

Identification of a molecular network regulated by multiple ASD high risk genes

Lei Wan, Guojun Yang, Zhen Yan*

Department of Physiology and Biophysics, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, 955 Main Street, Buffalo, NY 14203, United States

*Corresponding author. Department of Physiology and Biophysics, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, 955 Main Street, Buffalo, NY 14203, United States. E-mail: zhenyan@buffalo.edu

Abstract

Genetic sequencing has identified high-confidence ASD risk genes with loss-of-function mutations. How the haploinsufficiency of distinct ASD risk genes causes ASD remains to be elucidated. In this study, we examined the role of four top-ranking ASD risk genes, *ADNP*, *KDM6B*, *CHD2*, and *MED13*, in gene expression regulation. ChIP-seq analysis reveals that gene targets with the binding of these ASD risk genes at promoters are enriched in RNA processing and DNA repair. Many of these targets are found in ASD gene database (SFARI), and are involved in transcription regulation and chromatin remodeling. Common gene targets of these ASD risk genes include a network of high confidence ASD genes associated with gene expression regulation, such as *CTNNB1* and *SMARCA4*. We further directly examined the transcriptional impact of the deficiency of these ASD risk genes. Our mRNA profiling with qPCR assays in cells with the knockdown of *Adnp*, *Kdm6b*, *Chd2* or *Med13* has revealed an intricate pattern of their cross-regulation, as well as their influence on the expression of other ASD genes. In addition, some synaptic genes, such as *Snap25* and *Nrxn1*, are strongly regulated by deficiency of the four ASD risk genes, which could be through the direct binding at promoters or indirectly through the targets like *Ctnnb1* or *Smarca4*. The identification of convergent and divergent gene targets that are regulated by multiple ASD risk genes will help to understand the molecular mechanisms underlying common and unique phenotypes associated with haploinsufficiency of ASD-associated genes.

Keywords: autism; chromatin remodellers; transcription regulators; autism risk genes

Introduction

Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental disorder with the manifestation of core symptoms, such as social deficits and repetitive behaviors. Genetic sequencing has identified ~100 high-confidence ASD risk genes with loss-of-function mutations [1, 2]. Most of these ASD genes fall into two major functional groups: (1) gene expression regulation (GER), including chromatin regulators and transcription factors, (2) neuronal communication (NC), including synaptic function [3]. While genetic studies have pointed to potential molecular mechanisms underlying ASD in general, we still do not know how the haploinsufficiency of distinct ASD risk genes leads to a set of similar clinical outcomes. Many key questions remain to be investigated. For example, do ASD risk genes in GER group regulate each other? Do ASD risk genes in GER group have common or unique targets? Are there connections between the two groups (GER and NC) of ASD risk genes? In this study, we attempted to address these questions by focusing on four top-ranking ASD risk genes involved in chromatin and transcription regulation, *ADNP*, *KDM6B*, *CHD2* and *MED13* [4].

ADNP, which encodes activity-dependent neuroprotective protein (ADNP), contains 9 N-terminal zinc-fingers and a C-terminal homeobox domain, suggesting transcription factor activity. It is involved in gene silencing through interacting with heterochromatin protein 1 (HP1) for H3K9me3 modification, establishing inaccessible chromatin around its DNA-binding sites, and

counteracting chromatin looping at CTCF sites [5–7]. *ADNP* is also involved in the regulation of microtubules, autophagy, neural differentiation, synaptic function and cognitive behaviors [2, 8–11].

KDM6B encodes a histone demethylase that specifically demethylates the repressive epigenetic marks, di- or tri-methylated lysine 27 of histone H3 (H3K27me2/3). Through its demethylation activity, *KDM6B* influences Hox gene expression and animal body patterning, cellular differentiation and development, and inflammatory responses [12–14].

CHD2 encodes a member of the Chromodomain-helicase-DNA-binding protein (CHD) family proteins characterized by the presence of chromo (chromatin organization modifier) domains and SNF2-related helicase/ATPase domains. CHD family proteins alter gene expression by binding to gene promoters, controlling access to DNA for the transcriptional apparatus via chromatin remodeling [15–17]. *Chd2* haploinsufficiency leads to broad transcriptional changes, aberrant cortical network function and impaired long-term memory [18].

MED13 encodes a component of the mediator complex, a transcriptional coactivator complex required for gene expression. Recruited by various transcriptional activators and nuclear receptors, the mediator complex plays a key role in facilitating initial steps of gene expression by interacting with RNA polymerase II and promoting the formation of transcriptional pre-initiation complex [19, 20]. The biological role of *MED13* in neuronal development and brain function is largely unknown.

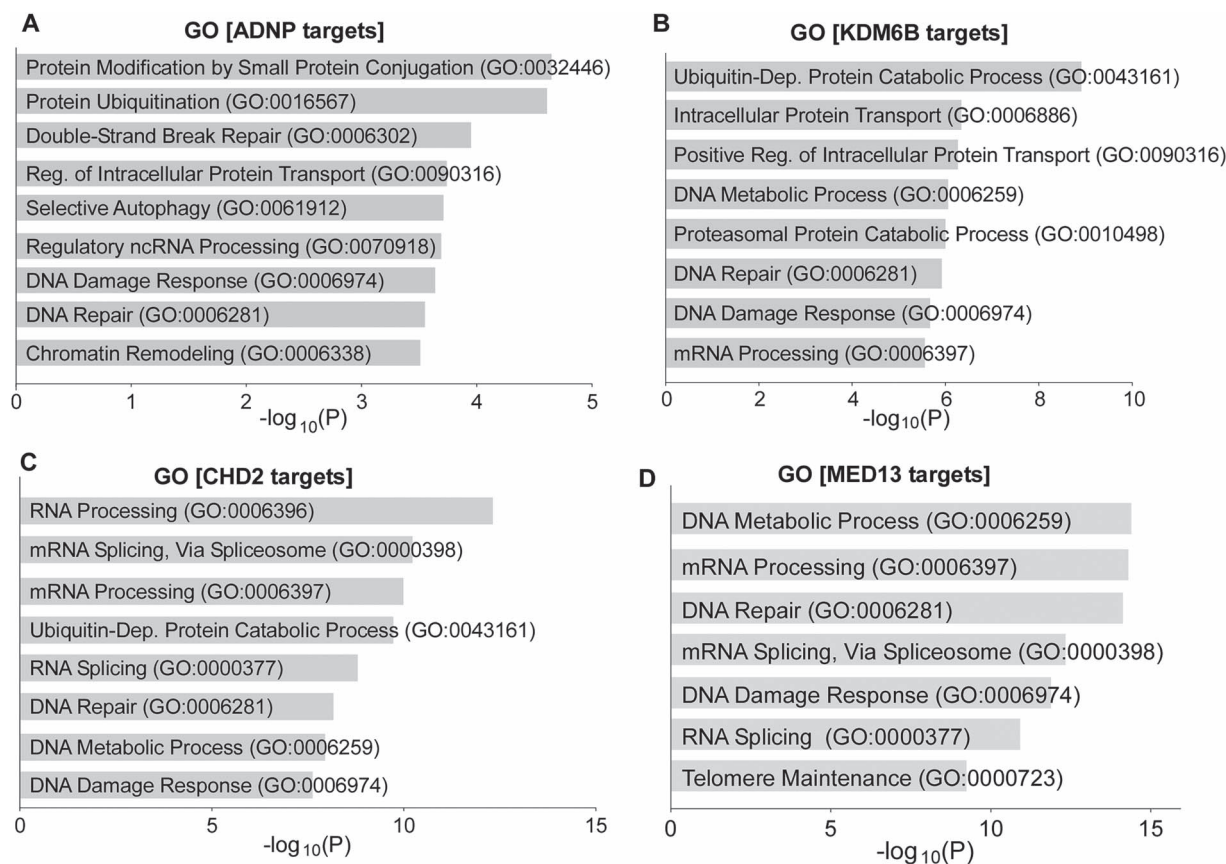


Figure 1. Enriched biological pathways of gene targets of four ASD risk genes. (A–D), Bar graphs showing the gene ontology (GO) pathways of gene targets of ADNP (A), KDM6B (B), CHD2 (C) and MED13 (D) identified from their binding at gene promoters in ChIPseq data.

To better understand the role of ADNP, KDM6B, CHD2, and MED13 in ASD, we performed ChIP-seq analysis of their gene targets to reveal the enriched pathways and networks. Based on the bioinformatic data, we further examined the transcriptional impact of knockdown of each of the four ASD risk genes. Results from this work provide a framework of the convergent and divergent targets that are directly or indirectly influenced by multiple ASD risk genes, which will help to understand the molecular mechanisms underlying common and unique phenotypes associated with haploinsufficiency of these ASD genes.

Results

ChIP-seq analysis reveals convergent gene targets of four ASD risk genes

To find out targets of the four ASD risk genes, we examined ChIP-seq data showing the occupancy of these transcriptional regulators at the promoter regions of genes across the whole genome. All the 2534 genes with ADNP binding at 5 Kb of TSS (transcription start site) and the top 2000 genes with KDM6B, CHD2, or MED13 binding at 1 Kb of TSS were selected for analyses (Supplementary Table 2). Gene Ontology (GO) enrichment assay (Supplementary Table 3) indicated that gene targets of ADNP (Fig. 1A) were enriched in protein ubiquitination, DNA repair, protein transport, and chromatin remodeling. Gene targets of KDM6B (Fig. 1B) were enriched in protein ubiquitination, protein transport, DNA repair and mRNA processing. Gene targets of CHD2 (Fig. 1C) were enriched in RNA processing and splicing, protein ubiquitination, and DNA repair. Gene targets of MED13 (Fig. 1D) were enriched in RNA processing and splicing, and DNA repair.

These data suggest that the four ASD risk genes are commonly targeting genes involved in RNA/DNA regulation.

Next, we compared gene targets of the four ASD risk genes with ASD gene database (SFARI) that contains 1128 genes in 4 Categories (Syndromic, High Confidence, Strong Candidate, Suggestive Evidence, Supplementary Tables 4 and 5). Among ADNP targets, 126 genes were found in ASD gene database (Fig. 2A), and they were enriched in transcription regulation and chromatin remodeling (Fig. 2B). Among KDM6B targets, 113 genes were found in ASD gene database (Fig. 2C), and they were enriched in transcription regulation and chemical synaptic transmission (Fig. 2D). Among CHD2 targets, 116 genes were found in ASD gene database (Fig. 2E), and they were enriched in transcription regulation, chromatin remodeling, and DNA damage response (Fig. 2F). Among MED13 targets, 98 genes were found in ASD gene database (Fig. 2G), and they were enriched in transcription regulation (Fig. 2H). These data suggest that the four ASD risk genes regulate many other ASD genes primarily in GER group.

We further constructed protein–protein interaction (PPI) networks of gene targets of the four ASD risk genes that belong to the 232 High Confidence (HC) ASD genes (SFARI). Among the 37 ADNP targets of ASD genes (Fig. 3A), CTNNB1, which encodes β -catenin (a dual function protein involved in cell–cell adhesion and gene transcription), was a hub gene with direct or indirect connections to histone modifiers (KMT2A, KDM3B, SETD5), chromatin remodelers (CHD3, CHD2, POGZ) and mediator complex component (MED13L). Among the 23 KDM6B targets of ASD genes (Fig. 3B), SMARCA4, which encodes a member of the SWI/SNF family of chromatin remodeler and transcription activator BRG1, was a hub gene connecting to other epigenetic enzymes (KMT2C,

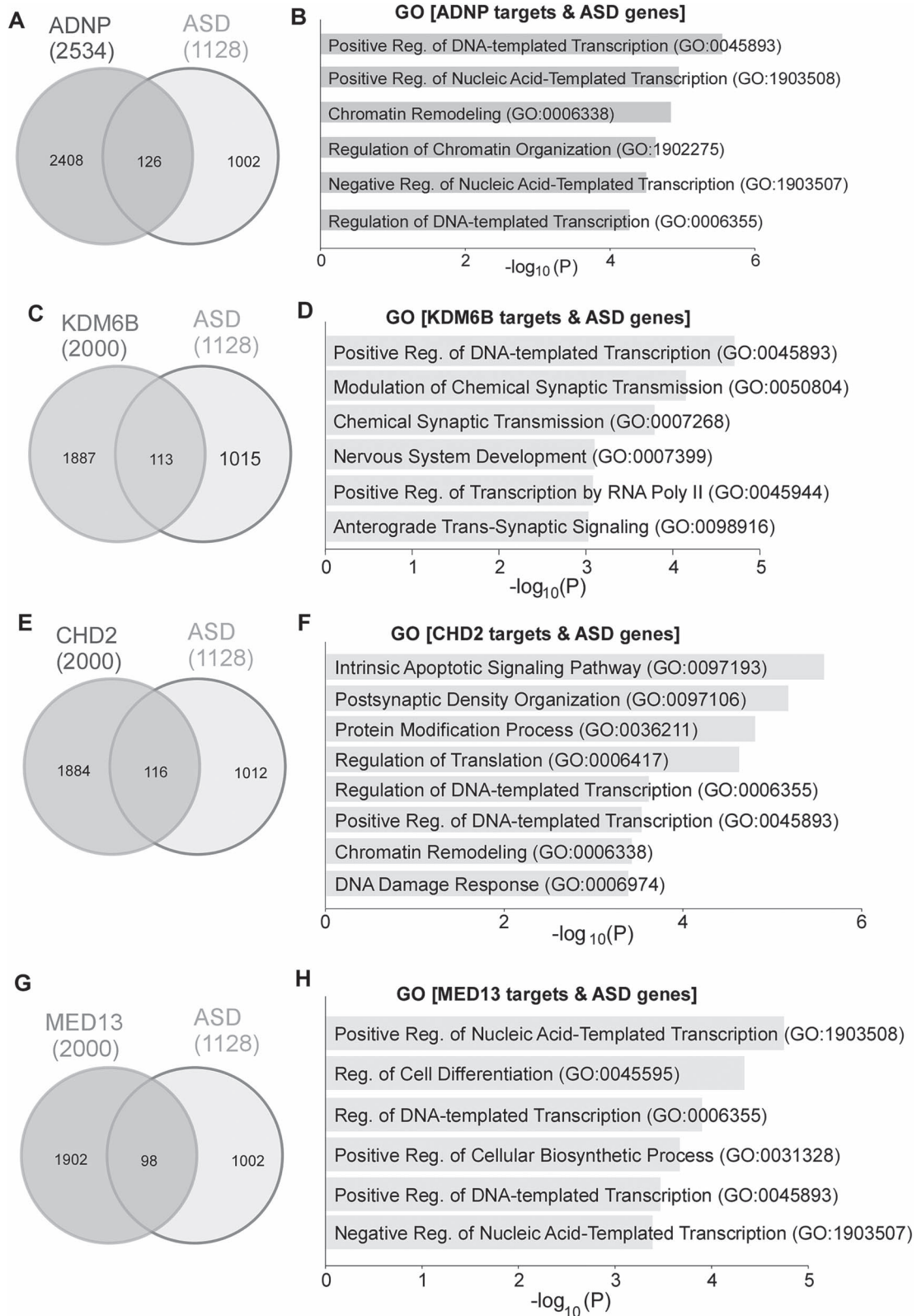


Figure 2. Overlapping of gene targets of four ASD risk genes with ASD genes. (A, C, E, and G) Venn diagrams showing the overlap of gene targets of ADNP (A), KDM6B (C), CHD2 (E) and MED13 (G) with ASD gene database (SFARI) that contains 1128 genes in all categories. (B, D, F, and H) Bar graphs showing the GO pathways of ASD-related gene targets of ADNP (B), KDM6B (D), CHD2 (F) and MED13 (H).

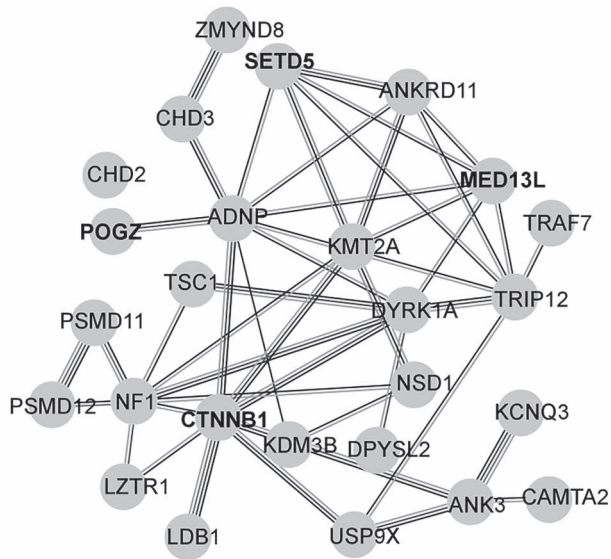
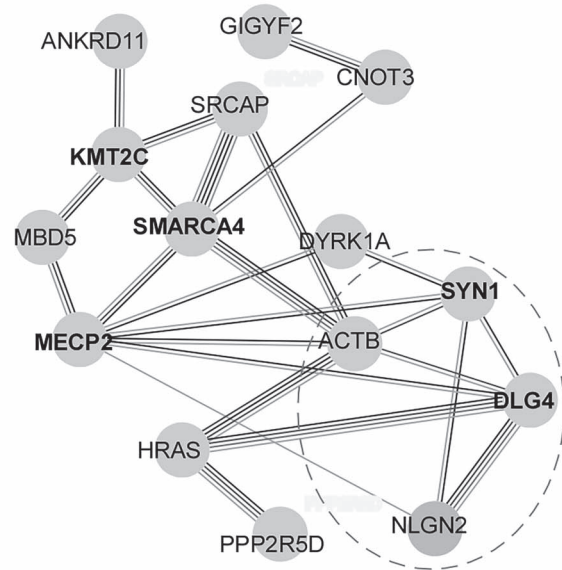
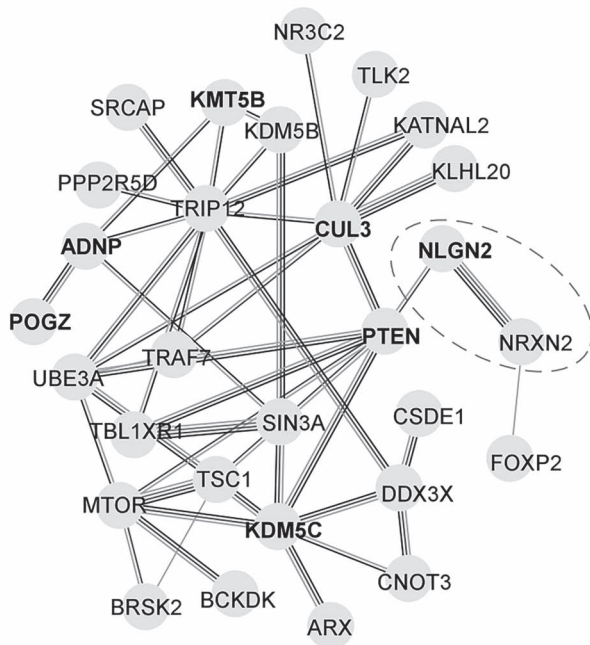
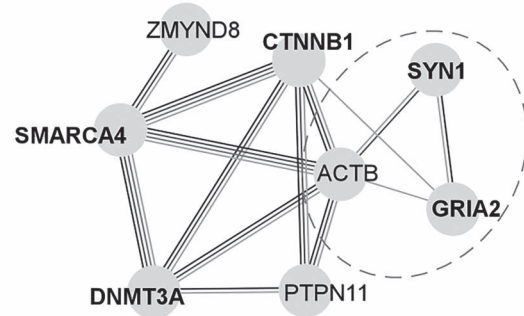
A PPI [ADNP targets & HC ASD genes]**B PPI [KDM6B targets & HC ASD genes]****C PPI [CHD2 targets & HC ASD genes]****D PPI [MED13 targets & HC ASD genes]**

Figure 3. Protein–protein interaction (PPI) network of ASD-related gene targets of four ASD risk genes. (A–D) Networks of targets of ADNP (A), KDM6B (B), CHD2 (C) and MED13 (D) overlapped with high-confidence (HC) ASD genes (SFARI). Edges represent protein–protein associations based on known or predicted interactions.

MECP2) and a set of synaptic genes, including NLGN2 (encoding Neuroligin-2, a neuronal surface protein at GABAergic inhibitory synapses), DLG4 (encoding PSD-95, a postsynaptic density protein at glutamatergic synapses), and SYN1 (encoding Synapsin I, a presynaptic vesicle protein). Among the 37 CHD2 targets of ASD genes, chromatin regulators (ADNP, POGZ, KMT5B, KDM5B, KDM5C) were predominant (Fig. 3C). Among the 17 MED13 targets of ASD genes (Fig. 3D), transcription regulators (CTNNB1, SMARCA4, DNMT3A) and synaptic proteins (SYN1, GRIA2) were

present. These data suggest that the four ASD risk genes regulate a network of ASD genes mainly involved in chromatin and transcription regulation.

To reveal the commonality of the four ASD risk genes, we examined the overlap of their gene targets that belong to ASD gene database (SFARI) (Supplementary Table 6). As shown in Fig. 4A, ADNP shared 26 ASD genes with CHD2, 24 ASD genes with KDM6B, and 16 ASD genes with MED13, while KDM6B shared 18 ASD genes with CHD2 and 19 ASD genes with MED13, and CHD2 shared 17

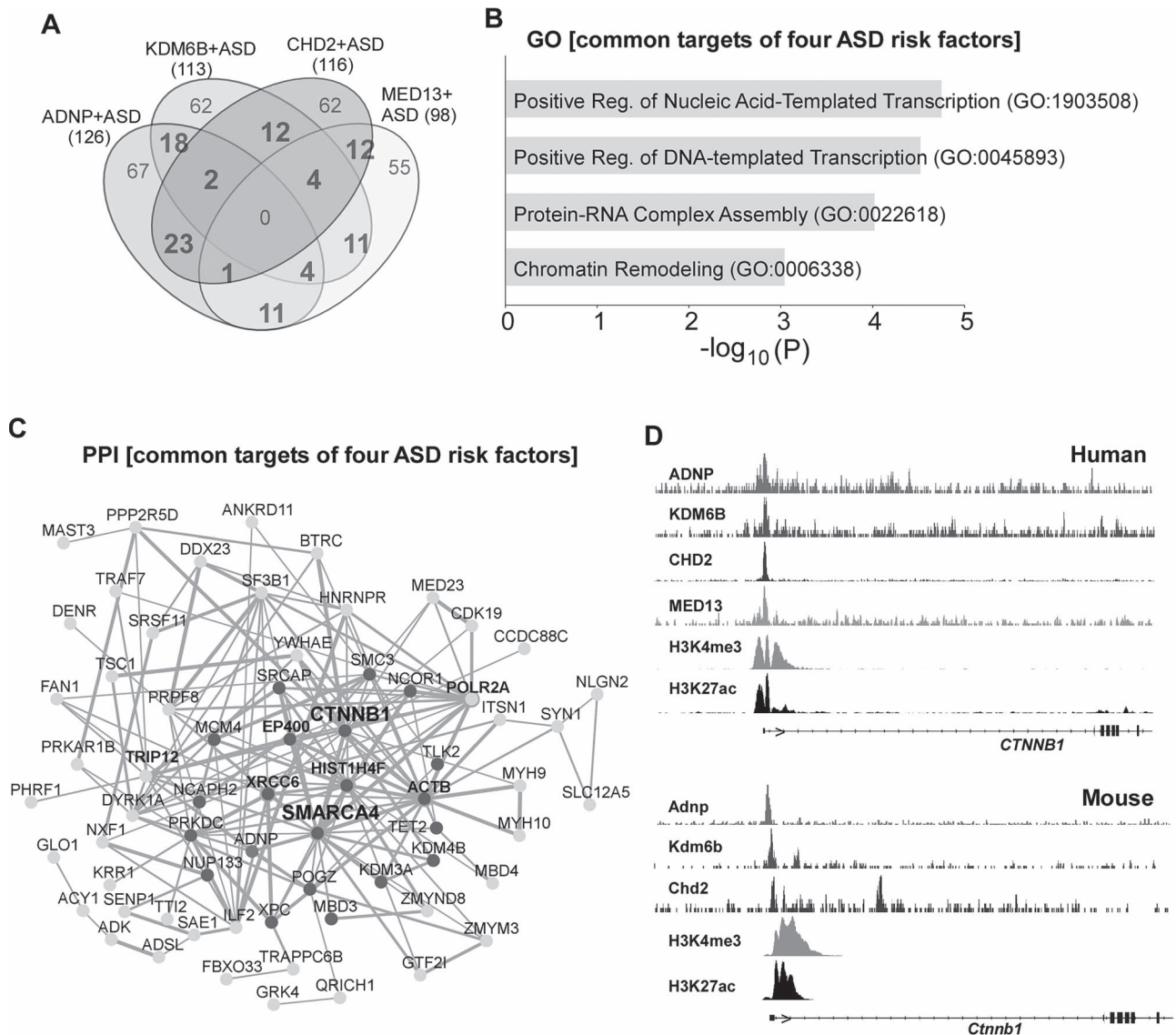


Figure 4. Common ASD-related targets of four ASD risk genes. (A) Venn diagram showing the overlap of ASD-related targets of ADNP, KDM6B, CHD2 and MED13. (B) Bar graphs of GO pathways of ASD-related targets that are common for at least two of the four ASD risk genes. (C) PPI network of the ASD-related common targets from B. Bold: hub proteins. Dark nodes: proteins involved in chromatin organization. (D) Human and mouse ChIP-seq data showing ADNP, KDM6B, CHD2 and MED13 enrichment at *Ctnnb1* promoter (marked by H3K4me3 and H3K27ac peaks).

ASD genes with MED13. GO analysis of common targets in ASD gene database that are shared by any two of the four ASD risk genes (Fig. 4B) indicated that they were enriched in transcription regulation and chromatin remodeling. PPI of these common gene targets (Fig. 4C) showed that CTNNB1 and SMARCA4 were hubs connecting to other chromatin and transcription regulators, such as ADNP, POGZ, KDM4B, KDM3A, EP400, and NCOR1. Genome browser of ChIP-seq landscape also confirmed the enrichment of ADNP, KDM6B, CHD2 and MED13 at the promoter of CTNNB1 (Fig. 4D). These data suggest that the four ASD risk genes converge on common ASD targets in GER group, such as CTNNB1 and SMARCA4, which could further regulate the transcription of more genes.

Alteration of genes by the deficiency of four ASD risk genes

To determine the transcriptional impact of haploinsufficiency of the four ASD risk genes, we designed shRNA against *Adnp*,

Kdm6b, *Chd2* and *Med13*, transfected each into N2A mouse cell line, followed by mRNA profiling with qPCR experiments. First, we examined whether these ASD risk genes regulate each other. As shown in Fig. 5A, *Adnp* knockdown induced a significant increase of *Med13* without altering *Kdm6b* and *Chd2*. *Kdm6b* knockdown reduced *Adnp* expression, but induced a significant increase of *Chd2* and a modest increase of *Med13* (Fig. 5B). *Chd2* knockdown substantially increased *Kdm6b*, and slightly reduced *Adnp* (Fig. 5C). Similarly, *Med13* knockdown significantly increased *Kdm6b* and decreased *Adnp* (Fig. 5D). These data have revealed an intricate cross-regulatory pattern of the four ASD risk genes: *Adnp* is reduced by deficiency of all the other three ASD risk genes; *Kdm6b* is elevated by deficiency of *Chd2* or *Med13*; *Chd2* is increased only by *Kdm6b* deficiency; and *Med13* is increased mainly by deficiency of *Adnp*.

Next, we examined whether these ASD risk genes regulate downstream targets identified in our ChIP-seq analysis. One of the common targets, *Ctnnb1*, was strongly increased by *Adnp*

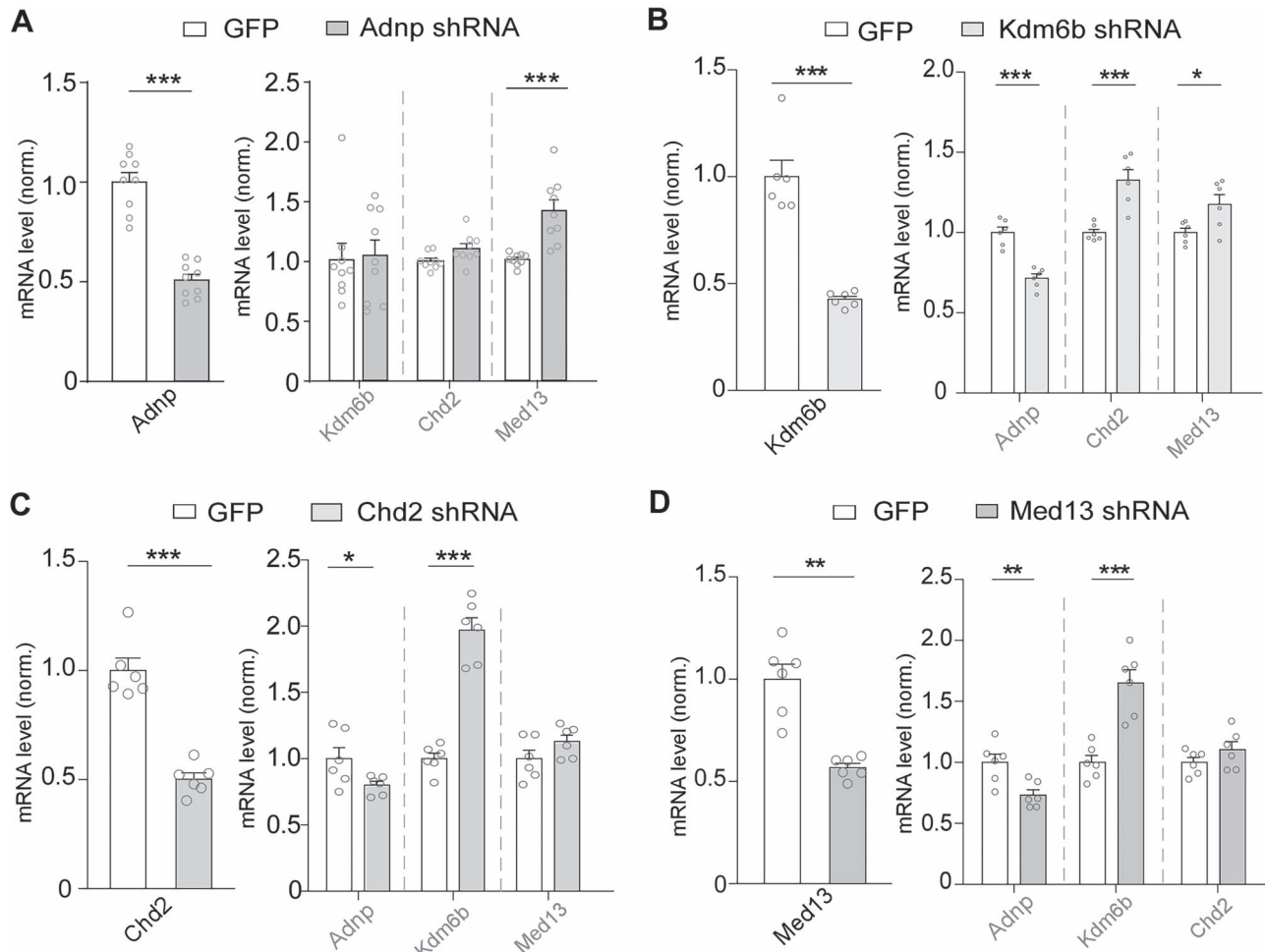


Figure 5. Cross regulation of four ASD risk genes. (A–D) Bar graphs showing the RT-qPCR data of *Adnp*, *Kdm6b*, *Chd2* and *Med13* mRNA levels in N2A cells transfected with shRNA against each of the four ASD risk genes. (A) *Adnp*, 49% knockdown, $P=0.0001$, *Med13*, 40% increase, $P=0.0004$; (B) *Kdm6b*, 57% knockdown, $P=0.0006$; *Adnp*, 38% decrease, $P=5E-05$; *Chd2*, 33% increase, $P=0.0007$; *Med13*, 18% increase, $P=0.03$; (C) *Chd2*, 50% knockdown, $P=6E-05$; *Adnp*, 20% decrease, $P=0.05$; *Kdm6b*, 97% increase, $P=3E-06$; (D) *Med13*, 43% knockdown, $P=0.002$; *Adnp*, 27% decrease, $P=0.008$; *Kdm6b*, 65% increase, $P=0.0003$, $n=6-9/\text{group}$; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, unpaired t-test.

knockdown (Fig. 6A), but was strongly decreased by *Med13* knockdown (Fig. 6D). Another common target, *Smarca4*, was modestly but significantly increased by deficiency of each of the four ASD risk genes (Fig. 6A–D). Several ASD genes in GER group identified as targets of the four ASD risk genes were also altered by their deficiency: *Adnp* knockdown significantly increased *Pogz* and *Med13l*, *Kdm6b* knockdown increased *Kmt2c*, *Chd2* knockdown decreased *Pten*, and *Med13* knockdown decreased *Dnmt3a*.

To find out whether there are connections between the two groups (GER and NC) of ASD risk genes, we also examined the impact of the four ASD risk genes on synaptic genes. While a few ASD genes in NC group were found to have the occupancy of these ASD risk genes at their promoters (Fig. 3), only *Dlg4* and *Gria2* were significantly elevated by *Kdm6b* knockdown (Fig. 6B). Since *Cttnb1* and *Smarca4* are common targets of the four ASD risk genes (Fig. 4), we speculated that synaptic genes might be affected by these transcription regulators. To test this, we examined *Snap25*, which encodes a core component of SNARE complex that is essential for transmitter release, and *Nrxn1*, which encodes presynaptic cell adhesion protein Neurexin-1 that is important for synaptic assembly. ChIP-seq landscape at *Snap25* promoter region showed no enrichment of ADNP, *Kdm6b* or *Chd2*, but prominent binding peaks of *Med13*, *Cttnb1* and *Smarca4* (Fig. 7A). Our qPCR assays

revealed that *Snap25* and *Nrxn1* were significantly decreased by knockdown of *Adnp*, but were strongly increased by knockdown of *Kdm6b*, *Chd2* or *Med13* (Fig. 7B–E). These data suggest that synaptic genes in NC group could be indirectly influenced by ASD risk genes probably via downstream targets in GER group (Fig. 7F).

Discussion

Given the predominance of chromatin and transcription regulators in ASD risk genes [1, 3, 4], a main pathogenic mechanism for the disease is likely related to gene dysregulation [21–24]. In this study, we have revealed the gene targets of four top-ranking ASD risk genes that represent different categories, ADNP (a transcription and chromatin regulator), KDM6B (a histone modifier), CHD2 (a chromatin remodeler), and MED13 (a mediator complex subunit). ADNP has been linked to gene silencing through multiple chromatin modifying mechanisms [5–7]. KDM6B removes repressive H3K27me3, but also has a H3K27 demethylase-independent role in chromatin remodeling in terminally differentiated cells: it acts as a link between T-box factors and the SMARCA4-containing SWI/SNF remodeling complex to regulate T-box family member-dependent gene expression [25]. *Chd2* deficiency leads to

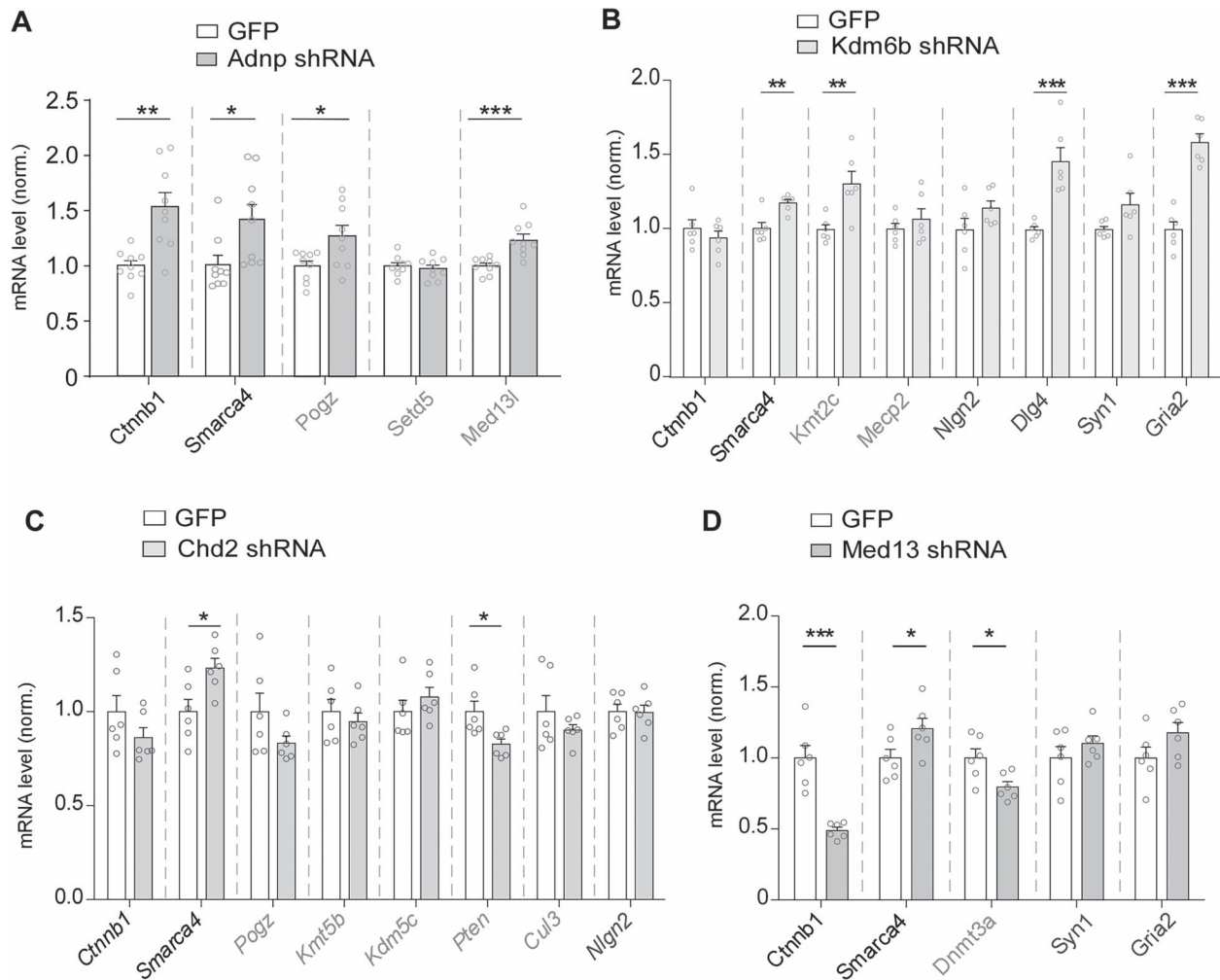


Figure 6. Regulation of gene targets by deficiency of four ASD risk genes. (A–D) Bar graphs showing the RT-qPCR data of key gene targets in N2A cells transfected with shRNA against each of the four ASD risk genes. (A) Adnp KD: *Ctnnb1*, 53% increase, $P = 0.001$, *Smarca4*, 41% increase, $P = 0.02$, *Pogz*, 27% increase, $P = 0.02$, *Med13l*, 24% increase, $P = 0.0008$; (B) Kdm6b KD: *Smarca4*, 17% increase, $P = 0.004$, *Kmt2c*, 31% increase, $P = 0.007$, *Dlg4*, 46% increase, $P = 0.0007$, *Gria2*, 59% increase, $P = 2E-05$; (C) Chd2 KD: *Smarca4*, 23% increase, $P = 0.02$, *Pten*, 17% decrease, $P = 0.02$; (D) Med13 KD: *Ctnnb1*, 51% decrease, $P = 0.0002$, *Smarca4*, 21% increase, $P = 0.05$, *Dnmt3a*, 20% decrease, $P = 0.02$, $n = 6-9/\text{group}$; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, unpaired t-test.

differential expression of transcription factors, epigenome regulators, synaptic molecules and Wnt pathway genes, which may contribute to the deficits in neural circuit development and long-term memory [18]. MED13 is a subunit in the mediator complex required for gene activation, but also forms a sub-complex with MED12, CDK8 and cyclin C, which can exert negative control over the transcriptional activators and switch to repress gene expression [26, 27]. Thus, the four ASD risk genes could control gene networks in a complex manner.

From CHIP-seq data, we have found that each of the four ASD risk genes binds to the promoter region of thousands of genes, and these targets are commonly enriched in RNA processing and DNA repair. Many of the target genes of the four ASD risk genes are transcription regulators found in ASD gene database (SFARI), suggesting that the four ASD risk genes can directly influence convergent networks of other ASD genes involved in gene expression regulation. Common gene targets of the four ASD risk genes are mostly transcription regulators and chromatin remodelers, such as *CTNNB1* and *SMARCA4*, providing additional avenues for the regulation of more genes.

While bioinformatic analyses have suggested the common and unique targets of the four ASD risk genes, we directly

examined the transcriptional impact of their deficiency. First, we have revealed that the four ASD risk genes regulate each other. The convergent reduction of Adnp by knockdown of Kdm6b, Chd2 or Med13 suggests that Adnp deficiency may play a role in ASD caused by different risk genes. The increased expression of Med13 (by Adnp knockdown), Chd2 (by Kdm6b knockdown), and Kdm6b (by Chd2 or Med13 knockdown) provides a potential compensatory effect for the haploinsufficiency of a single ASD risk gene.

Next, we have revealed that subsets of the high-confidence ASD genes in GER group with the occupancy of the four ASD risk genes at their promoters, including *Ctnnb1*, *Smarca4*, *Pogz*, *Kmt2c*, and *Pten*, are indeed transcriptionally regulated by deficiency of these ASD risk genes. Moreover, we have found that some synaptic genes, such as *Snap25* and *Nrxn1*, are strongly regulated by deficiency of the four ASD risk genes, which could be through the direct binding at promoters (e.g. for Med13) or indirectly through the targets like *Ctnnb1* or *Smarca4*.

The molecular network altered by multiple ASD risk genes could influence important biological processes. One of the convergent biological impact of the aberrant expression of ASD risk genes and their targets could be the dysregulation of

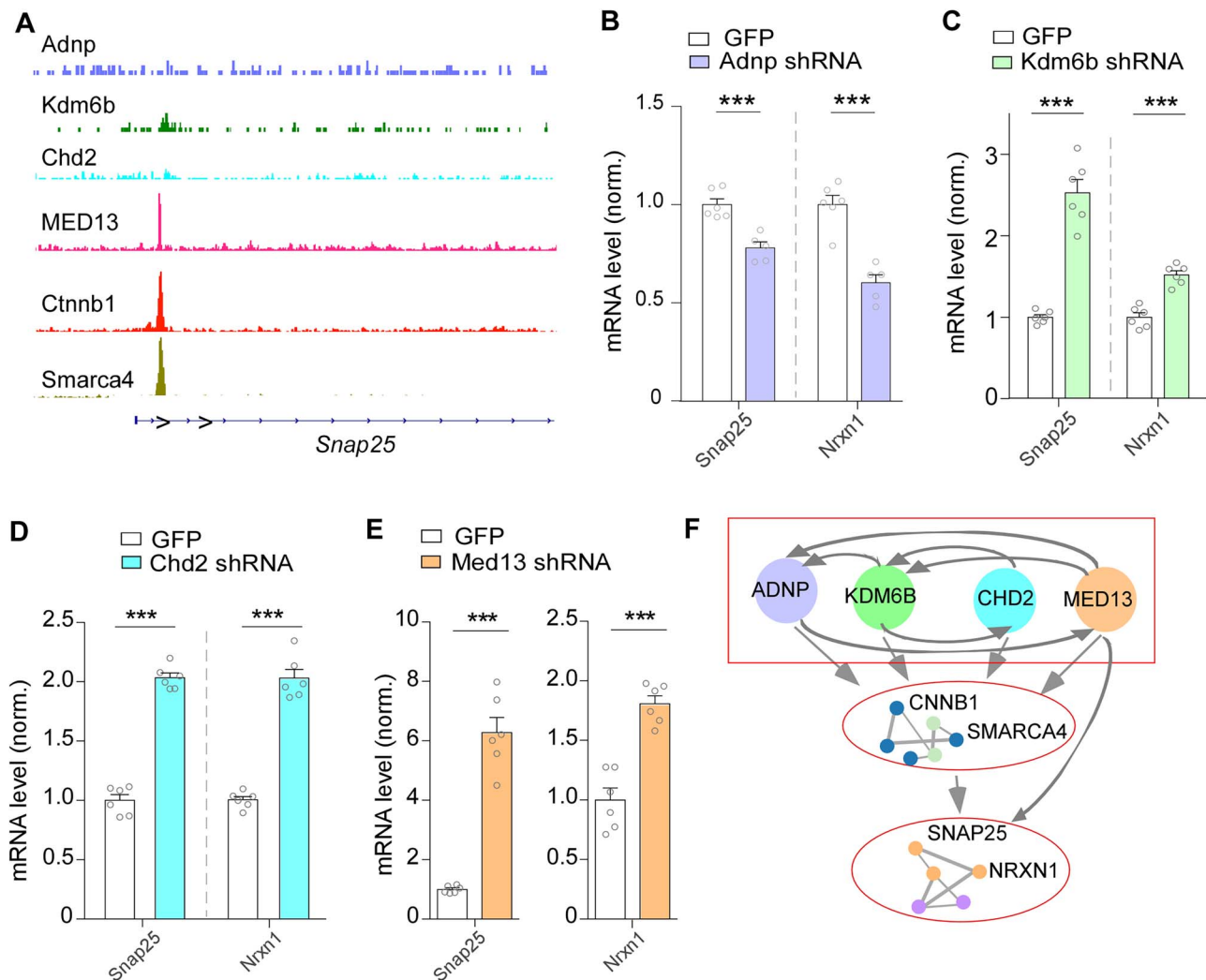


Figure 7. Alteration of synaptic genes in cells with the deficiency of four ASD risk genes. (A) ChIP-seq data showing the enrichment of Med13, Ctnnb1 and Smarca4, but not Adnp, Kdm6b and Chd2, at *Snap25* promoter. (B–E) Bar graphs showing the RT-qPCR data of *Snap25* and *Nrxn1* in N2A cells transfected with shRNA against the four ASD risk genes. (B) Adnp KD: *Snap25*, 22% decrease, $P=0.006$, *Nrxn1*, 40% decrease, $P=0.001$; (C) Kdm6b KD: *Snap25*, 153% increase, $P=2.99E-06$, *Nrxn1*, 52% increase, $P=3.21E-5$; (D) Chd2 KD: *Snap25*, 103% increase, $P=1.32E-08$, *Nrxn1*, 102% increase, $P=1.44E-07$; (E) Med13 KD: *Snap25*, 528% increase, $P=0.0001$, *Nrxn1*, 81% increase, $P=0.0001$, $n=6-9/\text{group}$; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, unpaired t-test. (F) A schematic model illustrating the regulatory networks of four ASD high risk genes, ADNP, KDM6B, CHD2 and MED13. They not only regulate each other, but also influence the transcription of many other ASD genes involved in gene expression regulation, such as CTNNB1 and SMARCA4, which can further alter the expression of synaptic genes, such as SNAP25 and NRXN1.

neuronal development. For example, ADNP, via the recruitment of HP1 and CHD4, regulates the expression of lineage-specifying genes, therefore governing cell fate plasticity [6]. ADNP also promotes neural induction and differentiation by enhancing Wnt signaling via stabilizing β -Catenin [10]. KDM6B regulates the differentiation of cerebellar granule neurons by inducing a mature neuronal gene expression program, which includes gene products required for functional synapse maturation [28]. CHD2 is required for differentiation of mouse embryonic stem cells [17] and neurogenesis in the developing cerebral cortex [29]. CHD2 deficiency leads to deficits in neuron proliferation [18]. Loss of SMARCA4 impairs cerebellar development, which is likely through an enhanced activity of Wnt/ β -catenin signaling [30]. Hence, ASD genes could profoundly influence the differentiation of neural progenitor cells into various types of neurons and glial cells by regulating the genes and pathways that determine the fate of neural stem cells in the developing brain.

Another convergent biological impact of the aberrant expression of ASD risk genes and their targets could be the dysregulation of synaptic function. For example, ADNP deficiency reduces dendritic spine density and alters synaptic gene expression [2]. Conditional knockout of KDM6B in excitatory neurons reduces spine density, synaptic vesicle number and synaptic activity by upregulating VGLUT1/2 through demethylating H3K27me3 at their promoters [31]. CTNNB1 regulates the coordinated changes in dendritic morphology and unitary excitatory synaptic strength [32]. The suppression of *Snap25* and *Nrxn1* by Adnp deficiency may contribute to the diminished synaptic structure and function in ADNP haploinsufficiency [2, 11]. On the other hand, the elevation of *Snap25* and *Nrxn1* by deficiency of Kdm6b, Chd2 or Med13 may contribute to the impairment of synaptic homeostasis and excitation/inhibition balance in ASD [33, 34].

Taken together, this study has revealed a molecular network regulated by multiple ASD risk genes. Dysregulation of these genes could collectively contribute to the aberrant behavioral

manifestations in ASD. Identification of convergent key genes and their connections could help develop more effective therapeutic targets for the complex brain disorder.

Materials and methods

Bioinformatic analysis of ChIP-seq data

ADNP, KDM6B, CHD2 and MED13 ChIP-seq data were acquired from available datasets in ChIP-Atlas [http://chip-atlas.org/target_gene]. ADNP ChIP-seq (GSE97945) was performed on mouse embryonic stem cells; KDM6B ChIP-seq (GSE36673, GSE38269) was performed on mouse embryonic brains; CHD2 ChIP-seq (GSE136499) was performed on mouse embryonic stem cells; MED13 ChIP-seq (GSE170110) was performed on human liver cells. ChIP-seq landscapes were generated using BigWig files and Integrative Genomics Viewer (IGV) Web App [<https://igv.org>] [35].

EnrichR [<https://maayanlab.cloud/Enrichr/>] [36, 37] was used to identify enriched categories of the gene targets of ADNP, KDM6B, CHD2 and MED13. ASD genes were downloaded from SFARI GENE with all the 4 gene scores (Cat. S: Syndromic; Cat. 1: High Confidence, Cat. 2: Strong Candidate, Cat. 3: Suggestive Evidence) on all chromosomes [<https://gene.sfari.org/database/human-gene/>]. InteractiveVenn [<http://www.interactivenn.net/>] [38] was used to detect overlapping genes among gene sets. String database [<https://string.db>] [39] was used to create protein-protein interaction (PPI) network for genes. CytoHubba [40] app in Cytoscape (version 3.9.0) [41] was used to detect top ranking hub genes using the Maximal Clique Centrality algorithm.

Gene knockdown and quantitative real-time RT-PCR

Short-hairpin RNA (shRNA) against mouse Adnp, Kdm6b, Chd2 or Med13 was designed using Millipore Sigma Pre-designed shRNA (<https://www.sigmaaldrich.com/US/en/semi-configure/shrna?activeLink=productSearch>). Adnp shRNA sequence was CCTACAGATACCCTACTCA. Kdm6b shRNA sequence was CCTCGTCATCTCAGTTCTCTA. Chd2 shRNA sequence was CAAGAACCATCACGATTTAAT. Med13 shRNA sequence was ACCGAAACTTTGTACGAATTG. The shRNA was cloned into a GFP-tagged adeno-associated virus (AAV) vector under the control of U6 promoter (Addgene, Cat. # 85741). The shRNA plasmid was transfected into 65–70% confluent mouse N2a cells using Lipofectamine 3000 (Invitrogen). After 48 hr transfection, RNA was extracted for qPCR experiments.

Total RNA was isolated from N2A cells using Trizol reagent (Invitrogen). Then iScript™ cDNA synthesis Kit (Bio-Rad) was used to obtain cDNA from the mRNA. Quantitative real time PCR was carried out using the iCycler iQ™ RealTime PCR Detection System and iQ™ Supermix (Bio-Rad) according to manufacturer's instructions. In brief, GAPDH was used as the housekeeping gene for quantitation of the expression of target genes in samples transfected with the shRNA plasmid for each of the four ASD risk genes or the GFP control AAV vector (Addgene, Cat. # 85741). Fold changes in the target genes were determined by: $\text{Fold change} = 2^{-\Delta(\Delta C_T)}$, where $\Delta C_T = C_T(\text{target}) - C_T(\text{GAPDH})$, and $\Delta(\Delta C_T) = \Delta C_T(\text{shRNA}) - \Delta C_T(\text{GFP})$. C_T (threshold cycle) is defined as the fractional cycle number at which the fluorescence reaches $10\times$ of the standard deviation of the baseline. A total reaction mixture of $20 \mu\text{l}$ was amplified in a 96-well thin-wall PCR plate (Bio-Rad) using the following PCR cycling parameters: 95°C for 3 min with 39 cycles of 95°C for 15 s, 60°C for 45 s, and followed by

melt curve analysis with $55\text{--}95^\circ\text{C}$, 0.5°C increment, and 5 s/step. Primers for all target genes are listed in [Supplementary Table 1](#).

Statistical analyses

Statistical analyses were performed with GraphPad Prism. Data were tested for normality before parametric analysis. Experiments with two groups were analyzed using two-tailed unpaired t-tests. All data points represent samples from distinct N2A cell cultures. All data are presented as the mean \pm SEM.

Acknowledgements

We thank Prachetas Jai Patel for his training of bioinformatic analyses.

Author contributions

L.W. and G.Y. performed cell culture, transfection and qPCR experiments and analyzed data. Z.Y. performed bioinformatic analyses, designed experiments, and wrote the paper.

Supplementary data

[Supplementary data](#) is available at *HMG Journal* online.

Conflict of interest statement. The authors declare no competing interests.

Funding

This work was supported by a grant from the National Institutes of Health (R01-NS127728) to Z.Y.

Data availability

Data will be made available on request.

References

1. Satterstrom FK, Kosmicki JA, Wang J. *et al.* Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell* 2020;**180**:568–584.e23.
2. Hacoheh-Kleiman G, Sragovich S, Karmon G. *et al.* Activity-dependent neuroprotective protein deficiency models synaptic and developmental phenotypes of autism-like syndrome. *J Clin Invest* 2018;**128**:4956–69.
3. De Rubeis S, He X, Goldberg AP. *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 2014;**515**:209–15.
4. Trost B, Thiruvahindrapuram B, Chan AJS. *et al.* Genomic architecture of autism from comprehensive whole-genome sequence annotation. *Cell* 2022;**185**:4409–4427.e18.
5. Mandel S, Rechavi G, Gozes I. Activity-dependent neuroprotective protein (ADNP) differentially interacts with chromatin to regulate genes essential for embryogenesis. *Dev Biol* 2007;**303**:814–24.
6. Ostapczuk V, Mohn F, Carl SH. *et al.* Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature* 2018;**557**:739–43.

7. Kaaij LJT, Mohn F, van der Weide RH. et al. The ChAHP complex counteracts chromatin looping at CTCF sites that emerged from SINE expansions in mouse. *Cell* 2019;**178**:1437–1451.e14.
8. Oz S, Kapitansky O, Ivashco-Pachima Y. et al. The NAP motif of activity-dependent neuroprotective protein (ADNP) regulates dendritic spines through microtubule end binding proteins. *Mol Psychiatry* 2014;**19**:1115–24.
9. Merenlender-Wagner A, Malishkevich A, Shemer Z. et al. Autophagy has a key role in the pathophysiology of schizophrenia. *Mol Psychiatry* 2015;**20**:126–32.
10. Sun X, Peng X, Cao Y. et al. ADNP promotes neural differentiation by modulating Wnt/ β -catenin signaling. *Nat Commun* 2020;**11**:2984.
11. Conrow-Graham M, Williams JB, Martin J. et al. A convergent mechanism of high risk factors ADNP and POGZ in neurodevelopmental disorders. *Brain* 2022;**145**:3250–63.
12. De Santa F, Totaro MG, Prosperini E. et al. The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell* 2007;**130**:1083–94.
13. Lan F, Bayliss PE, Rinn JL. et al. A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* 2007;**449**:689–94.
14. Agger K, Cloos PA, Christensen J. et al. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* 2007;**449**:731–4.
15. de Dieuleveult M, Yen K, Hmitou I. et al. Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. *Nature* 2016;**530**:113–6.
16. Gaspar-Maia A, Alajem A, Polesso F. et al. Chd1 regulates open chromatin and pluripotency of embryonic stem cells. *Nature* 2009;**460**:863–8.
17. Semba Y, Harada A, Maehara K. et al. Chd2 regulates chromatin for proper gene expression toward differentiation in mouse embryonic stem cells. *Nucleic Acids Res* 2017;**45**:8758–72.
18. Kim YJ, Khoshkhoo S, Frankowski JC. et al. Chd2 is necessary for neural circuit development and long-term memory. *Neuron* 2018;**100**:1180–1193.e6.
19. Nozawa K, Schneider TR, Cramer P. Core mediator structure at 3.4 Å extends model of transcription initiation complex. *Nature* 2017;**545**:248–51.
20. Plaschka C, Larivière L, Wenzek L. et al. Architecture of the RNA polymerase II-mediator core initiation complex. *Nature* 2015;**518**:376–80.
21. Qin L, Ma K, Wang ZJ. et al. Social deficits in Shank3-deficient mouse models of autism are rescued by histone deacetylase (HDAC) inhibition. *Nat Neurosci* 2018;**21**:564–75.
22. Qin L, Williams JB, Tan T. et al. Deficiency of autism risk factor ASH1L in prefrontal cortex induces epigenetic aberrations and seizures. *Nat Commun* 2021;**12**:6589.
23. Rapanelli M, Williams JB, Ma K. et al. Targeting histone demethylase LSD1 for treatment of deficits in autism mouse models. *Mol Psychiatry* 2022;**27**:3355–66.
24. Wang ZJ, Rein B, Zhong P. et al. Autism risk gene KMT5B deficiency in prefrontal cortex induces synaptic dysfunction and social deficits via alterations of DNA repair and gene transcription. *Neuropsychopharmacology* 2021;**46**:1617–26.
25. Miller SA, Mohn SE, Weinmann AS. Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. *Mol Cell* 2010;**40**:594–605.
26. Elmlund H, Baraznenok V, Lindahl M. et al. The cyclin-dependent kinase 8 module sterically blocks mediator interactions with RNA polymerase II. *Proc Natl Acad Sci USA* 2006;**103**:15788–93.
27. Knuesel MT, Meyer KD, Bernecky C. et al. The human CDK8 subcomplex is a molecular switch that controls mediator coactivator function. *Genes Dev* 2009;**23**:439–51.
28. Wijayatunge R, Liu F, Shpargel KB. et al. The histone demethylase Kdm6b regulates a mature gene expression program in differentiating cerebellar granule neurons. *Mol Cell Neurosci* 2018;**87**:4–17.
29. Shen T, Ji F, Yuan Z. et al. CHD2 is required for embryonic neurogenesis in the developing cerebral cortex. *Stem Cells* 2015;**33**:1794–806.
30. Moreno N, Schmidt C, Ahlfeld J. et al. Loss of Smar proteins impairs cerebellar development. *J Neurosci* 2014;**34**:13486–91.
31. Wang Y, Khandelwal N, Liu S. et al. KDM6B cooperates with tau and regulates synaptic plasticity and cognition via inducing VGLUT1/2. *Mol Psychiatry* 2022;**27**:5213–26.
32. Peng YR, He S, Marie H. et al. Coordinated changes in dendritic arborization and synaptic strength during neural circuit development. *Neuron* 2009;**61**:71–84.
33. Nelson SB, Valakh V. Excitatory/inhibitory balance and circuit homeostasis in autism Spectrum disorders. *Neuron* 2015;**87**:684–98.
34. Lee E, Lee J, Kim E. Excitation/inhibition imbalance in animal models of autism Spectrum disorders. *Biol Psychiatry* 2017;**81**:838–47.
35. Robinson JT, Thorvaldsdóttir H, Winckler W. et al. Integrative genomics viewer. *Nat Biotechnol* 2011;**29**:24–6.
36. Kuleshov MV, Jones MR, Rouillard AD. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016;**44**:W90–7.
37. Xie Z, Bailey A, Kuleshov MV. et al. Gene set knowledge discovery with Enrichr. *Curr Protoc* 2021;**1**:e90.
38. Heberle H, Meirelles GV, Da Silva FR. et al. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* 2015;**16**:169.
39. Szklarczyk D, Gable AL, Nastou KC. et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 2021;**49**:D605–12.
40. Chin C-H, Chen S-H, Wu H-H. et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014;**8**:S11.
41. Shannon P, Markiel A, Ozier O. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;**13**:2498–504.