Chromatin domain alterations linked to 3D genome organization in a large cohort of schizophrenia and bipolar disorder brains

Kiran Girdhar [1,2,3 ✉], Gabriel E. Hoffman [1,2,3], Jaroslav Bendl [1,2,3,4,5], Samir Rahman[1,2,3,4,5], Pengfei Dong [1,2,3,4,5], Will Liao [6], Mads E. Hauberg[1,2,4,5,7,8], Laura Sloofman[9], Leanne Brown[4], Olivia Devillers[4], Bibi S. Kassim[4], Jennifer R. Wiseman [4], Royce Park[4], Elizabeth Zharovsky[4], Rivky Jacobov[4], Elie Flatow[4], Alexey Kozlenkov[5], Thomas Gilgenast[10], Jessica S. Johnson[1,2,3], Lizette Couto[4], Mette A. Peters[11], Jennifer E. Phillips-Cremins [10], Chang-Gyu Hahn[12], Raquel E. Gur[13], Carol A. Tamminga [14], David A. Lewis [15], Vahram Haroutunian [4,5,16,17], PsychENCODE Consortium*, Stella Dracheva[4,5,17], Barbara K. Lipska[18], Stefano Marenco [18], Marija Kundakovic [4,5,16,19], John F. Fullard [1,2,3,4,5], Yan Jiang [4,5,16,20], Panos Roussos [1,2,3,4,5,17,21,22 ✉] and Schahram Akbarian [4,5,16,22 ✉]

Chromosomal organization, scaling from the 147-base pair (bp) nucleosome to megabase-ranging domains encompassing multiple transcriptional units, including heritability loci for psychiatric traits, remains largely unexplored in the human brain. In this study, we constructed promoter- and enhancer-enriched nucleosomal histone modification landscapes for adult prefrontal cortex from H3-lysine 27 acetylation and H3-lysine 4 trimethylation profiles, generated from 388 controls and 351 individuals diagnosed with schizophrenia (SCZ) or bipolar disorder (BD) ($n = 739$). We mapped thousands of *cis*-regulatory domains (CRDs), revealing fine-grained, 10^4–10^6-bp chromosomal organization, firmly integrated into Hi-C topologically associating domain stratification by open/repressive chromosomal environments and nuclear topography. Large clusters of hyper-acetylated CRDs were enriched for SCZ heritability, with prominent representation of regulatory sequences governing fetal development and glutamatergic neuron signaling. Therefore, SCZ and BD brains show coordinated dysregulation of risk-associated regulatory sequences assembled into kilobase- to megabase-scaling chromosomal domains.

Chromosomal organization scales from nucleosomes, or 1.47×10^2 bp of DNA wrapped around a histone octamer, to functional and structural domains extending across 10^3–10^7 bp, with highly interdependent regulation across scales. This includes transcription-associated nucleosomal histone modifications in fibroblasts and peripheral myeloid cells, including mono- and tri-methyl-H3K4 (H3K4me) and acetyl-H3K27 (H3K27ac), which are tightly linked to chromatin structures defined by local chromosomal conformations, including the megabase (Mb)-scaling

'self-folded' topologically associating domains (TADs) and other features of three-dimensional (3D) genome organization[1]. Whether such type of acetyl- and methyl-histone-defined higher-order chromatin exists in the human brain, including cell-type-specific regulation and disease-associated alterations, remains unexplored. To date, virtually all conventional brain epigenomic maps present transcriptional histone marks (including H3K4me and H3K27ac) as isolated 'peaks' confined to short nucleosomal arrays, typically covering an average of 3.6–3.8 kilobases (kb)[2,3] in the human brain,

with only a very small portion of peaks showing some degree of confluence by merging into super-enhancers important for cell-specific gene expression programs[4].

Interestingly, however, regulators of nucleosomal histone modifications, including H3K4me3 and H3K27ac, confer heritable risk for SCZ and related co-heritable traits, including BD, by genome-wide association and exome sequencing[5–7]; and furthermore, in the adult human frontal lobe, SCZ and BD risk loci are enriched for active neuronal promoters and enhancers and other regulatory elements tagged by open chromatin-associated histone marks[2,6,8,9]. Unfortunately, representative genome-scale histone modification studies in diseased brain disease are lacking. It is not known whether changes in acetylation and methylation landscapes affect the general population of individuals with SCZ and individuals with BD[10] and whether such type of alteration could reveal broader changes in chromosomal organization beyond the classical 'peak-by-peak'-based analysis for nucleosomal histone modifications.

In this study, we generated 739 ($n_{H3K4me3} = 230$ and $n_{H3K27ac} = 260$ from neurons and $n_{H3K27ac} = 249$ from bulk tissue) chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) libraries from prefrontal cortex (PFC) of adult SCZ, BD and control brains. Using population-scale correlational analysis and cell-type-specific chromosomal conformation mapping, we defined acetylation and methylation landscapes by the coordinated regulation of sequentially arranged histone peaks constrained by local chromosomal conformations and nuclear topographies. We report widespread disease-associated alterations affecting the neuronal H3K27ac acetylome but not the H3K4me3 methylome. On a genome-wide scale, hundreds of kilobase- to megabase-scale chromosomal domains are altered in disease, with converging alignments by genetic risk, cell type, developmental function, nuclear topography and active versus repressive chromosomal environments. Our findings, reproducible across two independent brain cohorts, identify higher-order chromatin alterations representative of the broader population of individuals with SCZ and individuals with BD and link cognitive disease to altered organization of neuronal genomes in the PFC.

Results

Acetyl-histone peaks show disease-specific dysregulation. We first generated 490 ChIP-seq genome-wide maps of H3K4me3 and H3K27ac from ~3–5×10⁵ neuronal NeuN⁺ nuclei/samples isolated from dorsolateral PFC via fluorescence-activated nuclear sorting (FANS), from 321 demographically matched SCZ and non-psychiatric control brains that are part of the CommonMind Consortium collection[11–13] (hereafter referred to as Study-1). We then generated an additional set of H3K27ac ChIP-seq libraries ($n = 249$) prepared from unsorted nuclei extracted from bulk dorsolateral PFC tissue of SCZ, BD and control brains, contributed by the Human Brain Collection Core (HBCC) at the National Institute of Mental Health (hereafter referred to as Study-2) (Fig. 1a, Table 1 and Supplementary Table 1a,b).

Samples were processed with our in-house version of the encode ChIP-seq pipeline[2]. H3K4me3 NeuN⁺ peaks had a narrower genomic coverage (3.1%; mean peak width ~1,425 bp) compared to H3K27ac (genomic coverage 12.8% in NeuN⁺ and 17.1% in bulk tissue; Fig. 1a; mean peak width ~3,364 bp NeuN⁺ and ~3,583 bp bulk tissue) (see quality control metrics in Supplementary Fig. 1a and consensus peak sets in Supplementary Table 2a–c), with more than 60% of H3K4me3 and more than 75% of H3K27ac peaks distributed among distal intergenic, exonic, intronic and untranslated region (UTR) elements while remaining annotate as promoters peaks (Fig. 1a). Notably, each dataset showed high concordance (Jaccard similarity coefficients ~0.7) to previously generated PFC NeuN⁺ and bulk tissue H3K4me3 and H3K27ac datasets of brains not included in the present study (Supplementary Fig. 1b).

After various technical factors related to tissue processing and sequencing were regressed out (Methods and Supplementary Fig. 1c), we obtained three sets of normalized histone peak activity matrices (64,254 peaks × 230 H3K4me3 NeuN⁺, 114,136 peaks × 260 H3K27ac NeuN⁺ from Study-1 and 143,092 peaks × 249 H3K27ac Tissue from Study-2; Supplementary Table 3a–c). Furthermore, for each Study-2 sample, cell type heterogeneity in bulk tissue was adjusted by estimating the proportion of oligodendrocytes (OLIGs) and glutamatergic (Glu) and GABAergic (GABA) neurons using cell-type-specific ChIP-seq data from an independent reference set[14].

We explored H3K4me3 NeuN⁺ (mostly promoter-associated) 'peak'-based epigenomic aberrations in SCZ Study-1. Surprisingly, none of the 64,254 peaks (Supplementary Table 3a) survived multiple testing corrections after differential (cases versus controls) analysis, indicating that this methylation mark is not consistently affected. Next, we evaluated H3K27ac (promoter- and enhancer-associated) peaks. Altogether, 11,471 of the 114,136 H3K27ac NeuN⁺ peaks were dysregulated (false discovery rate (FDR) 5%) in SCZ Study-1 (Supplementary Table 3b), and, similarly, 5,656 of the 143,092 H3K27ac Tissue peaks were significantly affected in SCZ Study-2 (Fig. 1b and Supplementary Table 3c), with 559 dysregulated peaks present in both studies (Supplementary Fig. 2a). However, there was a significant correlation between cases–controls effect sizes of SCZ Study-1 H3K27ac NeuN⁺ at FDR 5% and SCZ Study-2 H3K27ac Tissue peaks ($n = 9,951$ peaks, Spearman's $\rho = 0.36$, $P = 4.9 \times 10^{-295}$) (Methods and Fig. 1c). Having shown that histone acetylation changes in SCZ PFC are broadly reproducible across independent brain collections (Supplementary Table 1), we next combined the differential histone peak effects sizes and P values from Study-1 H3K27ac NeuN⁺ and Study-2 H3K27ac Tissue datasets (Methods), yielding a consensus set of 46,294 H3K27ac Meta NeuN⁺ peaks each with 90% overlap of peak extension across the two studies. Of these, 6,219 peaks were dysregulated in SCZ (FDR 5%) (Fig. 1d and Supplementary Table 3d).

We applied a similar differential analysis workflow to determine BD-specific epigenomic aberrations in Study-2 H3K27ac Tissue and identified 1,809 of 143,092 dysregulated peaks (FDR 5%) (Supplementary Table 3e), with 630 dysregulated peaks shared with H3K27ac Tissue SCZ and 158 shared with H3K27ac NeuN⁺ SCZ (Supplementary Fig. 2b). Furthermore, there was significant correlation ($n = 5,656$ peaks, Spearman's $\rho = 0.87$, $P < 0.05$) of BD versus control effect sizes with SCZ versus control effect sizes at FDR 5% in H3K27ac Tissue within the study as well as across studies ($n = 9,951$ peaks, Spearman's $\rho = 0.19$, $P = 1.9 \times 10^{-78}$) with H3K27ac NeuN⁺ (Supplementary Fig. 3a), suggesting shared epigenomic dysfunction in these two common types of psychiatric disorders.

Indeed, gene set enrichment analysis of dysregulated peaks was consistent for immune responses across our SCZ and BD cohorts, and neuronal (including neuron development) signaling and synaptic plasticity pathways ranked top among gene ontologies in SCZ-sensitive H3K27ac peaks (Supplementary Fig. 4). A representative example for H3K27ac peak-based alterations in our SCZ PFC datasets (Study-1, Study-2 and Meta) includes the 0.25-Mb-wide *STX1A* psychiatric susceptibility locus, encoding a regulator of synaptic vesicle docking[15] (Fig. 1e). Finally, we compared our H3K27ac NeuN⁺ (tissue) peaks in PFC showing more than 10% sequence overlap with PFC tissue open chromatin regions (OCRs)[16] generated from a cohort partially congruent with Study-1. Using π1 statistics, the proportion of true-positive SCZ-sensitive H3K27ac peaks that overlapped with OCRs ranged from 27.3% (tissue) to 30.6% (NeuN⁺).

Because most of our diseased brains were exposed to anti-psychotic drugs (APD) before death (Supplementary Table 1a,b), we assessed the potential effect of medication by studying the subset of $n = 116$ (36 = yes, 80 = no) Study-1 cases with documented
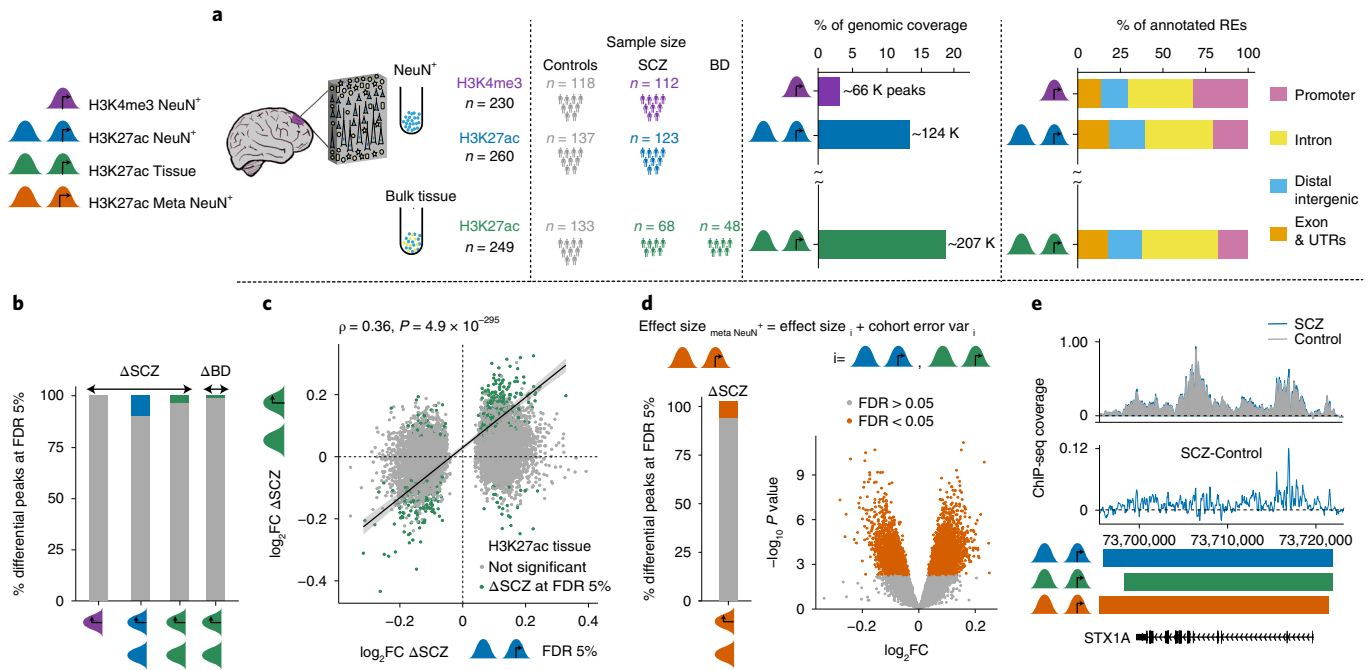
Fig. 1 | Histone peak profiling in 739 ChIP-seq datasets from two studies consisting of SCZ, BD and control subjects. a, Left: Datasets and studies: Study-1, FANS-isolated PFC NeuN+ nuclei: H3K4me3 (purple) and H3K27ac (blue); Study-2, total (non-sorted) tissue PFC nuclei: H3K27ac (green) and Meta-analysis PFC H3K27ac NeuN+ (orange). Middle: Bar plot; genomic coverage (%) of each ChIP-seq dataset. Numbers of subjects as indicated and bar plot to show % of regulatory elements in each dataset. **b**, Bar plots (colored sectors marking significantly different peaks) showing the proportion of differentially regulated histone peaks in our four datasets of case–control comparisons, as indicated. See also Supplementary Fig. 2 for overlap of disease-sensitive peaks across the various H3K27ac datasets. **c**, Spearman correlation of effect sizes of H3K27ac NeuN+ peaks altered (Δ) in SCZ Study-1 subjects compared to effect sizes for corresponding peaks in SCZ Study-2 subjects with bulk PFC tissue as input. Green dots mark altered peaks in H3K27ac tissue at FDR < 5%. P value (ρ), Spearman's rank correlation test. **d**, Meta-analysis of H3K27ac NeuN+ and H3K27ac Tissue. The bar shows differential SCZ-specific peaks of H3K27ac Meta NeuN+ at FDR < 5%. Volcano plot of differentially modified H3K27ac Meta NeuN+ peaks by fixed effect model. Orange dots represent peaks with FDR < 5%. P values for t-test of the limma pipeline, with FDR correction (Benjamini–Hochberg) across all peaks. **e**, Top: Visualization of the hyper-acetylated region covering the *STX1A* gene. H3K27ac NeuN+ ChIP-seq landscape from controls (gray) and SCZ subjects (blue). Bottom: Differential peak profile; horizontal bars demarcate linear extension of disease-sensitive peak population for each of the three H3K27ac datasets, as indicated: H3K27ac NeuN+ Study-1 (blue), H3K27ac Tissue Study-2 (green) and H3K27ac NeuN+ meta-analysis (orange). RE, regulatory element.

**Table 1 | Summary of metadata of ChIP-Seq samples in Study-1 and Study-2**

Study	Histone mark	PFC cell type	Brain bank	Diagnosis (n)	Sex (n)	Age (years)	Ethnicity (n)
Study-1	H3K4me3	NeuN+	MSSM PENN PITT	112 SCZ, 118 Ctrl	92 F, 138 M	66.33 ± 18.42	183 EU, 34 AA, 2 EA, 10 HISP, 1 Multiracial
	H3K27ac			123 SCZ, 137 Ctrl	98 F, 162 M	67.32 ± 18.02	203 EU, 38 AA, 2 EA, 16 HISP, 1 Multiracial
Study-2	H3K27ac	Tissue	HBCC	68 SCZ, 48 BD, 133 Ctrl	88 F, 161 M	42.27 ± 17.59	121 EU, 117 AA, 6 EA, 5 HISP

AA, African American; Ctrl, Control; EU, European; EA, East Asian; HISP; Hispanic; F, female; M, male.

exposure to typical (D2-like receptor antagonists) and $n = 117$ ($52 =$ yes, $65 =$ no) atypical/mixed receptor profile APD in the month before death. However, disease-associated H3K27ac changes showed almost null (atypical, Spearman's ρ = 0.0006, $P = 0.0031$) or negative (typical, Spearman's ρ = −0.27, $P < 2.2 \times 10^{-16}$) correlation with APD. Therefore, medication is not a driver (atypical APD) or even anti-correlated (typical APD) with H3K27ac alterations in diseased PFC NeuN+ (Supplementary Fig. 5).

Hyper-acetylated peaks are enriched for SCZ risk variants. To better understand these disease-associated aberrations in PFC H3K27ac peaks and their link with directionality in acetylation,

we stratified peak alterations into hyper-acetylation 'ΔSCZ↑' and hypo-acetylation 'ΔSCZ↓' based on log₂ fold change (FC) (cases vs. controls) >0 and <0, respectively, and computed their enrichment for genetic variants associated with SCZ and related psychiatric traits using stratified linkage disequilibrium (LD) score regression[17]. Interestingly, SCZ heritability coefficients were driven significantly by the group of hyper-acetylated, but not hypo-acetylated, peaks in all three of our SCZ case–control comparisons, an effect particularly striking for the H3K27ac Meta NeuN+ dataset (Fig. 2a). These changes were highly specific to psychiatric traits because non-psychiatric traits, such as height, or medical conditions, including autoimmune and cardiac disease, completely lacked association
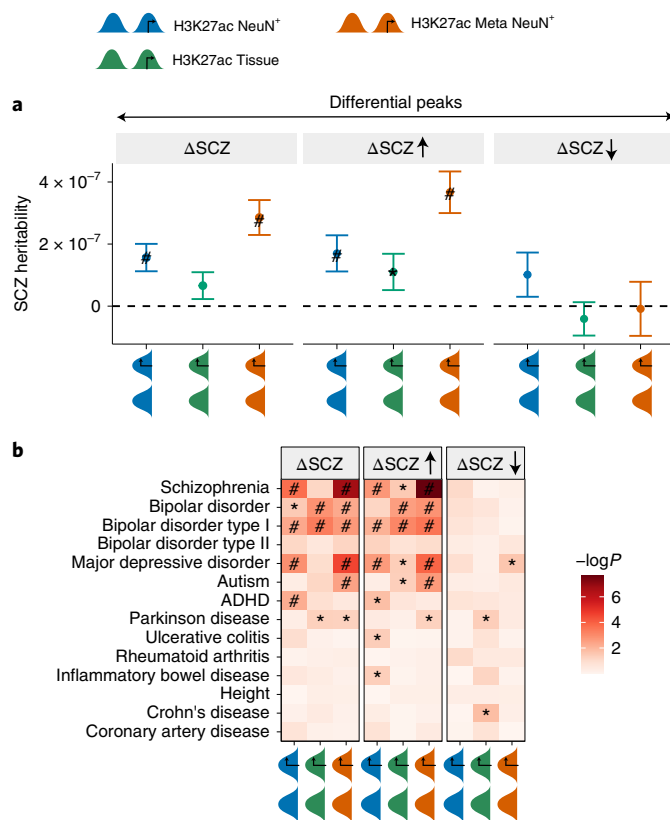
Fig. 2 | Enrichment of common SCZ risk variants in dysregulated peaks in NeuN⁺ and bulk tissue. a, SCZ heritability coefficients of genetic variants overlapping histone peaks from Study-1 H3K27ac NeuN⁺, Study-2 H3K27ac Tissue and H3K27ac Meta NeuN⁺ stratified by (1) 'ΔSCZ': dysregulated peaks ($n = 3,360$, 5,656 and 6,219 peaks, respectively); (2) 'ΔSCZ↑': hyper-acetylated dysregulated peaks ($n = 1,918$, 2,681 and 4,031 peaks, respectively); and (3) 'ΔSCZ↓': hypo-acetylated dysregulated peaks ($n = 1,442$, 2,975 and 2,188 peaks, respectively) with $\log_2 FC$ (SCZ versus controls) > 0 and < 0, respectively. Error bars represent standard error in SCZ heritability from LD score regression. **b**, Heat map of enrichment P values of brain-related GWAS traits. The overlap of peaks with genetic variants was assessed using LD score regression. '#': significant for enrichment in LD score regression after FDR correction of multiple testing across all tests in the plot (Benjamini–Hochberg test); '*': nominally significant for enrichment.

with our disease-associated PFC peaks (Fig. 2b and Supplementary Table 4).

To further assess enrichment of common variants of SCZ and other behavioral traits by the type of regulatory element, we stratified peak alterations into promoters ($< \pm 3$ kb from transcriptional start site (TSS)) and enhancers ($> \pm 3$ kb from TSS). The coefficient of SCZ heritability was of higher magnitude in enhancers than promoters, an effect specific to hyper-acetylated peaks (Supplementary Fig. 6a). Next, with genomic coverage of dysregulated BD peaks underpowered to run LD score regression, we instead annotated differentially acetylated BD Study-2 peaks to genes and checked for enrichment in SCZ and BD genome-wide association study (GWAS) variants. We found a significant association of hyper-acetylated 'ΔBD↑' peaks with SCZ (but not BD) genetics (Multimarker Analysis of GenoMic Annotation (MAGMA) $P < 0.05$) (Supplementary Fig. 6b). Notably, this effect was again significant for (SCZ) risk-associated enhancers in ΔBD and ΔBD↑, in contrast to dysregulated promoters in ΔBD, ΔBD↑ and ΔBD↓ peaks (Supplementary Fig. 6c). Therefore,

genetic risk for SCZ tracks genomic loci that are hyper-acetylated in diseased SCZ and BD PFC, an effect consistent across all our disease cohorts.

Histone peak correlations reveal chromosomal architecture.
After identifying alterations in the activity of PFC histone peaks, we investigated the effect of disease on the structural organization of PFC chromatin by characterizing the modular architecture of coordinated histone peaks in the brain epigenome. We hypothesized that the structure of coordinated histone peaks could be particularly important in disease context. This hypothesis is plausible, given recent reports from peripheral cells with coordinated regulation of multiple *cis*-regulatory elements sequentially organized along the linear genome[18]. Additionally, we observed that pairwise correlation between PFC histone peaks within chromosomal loopings in Hi-C NeuN⁺ from an independent set of PFCs ($n = 6$; three female and three male)[19,20] was substantially higher as compared to peaks of equivalent distance located outside of chromosomal loop contacts (Supplementary Fig. 8a), indicating the presence of correlation structure in histone peaks within Hi-C defined loops. Furthermore, studies on hundreds of lymphoblastoid and fibroblast cultures, leveraging 'population-scale' inter-individual correlations between histone peaks, successfully uncovered cis-regulatory domains (CRDs), with spatial clustering of CRD histone peaks ranging from 10^4 to 10^6 bp of linear genome and integrated into local chromosomal conformation landscapes[21,22]. Similar approaches have been applied to OCRs in Alzheimer's postmortem brains[19].

In this study, we developed a systematic workflow (Methods and Supplementary Fig. 7) by combining the previously developed software decorate[23] with additional steps of statistical analyses to identify CRDs on our population-scale H3K27ac and H3K4me3 datasets encompassing 739 PFC ChIP-seq libraries. The pipeline applied adjacency-constrained hierarchical clustering[23,24], across each of our three ChIP-seq datasets (H3K4me3 NeuN⁺, H3K27ac NeuN⁺ and H3K27ac Tissue), to identify sequentially aligned clusters of peaks as a strongly correlated structure (Methods and Supplementary Fig. 8b,c). Altogether, 39% (H3K4me3 NeuN⁺), 65% (H3K27ac NeuN⁺) and 68% (H3K27ac Tissue) of peaks assembled into 2,721, 6,389 and 8,239 CRDs, respectively (Fig. 3a and Supplementary Table 5a–c), with H3K27ac (H3K4me3) CRDs encompassing an average of ~11.7 (~9.3) histone peaks (Supplementary Fig. 8d).

Comparison of Study-1 and Study-2 H3K27ac CRDs showed higher similarity (Jaccard $J = 0.39$) as compared to Study-1 H3K4me3 CRDs and H3K27ac CRDs ($J = 0.22$) (Fig. 3b). Furthermore, 78–79% of H3K27ac CRD peaks were putative enhancers (that is, $> \pm 3$ kb from TSS) in contrast to ~61% of H3K4me3 CRD peaks (Supplementary Fig. 8e). Promoters comprised the remaining peak populations and, in H3K27ac (H3K4me3) CRDs, were linked with an average of ~4 (1.6) enhancer peaks.

Next, we wanted to explore the potential link between our CRD and higher-order chromosomal conformations, such as the TADs computed from Hi-C libraries from ensembles of PFC NeuN⁺ nuclei (Methods). Mb-scaling TADs, and the smaller sub-TADs hierarchically nested into them, are thought to represent a type of conformation defined by dynamic chromosomal loop extrusions of individual chromatin fibers, constrained by strong boundary elements at TAD peripheries and weaker in-TAD boundaries demarcating sub-TADs[25,26].

Indeed, visual examination of Hi-C maps and H3K27ac CRD structure reveals CRDs located within TADs. A representative example (Fig. 3c) shows the 2-Mb *GATB* (glutamyl-TRNA amidotransferase subunit B) locus linked to cognitive traits and educational attainment[27]. Notably, CRDs, with a median length of 120–168 kb (Supplementary Fig. 8f), were significantly more likely to be inside TADs as compared to any random sequence of the same width as CRDs (Fisher's exact test: odds ratio (OR) > 1, $P < 0.05$), an
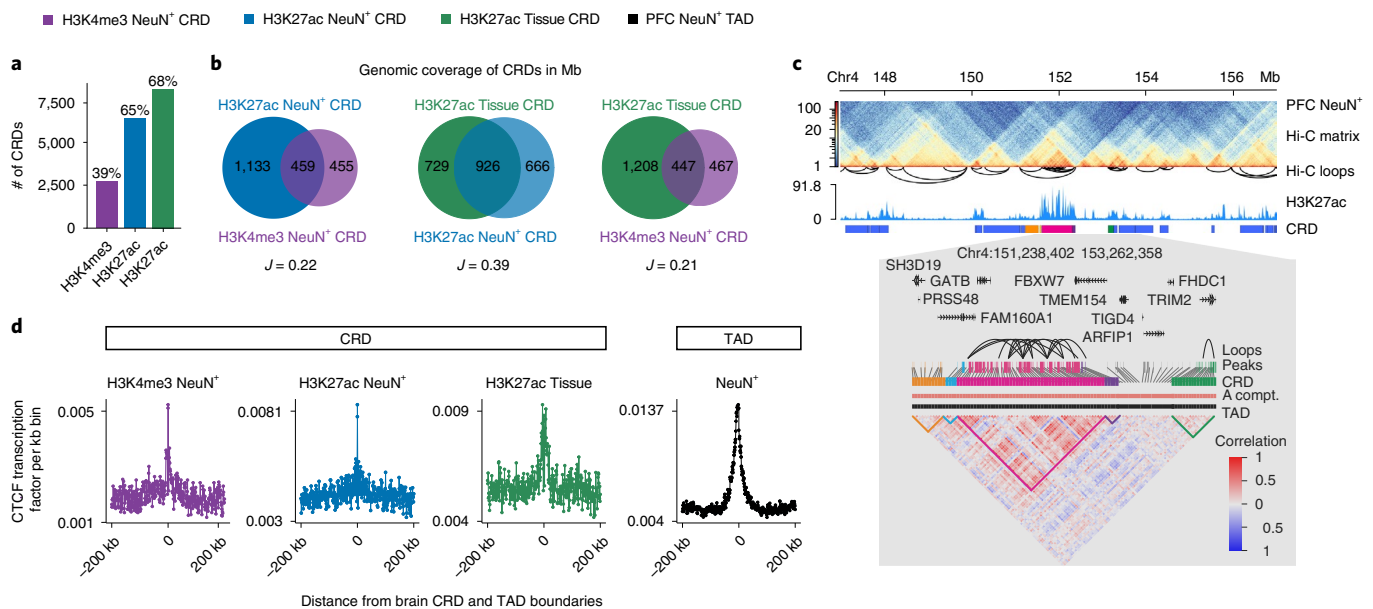
Fig. 3 | PFC histone CRDs reveal sub-TAD chromosomal organization. a, CRD analyses were conducted separately for each of our three ChIP-seq datasets (H3K4me3 NeuN+, H3K27ac NeuN+ and H3K27ac Tissue) (see also Fig. 1a,b). The numbers next to bars indicate the proportion of total peak population integrated into CRD structures. **b**, Venn diagrams summarizing genome-wide sequences (in Mb) integrated into CRD structures, including overlap and Jaccard similarity index between different histones and cell populations (NeuN+ or tissue). **c**, Top: Representative 10-Mb window of chromosome 4 showing PFC NeuN+ Hi-C TAD, chromosomal loop and H3K27ac landscape, including CRD structure. Bottom (shaded in gray color): higher-resolution (2-Mb) peak-to-CRD assignments and peak correlational structure expressed as an interaction matrix. **d**, Neuronal CTCF chromatin occupancies (y axis, using CTCF ENCODE reference ChIP-seq from H1 stem-cell-differentiated neuronal culture) in relation to distance from CRD (colored graphs) and TAD (black graph) boundaries.

effect that was particularly pronounced (OR ~3–4) for acetylated CRDs (Supplementary Fig. 9). A detailed analysis revealed that up to 77.4% (81.4%) and 94.3% (95.7%) of acetylated CRDs were within sub-TADs and TADs, respectively in PFC NeuN+ (Tissue). Also, a substantial proportion of 81.5% (83.5%) of sub-TADs and 59.1% (64.1%) of TADs covered one full-length H3K27ac CRD in PFC NeuN+ (Tissue) (Supplementary Fig. 10). Interestingly, acetylated CRDs consistently showed, both in H3K27ac NeuN+ (Study-1) and H3K37ac Tissue (Study-2), maximum density at the center of sub-TADs and TADs (Supplementary Fig. 11). In contrast, methylated CRDs were enriched at TAD boundaries (Supplementary Fig. 12), resonating with earlier reports on H3K4me3 enrichment at TAD boundaries[1,28,29]. Furthermore, both histone CRD and TAD borders were strongly enriched for occupancies of the structural protein CTCF (Fig. 3d), affirming that CRD modules are heavily constrained by the boundaries of their local TAD. Taken together, our studies reveal CRDs as structural units inserted into TADs of the adult PFC, with H3K27ac CRDs primarily representing enhancer-associated transcriptional domains localizing toward TAD centers, whereas the topology of H3K4me3 CRDs indicates a more diverse function at TAD peripheries.

Reproducible alterations of acetylated CRDs in diseased PFC. Having shown that individual histone peaks organize into CRDs as structural subunits within chromosomal domains, we then wanted to explore genome-wide CRD alterations in diseased brain. To this end, we applied a two-step stage-wise statistical test (Methods and Supplementary Fig. 7) to identify dysregulated CRDs (ΔCRDs) and dysregulated histone peaks (ΔCRD$_{\Delta Peaks}$) inside them. There were 1,010 of 6,389 (15.8%) significantly hyper-acetylated and 953 (14.9%) hypo-acetylated ΔCRDs at FDR 5% in PFC NeuN+ SCZ Study-1, with proportions of ΔCRDs somewhat lower in PFC tissue SCZ Study-2, with 563 (6.8%) hyper-acetylated and 521 (6.3%)

hypo-acetylated ΔCRDs (Fig. 4a and Supplementary Table 6). However, there was significant correlation between SCZ ΔCRDs, quantified as log$_2$FC of peaks inside ΔCRDs, from SCZ Study-1 and Study-2 ($\rho = 0.28$, $P = 4.4 \times 10^{-55}$, from 375 (2,138) and 367 (2,645) CRD (peaks) of Study-1 and Study-2, respectively; Supplementary Fig. 12a). Similarly, we counted 203 (2.5%) hyper-acetylated and 251 (3.0%) hypo-acetylated ΔCRDs for PFC tissue BD (Study-2), with significant correlation ($\rho = 0.69$, $P = 1.1 \times 10^{-266}$, from 126 (1,918) CRD (peaks); Supplementary Fig. 12b) between log$_2$FC of peaks inside ΔCRDs of SCZ and BD of Study-2. Furthermore, ~10–12% of in-ΔCRD H3K27ac peaks were significantly dysregulated (ΔCRD$_{\Delta Peaks}$) (Fig. 4a and Supplementary Fig. 13).

To determine SCZ genetic variant enrichment in ΔCRDs, we applied LD score regression analysis and found higher SCZ heritability coefficients in hyper-acetylated ΔCRDs over hypo-acetylated ΔCRDs and over all ΔCRDs, a highly consistent effect across Study-1 and Study-2 (Fig. 4b). Because the genomic coverage of BD ΔCRDs was insufficient for LD score computation, we, instead, estimated enrichment for common risk variants in ΔCRD genes using MAGMA and, as for the SCZ-sensitive ΔCRDs, observed a significantly higher coefficient of genetic association for SCZ (not BD) in hyper-acetylated BD ΔCRDs as compared to the total set of ΔCRDs and no genetic association in hypo-acetylated ΔCRDs (Supplementary Fig. 14). Interestingly, SCZ and BD ΔCRD$_{\Delta Peaks}$ showed enrichment for neuronal signaling and metabolic functions (Supplementary Table 7), with even higher heritability coefficients compared to all Study-1 and Study-2 ΔCRDs (Fig. 4b and Supplementary Fig. 14).

Next, to assess a potential link between local 'peak'-level (from Fig. 1) and CRD-level dysregulation, we first evaluated the odds of peaks to be in-CRD, with focus on disease-associated H3K27ac differences using a binomial generalized linear model (GLM) (Supplementary Fig. 15). Strikingly, in all three datasets (Study-1
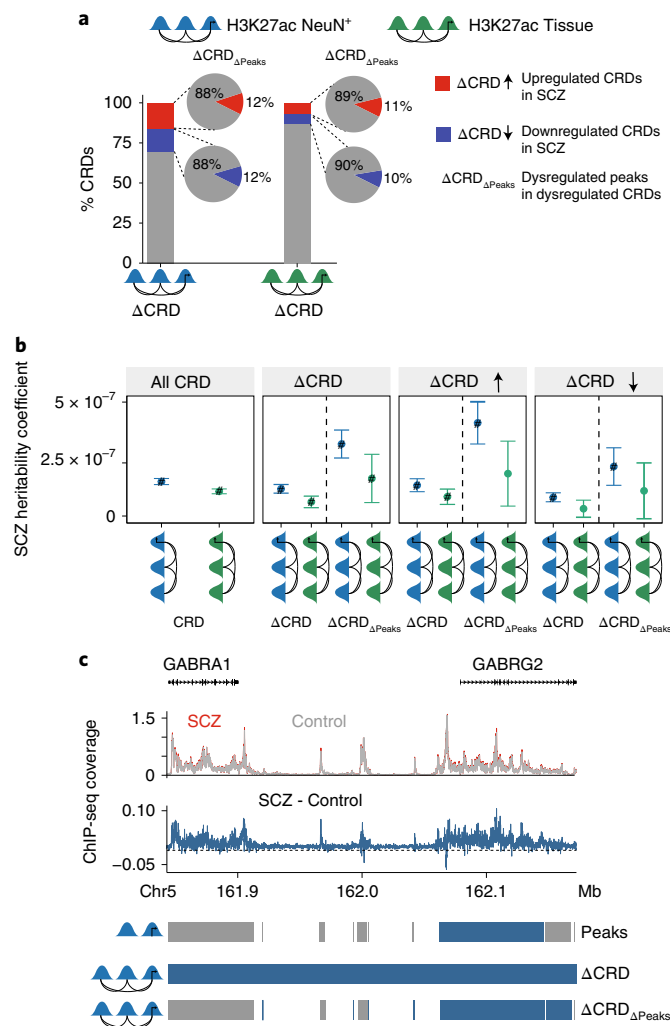
Fig. 4 | Acetylated CRDs dysregulated in SCZ. a, Proportional representation of SCZ-sensitive H3K27ac NeuN+ ΔCRDs stratified by hypo-acetylation (blue), hyper-acetylation (red) and not dysregulated (gray). Pie chart shows the proportion of dysregulated histone peaks ΔCRD$_{ΔPeaks}$ inside ΔCRD. **b**, SCZ heritability coefficients shown separately for H3K27ac NeuN+ (blue) and H3K27ac Tissue (green), as indicated, by (1) 'All CRD': all peaks inside CRD ($n = 114{,}123$ and $143{,}092$ peaks in CRD); (2) 'ΔCRD': dysregulated CRD ($n = 28{,}866$ and $15{,}787$ peaks in ΔCRD and $3{,}507$ and $1{,}673$ peaks in ΔCRD$_{ΔPeaks}$); (3) 'ΔCRD↑': hyper-acetylated with mean log$_2$FC (SCZ versus controls) > 0 ($n = 14{,}710$ and $7{,}770$ peaks in ΔCRD↑ and $1{,}825$ and 873 peaks in ΔCRD↑$_{ΔPeaks}$); and (4) 'ΔCRD↓': hypo-acetylated with mean log$_2$FC (SCZ versus controls) < 0 ($n = 14{,}156$ and $8{,}017$ peaks in ΔCRD↓ and $1{,}682$ and 800 peaks in ΔCRD↓$_{ΔPeaks}$) classified on the x axis as ΔCRD for all peaks and ΔCRD$_{ΔPeaks}$ for only dysregulated histone peaks. The overlap of peaks within the dysregulated CRDs in clusters with SCZ risk variants was assessed using LD score regression. '#': significant for enrichment in LD score regression after FDR correction of multiple testing across all tests in the plot (Benjamini–Hochberg, multiple testing $P < 0.05$); '*': nominally significant for enrichment ($P < 0.05$). Error bars show standard error in SCZ heritability from LD score regression. **c**, Representative example of a genomic region that spans ten peaks. Horizontal bars (blue, FDR 5%; gray, NS) mark peak-level analysis with a single differential peak (top row), hyper-acetylated CRD (middle row) and ΔCRD$_{ΔPeaks}$ (bottom row).

PFC NeuN+ SCZ, Study-2 PFC tissue SCZ and BD), disease-sensitive peaks were significantly more likely (OR > 1, $P < 0.05$) to fall inside CRDs than outside of CRDs (Supplementary Fig. 15). Moreover,

dysregulated peaks showed a strong tendency toward in-ΔCRD clustering (Poisson-based GLM model, OR = 1.65–3.48, $P < 0.05$) (Supplementary Fig. 15). Furthermore, differentially expressed genes across SCZ and controls from the CommonMind Consortium RNA sequencing cohort[30] were more likely to be in-ΔCRDs than non-dysregulated CRDs (Poisson-based GLM model, OR = 1.3–1.5, $P < 0.05$) (Supplementary Fig. 15). These findings strongly suggest that, in diseased SCZ PFC, alterations in histone acetylation manifest in a domain-specific manner encompassing an array of peaks, potentially affecting transcription. A representative hyper-acetylated ΔCRD (Fig. 4c) shows 0.4 Mb of the chr. 5 GABA receptor *GABRA1/GABRG2* gene cluster and risk locus, encompassing ten H3K27ac peaks, including 1 (6) hyper-acetylated peaks from Fig. 1 ΔPeak (Fig. 4a ΔCRD$_{ΔPeaks}$).

Dysregulated CRDs are aligned by chromosomal organization. Having shown that SCZ/BD PFC harbors alterations in structural domains, or dysregulated CRDs, we then asked whether disease-sensitive CRDs show evidence for coordinated ('trans-CRD') regulation in higher chromatin structure. We quantified each ΔCRD as the mean of in-CRD H3K27ac peak levels followed by its correlation as (diseased) CRD contact matrix (m CRDs × m CRDs). Indeed, principal component analysis (PCA) of the CRD contact matrix (Supplementary Fig. 16) revealed stratifications by the HiC-defined A and B compartments along with hyper-acetylation and hypo-acetylation across component-1. This finding suggested that ΔCRDs are aligned by directionality (hyper-acetylation versus hypo-acetylation) and chromatin structure, including 'A' permissive versus 'B' repressed/condensed compartments.

For a more detailed analysis on ΔCRD stratification, we applied the Bayesian information criterion (BIC) to identify the optimal number of clusters in every CRD contact matrix (k-means; Methods). We identified $k = 3$ in SCZ PFC NeuN+ Study-1 and $k = 2$ in SCZ, BD PFC tissue Study-2 as the optimal number of clusters (Supplementary Fig. 17a–c). We then created a resource of functional annotation of CRDs, including (1) cell-type-specific PFC reference sets, including H3K27ac for Glu projection neurons, GABA interneurons and OLIGs (Supplementary Table 8)[14]; (2) NeuN+ Hi-C chromosomal A and B compartments[19]; and (3) developmental (fetal versus adult) stage, defined from the epigenetic trajectory of human cortical development[31] (Methods; see Supplementary Fig. 18a for the distribution of annotated CRDs).

Of note, two out of three clusters in the Study-1 CRD contact matrix (Fig. 5a) were primarily composed of hyper-acetylated ΔCRDs representing Glu projection neurons, with chromosomal A:B compartmentalization further differentiating into cluster-1 A:B~1:2 and cluster-3 A:B~2:1. In striking contrast, cluster-2, overwhelmingly composed of hypo-acetylated ΔCRDs (91.2%), showed a ten-fold over-representation of interneuron-specific 'GABAergic' CRDs (Fig. 5b). Furthermore, cluster-3, which showed the highest proportion of A-compartment, showed an overall higher magnitude of gene expression as compared to cluster-1 and cluster-2 (Supplementary Fig. 20).

Similarly, Study-2 SCZ-specific and BD-specific ΔCRD contact matrices again showed stratification by hyper-acetylation versus hypo-acetylation and chromosomal compartmentalization as A versus B (Supplementary Fig. 19a,b). Analysis of enrichment of SCZ GWAS variants in dysregulated CRDs (using LD score) by annotated CRDs in each cluster revealed that the coefficient of SCZ heritability, as determined by LD score, was highest in magnitude for fetal as compared to adult annotated CRDs. This finding, consistent across Study-1 (coefficient = $2.6 × 10^{-7} ± 9.1 × 10^{-8}$, $P = 2.6 × 10^{-3}$ in cluster-1; Fig. 5c) and Study-2 (coefficient = $1.4 × 10^{-7} ± 9.9 × 10^{-8}$, $P = 0.085$ in cluster-2; Supplementary Fig. 21a), indicates the presence of
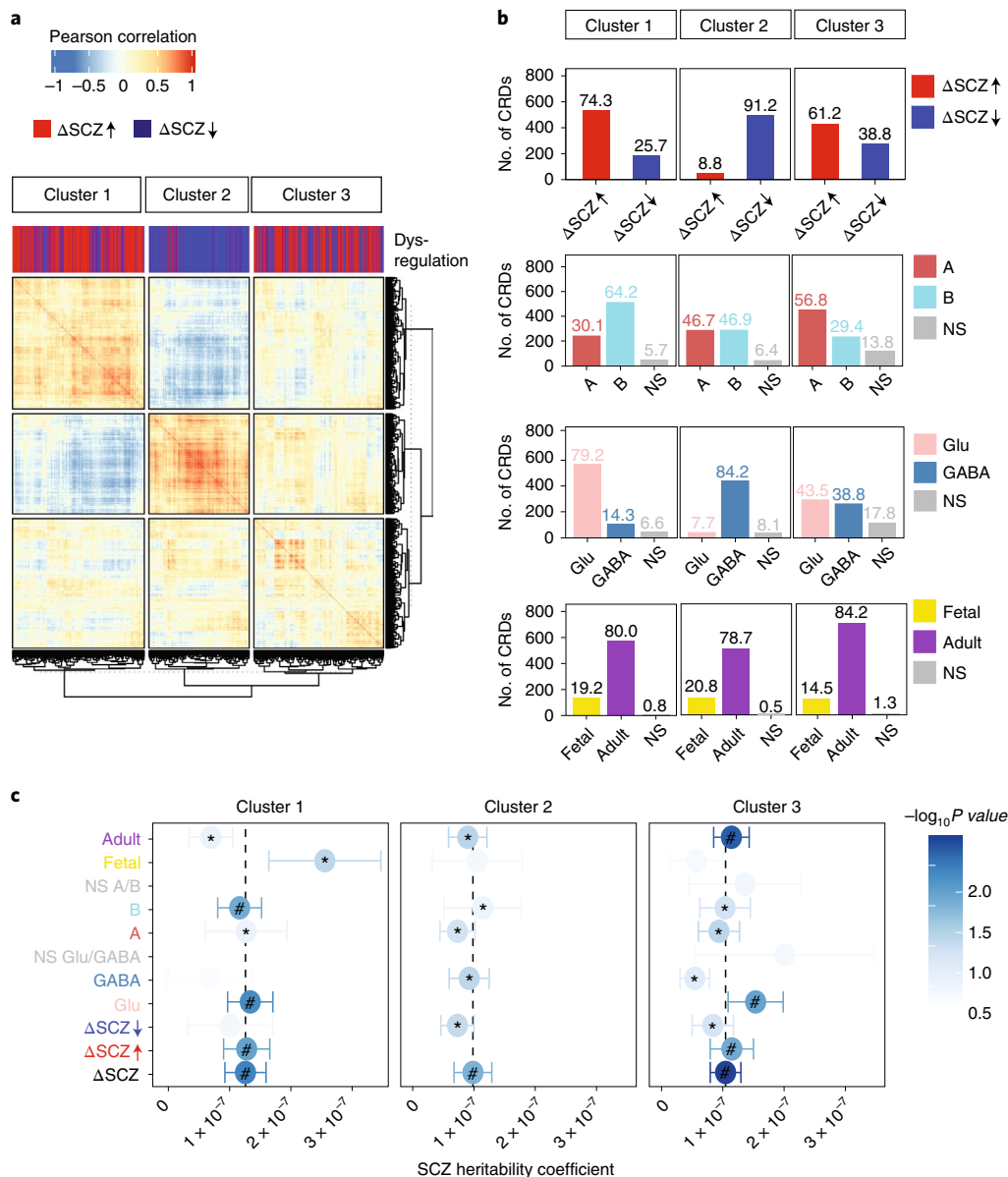
Fig. 5 | Fingerprinting disease-sensitive CRDs. a, CRD contact matrix of SCZ-sensitive H3K27ac NeuN[+] CRDs clustered into three large clusters; notice the striking separation of cluster-1 and cluster-3 representing hyper-acetylated H3K27ac CRDs in red and cluster-2 overwhelmingly defined by hypo-acetylated H3K27ac CRDs in navy. **b**, Composition of annotated CRDs by cell type (GABA in light blue, Glu in pink), compartments (A in cyan, B in indian red), dysregulation (hypo-acetylation versus hyper-acetylation in red and navy, respectively) and development (fetal in yellow, adult in purple). For every annotation, not significant or not annotated CRDs are shown in gray. **c**, Coefficients of heritability of SCZ by cluster and annotation. The overlap of peaks within the dysregulated CRDs in clusters with genetic variants was assessed using LD score regression. *P* values are from LD score regression. '#': significant for enrichment in LD score regression after Benjamini–Hochberg FDR correction for multiple testing across all tests in the plot (FDR < 5%); '*': nominally significant for enrichment (*P* < 0.05). Error bars show standard error in SCZ heritability from LD score regression.

neurodevelopmental signatures in SCZ, an effect that was particularly strong in PFC NeuN[+]. Furthermore, in PFC NeuN[+] Study-1, there was a strong cell-specific effect with higher heritability for Glu (as compared to GABA) annotated CRDs in hyper-acetylated cluster-1 and cluster-3. However, due to the additional signal from non-neuronal cell types, Study-2 SCZ PFC tissue lacked clear cell-specific heritability coefficients (Supplementary Fig. 21a), whereas, in BD PFC tissue, genetic association of BD risk genes (MAGMA) was observed specifically for Glu CRDs in hyper-acetylated cluster-2 (Supplementary Fig. 21b). Furthermore, SCZ heritability was present in both A/B compartment (Study-1) or A only (Study-2) (Fig. 5c and Supplementary Fig. 21a).

Nuclear topography of hyper-acetylated CRDs. Having shown that histone CRDs comprise a type of structural subunit embedded within the chromosomal TADs, with stratification of disease-sensitive CRDs aligning with facilitative versus repressive chromosomal environment and hyper-acetylation versus hypo-acetylation, we then explored nuclear topography and spatial 3D genome organization of the dysregulated CRDs. We used TAD coordinates of PFC NeuN[+] Hi-C reference sets in Chrom3D, a Monte Carlo-type algorithm for spherical genome modeling[32,33]. Indeed, pairwise Euclidean distances between TAD coordinates of PFC NeuN[+] that overlapped with the genomic coordinates of A-compartment-rich disease clusters defined by hyper-acetylated
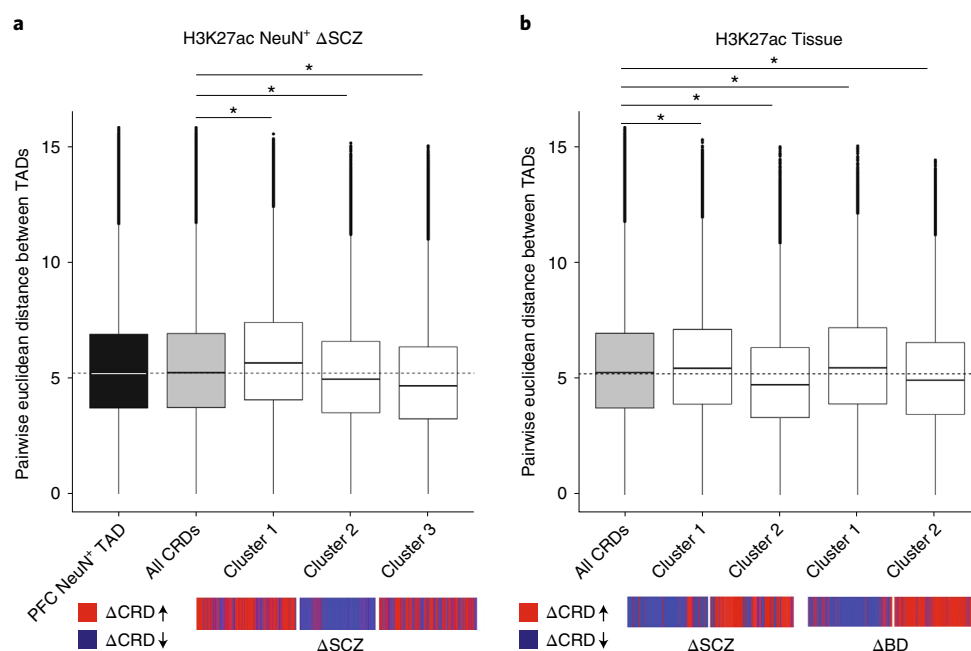
Fig. 6 | Spatial organization of diseased CRDs in the virtual Chrom3D model of the neuronal nucleus. a, Box plots show for SCZ H3K27ac NeuN+ cohort, showing for diseased ΔCRDs the pairwise Euclidean distance of PFC NeuN+ TAD, and all CRDs, and diseased ΔCRDs from cluster-1, cluster-2 and cluster-3 (from Fig. 5) with $n = 11,103$, 8,690 and 9,073 peaks for each cluster, respectively, as indicated. **b**, Box plots for H3K27ac Tissue cohort, separated by SCZ-sensitive and BD-sensitive ΔCRDs, with $n = 8,553$ in cluster-1 and 7,234 peaks in cluster-2 (from Fig. 5). Bar plots in the bottom in red and navy show the dysregulation status of clusters. The center shows the median; the box shows the interquartile range; whiskers indicate the highest and lowest values within 1.5× the interquartile range; and potential outliers from this are shown as dots. * represents $P < 0.05$, and P values are estimated using the Wilcoxon test.

CRDs revealed significantly higher TAD proximity and connectivity when compared to Chrom3D connectivity of all CRDs ($P < 0.05$) (Fig. 6a). This 3D genome phenotype was remarkably consistent across all three disease cohorts, including cluster-3 from SCZ PFC NeuN+ Study-1 and cluster-2 from SCZ and BD PFC Tissue Study-2, respectively (Fig. 6b). Therefore, diseased CRDs show distinct differences in spatial organization, including high chromosomal interactions between the TADs from hyper-acetylated clusters.

Discussion

This study mapped active promoter-associated and enhancer-associated histone methylation and acetylation profiles in PFC of 563 brain donors, providing, to our knowledge, the largest histone modification dataset for SCZ and BD to date. Our analyses based on histone peaks linked hyper-acetylation to regulatory sequences for neuronal signaling and development, and to SCZ genetic risk, with hyper-acetylated enhancers disproportionally enriched (compared to promoters) for risk-associated variants. These findings strongly suggest that epigenomic alterations in SCZ and BD brain are tracking the underlying genetic risk architecture.

Our finding that acetylated chromatin shows disease-sensitive changes in PFC neurons is interesting given that the frontal lobe of SCZ and BD subjects is reportedly affected by alterations in histone deacetylase enzyme (HDAC) activity, according to in vivo imaging[34,35] and postmortem expression[36] studies. Likewise, in the animal model, transgene-derived HDAC expression in PFC neurons alters cognition and behavior[37,38], and, furthermore, negative interference with PFC HDAC expression and activity exerts a therapeutic effect in psychosis[39–41]. In addition, we found that, although dysregulation of H3K27ac acetylation in (adult) PFC is representative for the broader population of individuals diagnosed with SCZ and BD, altered histone methylation, or at least H3K4me3, is not.

However, regulation of H3K4 methylation is highly dynamic during the extended period of human PFC development and maturation[42], and, furthermore, according to animal systems modeling disrupted fetal development in SCZ, brain-specific alterations in H3K4me3 are transient and antecede the emergence of defective cognition and behavior in the adult[43]. Therefore, considering the neurodevelopmental etiology of common psychosis, including SCZ and BD, it is possible that the PFC of our (adult) disease cases was transiently affected by H3K4 methylation changes during a much earlier (including prenatal) period and could include many of the regulatory H3K4me3-tagged sequences that are associated with heritable risk[2] (Supplementary Fig. 6).

In the second part of this study, we constructed CRD chromosomal domains by estimating the inter-individual correlations between histone peaks. We show that acetylated and methylated CRDs are firmly embedded into TAD and sub-TAD (self-folded) domains of chromosomal conformations but at much finer resolution. Histone CRDs, like the TADs and their nested sub-TADs, showed enrichment of CTCF structural protein at domain boundaries. This finding, together with our observation that H3K27ac-CRDs, comprised of arrays of active enhancers and promoters, are primarily located in the TAD center, whereas H3K4me3-CRDs tend to locate toward the TAD periphery, underscores that CRDs are a type of chromosomal modular unit linked to transcriptional activity and organized on a smaller scale than the chromosomal conformation-defined TADs. The CRD concept could open new avenues in neurogenomics, with the combination of Hi-C and CRD analyses, as presented here, offering novel insights into the finer-grained architecture of functional organization of chromosomes.

Notably, we found that regulatory sequences affected by H3K27ac 'peak' alterations in diseased PFC from SCZ and BD subjects were much more likely to origin within CRDs as compared

to isolated peaks positioned outside of CRDs. This observation strongly speaks to the functional significance of acetylated CRDs for human cognition and behavior. Along these lines, our two-stage analysis, $\Delta CRD_{\Delta Peaks}$ in $\Delta CRDs$, confirmed that heritability risk for SCZ was highest for hyper-acetylated H3K27ac peaks located in diseased CRDs. Because hyper-acetylated CRDs were strongly enriched for regulatory sequences linked to excitatory (projection) neurons, and also harbored H3K27ac peaks with high coefficient of heritability in fetal annotated CRDs, such type of cluster-specific fingerprint could signal functional importance for many acetylated chromatin domains early in the disease process. These findings broadly resonate with the notion that enhancers and other *cis*-regulatory sequences of the fetal brain are disproportionally over-represented among the set of common risk variants linked to SCZ[6,44,45]. Therefore, it is plausible to hypothesize that a subset of hyper-acetylated CRDs in diseased PFC neurons are vestiges of an early occurring neurodevelopmental disease process. Such type of epigenetic pathology in developing PFC could extend beyond the level of histone acetylation, given that alterations in DNA cytosine methylation profiles in adult SCZ PFC frequently encompass regulatory sequences defined by dynamic methylation drifts during the transition from the prenatal to the postnatal period[44].

Furthermore, the strong Glu-neuron-specific fingerprint in our hyper-acetylated CRD clusters is in excellent agreement with recent single-nucleus-level transcriptome profilings, reporting upregulation of Glu-neuron-specific expression modules in cortical layers of SCZ PFC[9,46] in addition to increased composite measures for Glu transcripts in PFC of SCZ subjects[47].

The final part of our analyses was focused on spatial genome organization of disease-associated CRDs, revealing, inside the virtual 3D sphere of a PFC neuron nucleus, an overall increased connectivity of TADs harboring hyper-acetylated CRDs, an effect that was particular pronounced for CRDs (clusters) stratified by a high proportion of the 'A' facilitative chromosomal compartment. This includes an overall higher inter-domain connectivity score in the Chrom3D simulated nuclear sphere. Which types of molecular mechanisms could drive the nuclear topography of disease-relevant chromosomal domains, including the structural convergence of functionally inter-related hyper-acetylated domains, as reported here? Interestingly, chromosomal contacts in brain and other tissues preferentially occur between loci targeted by the same transcription factors[48], with convergence on intra-chromosomal and inter-chromosomal hubs sharing a similar regulatory architecture among the interconnected enhancers[49–52]. In any case, based on the work presented here, we propose a longitudinal 3D or '4D nucleome' model for the epigenomics of SCZ and BD. According to this model, H3K27ac peaks that became dysregulated as early as in the fetal period of (PFC) development could subsequently serve as 'seed points', ultimately spreading epigenomic dysregulation, specifically hyper-acetylation, across an entire functional chromosomal domain or H3K27ac CRD. Therefore, SCZ and BD could ultimately be the manifestation of hyper-acetylation events progressing from risk-associated histone peaks to their chromosomal domains and eventually further spread by nuclear topography.

From this work, we provide unique resources of (1) dysregulated histone peaks in a large brain cohort of SCZ and BD and (2) genome-wide CRDs that delimit the highly connected histone peaks from unconnected histone peaks. We also provide (3) a workflow to integrate epigenomics into 3D nuclear organization from *cis* to *trans* level of interactions between histone peaks to investigate the effect of disease at population scale and (4) domain-specific disease-sensitive peaks ($\Delta CRD_{\Delta Peaks}$) as critical seed points affecting gene regulation. We expect that these resources will provide a roadmap for future studies with even larger cohorts of SCZ and BD brains, aimed at gaining a deeper understanding of the emerging link between circuit-specific dysfunction and genome organization

in SCZ and BD. These could include polygenic liabilities affecting distinct dimensions of psychosis, including disorganization of thought process, delusions and hallucinations, and social withdrawal and other negative symptoms[53].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41593-022-01032-6.

References

1. Dixon, J. R. et al. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
2. Girdhar, K. et al. Cell-specific histone modification maps in the human frontal lobe link schizophrenia risk to the neuronal epigenome. *Nat. Neurosci.* **21**, 1126–1136 (2018).
3. Cheung, I. et al. Developmental regulation and individual differences of neuronal H3K4me3 epigenomes in the prefrontal cortex. *Proc. Natl Acad. Sci. USA* **107**, 8824–8829 (2010).
4. Khan, A., Mathelier, A. & Zhang, X. Super-enhancers are transcriptionally more active and cell type-specific than stretch enhancers. *Epigenetics* **13**, 910–922 (2018).
5. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat. Neurosci.* **18**, 199–209 (2015).
6. Roussos, P. et al. A role for noncoding variation in schizophrenia. *Cell Rep.* **9**, 1417–1429 (2014).
7. Takata, A. et al. Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron* **82**, 773–780 (2014).
8. Fullard, J. F. et al. An atlas of chromatin accessibility in the adult human brain. *Genome Res.* **28**, 1243–1252 (2018).
9. Hauberg, M. E. et al. Common schizophrenia risk variants are enriched in open chromatin regions of human glutamatergic neurons. *Nat. Commun.* **11**, 5581 (2020).
10. Smigielski, L., Jagannath, V., Rössler, W., Walitza, S. & Grünblatt, E. Epigenetic mechanisms in schizophrenia and other psychotic disorders: a systematic review of empirical human findings. *Mol. Psychiatry* **25**, 1718–1748 (2020).
11. Fromer, M. et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat. Neurosci.* **19**, 1442–1453 (2016).
12. Hoffman, G. E. et al. CommonMind Consortium provides transcriptomic and epigenomic data for schizophrenia and bipolar disorder. *Sci. Data* **6**, 180 (2019).
13. Hauberg, M. E. et al. Differential activity of transcribed enhancers in the prefrontal cortex of 537 cases with schizophrenia and controls. *Mol. Psychiatry* **24**, 1685–1695 (2019).
14. Kozlenkov, A. et al. A unique role for DNA (hydroxy)methylation in epigenetic regulation of human inhibitory neurons. *Sci. Adv.* **4**, eaau6190 (2018).
15. Wong, A. H. C. et al. Association between schizophrenia and the syntaxin 1A gene. *Biol. Psychiatry* **56**, 24–29 (2004).
16. Bryois, J. et al. Evaluation of chromatin accessibility in prefrontal cortex of individuals with schizophrenia. *Nat. Commun.* **9**, 3121 (2018).
17. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
18. Madani Tonekaboni, S. A., Mazrooei, P., Kofia, V., Haibe-Kains, B. & Lupien, M. Identifying clusters of cis-regulatory elements underpinning TAD structures and lineage-specific regulatory networks. *Genome Res.* **29**, 1733–1743 (2019).
19. Bendl, J. et al. The three-dimensional landscape of chromatin accessibility in Alzheimer's disease. Preprint at https://www.biorxiv.org/content/10.1101/2021.01.11.426303v1 (2021).
20. Dong, P. et al. Population-level variation of enhancer expression identifies novel disease mechanisms in the human brain. Preprint at https://www.biorxiv.org/content/biorxiv/early/2021/06/11/2021.05.14.443421.full.pdf (2021).
21. Delaneau, O. et al. Chromatin three-dimensional interactions mediate genetic effects on gene expression. *Science* **364**, eaat8266 (2019).

22. Waszak, S. M. et al. Population variation and genetic control of modular chromatin architecture in humans. *Cell* **162**, 1039–1050 (2015).

23. Hoffman, G. E., Bendl, J., Girdhar, K. & Roussos, P. decorate: differential epigenetic correlation test. *Bioinformatics* **36**, 2856–2861 (2020).

24. Ambroise, C., Dehman, A., Neuvial, P., Rigaill, G. & Vialaneix, N. Adjacency-constrained hierarchical clustering of a band similarity matrix with application to genomics. *Algorithms Mol. Biol.* **14**, 22 (2019).

25. Beagan, J. A. & Phillips-Cremins, J. E. On the existence and functionality of topologically associating domains. *Nat. Genet.* **52**, 8–16 (2020).

26. Nuebler, J., Fudenberg, G., Imakaev, M., Abdennur, N. & Mirny, L. A. Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc. Natl Acad. Sci. USA* **115**, E6697–E6706 (2018).

27. Kichaev, G. et al. Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75 (2019).

28. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).

29. Lazar, N. H. et al. Epigenetic maintenance of topological domains in the highly rearranged gibbon genome. *Genome Res.* **28**, 983–997 (2018).

30. Hoffman, G. E. et al. Sex differences in the human brain transcriptome of cases with schizophrenia. *Biol. Psychiatry* **91**, 92–101 (2022).

31. Li, M. et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018).

32. Paulsen, J. et al. Chrom3D: three-dimensional genome modeling from Hi-C and nuclear lamin-genome contacts. *Genome Biol.* **18**, 21 (2017).

33. Paulsen, J., Liyakat Ali, T. M. & Collas, P. Computational 3D genome modeling using Chrom3D. *Nat. Protoc.* **13**, 1137–1152 (2018).

34. Tseng, C.-E. J. et al. In vivo human brain expression of histone deacetylases in bipolar disorder. *Transl. Psychiatry* **10**, 224 (2020).

35. Gilbert, T. M. et al. PET neuroimaging reveals histone deacetylase dysregulation in schizophrenia. *J. Clin. Invest.* **129**, 364–372 (2019).

36. Schroeder, F. A. et al. Expression of *HDAC2* but not *HDAC1* transcript is reduced in dorsolateral prefrontal cortex of patients with schizophrenia. *ACS Chem. Neurosci.* **8**, 662–668 (2017).

37. Bahari-Javan, S. et al. HDAC1 links early life stress to schizophrenia-like phenotypes. *Proc. Natl Acad. Sci. USA* **114**, E4686–E4694 (2017).

38. Jakovcevski, M. et al. Prefrontal cortical dysfunction after overexpression of histone deacetylase 1. *Biol. Psychiatry* **74**, 696–705 (2013).

39. Schroeder, F. A., Lin, C. L., Crusio, W. E. & Akbarian, S. Antidepressant-like effects of the histone deacetylase inhibitor, sodium butyrate, in the mouse. *Biol. Psychiatry* **62**, 55–64 (2007).

40. de la Fuente Revenga, M. et al. HDAC2-dependent antipsychotic-like effects of chronic treatment with the HDAC inhibitor SAHA in mice. *Neuroscience* **388**, 102–117 (2018).

41. Thomas, E. A. Histone posttranslational modifications in schizophrenia. *Adv. Exp. Med. Biol.* **978**, 237–254 (2017).

42. Shulha, H. P., Cheung, I., Guo, Y., Akbarian, S. & Weng, Z. Coordinated cell type–specific epigenetic remodeling in prefrontal cortex begins before birth and continues into early adulthood. *PLoS Genetics* **9**, e1003433 (2013).

43. Connor, C. M. et al. Maternal immune activation alters behavior in adult offspring, with subtle changes in the cortical transcriptome and epigenome. *Schizophr. Res.* **140**, 175–184 (2012).

44. Jaffe, A. E. et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. *Nat. Neurosci.* **19**, 40–47 (2016).

45. Hannon, E. et al. Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nat. Neurosci.* **19**, 48–54 (2016).

46. Ruzicka, W. B. et al. Single-cell dissection of schizophrenia reveals neurodevelopmental-synaptic axis and transcriptional resilience. Preprint at https://www.medrxiv.org/content/10.1101/2020.11.06.20225342v1 (2020).

47. Dienel, S. J., Enwright, J. F., Hoffman, G. D. & Lewis, D. A. Markers of glutamate and GABA neurotransmission in the prefrontal cortex of schizophrenia subjects: disease effects differ across anatomical levels of resolution. *Schizophr. Res.* **217**, 86–94 (2020).

48. Bonev, B. et al. Multiscale 3D genome rewiring during mouse neural development. *Cell* **171**, 557–572 (2017).

49. Lomvardas, S. et al. Interchromosomal interactions and olfactory receptor choice. *Cell* **126**, 403–413 (2006).

50. Quinodoz, S. A. et al. Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* **174**, 744–757 (2018).

51. Khanna, N., Hu, Y. & Belmont, A. S. HSP70 transgene directed motion to nuclear speckles facilitates heat shock activation. *Curr. Biol.* **24**, 1138–1144 (2014).

52. Ahanger, S. H. et al. Distinct nuclear compartment-associated genome architecture in the developing mammalian brain. *Nat. Neurosci.* **24**, 1235–1242 (2021).

53. Legge, S. E. et al. Associations between schizophrenia polygenic liability, symptom dimensions, and cognitive ability in schizophrenia. *JAMA Psychiatry* **8**, 1143–1151 (2021).

Methods

Brains (postmortem). All tissue donors of Study-1 were from the Icahn School of Medicine at Mount Sinai (MSSM), University of Pennsylvania (PENN) and University of Pittsburgh (PITT) brain banks. All tissue donors of Study-2 were from the HBCC at the National Institute of Mental Health. Demographics of the brain cohort, toxicology and neuropathology reports are summarized in Table 1 and Supplementary Table 1. No statistical methods were used to pre-determine sample sizes.

ChIP-seq library preparation and sequencing. From the total set of 739 histone ChIP-seq datasets presented here, 28% (100 controls and 109 SCZ cases) were included in a recent PsychENCODE genomics reference paper for the adult human brain[54]; the remaining 530 ChIP-seq datasets had not been presented before.

Nuclei were extracted from approximately 300-mg aliquots of frozen frontal (dorsolateral prefrontal) cortex tissue, immuno-tagged with Anti-NeuN Alexa 488 (cat. no. MAB377X, EMD Millipore) antibody that robustly stains human cortical neuron nuclei[55,56] for subsequent FANS. Next, chromatin of sorted nuclei was digested with micrococcal nuclease and subsequently pulled down with anti-histone antibodies, followed by library preparation and sequencing. Two histone antibodies, anti-H3K4me3 (cat. no. 9751BC, lot no. 7, Cell Signaling Technology) and anti-H3K27ac (cat. no. 39133, lot no. 01613007, Active Motif) were used for immunoprecipitation. Antibody specificity was tested using peptide-binding assays and immunoblotting of nuclear extracts from human postmortem cortical tissue. A commercially available histone H3 peptide array (cat. no. 16–667, Millipore) containing 46 peptides representing 46 different histone H3 post-translational modifications was used as previously described[55]. All procedures were performed as described in the recent PsychENCODE methods paper, providing a detailed description of the protocol[55]. For each cell-type-specific ChIP assay, a minimum of 400,000 sorted neuronal (NeuN+) nuclei was required as starting material. For selected gene promoters, ChIP-PCR was conducted to validate cell-type-specific peak profiles. Furthermore, quality controls for nuclei after fluorescence-activated cell sorting (FACS) included visual inspection under the microscope as described[55]. Of note, owing to our stringent FACS gating criteria with maximized specificity (not sensitivity), 100% of sorted nuclei in the neuronal fraction showed green fluorescence, confirming NeuN+ status, whereas 100% of sorted nuclei in the non-neuronal fraction showed only blue DAPI stain, confirming NeuN– status. Additional ChIP-seq studies were conducted with homogenized dorsolateral PFC as input. To this end, frozen human postmortem brain tissue (approximately 20–200 mg) was homogenized in lysis buffer, and the total nuclei were purified. The nuclei solution was resuspended in 300 µl of douncing buffer, treated with 2 µl of micrococcal nuclease (0.2 U µl⁻¹) for 5 minutes at 28 °C, followed by 30 µl of 500 mM of EDTA to stop the reaction. After this initial procedure for nuclei preparation and digestion, the sample was processed in the same manner as described for the nuclei samples sorted by FACS.

Randomization and blinding. To avoid batch effects and other confounds, samples underwent repeated rounds of randomization, including (1) ChIP and (2) library preparation. Blinding was not relevant to this study; analysts were aware of data generation, processing and donor metadata.

Adapter sequences removal. First, the raw FASTQ files were corrected for adaptor paired-end sequences using trimming tool called Trimmomatic (version 0.36)[57] with the following settings: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10:8:TRUE, LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15 and MINLEN:36.

Alignment, filtering, quality control and consolidation of BAM files. Trimmed FASTQ files from each study were aligned to hg38 (GRCh38) human genome using the Burrows–Wheeler Aligner (BWA-0.7.8-r455) method with default settings[58]. The output files were exported as BAM files. For quality control steps of BAM files, we implemented the ENCODE pipeline workflow, which is as follows: (1) remove unmapped reads, mates and low-quality mapping reads (mapq = 30); (2) remove orphan reads and reads that were mapped to different chromosomes; and (3) remove PCR duplicates using the picard (version 2.2.4) tool (http://broadinstitute.github.io/picard).

All BAM files from the above step were tested for ENCODE quality control parameters for ChIP-seq files: normalized strand coefficient (NSC > 1.0) and relative strand coefficient (RSC > 1) using phantompeakqualtools (version 2.0)[59]. Supplementary Fig. 1a shows the frequency of NSC and RSC of samples from Study-1 and Study-2. We provide the NSC and RSC of each sample (see 'Data availability').

After filtering out the BAM files based on ChIP-seq quality control parameters, we prepared the files for the next step, which is a consolidation of BAM files separately for each dataset. The objective was to subsample each ChIP-seq library to a fixed number of mapped reads and consolidate the subsampled libraries into one file. To obtain a fixed number of mapped reads for subsampling of BAM files, we took the minimum number of mapped reads from each study: H3K4me3 NeuN+ = 12 M, H3K27ac NeuN+ = 22 M and H3K27ac Tissue = 23 M. We obtained (median) ~30, ~60 and ~59 million mapped paired-end reads (2 × 75 bp) for H3K4me3 NeuN+, H3K27ac NeuN+ and H3K27ac Tissue, respectively

(Supplementary Fig. 1a). A similar procedure was followed to create a consolidated input–control file for NeuN+ Study-1 and tissue Study-2.

Mislabeling and contamination of samples check. For samples mismatch and contamination check, we used the QTLtools (version 1.3) Match BAM to VCF (MBV)[60] option. MBV takes as input a VCF file containing the genotype data for Study-1 and Study-2 samples and a mapped BAM file from the above section (see the 'Alignment, filtering and consolidation of BAM files' section). We did this step using the merged VCF file of genotypes of Study-1 and Study-2 separately. None of our samples was mismatched or contaminated. We provide a summary of QTLtools (version 1.3) MBV[60] results of all samples (see 'Data availability').

Peak calling. Narrow peak regions were called on a consolidated file of H3K4me3 histone mark dataset using macs2 (version 2.2.6)[61] with Poisson P = 0.01 with —keep-dup all —nomodel —extsize = 150. Similarly, broad peak regions were called on Study-1 and Study-2 consolidated files of H3K27ac histone mark datasets using macs2 (version 2.2.6)[61] with P value cutoff = 0.01, —extsize = 150. We used a NeuN+ consolidated input control and a tissue consolidated control file separately as control inputs for peak calling on each study. All peaks were filtered from blacklisted[62] region peaks for downstream analysis.

Quantification of ChIP-seq signal. ChIP-seq signal was quantified for every sample and every consensus peak obtained from the above section using featureCounts (version 1.5.0) software[63]. The objective is to count the number of reads overlapping the genomic coordinates of peaks. This step results into a matrix of $m_{peaks} \times n_{samples}$, 66,163 peaks × 230 H3K4me3 NeuN+, 124,054 peaks × 260 H3K27ac NeuN+ and 207,866 peaks × 249 H3K27ac Tissue (Supplementary Table 2a–c).

From these matrices, peaks with the low expression were filtered out using counts per million (CPM) of histone peaks >1 in at least 10% of samples as a threshold resulting into 64,254 peaks × 230 H3K4me3 NeuN+, 114,136 peaks × 260 H3K27ac NeuN+ and 143,092 peaks × 249 H3K27ac Tissue (Supplementary Table 3). Next, the read counts were corrected for library size using the trimmed mean of M-values (TMM) method from edgeR library[64] and converted into the voom-normalized matrices.

Estimation of proportion of cell types in H3K27ac Tissue. To account for cell type heterogeneity in H3K27ac Tissue samples, we estimated the proportion of Glu, GABA and OLIG using dtangle (version 2.0.9) software[65]. Each tissue was modeled as a mixture of Glu, GABA and OLIG. The reference samples and peaks of Glu, GABA and OLIG were created using the previously published H3K27ac dataset from the PFC brain region[14]. We provide a vector of percent of cell types for each sample in the metadata table (see 'Data availability').

Covariates model selection. To estimate the technical and biological noise at sample level, we employ a two-step approach.

1. We first identify the number of principal components using the PCA method on the normalized read counts to identify the number of components that had variance of at least 1% of variance in the data. For each dataset, we take the correlation of all technical and biological covariates with the identified principal components and shortlisted the ones with FDR < 20%.
2. BIC: To identify the optimal number of covariates to have a good average model of histone peaks expression, we apply a BIC approach[66], which introduces a penalty term for the number of parameters in the model. We start with 'Diagnosis+Sex' as a base model and test all covariates one by one that were identified in the PCA step. Selection criterion of a covariate in the model is that at least 5% of peaks should have ($BIC_{Diagnosis+Gender+Covariate} - BIC_{Diagnosis+Gender}$) per histone peak ≥ 2.

Other covariates are added sequentially in this model until they fail to meet the criterion of BIC threshold. The following are the covariates that were used to correct the voom-normalized matrices for each study:

$$H3K4me3\ NeuN^+ = Sex + Diagnosis + GC\ content + FRIP + Magnetic\ Beads$$

$$H3K27ac\ NeuN^+ = Sex + Diagnosis + GC\ content + FRIP + Age$$

$$H3K27ac\ Tissue = Sex + Diagnosis + Neurons\ proportion$$
$$+ GC\ content + FRIP + FRIP^2 + Age$$

FRIP, fraction of reads in peaks.

Supplementary Fig. 1c shows the distribution of variance explained by each covariate in the ChIP-seq peaks activity matrix of each study. For a complete list of covariates, see the metadata table in the 'Data availability' section.

Annotating ChIP-seq histone peaks regions. Genes and genomic context: The Ensembl 95 genes were used for all analyses in this paper. To annotate the genomic

region of a histone peak as TSS, exon, 5′ UTR, 3′ UTR, intronic or intergenic, we used ChIPSeeker (version 1.18.0)[67]. The transcript database used for the annotation is 'TxDb.Hsapiens.UCSC.hg38.knownGene'. We used a threshold of ±3 kb distance from TSS of a gene for promoter annotation. Figure 1a shows the distribution of peaks annotated to categories (1) promoters, (2) introns, (3) distal intergenic and (4) exon and UTRs using the hg38 transcript database imported using the ChIPSeeker package.

Overlap with previously published datasets. We calculated the Jaccard index to measure the concordance of histone peaks in Study-1 and Study-2 with existing datasets of REP[68] and EpiMap[2]. Jaccard index is measured as the intersection of base pairs divided by the union of base pairs. Supplementary Fig. 1b shows the pairwise similarity of datasets REP and EpiMap and Study-1 and Study-2.

Peaks analysis. Disease differential analysis. To identify SCZ- and BD-sensitive peaks, we performed differential analysis on covariates-corrected ('Covariates model selection' section) matrices from H3K4me3 NeuN+, H3K27ac NeuN+ and H3K27ac Tissue using the limma (version 4.1)[69] pipeline. Supplementary Table 3 provides differential analysis results from the above-mentioned studies. Supplementary Fig. 2a shows the overlap of SCZ-sensitive H3K27ac NeuN+ peaks that overlap at least one base pair with H3K27ac Tissue peaks, whereas Supplementary Fig. 2b shows the overlap of SCZ- and BD-sensitive peaks in H3K27ac Tissue.

Meta analysis of H3K27ac NeuN+ and H3K27ac Tissue. Next, we combined the differential analysis results from H3K27ac NeuN+ and H3K27ac Tissue to obtain the consensus peaksets using fixed effect analysis[70]. We first created the consensus peakset by taking the set of H3K27ac NeuN+ peaks that had at least 90% overlap of its width with H3K27ac NeuN+ peaks. Then, we took the differential analysis table of overlapping peaksets of both NeuN+ and tissue to run fixed effect analysis using the rma function from the R metafor package (version 2.0)[70]. Figure 1d shows the proportion of differential peaks and the volcano table of rma analysis.

Pathway analysis of histone peaks. To interpret the disease-specific signatures in dysregulated H3K27ac NeuN+, Meta NeuN+ and Tissue peaks, we used the Genomic Regions Enrichment of Annotations Tool (GREAT) approach to assign peaks to genes. We examined the biological function of nearby genes for these non-overlapping peak regions using GREAT[71]. The settings for GREAT are as follows: proximal 5.0 kb upstream, 5.0 kb downstream and plus Distal: up to 100 kb. Supplementary Fig. 4 shows the pathway enrichment of SCZ dysregulated peaks from H3K27ac NeuN, Meta NeuN+ and Tissue and BD dysregulated peaks from H3K27ac Tissue.

Anti-psychotics differential analysis. To estimate the variance explained by anti-psychotic treatment in SCZ- and BD-sensitive peaks, we performed differential analysis on covariates corrected ('Covariates model selection' section) matrices from H3K27ac NeuN+ using the limma (version 4.1)[69] pipeline. We had information on anti-psychotics-typical (for example, Haldol and chlorpromazine) from 116 patients with SCZ (36 = yes, 80 = no) and anti-psychotics-atypical (for example, risperidone and clozapine) from 117 patients with SCZ (52 = yes, 65 = no). We conducted differential analysis in an expression matrix from 116 (117) patients with SCZ across (1) anti-psychotics-typical (yes) versus no anti-psychotics-typical and (2) anti-psychotics-atypical (yes) versus no anti-psychotics-typical. Supplementary Table 3 provides the results from differential analysis.

LD score enrichment analysis. To estimate the enrichment of brain-related and non-brain-related GWAS in all identified histone peaks and disease-sensitive peaks from H3K27ac NeuN+, Meta NeuN+ and Tissue, we used LD score partitioned heritability (version 1.0.0)[17]. Fig. 2, Supplementary Table 4 and Supplementary Fig. 6 show the LD score enrichments of SCZ-sensitive peaks from H3K27ac NeuN+, Meta NeuN+ and Tissue.

For the traits, we used the European-only version of the summary statistics when available. As a consequence, all GWAS results were based on individuals of European ancestry. The broad major histocompatibility complex region (hg19:chr6:25–35 Mb) was excluded owing to its extensive and complex LD structure, but, otherwise, default parameters were used for the algorithm. We ran LD score analyses only with sets of histone peaks covering 0.05% or more of the human genome.

MAGMA association trait analysis. Owing to the genomic coverage of BD-associated peaks less than 0.05%, we used MAGMA[72] version 1.06b to measure the association with SCZ risk peaks.

CRD. Genome-wide CRD calling. We identified CRDs separately on H3K4me3 NeuN+, H3K27ac NeuN+ and H3K27ac Tissue by leveraging the inter-individual correlations of samples. Here, we discuss in detail the stepwise workflow of CRD calling and identification of disease-specific CRDs as shown in Supplementary Fig. 7.

Removal of low correlation structure. We first corrected for global effects of covariates to retain the correlation structure using probabilistic estimation of expression residuals (PEER) residualization[73] of histone peaks normalized

expression from each study. A total of 18 PEER-corrected matrices $m_{peak\_PEER\_i} \times n_{samples}$ ($i = \{1, 5, 10, 15, 20, 25\}$) were produced (six PEER-corrected H3K4me3 NeuN+, six PEER-corrected H3K27ac NeuN+ and six H3K27ac Tissue). CRDs were called on 18 matrices individually using the following R functions from the decorate (version 1.0.14)[23] package:

$$CRDlist = runOrderedClusteringGenome$$

$$\left(m_{peak\_PEERj} \times n_{samples}, peaks, method.corr = ''Spearman'' \right)$$

$$CRDClusters = createClusters \left(CRDlist, method = ''meanClusterSize'', \right.$$

$$meanClusterSize = c \left(10, 25, 50, 80, 100 \right) \left. \right)$$

$$CRDScore = scoreClusters \left(CRDlist, CRDClusters \right)$$

The output from the above-mentioned commands were 18 CRDscore objects (six for each study dataset) containing a table of histone peaks assigned to CRDs, their mean correlation and lead eigen factor (LEF). The LEF of a CRD is a fraction of variance explained by the first eigenvalue of the correlation matrix $[m \times m]$ of histone peaks located within a CRD. Larger LEF values (that is, >10%) can be interpreted as strongly correlated peaks, whereas smaller values correspond to weaker correlations of peaks located within a CRD. Filtering out the CRDs with weaker correlations is an important step because it substantially reduces the burden of multiple testing in differential CRD analysis.

CRD filtering and merging. To filter out CRDs with weaker correlations, histone peaks positions were shuffled per chromosome for all samples to create permuted matrices $m_{Permutation\_j\_peaks\_PEER\_i} \times n_{samples}$ (where $i = \{1, 5, 10, 15, 20, 25\}$ and $j = 1–10$). A total of 180 PEER-corrected matrices $m_{peak\_PEER\_i} \times n_{samples}$ ($i = \{1, 5, 10, 15, 20, 25\}$) were produced (10 permutations × 6 PEER-corrected matrices × 3 datasets; $m_{Permutation\_j\_peaks\_PEER\_i} \times n_{samples}$). CRDs were called on 180 matrices individually using the following R functions from the decorate package:
CRD calling on 180 matrices (10 permutations × 6 PEER-corrected matrices × 3 datasets; $m_{Permutation\_j\_peaks\_PEER\_i} \times n_{samples}$) on permuted matrices followed the same workflow as explained. Lastly, LEF_{cutoff} was obtained as vectors were combined from ten $CRDScore$ lists obtained from CRD calling on $m_{Permutation\_j\_OCR\_PEER\_i} \times n_{samples}$ ($i = 1–10$):

$$LEF_{Permuted} = [LEF_1, LEF_2, LEF_3, ..., LEF_{10}], where LEF_i = N \times 1$$

$$LEF_{cutoff} = Pr \left(LEF_{Permuted} = threshold \right), where threshold = 0.90$$

The final table of CRDs was obtained by keeping all CRDs with $LEF_{measured} > LEF_{cutoff}$ in $CRDScore$ list of $m_{peaks\_PEER\_i} \times n_{samples}$. Supplementary Fig. 8b shows an example of the distribution of $LEF_{measured}$ and $LEF_{permuted}$.

Next, overlapping CRDs of different sizes were merged to obtain discrete CRDs for downstream analysis. To decide the optimal number of PEER factors, we measured the LEF_{cutoff} of CRDs called on the input matrix histone peaks matrix residualized by various numbers of PEER factors; we tested {1, 5, 10, 15, 20, 25} PEER factors (Supplementary Fig. 8c). The number of peaks within the CRDs are shown in Supplementary Fig. 8d, and the final lists of coordinates of CRDs of Study-1 H3K4me3 NeuN+, H3K27ac NeuN+ and H3K27ac Tissue are provided in Supplementary Table 5.

In silico biological validation of CRDs. To validate 3D interactions captured by CRDs with Hi-C dataset, we used the CTCF ChIP-seq peak list from ENCODE human neural cells[74] (see 'Data availability' section). We quantified the density of CTCF sites in 200 bins (each bin size equals to 1 kb) around CRD boundaries (Fig. 3e). To quantify how many in silico 3D interactions captured by CRDs are within the 3D interactions measured as TADs from PFC NeuN+ Hi-C experiments[19], we measured the number of CRDs overlapping with PFC NeuN+ Hi-C TADs stratified by number of TADs ($N = \{0, 1, 2, 3, \geq4\}$). Next, we measured how many PFC NeuN+ Hi-C TADs are within the CRDs stratified by the number of CRDs ($N = \{0, 1, 2, 3, \geq4\}$). We measured the correlation of histone peaks inside the Hi-C loops and outside the Hi-C loops to show that peaks inside the Hi-C loops have more correlation than the peaks outside the Hi-C loops (Supplementary Fig. 10).

Glu, GABA and OLIG ChIP-seq data. The data were obtained from ref. [14]. H3K27ac peaks were called with DFilter[75] using the following parameters: '-f=bam -pe -ks=60 -lpval=4'. For each cell type, H3K27ac peak lists for replicate samples were then overlapped using a custom R script, and peaks that were present in at least half of replicates were preserved for further analysis (peak numbers: 44,519 in GABA neurons, 46,580 in Glu neurons and 45,963 in OLIG cells). Peaks detected in Glu, GABA and OLIG cells were further overlapped using the bedtools package to obtain Glu-specific (19,697), GABA-specific (16,297) and OLIG-specific (26,975) peaks (Supplementary Table 8).

Annotation of CRDs. CRDs were annotated to (1) fetal/adult category as fetal and adult[31]; (2) cell types as Glu, GABA and OLIG from Supplementary Table 8 (refs. [14,76]); (3) active compartment as A and outside A compartment; and (4) inactive compartment as B and outside B compartment using the PFC NeuN+ HiC data[19]. Every CRD was assigned to a specific category, if the fraction of peaks coverage in a CRD in a given assay matches the testing dataset (using the data resources as explained above) and is significantly different from the fraction of peaks coverage in all other CRDs as a background dataset using Fisher's exact test at $P < 0.05$.

For fetal/adult annotation (1), we ran the annotation test to assess whether a given CRD in an assay is enriched for H3K27ac fetal-specific peaks versus all other CRDs. From this test, all CRDs with OR > 1 and OR < 1 are annotated as fetal and adult, respectively, whereas all non-significant (NS) CRDs are annotated as NS fetal.

For cell type annotation (2) in Study-1 from the neuronal assay, we ran the annotation test for Glu and GABA only. All NS CRDs from this test are annotated as NS GABA/GLU. In Study-2 H3K27ac tissue, we first assessed whether a given CRD has enrichment in OLIG or GABA/Glu cell types. CRDs that were significantly enriched for OLIGs are annotated as OLIG and not significant as NS OLIG. Next, we took CRDs that show GABA/Glu cell types and assessed their enrichment in GABA versus Glu cell type as explained above. Overall, we obtained five categories here: (1) GABA, (2) Glu, (3) NS GABA/Glu, (4) OLIG and (5) NS OLIG.

For chromosomal environment annotation (3), a fraction of full CRD coverage was used for the annotation test instead of using coverage of peaks within the CRD, and all non-significant CRDs from this test are annotated as NS A/B. Supplementary Fig. 18 shows the final counts of CRDs annotated to each category of cell type, development and compartments, and Supplementary Table 6 lists the annotation of CRDs from three disease groups.

Disease-specific CRDs. In this section, we show how differential analysis of CRDs was done and how we identified the relationship between structure and activity of CRDs.

Differential CRD analysis. We applied a two-stage testing procedure using the stageR package[77] that identifies significant CRDs using aggregated CRD-level P values in the screening stage, and, in the confirmation stage, individual hypotheses are assessed to determine dysregulated peaks for CRDs that pass the screening stage. Hence, it has the advantage of improving the resolution in stage 2 by providing dysregulated peaks in dysregulated CRDs from stage 1.

For stage 1, we used the peak differential analysis table as input (Supplementary Table 3) and aggregated P values using equation 1. Here, we show the calculation of P value and $\log_2$FC for one CRD (CRD_x) that is linked to k peaks.

$$CRD_x = [Peak_1, Peak_2, Peak_3, ..., Peak_k], \text{ where } CRD_x \text{ has } k \text{ histone peaks}$$

$$M = min\{Pvalue_{Peak=i}; i = 1, 2, 3, ..., k\}$$

$$PValue_{CRD_x} = 1 - \left(1 - M^k\right) \quad (eq. 1)$$

$$\log_2 FC_{CRD_x} = \frac{1}{k}\sum_{i=1}^{k} \log_2 FC(Peak_i)$$

Supplementary Table 6 is the final differential analysis table of SCZ H3K27ac NeuN+, SCZ and BD H3K27ac Tissue.

Model fitting and hypothesis testing. To assess the link between the disease-associated peaks from 'disease differential peak analysis' and CRDs, we fit the logistics regression to predict the status of peak inside or outside CRD using t statistics from the H3K27ac 'Disease differential analysis' section. We ran this regression in R using the equations below:

$$glm\left(Peaks \text{ } inside \text{ } CRD \sim Peaks \text{ } Tstat^2, data = df, family = binomial\right)$$

$$Peaks \text{ } inside \text{ } CRD = 0/1 \text{ } \{peaks \text{ } outside/inside \text{ } CRD\}$$

$$Peaks \text{ } T \text{ } stat^2 = square \text{ } of \text{ } T \text{ } statistics \text{ } from \text{ } differential \text{ } table \text{ } of \text{ } H3K27ac \text{ } peaks$$

$$df = cbind\left(Peaks \text{ } Inside \text{ } CRD, Peaks \text{ } T \text{ } stat^2\right) \{data \text{ } frame\}$$

To test if differential peaks tend to be clustered inside dysregulated CRDs, we applied Poisson regression in which the predictor variable is the number of differential peaks from 'Disease differential analysis' inside dysregulated CRDs

from 'Differential CRD analysis' accounting for the number of peaks inside CRDs as an offset in the equation.

$$glm\left(N_{\Delta peaks} \sim offset\left(log\left(N_{peaks \text{ } in \text{ } CRD}\right)\right) + N_{\Delta CRD}, data = df, family\right.$$
$$= poisson\left(link = log\right))$$

$$N_{\Delta peaks} = counts \text{ } of \text{ } differential \text{ } peaks \text{ } in \text{ } every \text{ } CRD \text{ } \{m \text{ } x \text{ } 1\}$$

$$N_{peaks \text{ } in \text{ } CRD} = number \text{ } of \text{ } peaks \text{ } inside \text{ } CRD \text{ } \{m \text{ } x \text{ } 1\}$$

$$N_{\Delta CRD} = 0/1 \text{ } for \text{ } non - dysregulated / dysregulated \text{ } CRD \text{ } \{m \text{ } x \text{ } 1\}$$

$$df = cbind\left(N_{\Delta peaks}, N_{peaks \text{ } in \text{ } CRD}, N_{\Delta CRD}\right) \{data \text{ } frame\}$$

Next, we tested if differential genes tend to be clustered inside dysregulated CRDs to affect the gene regulation. To do this, we first annotated CRDs to genes by taking the genes that reside inside CRDs. After that, we applied Poisson regression in which the predictor variable is the number of differential genes using the differential analysis table of the CommonMind Consortium RNA sequencing cohort. The R functions used are explained below.

$$glm\left(N_{\Delta genes} \sim offset\left(log\left(N_{genes \text{ } in \text{ } CRD}\right)\right) + N_{\Delta CRD}, data = df, family\right.$$
$$= poisson\left(link = log\right))$$

$$N_{\Delta genes} = counts \text{ } of \text{ } differential \text{ } genes \text{ } in \text{ } every \text{ } CRD \text{ } \{m \text{ } x \text{ } 1\}$$

$$N_{genes \text{ } in \text{ } CRD} = No \text{ } of \text{ } genes \text{ } in \text{ } every \text{ } CRD \text{ } \{m \text{ } x \text{ } 1\}$$

$$N_{\Delta CRD} = 0/1 \text{ } for \text{ } non\text{-}dysregulated / dysregulated \text{ } CRD \text{ } \{m \text{ } x \text{ } 1\}$$

$$df = cbind\left(N_{\Delta genes}, N_{genes \text{ } in \text{ } CRD}, N_{\Delta CRD}\right) \{data \text{ } frame\}$$

CRD contact matrix. Next, we quantified the expression of CRD for m $_{CRDs}$ and n $_{samples}$ as CRD contact matrix by taking the mean of peaks that are within the CRD per sample as shown in the equations below.

$$CRD \text{ } expression \text{ } matrix = \sum_{i=1}^{k} CPM_{peaks}$$

$$CRD \text{ } contact \text{ } matrix = corr\left(CRD \text{ } expression \text{ } matrix \text{ } X \text{ } CRD \text{ } expression \text{ } matrix\right)$$

We applied k-means clustering[78] on the disease-sensitive CRD contact matrix and evaluated the optimal number of clusters using the equations below.

$$BIC = D + log(n) * m * k$$

$$D = sum \text{ } of \text{ } squares \text{ } within \text{ } clusters; n = number \text{ } of \text{ } CRDs, m = number \text{ } of \text{ } CRDs$$

$$k = number \text{ } of \text{ } clusters \text{ } in \text{ } kmeans \text{ } function$$

Supplementary Fig. 17 shows the BIC value from k-means clustering as a function of k(1:10) for CRD contact matrix SCZ-sensitive H3K27ac NeuN+, H3K27ac Tissue and BD-sensitive H3K27ac Tissue.

LD score enrichment analysis of CRDs. To estimate the enrichment of brain-related and non-brain-related GWAS in all identified CRD and disease-sensitive CRD, we tested the genomic regions of histone peaks within the CRDs from H3K27ac NeuN+, Meta NeuN+ and Tissue and applied LD score partitioned heritability (version 1.0.0)[17] as explained in the LD score enrichment analysis of peaks section. Figure 5 and Supplementary Fig. 21 show the LD score enrichments of SCZ- and BD-sensitive CRDs from H3K27ac NeuN+ and H3K27ac Tissue.

Modeling chromatin conformation in 3D. Hi-C data from PFC NeuN+ were used to infer chromatin conformation structure in 3D. We used bulk Hi-C data from PFC NeuN+ from four adults. Primary processing was performed with the HiC-Pro pipeline[79] at 50-kb and 1-Mb resolution. To improve the accuracy of 3D modeling, we combined data from different donors for the PFC NeuN+ to increase sequencing depths. Contact matrices produced by HiC-Pro were converted to cooler format using HiCExplorer[80], balanced using the cooler suite of tools[81], excluding the ENCODE version 3 blacklisted regions (https://www.encodeproject.

org/files/ENCFF356LFX/) from balancing with the 'cooler balance —blacklist' parameter. TADs were called at 50 kb using the 'diamond-insulation' algorithm implemented in the cooltools suite (https://cooltools.readthedocs.io/en/latest/). Hi-C contact matrices and TAD calls were pre-processed to 'gtrack' files as input to Chrom3D, as previously described[32,33]. For more details on HiC data generation and processing, see the Methods section on HiC[19]. We restricted our analysis to diploid autosomal interactions, 50 kb for intra-chromosomal and 1 Mb for inter-chromosomal. Chrom3D was run with a nucleus radius of 5.0 for 2×10^6 iterations, '—radius 5.0 —iterations 2000000'. *xyz* coordinates were parsed from the output 'cmm' files.

We took the coordinates of SCZ-sensitive H3K27ac NeuN$^+$ and Tissue CRDs and BD-sensitive H3K27ac Tissue CRDs and overlapped with the PFC NeuN$^+$ TAD coordinates obtained from above. To test the presence of localized SCZ- or BD- sensitive CRDs in the 3D genome, we measured the pairwise 3D distance of the TADs that overlapped with diseased CRDs stratified by clusters as shown in Fig. 6.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
Raw data (FASTQ files) and processed data (BigWig files, metadata, peaks and raw and normalized count matrices) have been deposited in Synapse under synID syn25705564. Browsable UCSC genome browser tracks of our processed ChIP-seq data are available as a resource at EpiDiff Phase 2 http://genome.ucsc.edu/s/girdhk01/EpiDiff_Phase2.
External validation sets used in the study are as follows: H3K27ac ChIP-seq fetal-specific peaks: spatio-temporal enrichment of H3K27ac peaks table from http://development.psychencode.org/#; RoadMap Epigenome Project H3K27ac, H3K4me3 tissue ChIP-seq peaks, chromHMM states on E073 and fetal male E081 and fetal female E082 (https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/coreMarks/jointModel/final/); and CTCF ChIP-seq on human neural cell (Gene Expression Omnibus GSE127577). TruSeq3-PE.fa file was downloaded from the adaptor folder under the Trimmomatic repository: https://github.com/timflutre/trimmomatic/blob/master/adapters/TruSeq3-PE.fa.
The source data described in this manuscript are available via the PsychENCODE Knowledge Portal (https://psychencode.synapse.org/). The PsychENCODE Knowledge Portal is a platform for accessing data, analyses and tools generated through grants funded by the National Institute of Mental Health PsychENCODE program. Data are available for general research use according to the following requirements for data access and data attribution: https://psychencode.synapse.org/DataAccess.

Code availability
All publicly available software used is noted in the Methods. We used decorate software to call CRDs: https://github.com/GabrielHoffman/decorate.

References
54. Wang, D. et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018).
55. Kundakovic, M. et al. Practical guidelines for high-resolution epigenomic profiling of nucleosomal histones in postmortem human brain tissue. *Biol. Psychiatry* **81**, 162–170 (2017).
56. Jiang, Y., Matevossian, A., Huang, H.-S., Straubhaar, J. & Akbarian, S. Isolation of neuronal chromatin from brain tissue. *BMC Neurosci.* **9**, 42 (2008).
57. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
58. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
59. Landt, S. G. et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res.* **22**, 1813–1831 (2012).
60. Fort, A. et al. MBV: a method to solve sample mislabeling and detect technical bias in large combined genotype and sequencing assay datasets. *Bioinformatics* **33**, 1895–1897 (2017).
61. Zhang, Y. et al. Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
62. Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019).
63. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
64. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
65. Hunt, G. J., Freytag, S., Bahlo, M. & Gagnon-Bartsch, J. A. dtangle: accurate and robust cell type deconvolution. *Bioinformatics* **35**, 2093–2099 (2019).
66. Neath, A. A. & Cavanaugh, J. E. The Bayesian information criterion: background, derivation, and applications. *WIREs Computational Statistics* https://doi.org/10.1002/wics.199 (2011).
67. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
68. Ernst, J. & Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **12**, 2478–2492 (2017).
69. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
70. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48 (2010).
71. McLean, C. Y. et al. GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
72. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
73. Stegle, O., Parts, L., Piipari, M., Winn, J. & Durbin, R. Using probabilistic estimation of expression residuals (PEER) to obtain increased power and interpretability of gene expression analyses. *Nat. Protoc.* **7**, 500–507 (2012).
74. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
75. Kumar, V. et al. Uniform, optimal signal processing of mapped deep-sequencing data. *Nat. Biotechnol.* **31**, 615–622 (2013).
76. Kozlenkov, A. et al. Substantial DNA methylation differences between two major neuronal subtypes in human brain. *Nucleic Acids Res.* **44**, 2593–2612 (2016).
77. Van den Berge, K., Soneson, C., Robinson, M. D. & Clement, L. stageR: a general stage-wise method for controlling the gene-level false discovery rate in differential expression and differential transcript usage. *Genome Biol.* **18**, 151 (2017).
78. Forgy, E. Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics* **21**, 768–780 (1965).
79. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).
80. Ramírez, F. et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).
81. Abdennur, N. & Mirny, L. A. Cooler: scalable storage for Hi-C data and other genomically labeled arrays. *Bioinformatics* **36**, 311–316 (2020).

Author contributions
Wet lab work, including tissue processing, sorting of nuclei and ChIP-seq and Hi-C library generation: Y.J., L.B., M.K., E.Z., R.J., J.R.W., R.P., B.S.K., L.C., O.D., S.R., J.F., E.F. and A.K. Data processing and coordination: Y.J., M.K., M.A.P. and J.S.J. Bioinformatics and computational genomics: K.G., G.E.H., J.B., S.R., T.G., J.P.-C., P.D., W.L., M.E.H., L.S. and L.C. Provision of brain tissue and resources: C.A.T., S.M., B.K.L., D.A.L., V.H., C.-G.H., R.E.G., S.D. and P.C. Conception of study and design: P.R., S.A., K.G., G.E.H. and J.B. Writing of the paper: K.G., S.A. and P.R.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Kiran Girdhar, Panos Roussos or Schahram Akbarian.

nature portfolio

Corresponding author(s): Girdhar, Akbarian, Roussos

Last updated by author(s): Feb 9, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BD FACSdivca 8.0.1 , ImageJ (version 1.51).
Data analysis	Tools and R packages used: BEDTools 2.25.0, bedGraphToBigwig 4.0.0, ChIPSeeker 1.18.0, dTangle v2.0.9, edgeR 3.22.3, fastqc 0.11.2, GATK 3.5.0, GenomicFeatures 1.14.8, HiC-Pro 3.0.0, Juicer / HiCCUPS 1.6, LDSC 1.0.1, limma v4.1, MACS 2.1, PEER 1.3, PICARD 2.2.4, phantompeakqualtools 2.0, plink 1.90, QTLtools 1.3.1, R 3.5.3, RSubread 1.22.0, samtools 0.1.19, bwa -0.7.8-r455, Trimmomatic 0.36, variancePartition 1.17.9 All software are freely available.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Raw (FASTQ files) and processed data (peaks, differential analysis and CRD coordinates) have been deposited in synapse under synID syn25705564 (https://www.synapse.org/#!Synapse:syn25796149)

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. Sequencing was done aon all available brains (N=739) from the Common Minds Consortium
Data exclusions	We excluded overall ~5% (N=37 out of 739) of the samples based on ENCODE quality controls metrics and sample outliers test using PCA plot.
Replication	We profiled H3K27ac from bulk tissue homogenate generated from N=249 (133 Control, 68 SCZ, 48 BD) samples as a replication dataset. All the findings were tested using statistical methods and were adjusted for multiple testing, either using Bonferroni or FDR.
Randomization	To avoid batch effects and other confounds, samples underwent repeated rounds of randomization, including (i) chromatin immunoprecipitation procedures and (ii) library preparation.
Blinding	Blinding was not relevant to this study, analysts were aware of data generation, processing and donor metadata.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
☐	☒ Antibodies	☐	☒ ChIP-seq
☒	☐ Eukaryotic cell lines	☐	☒ Flow cytometry
☒	☐ Palaeontology and archaeology	☒	☐ MRI-based neuroimaging
☒	☐ Animals and other organisms		
☒	☐ Human research participants		
☒	☐ Clinical data		
☒	☐ Dual use research of concern		

Antibodies

Antibodies used	antihistone antibodies: anti-H3K4me3 (Cat# 9751BC, lot 7; Cell Signaling, Danvers, MA) and anti-H3K27ac (Cat# 39133, Lot# 31814008; Active Motif, Carlsbad, CA). See detailed description of the protocol (PMID:27113501) that includes details on dot blot with synthetic peptides and other QC.
Validation	Specificity was checked with histone H3 peptide array (Cat# 16-667; Millipore). Anti-NeuN-Alexa488 (Cat# MAB377X, EMD Millipore) antibody to stain nuclei, specificity was checked by microscopy. Quality controls for FACS-sorted nuclei were described in Kundakovic, M. et al. Practical Guidelines for High-Resolution Epigenomic Profiling of Nucleosomal Histones in Postmortem Human Brain Tissue. Biol. Psychiatry 81, 162–170 (2017).

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	Raw (FASTQ files) and processed data (peaks, differential analysis and CRD coordinates) have been deposited in synapse under synID syn25705564 (https://www.synapse.org/#!Synapse:syn25705564)
Files in database submission	Raw data(FASTQ) and Processed data(H3K4me3 NeuN+ peaks, H3K27ac NeuN+ peaks, H3K27ac-Tissue peaks BED files)
Genome browser session (e.g. UCSC)	Browsable UCSC genome browser tracks of our processed ChIP-seq data are available as a resource http://genome.ucsc.edu/s/girdhk01/EpiDiff_Phase2

Methodology

Replicates	We generate H3K27ac=249 from bulk tissue.
Sequencing depth	median of sequencing depth is 59 million mapped paired-end reads
Antibodies	antihistone antibodies: anti-H3K27ac (Cat# 39133, Lot# 31814008; Active Motif, Carlsbad, CA)
Peak calling parameters	narrow: macs2 Poisson p-value = 0.01 with --keep-dup all --nomodel --extsize = 150, Broad: macs2 P-value cutoff = .01, --extsize = 150.
Data quality	We used encode quality metrics like RSC(>1.0), NSC (>1.0) and NRF(>)0.8 to ensure data quality of samples.
Software	Alignment of FASTQ files:Burrows-Wheeler Aligner Filtering of bam files: Picard, samtools QC parameters: SPP, phantompeakcall Peak calling and quantifying chip-seq counts in peaks: macs2, bedtools, featureCounts Covariate Analysis: variancePartition +BIC Peak annotation: CHIPseeker Differential modification analysis: R (libraries: edgeR, diffbind) CRD calling: Decorate genetics enrichment analysis: LDSc heritability

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Nuclei extraction from approximately 300mg aliquots of frozen prefrontal cortex tissue, immunotagging with NeuN antibody for subsequent fluorescence-activated nuclei sorting, chromatin digestion of sorted nuclei with micrococcal nuclease and subsequent pulldown with anti-H3K27ac and anti-H3K4me3 antibodies, library preparation and sequencing.
Instrument	FACSAria [BD Biosciences]
Software	BD FACSDiva Software
Cell population abundance	We report in the Methods section that for each cell type, a minimum of 400,000 nuclei was required as starting material.
Gating strategy	Separation of NeuN+ and NeuN- nuclei in FITC-A and Texas-Red (from DAPI-A), with windows picked to maximize specificity.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.