Supplemental Information

Identification of Human Neuronal Protein Complexes Reveals Biochemical Activities and Convergent Mechanisms of Action in Autism Spectrum Disorders

Jingjing Li[1], Zhihai Ma[1], Minyi Shi[1], Ramy H. Malty[4], Hiroyuki Aoki[4], Zoran Minic[4], Sadhna Phanse[4], Ke Jin[4,5], Dennis P. Wall[2,3], Zhaolei Zhang[5], Alexander E. Urban[1,2], Joachim Hallmayer[2], Mohan Babu[4,#] and Michael Snyder[1,#]

Affiliation: [1]Department of Genetics, Stanford Center for Genomics and Personalized Medicine, [2]Department of Psychiatry & Behavioral Sciences, Stanford University School of Medicine, [3]Department of Pediatrics, Stanford, California, 94305 USA
[4]Department of Biochemistry, Research and Innovation Centre, University of Regina, Regina, Saskatchewan S4S 0A2, Canada
[5]Banting and Best Department of Medical Research, Terrence Donnelly Center for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1, Canada

[#] Correspondence should be addressed to: mpsnyder@stanford.edu or mohan.babu@uregina.ca

Supplemental Tables (EXCEL data files)

Table S1. The ASD candidate genes in this study (related to Fig. 2 and 3). The merged gene list, and the separate gene lists (grouped by mutation types) are shown in each column. The sources of the data are also indicated.

Table S2. Protein complexes involving at least one ASD-associated subunit (related to Fig. 2). The first datasheet contains all the complexes with at least one subunit associated with ASD. Each row is for a complex. The second datasheet is for the enriched MGI phenotypes for subunits co-complexed with ASD candidate genes (the proband column). The phenotypic terms were also tested on the subunits co-complexed with non-ASD genes (the sibling column).

Table S3. Gene expression analysis for the differentiated SH-SY5Y neuronal cells (related to Fig.4). The first datasheet is the enriched gene ontology terms for the up-regulated genes upon SH-SY5Y differentiation. The second datasheet is the RNA-Seq data (FPKM output of Cuffdiff) for all the genes before and after neuronal differentiation.

Table S4. **All the antibodies used in this study (related to Fig.4).**

Table S5. The identified prey proteins in the differentiated SH-SY5Y neuronal cells (related to Fig.5). The first datasheet contains each co-purifying bait proteins with each bait protein. The spectral counts of each prey protein in two biological replicates as well as their occurrences in each of the four control IP experiments are listed. The fold changes between the IP and control experiments are also shown. The statistical confidence for each prey protein identification is indicated by the SaintScores. The second datasheet lists the 19 genes previously identified in the PSD proteome.

Table S6. The identified 35 co-purifying prey proteins with HDAC2 in the differentiated SH-SY5Y neuronal cells (related to Fig.4, 5). The data format is the same as that in Table S5.

Supplemental Figure S1

related to Fig. 2

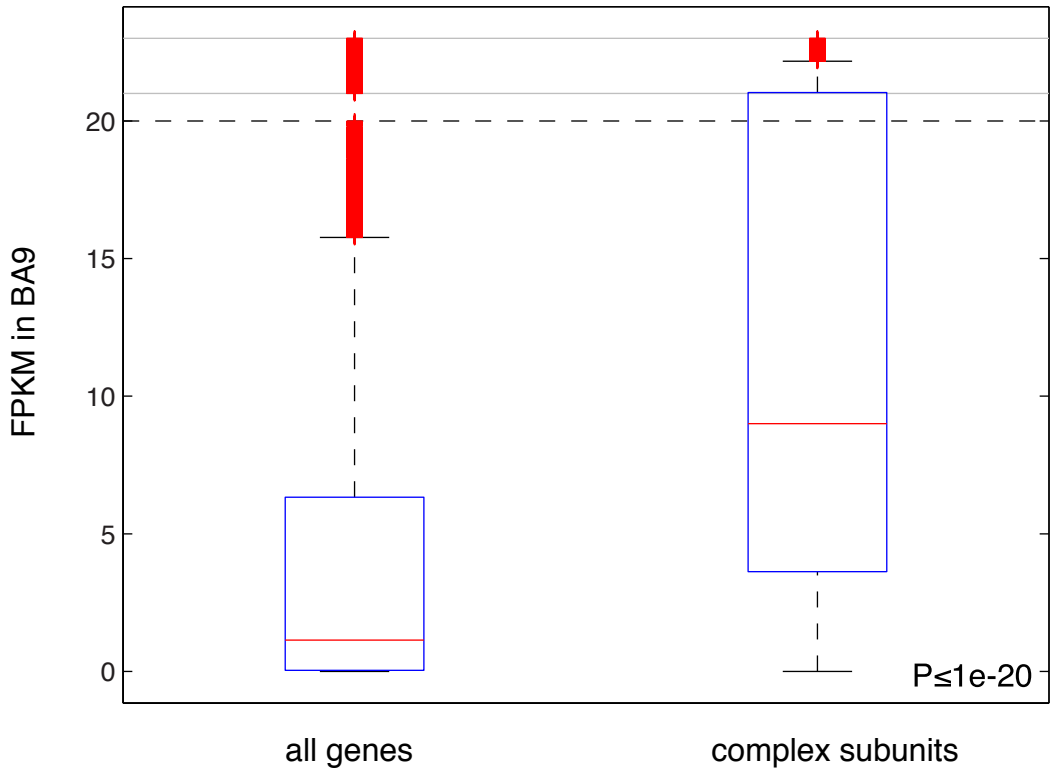

Fig.S1. The increased expression of the protein complex subunits in the dorsolateral dorsolateral prefrontal cortex (Brodmann area 9) from a control subject. Gene expression in BA9 was previously determined by RNA-Seq analysis.

Supplemental Figure S2
related to Fig. 3

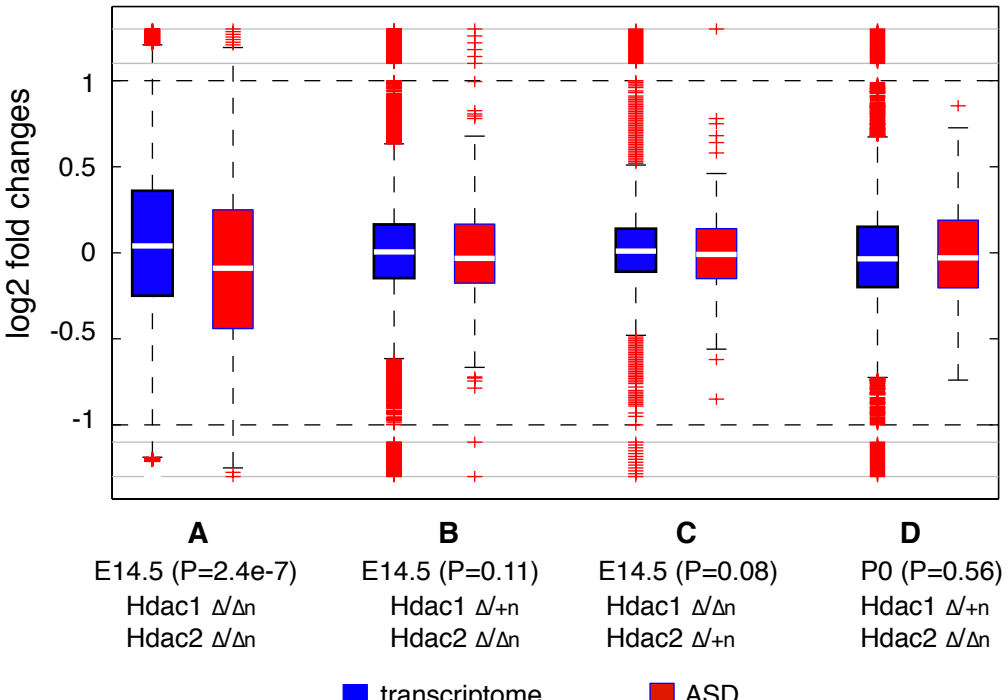

Fig. S2. Hdac1/2 double-knockout (A) down-regulates ASD-associated genes, whereas expressing a single HDAC1 (B) or HDAC2 (C) allele at E14.5 or P0 (D) up-regulates the ASD-implicated genes to a level similar with the transcriptome background.

Supplemental Figure S3
related to Fig. 4

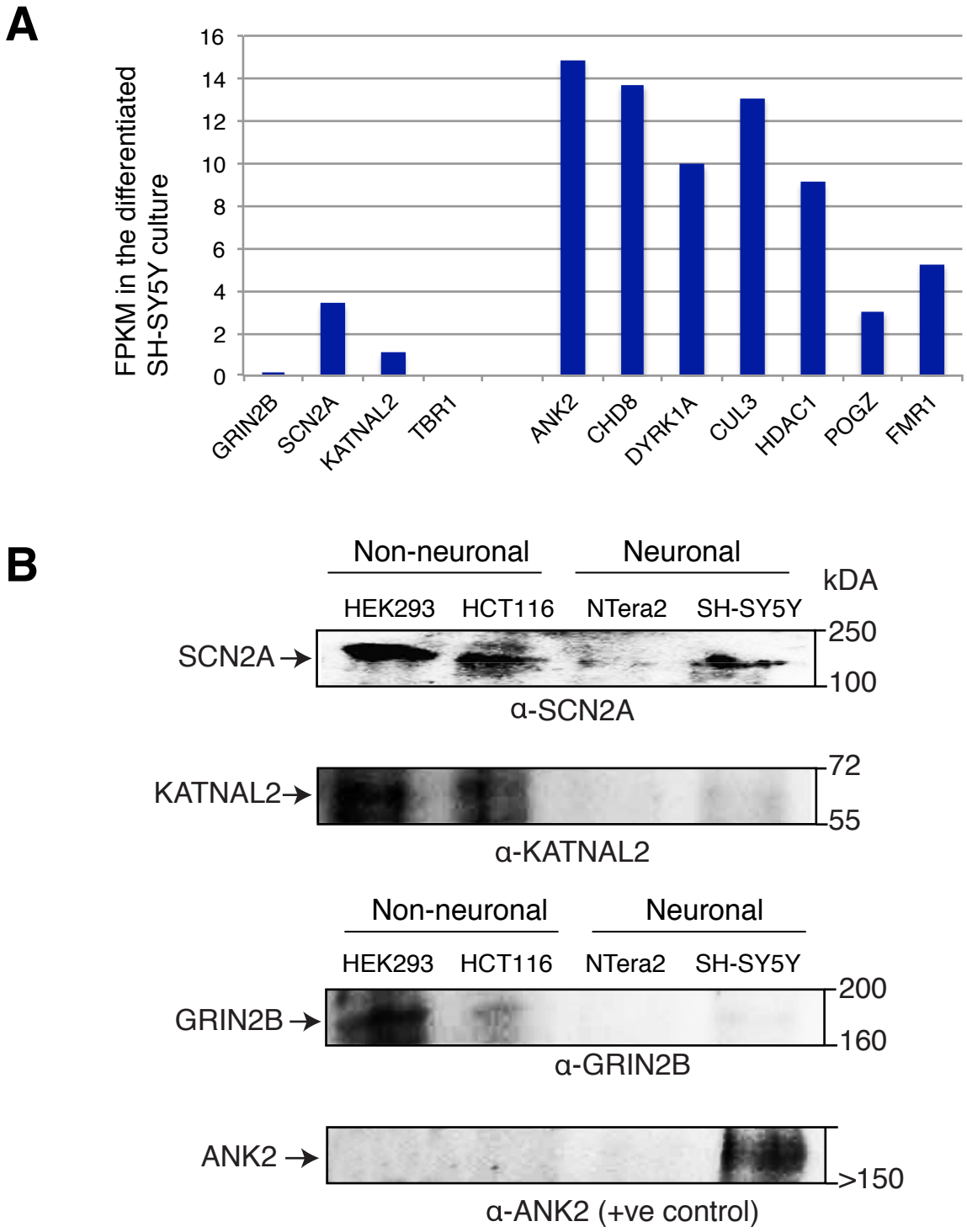

Fig. S3. Expression analysis of the four bait proteins (GRIN2B, SCN2A, TBR1 and KATNAL2) not recovered in IP-MS experiments in the differentiated SH-SY5Y neuronal culture. (A) RNA-sequencing revealed the low expression of the four proteins in the differentiated SH-SY5Y neuronal culture, whereas the other seven bait proteins on average showed high abundance in mRNA level. (B). Immunoblotting showing low abundance of the endogenous protein expression for GRIN2B, SCN2A KATNAL2 and ANK2 in two neuronal (differentiated SH-SY5Y and NTera2) relative to two non-neuronal (HEK293 and HCT116) cell lines. TBR1 was not tested due to its extremely low mRNA level (FPKM=0). ANK2, the bait protein successfully recovered from the IP-MS experiments, showed high expression in both mRNA and protein levels (relative to the non-neuronal cell lines).

Supplemental Experimental Procedures

Analysis of the ubiquitous human protein complexes

We employed a complete list of 622 soluble stable protein complexes from a recent proteome-wide mass spectrometric profiling study(Havugimana et al., 2012). The UniProt identifiers of the protein subunits were mapped onto human HGNC (HUGO Gene Nomenclature Committee) gene symbols. We examined mRNA expression of the protein complex subunits in the postmortem dorsolateral prefrontal cortex (Brodmann area, BA9) from a control subject in our earlier RNA-Seq study(Li et al., 2014), and observed that genes encoding subunits of the 622 complexes indeed displayed substantially higher expression (median FPKM>5) relative to the transcriptome background (P<1e-20, Wilcoxon rank-sum test, Fig. S1).

We considered 378 ASD-associated genes affected by different types of mutations from multiple studies (Table S1), including i) high-confidence ASD-associated de novo CNVs(Noh et al., 2013), ii) ASD-associated syndromic mutations (from SFARI annotations, Category S), iii) *de novo* loss-of-function mutations in ASD probands (note the union set was used from the exome-sequencing studies(Iossifov et al., 2012; Neale et al., 2012; O'Roak et al., 2012; Sanders et al., 2012; Willsey et al., 2013), and iv) nine high-confidence ASD-associated genes affected by *de novo* mutations from Willsey, *et al.*(Willsey et al., 2013), including *ANK2*, *CHD8*, *CUL3*, *DYRK1A*, *SCN2A*, *TBR1*, *GRIN2B*, *POGZ*, and *KATNAL2*. To determine their overall functional enrichment, an appropriate background control (proteins not co-complexed with the ASD candidates) was needed. We therefore assembled a control set based on two considerations: (1) the original study showed that the entire collection of these protein complexes were enriched for many functional categories(Havugimana et al., 2012), so the control proteins in this study should be sampled from the same set of protein complexes identified from the same study thereby avoiding potential sampling bias; (2) the vast majority of the ASD candidates in this study were from previous screens for *de novo* mutations (>86%), and many of the remaining genes (for syndromic ASDs) were also identified through *de novo* mutations (e.g. mutations in *FMR1*(De Boulle et al., 1993) and *MECP2*(Amir et al., 1999)). Because genes with varying functions have different rates of *de novo* mutations(Samocha et al., 2014), one cannot evenly permute the ASD candidates within the protein complexes for the purpose of experimental control. Instead, we compiled 592 genes affected by *de novo* loss-of-function mutations or exon-affecting CNVs from unaffected siblings or healthy control subjects in previous studies(Iossifov et al., 2012; Kirov et al., 2012; Levy et al., 2011; Neale et al., 2012; O'Roak

et al., 2012; Sanders et al., 2011; Sanders et al., 2012; Willsey et al., 2013), which mapped to 87 protein complexes. We excluded 20 complexes that overlapped between the 98 complexes harboring ASD-associated proteins and the 87 control protein complexes. We compared with the protein complex sizes (the number of subunits) between the two groups (after removing the 20 overlapping complexes), and found that their sizes were statistically similar (P=0.31, Wilcoxon rank-sum test). We removed the annotated ribosomal and spliceosomal subunits from our analysis, which were retrieved from the CORUM database(Ruepp et al., 2010). In the meantime,

The subunits co-complexed with the ASD gene candidates were compared to the subunits co-complexed with control proteins, where the ASD candidates and the control proteins were excluded from the analysis. ClueGO(Bindea et al., 2009) was used to perform the differential gene ontology analysis, where significant terms have adjusted P-values less than 0.05 (by Bonferroni step-down correction).

Throughout this manuscript, we also ued Enrichr(Chen et al., 2013) for phenotypic enrichment analysis. Briefly, this test was to identify the enriched phenotypic terms associated with individual mouse mutants, and the mammalian phenotypic terms are standardized by Mouse Genome Informatics. Specifically, we considered the mammalian phenotypes at the hierarchical level four, which allowed us to study more specific phenotypic terms. The significant phenotypic terms had adjusted P-values less than 0.05 (corrected by Benjamini-Hochberg's method).

Analysis of the role of HDAC1/2 in ASD

We examined the effects of HDAC1/2 double-knockout in embryonic mouse brain. Gene expression dataset for HDAC1/2 double-knockout was from a recent study (Hagelkruys et al., 2014). We followed the procedures in the original study, and considered gene differential expression if one of its associated probe-sets displayed statistically significant differential expression. For genes with multiple probe sets, we followed the procedures in the original study and considered the probe set with the strongest signal (Hagelkruys et al., 2014). Specially, if the probe sets mapping onto the same gene displaying consistent up- or down-regulation upon gene knockout, we considered the probe set showing the strongest differentiation signal, but we discarded genes whose probe sets showed inconsistent up- or down-regulation upon gene knockout(s). In addition to the HDAC1/2 double-knockout, single knockouts of HDAC1 (at E14.5 and P0) or HDAC2 (at E14.5) were also examined individually. The human-mouse gene orthology was determined by ENSEMBL Biomart (as of April, 2014), and only the one-to-one

orthologs (15,797 human-mouse orthologous pairs in total) were considered in this study. In addition to the ASD-implicated genes described above, we also analyzed genes implicated in intellectual disability (ID, 401 genes)(Parikshak et al., 2013), Schizophrenia (SZ, 499 genes)(Fromer et al., 2014; Purcell et al., 2014) and Alzheimer's disease (ALZ, 613 genes)(Bertram et al., 2007).

Cell culture and differentiation

SH-SY5Y cells (ECACC, Sigma-Aldrich) in 1:1 DMEM-F12 medium (Life technologies) supplemented with 10% fetal bovine serum (FBS), penicillin (50 u/mL), streptomycin (50 µg/mL), and L-glutamine (2 mM) were maintained at 37 °C in a saturated humidity atmosphere containing 95% air and 5% CO2. The SH-SY5Y cells were then differentiated into neuron-like cells by the addition of retinoic acid (RA, Sigma-Aldrich; R2625) and brain-derived neurotrophic factor (BDNF, eBioscience) essentially as previously described with minor modifications(Encinas et al., 2000). Briefly, for differentiation, cells were plated at a density of 4 x 10^4 / cm2 in medium containing 5% FBS along with aforementioned components. On the next day, medium containing 5% FBS and 10 µM RA was added to the cells and the medium was changed every day for a total of 7 days. On the 8th day, the medium was replaced with DMEM supplemented with penicillin (50 u/mL), streptomycin (50 µg/mL), L-glutamine (2 mM), and BDNF (20 ng/mL) but without FBS addition. Again the medium was changed every 2-3 days for an additional 9 days. The differentiated cells were detached by incubating the cells with versene at 37 °C for 10-15 min, followed by phosphate buffer saline wash, and centrifugation (twice) at 190 xg for 10 min. We analyzed the complexes isolated from differentiated SH-SY5Y neuron-like cells on day 16 (starting with undifferentiated cells on day 0). The differentiated cells were visually inspected for neuron-like morphology (Fig. 4A), followed by RNA-Seq analysis (two replicates each for both undifferentiated and differentiated cells) to confirm globally increased expression of neuronal markers(Cahoy et al., 2008) after differentiation (P=7.7e-5, Wilcoxon rank-sum test).

RNA-Seq procedures

We performed RNA-sequencing for the differentiated and undifferentiated SH-SY5Y cells. Two µg of total RNA each sample was subject to RNA-seq library preparation with ScriptSeq™ Complete Gold Kit from Epicentre (Cat. # SCL24EP, Madison, WI) following the manufacturer's instructions. In brief, ribosomal RNA was depleted from total RNA using Ribo-Zero magnetic beads, then the ribosomal RNA depleted RNA was purified and fragmented. Random primer tailed with Illumina adaptor was used to perform reverse transcription to get cDNA library. Adaptor sequence was added to the other end of

cDNA library with a Terminal-Tagging step. cDNA library was amplified with Illumina primers provided with this kit. The product was size selected (350~500 bp) from 2% agarose E-gels (Invitrogen) and sequenced in one lane per sample on Illumina's HiSeq 2000 platform. We sequenced two replicates for each of the differentiated and undifferentiated SH-SY5Y cells, and used cuffdiff (v.2.0.2) to identify the differentially expressed genes upon the induced differentiation. 258 genes with q value less than 0.05 and fold change (after differentiation) great than 2 were examined for their functional enrichment. DAVID (http://david.abcc.ncifcrf.gov) was then used to examine the gene ontology enrichment. The significantly up-regulated genes in the differentiated state displayed GO functional enrichment (by DAVID) for neuron differentiation (FDR=2.2e-3), neuron projection morphogenesis (FDR=3.2e-3) and development (FDR=4.1e-3), axonogenesis (FDR=6.2e-3) and synaptic transmission (FDR=4.6e-3), demonstrating neuron-like characteristics of these cells (Table S3).

IP-MS to identify interacting proteins

Affinity purification by immuno-precipitation combined with mass-spectrometry analysis (IP-MS) was used to identify interacting proteins in the SH-SY5Y neuronal differentiated culture. To identify interacting proteins, the differentiated SH-SY5Y neuron-like cells were harvested and lysed for rapid co-immunoprecipitation of the bait proteins, using antibodies with high binding specificity (Table S4).). Immunoprecipitated proteins were digested with trypsin, and peptide mixtures of proteins that co-immunoprecipitated with the bait proteins were analyzed by a high-resolution Orbitrap Elite mass spectrometer (Fig. 4A). All the experiments were performed using two different cell batches that were separately cultured, grown and differentiated (*i.e.* two biological replicates). Prior to cell lysis and IP, viable surviving neuron-like cells were cross-linked using the cell membrane-permeable bifunctional cross-linking reagent, dithiobis succinimidyl propionate (DSP, Thermo Scientific). The cross-linked cells were pelleted and lysed in RIPA buffer (150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 1% Na deoxycholate, 0.1% sodium dodecyl sulfate, 1% NP-40, and 1mM EDTA). Approximately 10 mg protein (measured using Bradford assay) with 50 µL µMACS protein A or G magnetic microbeads was incubated with protein-specific antibodies at 4°C for ~4 hours with gentle rotation. The sample mixtures purified using the magnetic column was washed three times with RIPA buffer containing detergent, followed by last two washes with detergent free RIPA buffer. Purified proteins are incubated for 30 min at room temperature with the proteolytic digestion mixture (2M urea, 50 mM Tris-HCl (pH 7.5), 1mM dithiothreitol, and immobilized trypsin (5 µg/mL)), and eluted further using the buffer containing 2M urea, 50 mM Tris-HCl (pH 7.5), and 5mM choloroacetamide. After continuing digestion of the eluted

mixture overnight, the peptides are desalted and zip-tipped for mass-spectrometry analysis (MS) analysis. The high-confidence matches of the resulting MS/MS spectra obtained from the Orbitrap Elite mass spectrometer was mapped to the reference human protein sequences using the SEQUEST database search engine with match quality evaluated against the STATQUEST algorithm(Kislinger et al., 2006). The information of all the antibodies used in this study can be found in Table S4.

Identify the bona fide protein interactions from IP-MS screens

The identified co-purifying proteins were filtered out at a peptide identification confidence 90% as well as with the most common background contaminants (compiled from in house purifications). With all these criteria, our initial screen identified 2812 potential protein interactions from the seven bait proteins. To exclude non-specific binding proteins, we additionally carried out four sets of independent negative control experiments, in which we used empty protein G beads (without antibodies), empty protein A beads (without antibodies), an anti-GFP antibody and a secondary goat IgG used in the co-immunoprecipitation steps, followed by the same IP-MS procedures. We then used SAINTexpress to build a Bayesian mixture model(Mellacheruvu et al., 2013; Teo et al., 2014), and derived the posterior probability of true interaction for each bait-prey pair by comparing their MS/MS spectra counts between the pull-down experiments and the four sets of negative control experiments. We ranked the posterior probabilities of all the 2812 bait-prey pairs. Using an independent set of known protein-protein interactions(Turinsky et al., 2011), we examined whether the interaction pairs assigned with higher posterior probability of true interaction were more likely to have known protein-protein interactions. We observed that the enrichment of known protein-protein interactions was positively correlated with our Bayesian confidence scores, and the highest enrichment signal was observed among the bait-prey interactions with posterior probabilities above 80% as true interactions (Fig. 4B, P=3.97e-6, Fisher's exact test). We thus retained 119 interactions above this threshold (Table S5), which was slightly more conservative than a threshold of 75% used in a recent study(Joshi et al., 2014).

Validation of the predicted IP-MS interactions

To confirm the physically bound interacting proteins by co-IP, the aforementioned IP procedure was followed, except that the co-purifying protein that bound to the beads was eluted by the addition of 25 µl Laemmli loading buffer at 95 °C, and the fractions separated on SDS-PAGE gel was transferred to PVDF membranes and probed with protein specific- and HRP conjugated secondary- antibodies. The co-purifying protein bound to the target bait on the blots were visualized using chemiluminescence.

Analysis of the interaction network

We analyzed the processed expression profiles from the BrainSpan data compiled from a recent study (Parikshak et al., 2013). We also used CORTECON expression atlas(van de Leemput et al., 2014) to examine co-expression of the interacting proteins during brain development. Expression measurements at the same time points for multiple probes of a single gene were averaged. Gene symbols of the interacting proteins were mapped to their respective Entrez ID, and only protein-coding genes (with Entrez IDs) were considered in this comparison.

To determine whether the IP-MS network has increased mutational burden among individuals with ASD, we examined previous published exome-sequencing dataset (Liu et al., 2013). This dataset contained sequenced variants from Broad Institute (with the Illumina platform) and from Baylor College of Medicine (BCM, with the SOLiD platform). Due to the missing information (quality scores) in the data generated from the Illumina platform, we were only able to limit to our analysis on 505 ASD cases and 491 matched control sequenced by BCM. We identified non-synonymous mutations identified only in the ASD patient cohort and only in to control cohort, and quantified their consequences using MutationTaster implemented by eXtasy (Sifrim et al., 2013). We considered the deleterious effects of the mutations with MutationTaster greater than 0.99. We also compiled 9201 genes whose CDS length and GC content are similar with these identified prey proteins (P=0.297 and 0.98 for CDS length and GC content, respectively, Wilcoxon rank-sum test). The information of CDS length and GC content for human genes was obtained from a recent publication (Georgi et al., 2013).

We further determined the role of FMRP/FMRP^{I304N}-mediated post-transcriptional regulation in regulating this identified protein interaction network. The confident FMRP (wild-type) binding sites were retrieved from the original PAR-CLIP study(Ascano et al., 2012). For each gene, the FMRP binding site density was the total number of distinct FMRP sites in a given gene normalized by the cDNA length of the gene. To determine the cDNA length, we compiled all the annotated cDNA sequences of human genes from ENSEMBL BioMart (as of February, 2014, only limited to the CCDS transcripts). For genes with multiple CCDS transcripts, the longest one was considered. A recent study compared the wild-type FMRP and the mutant FMRP^{I304N} PAR-CLIP sites in the genome-wide scale, and a hidden Markov model was developed and identified 9549 transcriptomic locations strongly bound by FMRP^{WT}, but not by FMRP^{I304N}(Wang et al., 2014). The genomic coordinates of these affected sites

were mapped onto the 3' UTRs and CDS of human RefSeq genes using BedTools (2.17.0), requiring strand specificity. In total, 1925 genes harbored at least one the affected site in their CDS or UTRs. To determine whether the number of 35 genes affected by $FMRP^{I304N}$ in our IP-MS network was not expected by chance, we compiled a collection of 2715 control genes with indistinguishable expression levels (P=0.62, Wilcoxon rank-sum test) and cDNA length (P=0.33, Wilcoxon rank-sum test) from the prey genes in the network. Expression data in the PAR-CLIP system was retrieved from the original study(Ascano et al., 2012). The same comparisons were also performed on ASD candidate genes from SFARI (https://gene.sfari.org, as of March, 2014) and ASD-associated genes in the earlier section for ubiquitous human protein complexes and genes with at least one exon affected by de novo CNV events identified in ASD probands(Levy et al., 2011; Pinto et al., 2014; Sanders et al., 2011). For each comparison the ASD-associated genes from different resources were all matched with their respective sets of control genes with indistinguishable gene expression and cDNA length.

We examined MECP2-mediated regulation from a recent study (Lanz et al., 2013), where eight ASD associated genes (Mecp2, Mef2a, Mef2d, Fmr1, Nlgn1, Nlgn3, Pten and Shank3) were individually knocked down using shRNAs (short hairpin RNAs) in murine cortical neurons, and the quadruplicate experiments confirmed that the knockdown efficiency achieved at least 75% expression reduction (Lanz et al., 2013). Fold changes of gene expression upon shRNA transfections were computed for each probe set in the original study (Lanz et al., 2013), and we averaged the data for different probe sets mapping onto the same genes.

References:

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nature genetics *23*, 185-188.

Ascano, M., Jr., Mukherjee, N., Bandaru, P., Miller, J.B., Nusbaum, J.D., Corcoran, D.L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M.*, et al.* (2012). FMRP targets distinct mRNA sequence elements to regulate protein expression. Nature *492*, 382-386.

Bertram, L., McQueen, M.B., Mullin, K., Blacker, D., and Tanzi, R.E. (2007). Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nature genetics *39*, 17-23.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.H., Pages, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091-1093.

Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A.*, et al.* (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. The Journal of neuroscience : the official journal of the Society for Neuroscience *28*, 264-278.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

De Boulle, K., Verkerk, A.J., Reyniers, E., Vits, L., Hendrickx, J., Van Roy, B., Van den Bos, F., de Graaff, E., Oostra, B.A., and Willems, P.J. (1993). A point mutation in the FMR-1 gene associated with fragile X mental retardation. Nature genetics *3*, 31-35.

Encinas, M., Iglesias, M., Liu, Y., Wang, H., Muhaisen, A., Cena, V., Gallego, C., and Comella, J.X. (2000). Sequential treatment of SH-SY5Y cells with retinoic acid and brain-derived neurotrophic factor gives rise to fully differentiated, neurotrophic factor-dependent, human neuron-like cells. Journal of neurochemistry *75*, 991-1003.

Fromer, M., Pocklington, A.J., Kavanagh, D.H., Williams, H.J., Dwyer, S., Gormley, P., Georgieva, L., Rees, E., Palta, P., Ruderfer, D.M.*, et al.* (2014). De novo mutations in schizophrenia implicate synaptic networks. Nature *506*, 179-184.

Georgi, B., Voight, B.F., and Bucan, M. (2013). From mouse to human: evolutionary genomics analysis of human orthologs of essential genes. PLoS genetics *9*, e1003484.

Hagelkruys, A., Lagger, S., Krahmer, J., Leopoldi, A., Artaker, M., Pusch, O., Zezula, J., Weissmann, S., Xie, Y., Schofer, C.*, et al.* (2014). A single allele of Hdac2 but not Hdac1 is sufficient for normal mouse brain development in the absence of its paralog. Development *141*, 604-616.

Havugimana, P.C., Hart, G.T., Nepusz, T., Yang, H., Turinsky, A.L., Li, Z., Wang, P.I., Boutz, D.R., Fong, V., Phanse, S.*, et al.* (2012). A census of human soluble protein complexes. Cell *150*, 1068-1081.

Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.H., Narzisi, G., Leotta, A.*, et al.* (2012). De novo gene disruptions in children on the autistic spectrum. Neuron *74*, 285-299.

Joshi, P., Greco, T.M., Guise, A.J., Luo, Y., Yu, F., Nesvizhskii, A.I., and Cristea, I.M. (2014). The functional interactome landscape of the human histone deacetylase family. Molecular Systems Biology *9*, 672-672.

Kirov, G., Pocklington, A.J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., Moran, J., Chambert, K., Toncheva, D., Georgieva, L.*, et al.* (2012). De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Molecular psychiatry *17*, 142-153.

Kislinger, T., Cox, B., Kannan, A., Chung, C., Hu, P., Ignatchenko, A., Scott, M.S., Gramolini, A.O., Morris, Q., Hallett, M.T.*, et al.* (2006). Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell *125*, 173-186.

Lanz, T.A., Guilmette, E., Gosink, M.M., Fischer, J.E., Fitzgerald, L.W., Stephenson, D.T., and Pletcher, M.T. (2013). Transcriptomic analysis of genetically defined autism candidate genes reveals common mechanisms of action. Molecular autism *4*, 45.

Levy, D., Ronemus, M., Yamrom, B., Lee, Y.H., Leotta, A., Kendall, J., Marks, S., Lakshmi, B., Pai, D., Ye, K.*, et al.* (2011). Rare de novo and transmitted copy-number variation in autistic spectrum disorders. Neuron *70*, 886-897.

Li, J., Shi, M., Ma, Z., Zhao, S., Euskirchen, G., Ziskin, J., Urban, A., Hallmayer, J., and Snyder, M. (2014). Integrated systems analysis reveals a molecular network underlying autism spectrum disorders. Mol Syst Biol *10*, 774.

Liu, L., Sabo, A., Neale, B.M., Nagaswamy, U., Stevens, C., Lim, E., Bodea, C.A., Muzny, D., Reid, J.G., Banks, E.*, et al.* (2013). Analysis of Rare, Exonic Variation amongst Subjects with Autism Spectrum Disorders and Population Controls. PLoS genetics *9*, e1003443.

Mellacheruvu, D., Wright, Z., Couzens, A.L., Lambert, J.-P., St-Denis, N.A., Li, T., Miteva, Y.V., Hauri, S., Sardiu, M.E., Low, T.Y.*, et al.* (2013). The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. Nature Methods *10*, 730-736.

Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V.*, et al.* (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature *485*, 242-245.

Noh, H.J., Ponting, C.P., Boulding, H.C., Meader, S., Betancur, C., Buxbaum, J.D., Pinto, D., Marshall, C.R., Lionel, A.C., Scherer, S.W.*, et al.* (2013). Network Topologies and Convergent Aetiologies Arising from Deletions and Duplications Observed in Individuals with Autism. PLoS genetics *9*, e1003523.

O'Roak, B.J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B.P., Levy, R., Ko, A., Lee, C., Smith, J.D.*, et al.* (2012). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature *485*, 246-250.

Parikshak, N.N., Luo, R., Zhang, A., Won, H., Lowe, J.K., Chandran, V., Horvath, S., and Geschwind, D.H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. Cell *155*, 1008-1021.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., Thiruvahindrapuram, B., Xu, X., Ziman, R., Wang, Z.*, et al.* (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. American journal of human genetics *94*, 677-694.

Purcell, S.M., Moran, J.L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., O'Dushlaine, C., Chambert, K., Bergen, S.E., Kähler, A.*, et al.* (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature *506*, 185-190.

Ruepp, A., Waegele, B., Lechner, M., Brauner, B., Dunger-Kaltenbach, I., Fobo, G., Frishman, G., Montrone, C., and Mewes, H.W. (2010). CORUM: the comprehensive resource of mammalian protein complexes--2009. Nucleic acids research *38*, D497-501.

Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnström, K., Mallick, S., Kirby, A.*, et al.* (2014). A framework for the interpretation of de novo mutation in human disease. Nature genetics *46*, 944-950.

Sanders, S.J., Ercan-Sencicek, A.G., Hus, V., Luo, R., Murtha, M.T., Moreno-De-Luca, D., Chu, S.H., Moreau, M.P., Gupta, A.R., Thomson, S.A.*, et al.* (2011). Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. Neuron *70*, 863-885.

Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L.*, et al.* (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature *485*, 237-241.

Sifrim, A., Popovic, D., Tranchevent, L.-C., Ardeshirdavani, A., Sakai, R., Konings, P., Vermeesch, J.R., Aerts, J., De Moor, B., and Moreau, Y. (2013). eXtasy: variant prioritization by genomic data fusion. Nature Methods *10*, 1083-1084.

Teo, G., Liu, G., Zhang, J., Nesvizhskii, A.I., Gingras, A.C., and Choi, H. (2014). SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. Journal of proteomics *100*, 37-43.

Turinsky, A.L., Razick, S., Turner, B., Donaldson, I.M., and Wodak, S.J. (2011). Interaction databases on the same page. Nature Biotechnology *29*, 391-393.

van de Leemput, J., Boles, Nathan C., Kiehl, Thomas R., Corneo, B., Lederman, P., Menon, V., Lee, C., Martinez, Refugio A., Levi, Boaz P., Thompson, Carol L.*, et al.* (2014). CORTECON: A Temporal Transcriptome Analysis of In Vitro Human Cerebral Cortex Development from Human Embryonic Stem Cells. Neuron *83*, 51-68.

Wang, T., Xie, Y., and Xiao, G. (2014). dCLIP: a computational approach for comparative CLIP-seq analyses. Genome Biol *15*, R11.

Willsey, A.J., Sanders, S.J., Li, M., Dong, S., Tebbenkamp, A.T., Muhle, R.A., Reilly, S.K., Lin, L., Fertuzinhos, S., Miller, J.A.*, et al.* (2013). Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell *155*, 997-1007.