publication-quality NGS Visualization Tool. *bioRxiv*,
965 845529 (2019). <https://doi.org/10.1101/845529>
- 966 77 An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74
967 (2012). <https://doi.org/10.1038/nature11247>
- 968 78 Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079
969 (2009). <https://doi.org/10.1093/bioinformatics/btp352>
- 970 79 Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data
971 analysis. *Nucleic Acids Research* **44**, W160-W165 (2016). <https://doi.org/10.1093/nar/gkw257>
- 972 80 de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set
973 analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).
974 <https://doi.org/10.1371/journal.pcbi.1004219>
- 975 81 Argelaguet, R. *et al.* Multi-Omics Factor Analysis—a framework for unsupervised integration of
976 multi-omics data sets. *Molecular Systems Biology* **14**, e8124 (2018).
977 [https://doi.org:https://doi.org/10.15252/msb.20178124](https://doi.org/https://doi.org/10.15252/msb.20178124)
- 978 82 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for
979 interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550
980 (2005). <https://doi.org/10.1073/pnas.0506580102>
- 981
982
983