Chemico-genetic Analysis of Native Autism Proteomes Reveals Shared Biology Predictive of Functional Modifiers

Yudong Gao[1], Matthew Trn[1], Daichi Shonai[2], Jieqing Zhao[3], Erik J. Soderblom[4], S. Alexandra Garcia-Moreno[5], Charles A. Gersbach[5,6], William C. Wetsel[7], Geraldine Dawson[8], Dmitry Velmeshev[9], Yong-hui Jiang[10], Laura Sloofman[11], Joseph D. Buxbaum[11,12], Scott H. Soderling[1,9,*]

[1] Department of Cell Biology, Duke University School of Medicine, Durham, NC, USA
[2] Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, NC, USA
[3] Department of Biology, Duke University, Durham, NC, USA
[4] Department of Cell Biology, Proteomics and Metabolomics Shared Resource, Duke University School of Medicine, Durham, NC, USA
[5] Department of Biomedical Engineering, Duke University, Durham, NC, USA.
[6] Center for Advanced Genomic Technologies, Duke University, Durham, NC, USA; Department of Cell Biology, Duke University Medical Center, Durham, NC, USA; Department of Surgery, Duke University Medical Center, Durham, NC, USA.
[7] Departments of Psychiatry and Behavioral Sciences, Cell Biology, and Neurobiology, Mouse Behavioral and Neuroendocrine Analysis Core Facility, Duke University School of Medicine, Durham, NC, USA.
[8] Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC, USA.
[9] Department of Neurobiology, Duke University School of Medicine, Durham, NC, USA
[10] Department of Genetics, Yale University School of Medicine, New Haven, Connecticut, USA.
[11] Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY, USA; Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
[12] The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA; Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
*Corresponding author. Email: Scott.Soderling@duke.edu

Abstract

One of the main drivers of autism spectrum disorder (ASD) are risk alleles within hundreds of genes, which may interact within shared biological processes through as-yet unclear mechanisms. Here we develop a high-throughput genome-editing-mediated approach to target 14 high-confidence ASD genes within the mouse brain for proximity-based proteomics of endogenous interactomes. The resulting interactomes are enriched for human genes dysregulated in the brain of ASD patients and reveal unexpected, but highly significant, interactions with other lower confidence ASD-risk gene products, positing new avenues to prioritize genetic risk. Importantly, the datasets are enriched for shared cellular functions and genetic interactions that may underlie the disorder. We test this notion by spatial proteomics and CRISPR-based regulation of expression in two ASD models, demonstrating new functional interactions that modulate mechanisms of their dysregulation. Together, our results reveal native protein-interaction networks in ASD, providing new inroads for understanding its cellular neurobiology.

Main Text

Autism spectrum disorder (ASD) is a neurodevelopmental condition with significant clinical heterogeneity and complex genetic etiology [1]. Decades of research have identified and curated an evolving list of gene mutations associated with ASD risk, many of which converge on pathways mediating synaptic/axonal functions and gene regulation [2-6], and exhibit cell-type specific expression patterns across the brain [7]. Hints of shared biology across these risk genes have been inferred from RNA-level gene expression analyses [8-10]. Nevertheless, molecular convergence may be best reflected at the protein product level, where they may form as-of-yet unknown physical complexes to orchestrate neural activities that are affected in ASD.

One hallmark of neurons is their distinctive sub-cellular compartmentation - such as the synapse and axonal initial segment (AIS) - that are pivotal for neurotransmission [11,12]. In addition, gene expression regulators, many of which reside within the nucleus, also orchestrate key milestones of neurodevelopment [6]. It is not surprising that ASD risk converges on genes encoding proteins associated with these sub-cellular compartments [13-17]. Identifying the ASD-associated proteome architecture of these compartments could define how seemingly diverse genetic mutations are functionally connected, thereby providing a new roadmap to reveal a converging ASD etiology. Current efforts to dissociate protein interactomes rely upon techniques such as immunoaffinity purification and proximity-based proteomic methods, including biotin-identification (BioID) [18-21]. However, due to inherent constraints of antibody-dependent or exogenous vector-driven overexpression strategies, access to the native interactomes of endogenously expressed ASD-risk proteins within brain tissue remains a significant challenge.

Previously, we published a two-vector CRISPR/Cas9 approach, Homology independent Universal Genome Engineering (HiUGE), that enables rapid endogenous neuronal gene knock-in for protein modification *in vivo* [22]. Here, we leveraged HiUGE and simplified it with a one-vector design to achieve a robust knock-in in brain tissue with an engineered biotin ligase, TurboID [23], for higher-throughput proteomics of endogenous protein complexes. Importantly, this strategy obviates the burden of producing transgenic mice by germline transmission. Coupling this approach with *in vivo* BioID (iBioID) [18], we targeted 14 genetic drivers of ASD (13 with SFARI gene score 1, one with gene score 2, also noted as syndromic) that were mapped to the synapse, AIS, or the nucleus. With this approach, we unraveled their native proximal proteomes from brain tissue for the first time to our knowledge. We identified 1115 unique proteins and 2923 proximity protein-protein interactions (PPI) associated with these ASD targets. Amongst them, 17% are proteins encoded by mouse homologs of SFARI genes[2], 9% overlap with differentially expressed genes (DEGs) found in ASD patient brain tissue [7], and 72% of the PPIs are not reported in STRING queries [24,25]. Notably, the interactomes contain products of many newly discovered ASD risk genes identified in recent studies [26-29] and reveal shared biological processes among ASD proteins that may be predictive of interactions influencing ASD phenotypes.

We tested this notion by identifying intersections between co-perturbed proteins and endogenous interactomes in two ASD mouse models (Syngap1 and Scn2a). In the Syngap1 model, we reveal that its binding with AnkS1b is disrupted by an ASD-associated Syngap1 mutation, and their interaction is essential for shaping neural activity during critical synaptogenesis periods. In the Scn2a model, we show that a patient-derived missense mutation results in repetitive behaviors and abnormal social interaction in mice. The mutation also downregulates a key Scn2a modulatory protein cluster discovered in its interactome and results in aberrant attenuation of neural activity. Strikingly, re-expression of this cluster rescues this ASD-associated electrophysiological impairment.

Together, our results establish a new scalable platform to map endogenous interactomes at the protein-level that are associated with genetic risks for neurodevelopmental conditions such as ASD. Our findings also reveal an intersectional approach to prioritize candidates based on proteomic co-perturbation. These data support a protein-centric model to reveal novel mechanisms of ASD etiology and potential mitigation approaches to this condition.

Results

Endogenous labeling of ASD-risk proteins using HiUGE

Previously we have used overexpression of bait proteins fused to various biotin ligases to discover interactomes from brain tissue [18,30]. Although this approach has been highly successful, it is known that the expression of proteins using artificial promoters may result in non-native interactions due to non-physiological expression levels and/or expression in inappropriate cell types. To overcome these issues, cell lines harboring CRISPR-edited knock-ins (KIs) of TurboID have been utilized [31]. Nonetheless, such approaches are limited by the inability to recapitulate the diversity of neuronal cell types that exist *in vivo*, and cultured cells cannot replicate many conditions governed by native neuropil interactions. To address this technological gap, we developed a new approach for iBioID experiments with in-frame TurboID fusions introduced into endogenous gene-coding regions by HiUGE genome editing (Fig. 1A). Using a highly expressed gene Tubb3 as a pilot example, we confirmed that the HiUGE-iBioID yields efficient *in vivo* KI and biotinylation in neurons across the brain (Fig. S1A).

Because ASD is driven by mutations in a large number of risk genes whose products are hypothesized to interact with each other in unknown functional ways, we next targeted 14 high-

confidence ASD-risk genes from the SFARI gene list (*Anks1b*, *Syngap1*, *Shank2*, *Shank3*, *Nckap1*, *Nbea*, *Ctnnb1*, *Lrrc4c*, *Iqsec2*, *Arhgef9*, *Ank3*, *Scn2a*, *Scn8a*, and *Hnrnpu*) expressed in neuronal compartments of the synapse, AIS, and nucleus [2,32,33]. Note that, due to packaging limits, many of these protein targets are too large to overexpress using conventional AAV methods, which further supports the need to label the endogenous copies of these genes. First, to confirm the proper localization of HiUGE-labeled targets, a highly immunogenic spaghetti monster (smFP) tag [34], similar in size to TurboID, was used to visualize fusion proteins. We found that HiUGE-labeled proteins were either properly localized to the synaptic sites, colocalizing with the Homer1 immunosignal (Fig. S1B-J, Q), or were restricted to the distinct features of the AIS and nucleus compartments (Fig. S1L-O). Having thus confirmed proper genome editing and correct fusion protein localization, we next fused each protein *in vivo* with TurboID-HA by injecting HiUGE AAV directly into Cas9 transgenic neonatal pup brains (P0-2), and biotinylated surrounding proteins by supplying biotin via intraperitoneal (i.p.) injections over 5 consecutive days starting at ~P21. Western blot analyses of the purified streptavidin-precipitations from the forebrain lysates collected at ~P26 detected epitope-tagged baits at the expected molecular masses, confirming correct TurboID fusion protein expression (Fig. S2A-G).

HiUGE-iBioID uncovers endogenous interactomes associated with ASD-risk proteins of diverse subcellular compartments

LC-MS/MS analysis of the 14 HiUGE-iBioID samples detected a total of 1115 proteins that are specifically enriched in the bait proteomes when compared to negative controls (fold change ≥ 2; FDR < 0.05) and filtered against the soluble TurboID background, with expected interactions faithfully captured (Fig. 1B, S6-19, Table S1,2). Importantly, 72% of the interactions detected were new, being absent from STRING queries (using a generous stringency interaction score of 0.15 [24,25]), likely a reflection of prior studies being conducted in non-native conditions or methods other than proximity proteomics. Gene ontology (GO) analyses revealed highly cohesive cellular functions corresponding to the known biology of the bait proteins, such as pathways associated with synaptic transmission (Anks1b, Syngap1, Shank2, Shank3, Nckap1, Nbea, Ctnnb1, Lrrc4c, Iqsec2, Arhgef9), voltage-gated channel activity (Ank3, Scn2a, Scn8a), and RNA processes (Hnrnpu), demonstrating high fidelity identification of local interactomes (Fig. S6-19). These networks are also consistent with the latest knowledge of the structures and functions of specific neuronal compartments. For example, the detection of Mical3 and Septin complexes with the AIS baits (Fig. S16-18, Table S2) echoes a previous study that suggested their roles in regulating cytoskeletal stability and polarized trafficking at the AIS [19].

Each interactome was further analyzed using Markov-clustering algorithm [35] to partition biologically relevant communities. This analysis revealed expected as well as new biologies for each bait protein. For example, neurobeachin (Nbea) (Fig. S11) is known to regulate the surface levels of ionotropic GABA and glutamate receptors as well serve as an A-Kinase (PKA) Anchoring Protein (AKAP) [36-39]. HiUGE-iBioID revealed that both the regulatory and catalytic subunits of PKA were detected. Clusters of the relevant receptors with significant enrichment for “GABA signaling pathway”, “Ionotropic glutamate receptor signaling”, and “Regulation of AMPA receptor activity” were also found, consistent with Nbea’s known functions. Of note, the “GABA signaling pathway” cluster was also significantly enriched for the terms “ASD” and “epilepsy”, both disorders Nbea is implicated in [40, 41]. Interestingly, GABA and glutamate surface levels are thought to be modulated by distinct pathways influenced by Nbea [42]. Consistent with this, clusters enriched for “AP-type membrane coat adaptor complex”, “COPI vesicle coat”, and “TRAPP complex” were discovered, suggesting these may be the trafficking processes that Nbea modulates. Moreover, the analysis suggests Nbea may influence other signaling pathways that have yet to be appreciated, including “G-protein-coupled GABA receptors” and “Voltage-gated potassium channel complexes”. Future experimental analyses will be needed to test these new interactome-derived hypotheses.

We also sought to determine if the interactomes showed a significant overlap with differentially expressed genes (DEGs) from specific cell types in ASD patients by cross-referencing a recent human tissue single-cell genomics study [7]. HiUGE-iBioID networks showed the highest level of overlap with ASD DEGs in layer 2/3 (L2/3) excitatory neurons (Fig. S3A), consistent with the study suggesting L2/3 neurons are significantly affected in ASD [7]. Notably, most networks were also significantly enriched for other ASD-risk genes (SFARI, and Fu et al., 2022 [26] - a recent large cohort study, Fig. 1C, S3B), demonstrating the convergence of ASD genetic susceptibilities at the protein interaction level in previously unknown ways. Larger ASD cohorts are needed to determine the potential genome-wide significance of moderate-risk genes [29]. Importantly, the HiUGE-iBioID networks of high-confidence ASD proteins formed physical communities with other proteins of moderate ASD risk (Table S2, overlap tab). These genes represent a resource of ASD candidates that likely should be prioritized in future studies of genetic contribution due to their possible functional interactions with high-confidence ASD genes. Furthermore, the interactomes contain numerous potentially druggable targets, including ~ 40 kinases and phosphatases, 6 of which are encoded by SFARI genes: Dyrk1a, Cdkl5, Pak1, Pak2, Ocrl, and Cask.

Finally, similarity clustering of the bait interactomes revealed subgroups that largely segregated according to their expected cellular compartments (Fig. 1D). Networks of the overlapping interactomes between two baits with the most significant SFARI gene enrichment (Syngap1 and Anks1b) and three similarity-clustered AIS baits (Ank3, Scn2a, and Scn8a)

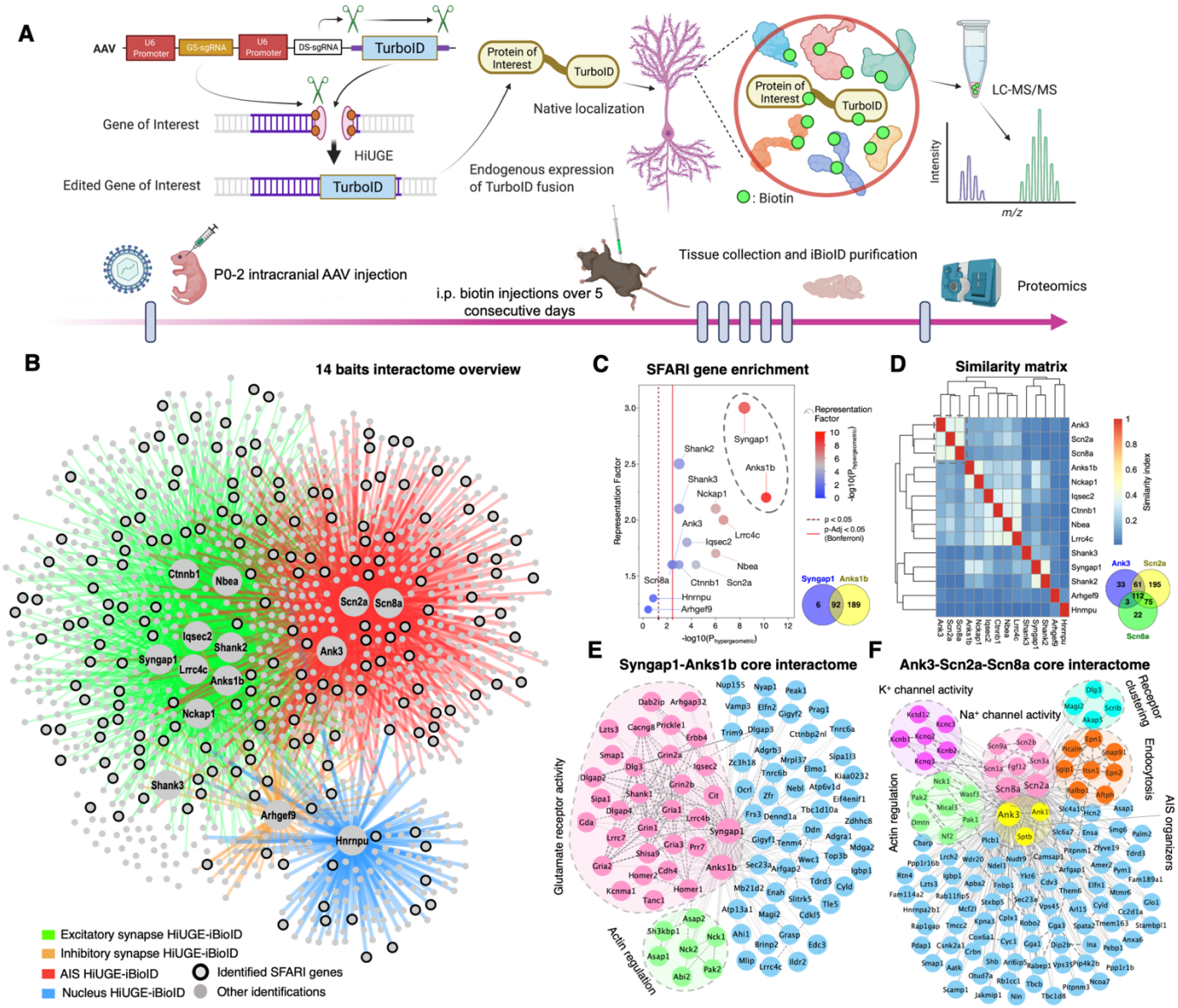


Fig. 1. HiUGE-iBioID reveals endogenous interactomes of 14 ASD-risk proteins

(**A**) Schematic illustration of HiUGE-iBioID and its workflow. GS-sgRNA: gene-specific-gRNA. DS-sgRNA: donor-specific-gRNA. (**B**) Overview of 14 interactomes that segregates according to expected bait functions. (**C**) Enrichment analysis of overlapping SFARI genes using hypergeometric probability. (**D**) Interactome clustering based on a similarity matrix. (**E**) Core interactome between Syngap1 and Anks1b that show the most significant SFARI gene overlap. (**F**) Core interactome amongst Ank3, Scn2a, and Scn8a that are clustered based on similarity.

revealed shared pathways that highlight glutamate receptor and voltage-gated channel activities respectively, and a module that is involved in actin regulation in both networks (Fig. 1E-F). These interactomes may contain proteins that function together with the ASD-associated baits, modulation of which may shine light on the complex genetic risk of ASD or be potential inroads for normalizing relevant phenotypes.

Syngap1 mutations lead to reshaping of its synaptic interactome and loss of Anks1b binding

We hypothesized that protein interactome discovery could inform genetic interactions relevant to the phenotypes underlying ASD. To test this proposition, we focused on the intersection between Syngap1 and Anks1b since of all Syngap1-interacting proteins identified by HiUGE-iBioID, 94% were also found in the Anks1b interactome, indicating that they likely reside within the same supramolecular complex. Markov cluster (MCL) analysis of their overlapping interactome identified a large PPI community (32 of 92 proteins), including Syngap1 and Anks1b, that is significantly enriched for pathways of “autistic disorder” and “glutamate receptor signaling”, indicating that this shared cluster may regulate excitatory synaptic transmission in ASD. Prior studies also indicated that Anks1b and Syngap1 both regulate synaptic activity and plasticity through NMDA-type glutamate receptors [43,44], are dispersed from the PSD in response to synaptic activity in a CaMKII-dependent manner [45-47], and are

associated with similar ASD-like phenotypes in haploinsufficiency models [48,49]. Hence, we sought to test whether Syngap1 and AnkS1b functionally interact in driving neuronal phenotypes.

We first analyzed the synaptic proteomes in wild-type (WT) and Syngap1 heterozygous (Syngap1-Het) mice [50-52] (Fig. 2A) to determine if Anks1b was influenced by haploinsufficiency of Syngap1. Fractionation steps were performed to isolate the synaptosomes [21,53,54] from the cortex and striatum. In Syngap1-Het mice, Anks1b was depleted (~60%) to a level similar as for Syngap1 in both cortical and striatal synaptosomes of Syngap1-Het mice (Fig. 2B, Table S3), supporting the notion of a very close functional interplay between Syngap1 and Anks1b not only in typical physiology, but also in ASD.

As truncating mutations of Syngap1 confer significant genetic risk for ASD [44,49], we next asked whether these mutations predicted to be pathogenic might lead to perturbation of its interactors, including Anks1b. When expressed in HEK293T cells, C-terminal truncation of human SYNGAP1 at amino acids (a.a.) 730 (C1-Trunc) and a.a. 848 (C2-Trunc) completely abolished binding with ANKS1B. In contrast, the interaction was retained with truncation at a.a. 1181 (C3-Trunc) or a truncation missing the N-term a.a. 2-361 (N-Trunc). These results suggest that the sequence surrounding the disordered region of SYNGAP1 (a.a. 848-1181) is crucial for the interaction with ANKS1B (Fig. S4). A pathological mutation found in a human patient (SYNGAP1-c.2214_2217del [55]), which results in a frame-shift and premature truncation in this region, also abolished the SYNGAP1-ANKS1B interaction (Fig. 2C). We next asked if Syngap1 truncation led to the remodeling of its interactome *in vivo*. Endogenously expressed smFP-labeled Syngap1 truncated at exon 13 (Syngap1:1-744-smFP) retained synaptic localization in cultured neurons, although the presence of mis-localization in the soma was detected as well (Fig. 2D). HiUGE-iBioID revealed that although some interactors were preserved in Syngap1:1-744 (e.g., Dlg3, Shisa9, Prickle1), most of its synaptic interactions were abolished, including Anks1b (Fig. 2E-G, Table S1). Thus, we confirmed that the region identified by immunoprecipitation using HEK293T cells is essential for the Syngap1-Anks1b interactions *in vivo* and that remodeling of synaptic protein complexes may be a mechanism associated with synaptopathy in Syngap1 loss-of-function (LOF) mutations.

Depletion of Ansk1b exacerbates electrophysiological abnormalities associated with Syngap1 LOF during *in vitro* neural development

We next sought to determine whether Syngap1 and Anks1b functionally interact by asking if further depletion of Anks1b would ameliorate or aggravate the phenotypes found in Syngap1 LOF mutants (Fig. 2H). AAV-mediated CRISPR disruption targeting Syngap1 at exon 13 and the first common exon of Anks1b (exon 15) led to a profound loss of their total protein (Fig. 2I). Using a multielectrode array (MEA) system to monitor electrophysiological activities longitudinally during neurodevelopment, we observed elevated firing rate and burst activities in Syngap1-gRNA treated neurons at DIV11, but not at DIV8 or DIV14 (Fig. 2J-M). This result phenocopies a previous report demonstrating that Syngap1 LOF atypically accelerates network activity specifically during synaptogenesis periods[56]. Interestingly, further depletion of Anks1b exacerbated the abnormal heightened neural activity linked to Syngap1 mutation (Fig. 2J-M), primarily during the period of synaptogenesis (DIV 11). Taken together, the data support a model in which Syngap1 and Anks1b physically and functionally interact within a common pathway to regulate neuronal activity development. Based on the phenotype and common biological pathways found in the overlapping interactomes, this effect may be due to the altered developmental trajectory of the glutamate receptor module found in both networks.

HiUGE-iBioID and spatial co-perturbation proteomics reveal targets that can restore spontaneous activity of neurons harboring a patient-derived Scn2a$^{+/R102Q}$ mutation.

Another critical rationale for interrogating endogenous interactomes is the prospect of identifying proteins functioning with the ASD-implicated targets that can be modulated to buffer or normalize phenotypes. Such candidates could be relevant targets for future drug development. To test this possibility, we focused on Scn2a, mutations of which are one of the most highly significant for association with ASD [3,57-59]. In recent years, whole exome sequencing (WES) from ASD patients identified a missense mutation, SCN2A-p.R102Q, that was previously reported [60,61] but uncharacterized. Clinical observations of one patient with this mutation included meeting DSM-5 diagnostic criteria for ASD (i.e., qualitative differences in social communication and the presence of restrictive interests/repetitive behaviors), minimal use of spoken language, disruptive and impulsive behaviors, sleep problems, and gastrointestinal concerns (reflux and constipation). A neurological exam including EEG did not reveal evidence of epilepsy. To test the mechanistic linkage of this mutation to ASD, we generated a Scn2a point-mutant heterozygous mouse model (*Scn2a$^{+/R102Q}$*) (Fig. 3A). Behavioral testing of the *Scn2a$^{+/R102Q}$* mice found that compared to WT littermates, the mutants presented with hyperactivity, reduced anxiety, excessive repetitive behaviors in self-grooming and the hole-board test, and reduced ultrasonic vocalizations (USV) during social interactions (Fig. 3B). It is known that mutations in Scn2a can result in either infantile epileptic encephalopathy (IEE) or ASD, depending on the nature of the mutational effect (gain-or-loss of function) [62]. Cortical neurons cultured from *Scn2a$^{+/R102Q}$* mice exhibited a significant attenuation in neural firing activity (~40% decrease at DIV14; $p < 0.001$, Fig. 3D), consistent with a role in ASD but not IEE. Next, we analyzed the spatial proteome of these

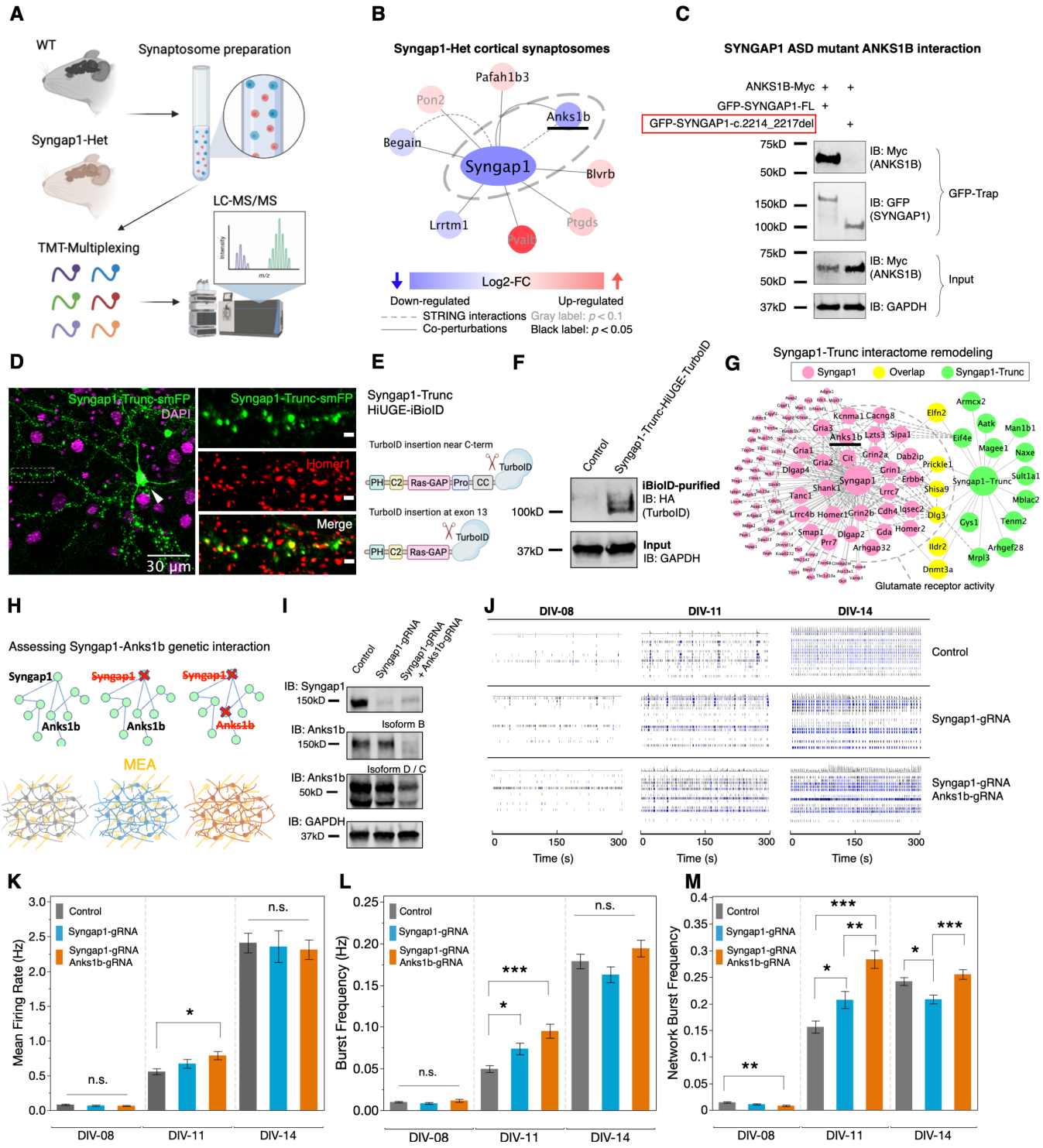


Fig. 2. Proteomic co-perturbation and functional convergence of Syngap1 and Anks1b

(**A**) Schematic illustration of the quantitative proteomic characterization of Syngap1-Het synaptosomes. (**B**) Proteomic alterations identified in the Syngap1-Het synaptosome. (**C**) Co-immunoprecipitation result showing loss of interaction with ANKS1B in frame-shifting c.2214_2217del SYNGAP1 mutation. (**D**) HiUGE labeling of truncated Syngap1 at exon 13 shows synaptic localization (boxed region) and aberrant somatic mis-localization (arrowhead). (**E**) Schematic illustration of labeling truncated Syngap1 with TurboID by targeting exon 13. (**F**) Western blot showing TurboID-HA labeled Syngap1 truncation at the expected molecular mass. (**G**) An interactome comparison showing loss of synaptic interactors, including Anks1b, in Syngap1 truncation. (**H**) Schematics assessing phenotypes of the Syngap1-Anks1b genetic interaction. (**I**) Western blot confirming disruption of Syngap1 and Anks1b expression. (**J**) Representative raster plots of neural activities at DIV-08, 11, and 14. (**K-M**) Neural metrics showing further depletion of Anks1b exacerbates the development of precocious neural activity associated with Syngap1-LOF. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; n.s.: non-significant. One-way ANOVA followed by *post-hoc* Tukey HSD tests (n = 36 wells).

mutants versus wildtype mice using fractionation steps of whole brain tissue adapted from the Localisation of Organelle Proteins by Isotope Tagging after Differential ultraCentrifugation (LOPIT-DC) procedure [63], as we have previously published [64]. Interestingly, we identified a down-regulation of a protein cluster associated with voltage-gated sodium channel (VGSC) activity in Scn2a$^{+/R102Q}$. This VGSC modulatory cluster included candidates intersecting with the Scn2a HiUGE-iBioID interactome: an auxiliary β subunit Scn1b [65], and an intracellular VGSC modulator, Fgf12 [66] (Fig. 3C; Table S3).

Based on the finding of the downregulated VGSC modulatory cluster, a rescue strategy to restore the electrophysiological deficits in Scn2a$^{+/R102Q}$ neurons was devised. First, lentiviral-mediated CRISPR-activation (CRISPRa) was used to upregulate endogenous Scn2a expression (Fig. S5A). This strategy has recently been used to phenotypically rescue Scn2a haploinsufficiency [67] but has not yet been tested in neurons harboring patient-derived missense mutations. This treatment transcriptionally activates Scn2a expression; however, it is expected to amplify both the WT and mutant allele. RT-PCR data from WT and Scn2a$^{+/R102Q}$ neurons showed normal transcript levels (Fig S5C), suggesting the reduced protein levels in the mutant are likely due to post-transcriptional effects such as protein destabilization. Thus, it was unclear if Scn2a-CRISPRa alone would fully rescue phenotypes, as the effect could potentially be diluted by the mutant allele. Indeed, CRISPRa treatment only partially rescued the Scn2a$^{+/R102Q}$ phenotype (Fig 3D, purple bars). A combinatory ("combo") treatment strategy was then tested by supplying additional SCN1B and FGF12 via AAV-mediated expression to augment all three down-regulated intersecting proteins identified in the VSGC modulatory cluster. MEA analysis revealed a full restoration of neural firing activity metrics by the combo treatment at DIV14 (Fig. 3D, S5B), confirming the hypothesis that this modulatory cluster is crucial for not only the Scn2a$^{+/R102Q}$ neuropathology but also for its phenotypic rescue. These results strongly indicate that intersectional proteomics between HiUGE-iBioID and spatial co-perturbation can be an informative approach to discover novel avenues for the functional rescue of disease phenotypes and potential therapeutic targets.

Discussion

Here we report a strategy combining the advantages of HiUGE and iBioID to resolve native interactomes associated with 14 ASD-associated proteins. The combination of the interaction data presented for Syngap1 and Anks1b and the Scn2a rescue results validate the HiUGE-iBioID method for discovering the functional links between ASD genetics and proteomics. The findings also emphasize an effective proteomic-driven systems-biology approach to discover hidden molecular etiology and potential treatment targets.

Compared to antibody-dependent or vector-driven overexpression methods, including recent reports of protein interactions in ASD [68,69], HiUGE-iBioID has four key benefits. First, the bait protein is expressed from the endogenous promoter with native cell-type specificity preserved at physiological levels. Second, the cells expressing the bait protein are within the context of the tissue, obviating perturbations to their native environment essential for development and cell physiology that can occur *in vitro*. Third, the selection of bait-protein is not limited by viral packaging capacity or available antibodies. Finally, the experiments are easily scalable in terms of time and resources when compared to the alternative of traditional transgenic animals. We have found that forebrain tissue collected from as few as two mice is sufficient as one biological replicate. Therefore, a typical mouse litter (~ 8 pups) is enough to analyze a bait interactome in biological triplicates. We anticipate this method can be optimized even further for large-scale applications, especially when coupled with multiplexed mass-spectrometry techniques such as isobaric labeling [70,71]. We envision further developments to our approach will include adapting *in silico* structure prediction tools (e.g. AlphaFold [72]), and intron-targeting strategies[73] to minimize potential disturbances introduced by TurboID fusions to native proteins.

A striking finding is that diverse ASD genes physically form protein interaction networks with each other and are often co-perturbed in single-gene models of ASD. This suggests divergent genetic mutations converge at the protein level to drive ASD neurobiology and provides a working paradigm to prioritize co-regulated candidates in physiology and atypical brain development. Furthermore, we demonstrate that high-confidence ASD genes interact with other lower-confidence ASD genes at the proteome level. Indeed, three proteins detected in the interactome dataset (Itsn1, associated with Scn2a, Scn8a, Ank3, and Shank3 baits; Nav3, associated with Scn2a bait; and Hnrnpul2, associated with Hnrnpu bait) were discovered subsequently as new ASD risk genes of genome-wide significance [29] during the preparation of this manuscript. These results highlight the possibility that additional genes - whose significance in ASD are as yet unknown - exist in the dataset. Thus, HiUGE-iBioID may be useful to prioritize lower-confidence ASD genes, which could either play a role in regulating core ASD drivers or serve as novel targets for pharmacological developments. In addition, we expect the endogenous ASD interactome data will stimulate a new impetus for predictive modeling of ASD genetics [74,75] and cell-type specific analyses harnessing single-cell proteomics [76-78].

Informed by the ASD interactome and proteomic co-perturbation results, we focused on a tightly co-regulated pair, Syngap1 and Anks1b, and elucidated a molecular mechanism and the functional significance of their interaction. The additive effects of Syngap1 and Anks1b deficits seen in the aberrant neural activity indicate that the downregulation of Ansk1b in Syngap1-Het mice is not a compensatory effect, but rather comprises a mechanism of disrupted brain development. The disordered region of Syngap1, identified as a critical domain mediating binding with Anks1b, contains proline-rich

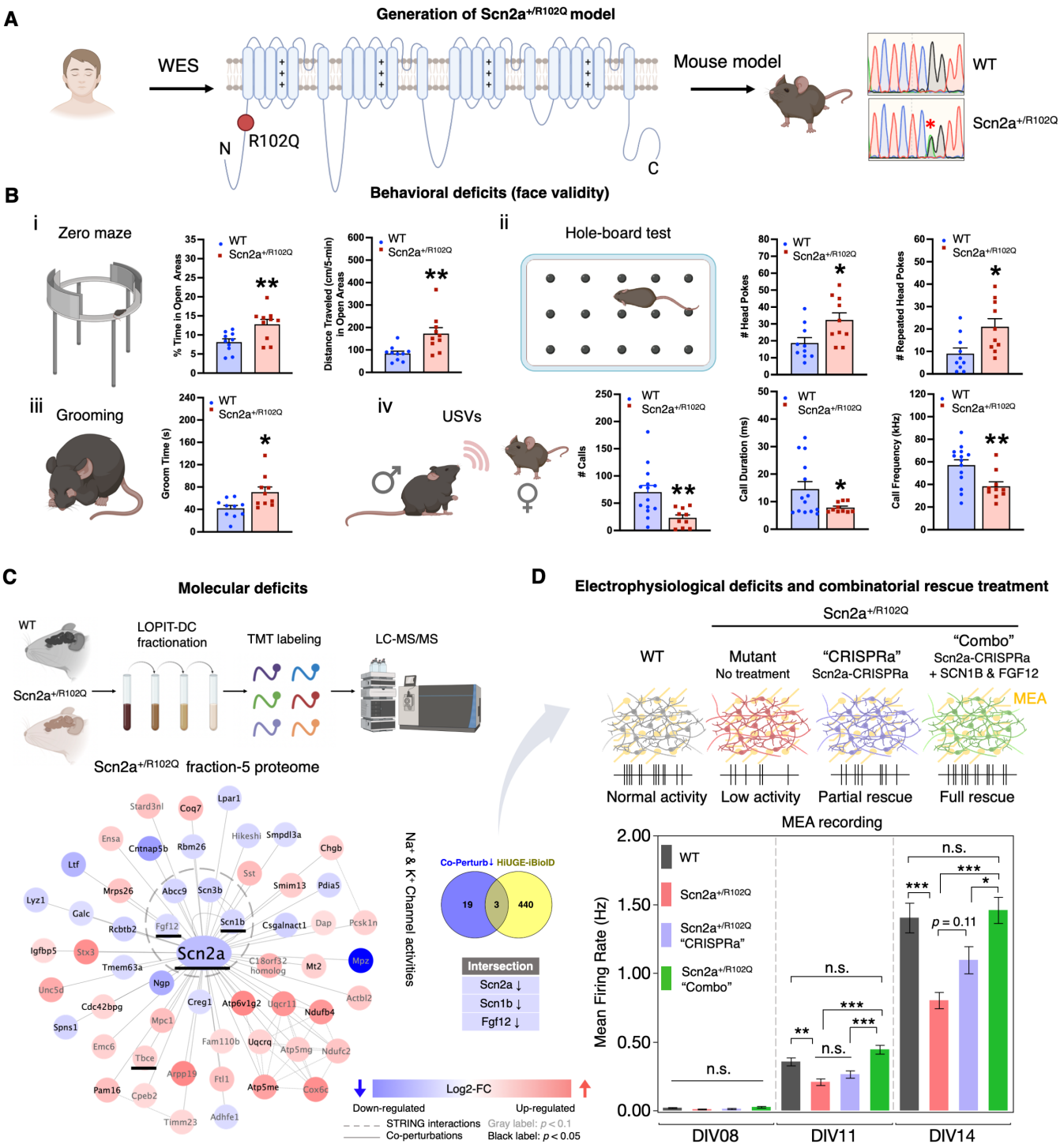


Fig. 3. Intersectional proteomics reveal hidden molecular etiology of a patient-derived Scn2a mutation

(**A**) Generation of a mouse model based on a clinically identified Scn2a missense mutation (R102Q) in ASD patients. WES: whole exome sequencing. (**B**) Behavioral face-validity of Scn2a$^{+/R102Q}$ mutants characterized by (**i**) zero maze as the percent time and distance traveled in the open areas, (**ii**) hole-board test as numbers of head pokes and repeated head-pokes, (**iii**) self-grooming, and (**iv**) ultrasonic vocalizations (USVs) as numbers of calls, call durations, and call frequencies during social interaction. *: $p < 0.05$, **: $p < 0.01$; independent samples t-tests, two-tailed. No difference was detected in the metrics of pre-social (baseline) responses. Statistics were summarized in Table S4. (**C**) Spatial proteomics reveals co-perturbations in Scn2a$^{+/R102Q}$ mutants. MCL analysis discovered a key cluster associated with voltage-gated channel activity that is down-regulated, including three targets that intersect with the Scn2a HiUGE-iBioID interactome. (**D**) Scn2a$^{+/R102Q}$ mutant neurons show attenuated activity with the MEA. Scn2a-CRISPRa treatment and a "Combo" treatment with additional expression of SCN1B and FGF12 show differential efficacy in restoring neural activity deficits. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; n.s.: non-significant. One-way ANOVA followed by *post-hoc* Tukey HSD tests (n = 48 wells).

stretches, a poly-histidine motif, and is phosphorylated at multiple sites [79,80]. This architecture suggests their interaction could be under activity-dependent kinase modulation. Within the Anks1b-Syngap1 core interactome, there are two identified kinases: Pak2, and Cdkl5, both of which are likely ASD-risk genes themselves [2]. Thus, it is intriguing to query whether they play a role in regulating synaptogenesis via acting on the Syngap1/Anks1b complex, and how they might contribute to the activity-dependent shuttling of Syngap1/Anks1b during plasticity.

Perhaps our most exciting finding is the discovery of a new paradigm for restoring *in vitro* neural activity deficits associated with a patient-derived SCN2A-p.R102Q mutation. We have identified three proteins (Scn2a, Scn1b, and Fgf12) within the VGSC complex that are critical for phenotypic rescue by intersecting the Scn2a HiUGE-iBioID interactome with co-perturbation spatial-proteomics in a mouse model harboring this mutation. A qPCR analysis revealed that the abundances of Scn2a, Scn1b, and Fgf12 mRNA in cultured $Scn2a^{+/R102Q}$ neurons are comparable to WT (Fig. S5C), suggesting the downregulation of these proteins occurs at the post-transcriptional level, likely due to destabilization of VGSC supramolecular assembly. The exact mechanisms as to how the loss of a positively charged arginine residue on the N-terminal intracellular tail of Scn2a affects VGSC assembly remains to be further determined.

Both β-subunits and FGF family proteins play critical roles in maintaining neuronal excitability through regulating VGSC kinetics [65,81]. Thus, their downregulation in $Scn2a^{+/R102Q}$ are likely contributing factors to VGSC channelopathy. These results also explain why Scn2a-CRISPRa treatment alone is insufficient to restore the deficits, especially since the treatment does not differentiate the functional allele from the point-mutant allele [67]. It appears that supplying additional SCN1B and FGF12 provides the molecular environment needed to restore functional VGSC complexes and reinstate the level of spontaneous neural activity. The *in vitro* phenotypic rescue presented here will require future testing *in vivo*. Although we did not observe an overt epileptogenic effect on the MEA, these potential adverse effects must be carefully assessed in future animal studies. In addition, since VGSCs are believed to contribute to neuronal excitability and plasticity at both pre- and post-synapses [57,82], future studies are needed to further dissect their unique subcellular effects.

Together, our results show that HiUGE-iBioID provides a new "molecular microscope" to reveal native proteomes in the brain with considerable ease and effectiveness. Combined with co-perturbation proteomics, our intersectional approach offers a robust mechanism to identify and prioritize candidates for discovering new biology and potential therapeutic targets. We anticipate that the framework constructed here will stimulate future studies in other neuropsychiatric models, transforming our knowledge of proteome organization in diverse fields of cellular neurobiology.

Acknowledgments

We would like to thank Drs. Tyler Bradshaw and Jamie Courtland for performing synaptosome and fractionation proteomics, and Dr. Ramona M. Rodgriguiz and Mr. Christopher R. Means for conducting the behavioral experiments. We would like to thank Arinze Okafor for assistance with the similarity matrix plot, Greg Waitt for developing the DPMSR Proteome Discoverer Data Visualization Tool, and Daniel J. Morone for assistance with lentiviral production and titration. We would also like to thank Dr. Jiechun Zhou for her assistance with mouse husbandry. Schematic illustrations were created using a licensed BioRender.com account.

Funding

Research reported in this publication was supported by a donation to the Duke Center for Autism and Brain Development, NIH R01MH111684 (SHS), UM1HG012053 (CAG), R01MH125236 (CAG), and in part by the Office of The Director, National Institutes of Health of the National Institutes of Health under Award Number S10OD024999 (M Arthur Moseley). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data and Materials Availability

Requests for data, resources, and reagents should be directed to and will be fulfilled by the Corresponding Author, Dr. Scott Soderling (scott.soderling@duke.edu).

Competing interests

SS and YG have a patent related to the HiUGE technology. The intellectual property was licensed to CasTag Biosciences. SS is a founder of CasTag Biosciences; Duke University as an institution holds equity in CasTag Biosciences. CAG is an inventor on patents and patent applications related to CRISPR-based gene activation, is a co-founder of Tune Therapeutics, Locus Biosciences, and Element Genomics, and is an advisor to Sarepta Therapeutics. GD Dr. Dawson is on the Scientific Advisory Boards of Akili Interactive, Inc, and Tris Pharma, is a consultant to Apple, Gerson Lehrman Group, and Guidepoint Global, Inc. GD has developed autism-related technology, data, and/or products that have been licensed to Apple, Inc. and Cryocell, Inc. and Dawson and Duke University have benefited financially. JDB is a consultant for Bridgebio.

References and Notes

1 Lord, C. *et al.* Autism spectrum disorder. *Nat Rev Dis Primers* **6**, 5, doi:10.1038/s41572-019-0138-4 (2020).
2 Banerjee-Basu, S. & Packer, A. SFARI Gene: an evolving database for the autism research community. *Dis Model Mech* **3**, 133-135, doi:10.1242/dmm.005439 (2010).
3 Satterstrom, F. K. *et al.* Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **180**, 568-584 e523, doi:10.1016/j.cell.2019.12.036 (2020).
4 Wang, T. *et al.* Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. *Nat Commun* **11**, 4932, doi:10.1038/s41467-020-18723-y (2020).
5 Zoghbi, H. Y. Postnatal neurodevelopmental disorders: meeting at the synapse? *Science* **302**, 826-830, doi:10.1126/science.1089071 (2003).
6 Ayhan, F. & Konopka, G. Regulatory genes and pathways disrupted in autism spectrum disorders. *Prog Neuropsychopharmacol Biol Psychiatry* **89**, 57-64, doi:10.1016/j.pnpbp.2018.08.017 (2019).
7 Velmeshev, D. *et al.* Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685-689, doi:10.1126/science.aav8130 (2019).
8 Menashe, I., Grange, P., Larsen, E. C., Banerjee-Basu, S. & Mitra, P. P. Co-expression profiling of autism genes in the mouse brain. *PLoS Comput Biol* **9**, e1003128, doi:10.1371/journal.pcbi.1003128 (2013).
9 Voineagu, I. *et al.* Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474**, 380-384, doi:10.1038/nature10110 (2011).
10 Quesnel-Vallieres, M., Weatheritt, R. J., Cordes, S. P. & Blencowe, B. J. Autism spectrum disorder: insights into convergent mechanisms from transcriptomics. *Nat Rev Genet* **20**, 51-63, doi:10.1038/s41576-018-0066-2 (2019).
11 Donato, A., Kagias, K., Zhang, Y. & Hilliard, M. A. Neuronal sub-compartmentalization: a strategy to optimize neuronal function. *Biol Rev Camb Philos Soc* **94**, 1023-1037, doi:10.1111/brv.12487 (2019).
12 Terenzio, M., Schiavo, G. & Fainzilber, M. Compartmentalized Signaling in Neurons: From Cell Biology to Neuroscience. *Neuron* **96**, 667-679, doi:10.1016/j.neuron.2017.10.015 (2017).
13 De Rubeis, S. *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**, 209-215, doi:10.1038/nature13772 (2014).
14 Santini, E. & Klann, E. Reciprocal signaling between translational control pathways and synaptic proteins in autism spectrum disorders. *Sci Signal* **7**, re10, doi:10.1126/scisignal.2005832 (2014).
15 Grant, S. G. Synaptopathies: diseases of the synaptome. *Curr Opin Neurobiol* **22**, 522-529, doi:10.1016/j.conb.2012.02.002 (2012).
16 Fujitani, M., Otani, Y. & Miyajima, H. Pathophysiological Roles of Abnormal Axon Initial Segments in Neurodevelopmental Disorders. *Cells* **10**, doi:10.3390/cells10082110 (2021).
17 Kruth, K. A., Grisolano, T. M., Ahern, C. A. & Williams, A. J. SCN2A channelopathies in the autism spectrum of neuropsychiatric disorders: a role for pluripotent stem cells? *Mol Autism* **11**, 23, doi:10.1186/s13229-020-00330-9 (2020).
18 Uezu, A. *et al.* Identification of an elaborate complex mediating postsynaptic inhibition. *Science* **353**, 1123-1129, doi:10.1126/science.aag0821 (2016).
19 Hamdan, H. *et al.* Mapping axon initial segment structure and function by multiplexed proximity biotinylation. *Nat Commun* **11**, 100, doi:10.1038/s41467-019-13658-5 (2020).
20 Bayes, A. *et al.* Characterization of the proteome, diseases and evolution of the human postsynaptic density. *Nat Neurosci* **14**, 19-21, doi:10.1038/nn.2719 (2011).
21 Distler, U. *et al.* In-depth protein profiling of the postsynaptic density from mouse hippocampus using data-independent acquisition proteomics. *Proteomics* **14**, 2607-2613, doi:10.1002/pmic.201300520 (2014).
22 Gao, Y. *et al.* Plug-and-Play Protein Modification Using Homology-Independent Universal Genome Engineering. *Neuron* **103**, 583-597 e588, doi:10.1016/j.neuron.2019.05.047 (2019).
23 Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* **36**, 880-887, doi:10.1038/nbt.4201 (2018).
24 Szklarczyk, D. *et al.* The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* **49**, D605-D612, doi:10.1093/nar/gkaa1074 (2021).
25 Snel, B., Lehmann, G., Bork, P. & Huynen, M. A. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res* **28**, 3442-3444, doi:10.1093/nar/28.18.3442 (2000).
26 Fu, J. M. *et al.* Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. *Nat Genet*, doi:10.1038/s41588-022-01104-0 (2022).
27 Beyreli, I., Karakahya, O. & Cicek, A. E. DeepND: Deep multitask learning of gene risk for comorbid neurodevelopmental disorders. *Patterns* **3**, doi:10.1016/j.patter.2022.100524 (2022).
28 Seiffert, S. *et al.* Modulating effects of FGF12 variants on NaV1.2 and NaV1.6 associated with Developmental and Epileptic Encephalopathy and Autism Spectrum Disorder. *bioRxiv* doi:10.1101/2021.12.03.471090 (2021).
29 Zhou, X. *et al.* Integrating de novo and inherited variants in 42,607 autism cases identifies mutations in new moderate-risk genes. *Nat Genet*, doi:10.1038/s41588-022-01148-2 (2022).
30 O'Neil, S. D. *et al.* Action potential-coupled Rho GTPase signaling drives presynaptic plasticity. *Elife* **10**, doi:10.7554/eLife.63756 (2021).
31 Stockhammer, A. *et al.* When less is more – Endogenous tagging with TurboID as a tool to study the native interactome of adaptor protein complexes. *bioRxiv*, doi:10.1101/2021.11.19.469212 (2022).

32 pfeliciano@simonsfoundation.org, S. C. E. a. & Consortium, S. SPARK: A US Cohort of 50,000 Families to Accelerate Autism Research. *Neuron* **97**, 488-493, doi:10.1016/j.neuron.2018.01.015 (2018).
33 Simons Vip, C. Simons Variation in Individuals Project (Simons VIP): a genetics-first approach to studying autism spectrum and related neurodevelopmental disorders. *Neuron* **73**, 1063-1067, doi:10.1016/j.neuron.2012.02.014 (2012).
34 Viswanathan, S. *et al.* High-performance probes for light and electron microscopy. *Nat Methods* **12**, 568-576, doi:10.1038/nmeth.3365 (2015).
35 Enright, A. J., Van Dongen, S. & Ouzounis, C. A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* **30**, 1575-1584, doi:10.1093/nar/30.7.1575 (2002).
36 Repetto, D. *et al.* Molecular Dissection of Neurobeachin Function at Excitatory Synapses. *Front Synaptic Neurosci* **10**, 28, doi:10.3389/fnsyn.2018.00028 (2018).
37 Gromova, K. V. *et al.* Neurobeachin and the Kinesin KIF21B Are Critical for Endocytic Recycling of NMDA Receptors and Regulate Social Behavior. *Cell Rep* **23**, 2705-2717, doi:10.1016/j.celrep.2018.04.112 (2018).
38 Wang, X. *et al.* Neurobeachin: A Protein Kinase A-Anchoring,beige/Chediak-Higashi Protein Homolog Implicated in Neuronal Membrane Traffic. *The Journal of Neuroscience* **20**, 8551-8565, doi:10.1523/jneurosci.20-23-08551.2000 (2000).
39 Nair, R. *et al.* Neurobeachin regulates neurotransmitter receptor trafficking to synapses. *J Cell Biol* **200**, 61-80, doi:10.1083/jcb.201207113 (2013).
40 Mulhern, M. S. *et al.* NBEA: Developmental disease gene with early generalized epilepsy phenotypes. *Ann Neurol* **84**, 788-795, doi:10.1002/ana.25350 (2018).
41 Castermans, D. *et al.* The neurobeachin gene is disrupted by a translocation in a patient with idiopathic autism. *J Med Genet* **40**, 352-356, doi:10.1136/jmg.40.5.352 (2003).
42 Farzana, F. *et al.* Neurobeachin Regulates Glutamate- and GABA-Receptor Targeting to Synapses via Distinct Pathways. *Mol Neurobiol* **53**, 2112-2123, doi:10.1007/s12035-015-9164-8 (2016).
43 Tindi, J. O. *et al.* ANKS1B Gene Product AIDA-1 Controls Hippocampal Synaptic Transmission by Regulating GluN2B Subunit Localization. *J Neurosci* **35**, 8986-8996, doi:10.1523/JNEUROSCI.4029-14.2015 (2015).
44 Gamache, T. R., Araki, Y. & Huganir, R. L. Twenty Years of SynGAP Research: From Synapses to Cognition. *J Neurosci* **40**, 1596-1605, doi:10.1523/JNEUROSCI.0420-19.2020 (2020).
45 Dosemeci, A., Toy, D., Burch, A., Bayer, K. U. & Tao-Cheng, J. H. CaMKII-mediated displacement of AIDA-1 out of the postsynaptic density core. *FEBS Lett* **590**, 2934-2939, doi:10.1002/1873-3468.12334 (2016).
46 Yang, Y., Tao-Cheng, J. H., Reese, T. S. & Dosemeci, A. SynGAP moves out of the core of the postsynaptic density upon depolarization. *Neuroscience* **192**, 132-139, doi:10.1016/j.neuroscience.2011.06.061 (2011).
47 Araki, Y., Zeng, M., Zhang, M. & Huganir, R. L. Rapid dispersion of SynGAP from synaptic spines triggers AMPA receptor insertion and spine enlargement during LTP. *Neuron* **85**, 173-189, doi:10.1016/j.neuron.2014.12.023 (2015).
48 Carbonell, A. U. *et al.* Haploinsufficiency in the ANKS1B gene encoding AIDA-1 leads to a neurodevelopmental syndrome. *Nat Commun* **10**, 3529, doi:10.1038/s41467-019-11437-w (2019).
49 Berryer, M. H. *et al.* Mutations in SYNGAP1 cause intellectual disability, autism, and a specific form of epilepsy by inducing haploinsufficiency. *Hum Mutat* **34**, 385-394, doi:10.1002/humu.22248 (2013).
50 Kim, J. H., Lee, H.-K., Takamiya, K. & Huganir, R. L. The Role of Synaptic GTPase-Activating Protein in Neuronal Development and Synaptic Plasticity. *The Journal of Neuroscience* **23**, 1119-1124, doi:10.1523/jneurosci.23-04-01119.2003 (2003).
51 Nakajima, R. *et al.* Comprehensive behavioral analysis of heterozygous Syngap1 knockout mice. *Neuropsychopharmacol Rep* **39**, 223-237, doi:10.1002/npr2.12073 (2019).
52 Komiyama, N. H. *et al.* SynGAP Regulates ERK/MAPK Signaling, Synaptic Plasticity, and Learning in the Complex with Postsynaptic Density 95 and NMDA Receptor. *The Journal of Neuroscience* **22**, 9721-9732, doi:10.1523/jneurosci.22-22-09721.2002 (2002).
53 Carlin, R. K., Grab, D. J., Cohen, R. S. & Siekevitz, P. Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities. *J Cell Biol* **86**, 831-845, doi:10.1083/jcb.86.3.831 (1980).
54 Schmeisser, M. J. *et al.* Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature* **486**, 256-260, doi:10.1038/nature11015 (2012).
55 Mignot, C. *et al.* Genetic and neurodevelopmental spectrum of SYNGAP1-associated intellectual disability and epilepsy. *J Med Genet* **53**, 511-522, doi:10.1136/jmedgenet-2015-103451 (2016).
56 Llamosas, N. *et al.* SYNGAP1 Controls the Maturation of Dendrites, Synaptic Function, and Network Activity in Developing Human Neurons. *J Neurosci* **40**, 7980-7994, doi:10.1523/JNEUROSCI.1367-20.2020 (2020).
57 Spratt, P. W. E. *et al.* The Autism-Associated Gene Scn2a Contributes to Dendritic Excitability and Synaptic Function in the Prefrontal Cortex. *Neuron* **103**, 673-685 e675, doi:10.1016/j.neuron.2019.05.037 (2019).
58 Ben-Shalom, R. *et al.* Opposing Effects on NaV1.2 Function Underlie Differences Between SCN2A Variants Observed in Individuals With Autism Spectrum Disorder or Infantile Seizures. *Biol Psychiatry* **82**, 224-232, doi:10.1016/j.biopsych.2017.01.009 (2017).
59 Sanders, S. J. *et al.* De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237-241, doi:10.1038/nature10945 (2012).
60 Wang, T. *et al.* De novo genic mutations among a Chinese autism spectrum disorder cohort. *Nat Commun* **7**, 13316, doi:10.1038/ncomms13316 (2016).
61 Callaghan, D. B. *et al.* Whole genome sequencing and variant discovery in the ASPIRE autism spectrum disorder cohort. *Clin Genet* **96**, 199-206, doi:10.1111/cge.13556 (2019).

62 Berecki, G. *et al.* Functional correlates of clinical phenotype and severity in recurrent SCN2A variants. *Commun Biol* **5**, 515, doi:10.1038/s42003-022-03454-1 (2022).
63 Geladaki, A. *et al.* Combining LOPIT with differential ultracentrifugation for high-resolution spatial proteomics. *Nat Commun* **10**, 331, doi:10.1038/s41467-018-08191-w (2019).
64 Courtland, J. L. *et al.* Genetic disruption of WASHC4 drives endo-lysosomal dysfunction and cognitive-movement impairments in mice and humans. *Elife* **10**, doi:10.7554/eLife.61590 (2021).
65 Brackenbury, W. J. & Isom, L. L. Na Channel beta Subunits: Overachievers of the Ion Channel Family. *Front Pharmacol* **2**, 53, doi:10.3389/fphar.2011.00053 (2011).
66 Turner, C. A., Watson, S. J. & Akil, H. The fibroblast growth factor family: neuromodulation of affective behavior. *Neuron* **76**, 160-174, doi:10.1016/j.neuron.2012.08.037 (2012).
67 Tamura, S. *et al.* CRISPR activation rescues abnormalities in SCN2A haploinsufficiency-associated autism spectrum disorder. *bioRxiv*, doi:10.1101/2022.03.30.486483 (2022).
68 Murtaza, N. *et al.*, doi:10.1101/2022.01.17.476220 (2022).
69 Pintacuda, G. *et al.*, doi:10.1101/2021.10.07.21264575 (2022).
70 Thompson, A. *et al.* Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **75**, 1895-1904, doi:10.1021/ac0262560 (2003).
71 Bachor, R., Waliczek, M., Stefanowicz, P. & Szewczuk, Z. Trends in the Design of New Isobaric Labeling Reagents for Quantitative Proteomics. *Molecules* **24**, doi:10.3390/molecules24040701 (2019).
72 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
73 Zhong, H. *et al.* High-fidelity, efficient, and reversible labeling of endogenous proteins using CRISPR-based designer exon insertion. *Elife* **10**, doi:10.7554/eLife.64911 (2021).
74 Lin, Y., Afshar, S., Rajadhyaksha, A. M., Potash, J. B. & Han, S. A Machine Learning Approach to Predicting Autism Risk Genes: Validation of Known Genes and Discovery of New Candidates. *Front Genet* **11**, 500064, doi:10.3389/fgene.2020.500064 (2020).
75 Brueggeman, L., Koomar, T. & Michaelson, J. J. Forecasting risk gene discovery in autism with machine learning and genome-scale data. *Sci Rep* **10**, 4569, doi:10.1038/s41598-020-61288-5 (2020).
76 Schoof, E. M. *et al.* Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat Commun* **12**, 3341, doi:10.1038/s41467-021-23667-y (2021).
77 Brunner, A. D. *et al.* Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Mol Syst Biol* **18**, e10798, doi:10.15252/msb.202110798 (2022).
78 Vistain, L. F. & Tay, S. Single-Cell Proteomics. *Trends Biochem Sci* **46**, 661-672, doi:10.1016/j.tibs.2021.01.013 (2021).
79 Walkup, W. G. t. *et al.* Phosphorylation of synaptic GTPase-activating protein (synGAP) by Ca2+/calmodulin-dependent protein kinase II (CaMKII) and cyclin-dependent kinase 5 (CDK5) alters the ratio of its GAP activity toward Ras and Rap GTPases. *J Biol Chem* **290**, 4908-4927, doi:10.1074/jbc.M114.614420 (2015).
80 Walkup, W. G. *et al.* A model for regulation by SynGAP-alpha1 of binding of synaptic proteins to PDZ-domain 'Slots' in the postsynaptic density. *Elife* **5**, doi:10.7554/eLife.16813 (2016).
81 Goldfarb, M. *et al.* Fibroblast growth factor homologous factors control neuronal excitability through modulation of voltage-gated sodium channels. *Neuron* **55**, 449-463, doi:10.1016/j.neuron.2007.07.006 (2007).
82 Johnson, K. W., Herold, K. F., Milner, T. A., Hemmings, H. C., Jr. & Platholi, J. Sodium channel subtypes are differentially localized to pre- and post-synaptic sites in rat hippocampus. *J Comp Neurol* **525**, 3563-3578, doi:10.1002/cne.24291 (2017).
83 Concordet, J. P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* **46**, W242-W245, doi:10.1093/nar/gky354 (2018).
84 Chan, K. Y. *et al.* Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat Neurosci* **20**, 1172-1179, doi:10.1038/nn.4593 (2017).
85 Schwammle, V., Hagensen, C. E., Rogowska-Wrzesinska, A. & Jensen, O. N. PolySTest: Robust Statistical Testing of Proteomics Data with Missing Values Improves Detection of Biologically Relevant Features. *Mol Cell Proteomics* **19**, 1396-1408, doi:10.1074/mcp.RA119.001777 (2020).
86 Hallett, P. J., Collins, T. L., Standaert, D. G. & Dunah, A. W. Biochemical fractionation of brain tissue for studies of receptor distribution and trafficking. *Curr Protoc Neurosci* **Chapter 1**, Unit 1 16, doi:10.1002/0471142301.ns0116s42 (2008).
87 Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628-2629, doi:10.1093/bioinformatics/btz931 (2020).
88 Sanson, K. R. *et al.* Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat Commun* **9**, 5416, doi:10.1038/s41467-018-07901-8 (2018).
89 Schmidt, R. *et al.* CRISPR activation and interference screens decode stimulation responses in primary human T cells. *Science* **375**, eabj4008, doi:10.1126/science.abj4008 (2022).
90 Gordon, M. G. *et al.* lentiMPRA and MPRAflow for high-throughput functional characterization of gene regulatory elements. *Nat Protoc* **15**, 2387-2412, doi:10.1038/s41596-020-0333-5 (2020).
91 Wang, X. *et al.* Synaptic dysfunction and abnormal behaviors in mice lacking major isoforms of Shank3. *Hum Mol Genet* **20**, 3093-3108, doi:10.1093/hmg/ddr212 (2011).
92 Pappas, A. L. *et al.* Deficiency of Shank2 causes mania-like behavior that responds to mood stabilizers. *JCI Insight* **2**, doi:10.1172/jci.insight.92052 (2017).

Supplementary Information

Materials and Methods

Animals

For all CRISPR-Cas9-related experiments, H11-Cas9 mice (Jackson Laboratory #28239) were used. The Syngap1-Het mouse model, originally described by Kim and colleagues [50], was a gift from Dr. Gavin Rumbaugh. The Scn2a$^{+/R102Q}$ mouse model of the human c.305G>A (p.R102Q) mutation [60,61] was created by the Duke Transgenic and Knockout Mouse Shared Resource. The Scn2a$^{+/R102Q}$ mice were generated using a heterozygous breeding scheme and genotyping was performed by sequencing the amplicon using the following primer set: Scn2a-s, acagacatggcggaaaacatgag; and Scn2a-as, agcaggagaggaaagaaagaagc. All procedures were performed with a protocol approved by the Duke University Institutional Animal Care and Use Committee in accordance with US National Institutes of Health guidelines.

Single-vector HiUGE TurboID knock-in

HiUGE plasmids were constructed based on the previously described method [22]. Briefly, a donor of HA-tagged TurboID coding sequence was flanked by DNA sequences that were specifically recognized by a synthetic donor-specific gRNA (DS-gRNA), inert to the host genome. The gene-specific gRNA (GS-gRNA) expression cassette (U6 promotor, GS-gRNA, and gRNA scaffold) was inserted in tandem to the DS-gRNA expression cassette permitting a single-vector delivery. GS-gRNAs were designed using CRISPOR [83], and a pair of 23-24mer oligonucleotides were annealed and ligated into the *SapI* site of the GS-gRNA expression cassette. The genomic target sequences for the baits were: Anks1b: cggcgggtatcagaaaatcgtgg, Syngap1: ctgggtgggctagtggtctgcgg, Shank2: aacagctgctggacagataaggg, Shank3: cgtgcgctcaggcagctggatgg, Nckap1: gcatttctcagcaacacataagg, Nbea: ggcattatgagcatcagaacagg, Ctnnb1: ttacaggtcagtatcaaaccagg, Lrrc4c: gagttcattcggatcaataacgg, Iqsec2: accacagtgctgatccggcttgg, Arhgef9: cattctggcaaaacttcagtagg, Ank3: gaagaaggaaatccggaacgtgg, Scn2a: ggacaaggggaaagatatcaggg, Scn8a: tcagggagtccaagtgctagagg, and Hnrnpu: tggagtcagcattatcaccaagg. To endogenously express soluble TurboID as a survey for background detection, 3'-UTRs of Syngap1, Scn2a, and Hnrnpu were targeted with an internal ribosome entry site (IRES)-TurboID-HA donor. The genomic target sequences were: Syngap1: aggaggtctgtgacgctgggtgg, Scn2a: agtttggcatagacctcctgagg, and Hnrnpu: aaacagtcgacttcttgtgaagg. Exon 13 of Syngap1 was targeted for Syngap1 truncation and disruption: acggactcggtctcagcccatgg.

AAV preparation

AAV was prepared following previously described methods [18,22]. Briefly, concentrated AAV virus was produced in HEK293T cells grown in six 15cm dishes by triple-transfection with 15μg HiUGE vector, 30μg pADdeltaF6, and 15μg serotype plasmid (pUCmini-iCAP-PHP.eB, a gift from Viviana Gradinaru [84], Addgene plasmid #103005). Three days following transfection, cells were lysed and virus was concentrated using an Optiprep density gradient (Sigma #D1556). Small-scale AAV virus was produced in HEK293T cells grown in 12-well plates by triple-transfection with 0.4μg HiUGE vector, 0.8μg pADdeltaF6, and 0.4μg serotype 2/1 plasmids. Three days following transfection, the virus-containing medium was filtered through Costar Spin-X columns (Sigma #8162) as previously described [22].

AAV injection and biotin injection for HiUGE-iBioID

Neonatal (P0-2) H11-Cas9 mice were intracranially injected with the purified AAV (2μL per hemisphere, PHP.eB serotype, > 10^{10} GC / μL titer). Donor vector backbone AAV (empty gRNA) was used as negative control. Approximately 3 weeks after injection, mice received daily intraperitoneal (i.p.) injections of biotin (50mg/kg) over five consecutive days. Forebrain tissue was collected one day after the final injection and snap-frozen at -80°C until purification.

HiUGE-iBioID sample purification

For each replicate, forebrain tissue from two mice was combined, homogenized, sonicated in RIPA lysis buffer (Cell Signaling #9806) supplemented with cOmplete protease inhibitor cocktail (Sigma #11873580001), and centrifuged at 16,000 xg for 30min at 4°C. The supernatant lysate was desalted with Zebra columns 7K MWCO (ThermoFisher #89894 or #89892). The flow-through was combined with 150μL magnetic Strepavidin beads (Pierce #88816) and incubated at 4°C overnight. On the next

day, the beads were washed with the following steps: RIPA buffer 2 times, 1M KCl once, 0.1M Na_2CO_3 once, 2M urea in 10mM Tris-HCl once, and RIPA buffer 2 times. Biotinylated proteins were eluted by boiling the beads in 90μL 2X sample buffer, supplemented with 2.5mM biotin, and used for downstream LC-MS/MS and Western blot analyses.

HiUGE-iBioID LC-MS/MS analysis

Samples were spiked with undigested bovine casein at a total of either 120 or 240 fmol as an internal quality control standard. Next, samples were supplemented with 12.4 µL of 20% SDS, reduced with 10 mM dithiolthreitol for 30 min at 80 °C, alkylated with 20 mM iodoacetamide for 30 min at room temperature, then supplemented with a final concentration of 1.2% phosphoric acid and 723 µL of S-Trap (Protifi) binding buffer (90% MeOH/100mM TEAB). Proteins were trapped on the S-Trap micro cartridge, digested using 20 ng/µL sequencing grade trypsin (Promega) for 1 hr at 47 °C, and eluted using 50 mM TEAB, followed by 0.2% FA, and lastly using 50% ACN/0.2% FA. All samples were then lyophilized to dryness. Samples were resolubilized using 12 µL of 1% TFA/2% ACN with 25 fmol/µL yeast ADH.

Quantitative LC-MS/MS was performed on 2 µL (~17% of total sample) using a nanoAcquity UPLC system (Waters Corp) coupled to a Thermo Orbitrap Fusion Lumos high resolution accurate mass tandem mass spectrometer (Thermo) equipped with a FAIMSPro device via a nanoelectrospray ionization source. Briefly, the sample was first trapped on a Symmetry C18 20 mm × 180 µm trapping column (5 µl/min at 99.9/0.1 v/v water/acetonitrile), after which the analytical separation was performed using a 1.8 µm Acquity HSS T3 C18 75 µm × 250 mm column (Waters Corp.) with a 90-min linear gradient of 5 to 30% acetonitrile with 0.1% formic acid at a flow rate of 400 nanoliters/minute (nL/min) with a column temperature of 55C. Data collection on the Fusion Lumos mass spectrometer was performed for three difference compensation voltages (-40v, -60v, -80v). Within each CV, a data-dependent acquisition (DDA) mode of acquisition with a r=120,000 (@ m/z 200) full MS scan from m/z 375 – 1600 with a target AGC value of 4e5 ions was performed. MS/MS scans were acquired in the ion trap in rapid mode with a target AGC value of 1e4 and max fill time of 35 ms. The total cycle time for each CV was 0.66s, with total cycle times of 2 sec between like full MS scans. A 20s dynamic exclusion was employed to increase depth of coverage. The total analysis cycle time for each injection was approximately 2 hours.

Following UPLC-MS/MS analyses, data were imported into Proteome Discoverer 2.5 ("PD", Thermo Scientific Inc.) and individual LCMS data files were aligned based on the accurate mass and retention time of detected precursor ions ("features") using Minora Feature Detector algorithm in Proteome Discoverer. Relative peptide abundance was measured based on peak intensities of selected ion chromatograms of the aligned features across all runs. The MS/MS data was searched against the SwissProt M. musculus database and a common contaminant/spiked protein database (bovine albumin, bovine casein, yeast ADH, etc.), and an equal number of reversed- sequence "decoys" for false discovery rate determination. Sequest with Infernys enabled (v 2.5, Thermo PD) was utilized to produce fragment ion spectra and to perform the database searches. Database search parameters included fixed modification on Cys (carbamidomethyl) and variable modification on Met (oxidation). Search tolerances were 2ppm precursor and 0.8Da product ion with full trypsin enzyme rules. Peptide Validator and Protein FDR Validator nodes in Proteome Discoverer were used to annotate the data at a maximum 1% protein false discovery rate based on q-value calculations. Note that peptide homology was addressed using razor rules in which a peptide matched to multiple different proteins was exclusively assigned to the protein has more identified peptides. Protein homology was addressed by grouping proteins that had the same set of peptides to account for their identification. A master protein within a group was assigned based on % coverage.

Prior to imputation, a filter was applied such that a peptide was removed if it was not measured in at least 2 unique samples (50% of a single group). After that filter, any missing data missing values were imputed using the following rules; 1) if only a single signal was missing within the group of three, an average of the other two values was used or 2) if two out of three signals were missing within the group of three, a randomized intensity within the bottom 2% of the detectable signals was used. To summarize to the protein level, all peptides belonging to the same protein were summed into a single intensity. This protein value was then subjected to a robust mean normalization in which the top and bottom 10 of the signals were removed and then the remaining mean was made to be the same across all samples.

The results were then log2-transformed and analyzed using the PolySTest online tool [85]. Proteomic detection was defined as proteins identified by at least 2 peptides in LC-MS/MS. Protein abundance was considered significantly enriched if they met FDR < 0.05 (PolySTest), and fold change ≥ 2 (Log2-FC ≥ 1). Known biological and experimental contaminants were filtered, including Trypsin, Streptavidin, TurboID, Keratin, and Albumin. Further, proteins identified as significant 13 times or more, excluding bait

self-ID, across the panel of 14 baits were considered omnipresent contaminants and filtered. In addition, enriched proteins identified by endogenously expressed soluble TurboID were considered background detections and were filtered. Note that this high-stringency filter removed a few well-known interactors such as Dlg4, Shank2, and Shank3 from the dataset.

Synaptosomal preparation and proteomic analysis of Syngap1-Het mice.

Synaptosomal preparation was performed with four adult Syngap1-Het mice and four WT controls. Briefly, rapidly isolated brain tissue was sliced to 1mm sections using a brain matrix (Zivic Instruments), followed by dissection of cortical and striatal tissue. Tissue was homogenized in homogenization buffer (320 mM sucrose, 5 mM HEPES, pH 7.4) using a Dounce homogenizer. The lysate was centrifuged at 1,000 xg to remove cell debris and nuclei. The supernatant was further centrifuged at 12,000 xg to obtain a crude synaptosomal pellet and it was resuspended in Tris buffer (320 mM sucrose, 5 mM Tris/HCl, pH 8.1). Additional centrifugation in a sucrose density gradient (0.8/1.0/1.2 M) at 85,000 xg was performed to isolate purified synaptosome at the 1.0/1.2 interface. The purified synaptosomes were subjected to multiplexed LC-MS/MS quantification following tandem mass tags (TMT) isobaric labeling.

LOPIT-DC subcellular fractionation and proteomic analysis of Scn2a$^{+/R102Q}$ mice.

LOPIT-DC fractionation was performed with three adult Scn2a$^{+/R102Q}$ mice and three WT controls, following a previously described method [63,64,86]. Briefly, rapidly isolated brain tissue was added to isotonic TEVP homogenization buffer (320 mM sucrose, 10 mM Tris, 1 mM EDTA, 1 mM EGTA, 5 mM NaF, pH7.4), supplemented with cOmplete protease inhibitor cocktail (Sigma #11873580001). The tissue was homogenized for 15 passes in a 2mL Dounce homogenizer. The volume of the homogenate was brought up to a 5 mL volume with TEVP buffer, and then passed through a 0.5 mL ball-bearing homogenizer for two passes (14 μm ball, Isobiotec). Differential centrifugation steps were performed at 4°C sequentially at 200, 1000, 3000, 5000, 9000, 12,000, 15,000, 30,000, 79,000, and 120,000 xg. Fraction-5, determined to be enriched with Scn2a in a pilot experiment [64], was used for tandem mass tag (TMT)-multiplexed proteomic analysis.

TMT-multiplexed quantitative LC-MS/MS analysis.

Samples were supplemented with 100 µL of 8 M urea and probe sonicated. Protein concentrations were deteremined via Bradford Assay and ranged from 1.8 – 2.3 mg/mL. Samples were normalized to 120 µg using 8 M urea and spiked with undigested bovine casein at a total of either 120 or 240 fmol as an internal quality control standard. Next, they were supplemented with 13 µL of 20% SDS, reduced with 10 mM dithiolthreitol for 45 min at 32 °C, alkylated with 20 mM iodoacetamide for 45 min at room temperature, then supplemented with a final concentration of 1.2% phosphoric acid and 70 µL of S-Trap (Protifi) binding buffer (90% MeOH/100mM TEAB). Proteins were trapped on the S-Trap micro cartridge, digested using 100 ng/µL sequencing grade trypsin (Promega) for 1 hr at 47 °C, and eluted using 50 mM TEAB, followed by 0.2% FA, and lastly using 50% CAN/0.2% FA. All samples were then lyophilized to dryness.

For TMT labeling, each sample was resuspended in 120 µL 200 mM triethylammonium bicarbonate, pH 8.0 (TEAB). 40uL of each sample was combined to form an SPQC pool, which was then aliquoted into 3 SPQC pools. Fresh TMT10plex reagents (0.8 mg for each 10-plex reagent) were resuspended in 41 µL 100% acetonitrile (ACN) and was added to 75 µg of each sample. Samples were incubated for 1 hour at RT. After 1-hour reaction, 8 µL of 5% hydroxylamine was added and incubated for 15 minutes at room temperature to quench the reaction. Sample were combined then lyophilized to dryness.

For offline fractionation, samples were resuspended in 300uL 0.1% formic acid. 400ug was fractionated into 48 unique high pH reversed-phase fractions using pH 9.0 20 mM Ammonium formate as mobile phase A and neat acetonitrile as mobile phase B. The column used was a 2.1 mm x 50 mm BEH C18 (Waters) and fractionation was performed on an Agilent 1100 HPLC with G1364C fraction collector. Throughout the method, the flow rate was 0.4 mL/min and the column temperature was 55 °C. The gradient method was set as follows: 0 min, 3%B; 1 min, 7% B; 50 min, 50%B; 51 min, 90% B; 55 min, 90% B; 56 min, 3% B; 70 min, 3% B. 48 fractions were collected in equal time segments from 0 to 52 minutes, then concatenated into 12 unique samples using every 12th fraction. For instance, fraction 1, 13, 25, and 37 were combined, fraction 2, 14, 26, and 38 were combined, etc. Fractions were frozen and lyophilized overnight. Samples were resuspended in 50 µL 1%TFA/2% acetonitrile prior to LC-MS analysis.

Quantitative LC-MS/MS was performed on 2 µL (1 µg) of each sample, using a nanoAcquity UPLC system (Waters Corp) coupled to a Thermo Orbitrap Fusion Lumos high resolution accurate mass tandem mass spectrometer (Thermo) equipped with a FAIMSPro device via a nanoelectrospray ionization source. Briefly, the sample was first trapped on a Symmetry C18 20 mm × 180 µm trapping column (5 µl/min at 99.9/0.1 v/v water/acetonitrile), after which the analytical separation was performed using a 1.8 µm Acquity HSS T3 C18 75 µm × 250 mm column (Waters Corp.) with a 90-min linear gradient of 5 to 30% acetonitrile with 0.1% formic acid at a flow rate of 400 nanoliters/minute (nL/min) with a column temperature of 55C. Data collection on the Fusion Lumos mass spectrometer was performed for three difference compensation voltages (-40v, -60v, -80v). Within each CV, a data-dependent acquisition (DDA) mode of acquisition with a r=120,000 (@ m/z 200) full MS scan from m/z 375 – 1600 with a target AGC value of 4e5 ions was performed. MS/MS scans were acquired in the Orbitrap at r=50,000 (@ m/z 200) from m/z 100 with a target AGC value of 1e5 and max fill time of 105 ms. The total cycle time for each CV was 1s, with total cycle times of 3 sec between like full MS scans. A 45s dynamic exclusion was employed to increase depth of coverage. The total analysis cycle time for each sample injection was approximately 2 hours.

Data were imported into Proteome Discoverer 2.4 (Thermo Scientific Inc.) and individual LCMS data files were aligned based on the accurate mass and retention time of detected precursor ions ("features") using Minora Feature Detector algorithm in Proteome Discoverer. Relative peptide abundance was measured based on peak intensities of selected ion chromatograms of the aligned features across all runs. The MS/MS data was searched against the SwissProt M. musculus database (downloaded in Nov 2019), a common contaminant/spiked protein database (bovine albumin, bovine casein, yeast ADH, etc.), and an equal number of reversed- sequence "decoys" for false discovery rate determination. Mascot Distiller and Mascot Server (v 2.5, Matrix Sciences) were utilized to produce fragment ion spectra and to perform the database searches. Database search parameters included fixed modification on Cys (carbamidomethyl) and variable modification on Met (oxidation), Asn/Gln (deamindation), Lys (TMT6plex) and peptide N-termini (TMT6plex). Peptide Validator and Protein FDR Validator nodes in Proteome Discoverer were used to annotate the data at a maximum 1% protein false discovery rate based on q-value calculations. Note that peptide homology was addressed using razor rules in which a peptide matched to multiple different proteins was exclusively assigned to the protein has more identified peptides. Protein homology was addressed by grouping proteins that had the same set of peptides to account for their identification. A master protein within a group was assigned based on % coverage. To account for any missing data (from a misalignment, low quality peak, low signal to noise, etc.) missing values were imputed by using a randomized intensity within the bottom 2% of the detectable signals. The data was then intensity normalized using a trim-mean normalization in which the highest and lowest 10% of the signals from each sample was excluded and then the remaining average intensities of the proteins was made equal across all of the samples.

The results were then analyzed using the Duke Proteomics and Metabolomics Shared Resource (DPMSR) Proteome Discoverer Data Visualization Tool. Proteomic detection was defined as proteins identified by at least 2 peptides in LC-MS/MS following a 1% FDR correction. Protein abundance was considered significantly altered if they met p-value < 0.05 (two-tailed t-test), and abs(fold change) ≥ 1.2 (abs(Log2-FC) ≥ 0.263). Those with p-value < 0.1 and abs(fold change) ≥ 1.2 (abs(Log2-FC) ≥ 0.263) were considered indicative candidates.

Protein network visualization

Protein networks were constructed using Cytoscape with nodes representing enriched or dysregulated proteins identified by LC-MS/MS. Known interactions with high confidence (i.e., 0.7 score) between these nodes were queried from the full STRING database (https://string-db.org) [24,25] and plotted on the network figure. To assess the percentage of interactions not reported in STRING queries, a low confidence (0.15) threshold was used. The Markov Cluster Algorithm in STRING (inflation parameter: 3) was used to detect protein communities within each interactome, with additional adjustments made based on known protein functions.

Gene set enrichment analyses

Gene ontology (GO) enrichment of interactomes (including baits) was searched against a custom statistical domain of all identified brain proteins (9,686 unique proteins) from cumulative mouse brain proteomic studies in our lab (n=107), using ShinyGO [87]. Pathway size boundary was set at between 10 and 500 to exclude ambiguous terms for querying the Molecular Function pathway database. Default pathway size boundary was used for all other queries. Redundancy removal option was enabled. Overlaps of identified interactors with the SFARI gene list (2022 Q1 release) were calculated using Venny (Oliveros, J.C. (2007-2015) Venny. An interactive tool for comparing lists with Venn's diagrams.

https://bioinfogp.cnb.csic.es/tools/venny/index.html). Hypergeometric tests were performed using an online tool (http://www.nemates.org/MA/progs/overlap_stats.html). In addition, we successfully mapped 181/185 autism risk genes identified at FDR<0.05 in Fu et al., 2022 [26] to mouse orthologs, and examined their overlap with each set of proteins obtained from the 14 HiUGE-iBioID experiments, via hypergeometric testing (Github Link: git@github.com:lauragails/gene_baiting.git). Bonferroni adjustments were performed and thresholds for statistical significance were delineated on the graphs.

For cell-type specific representations, the list of genes for each bait were intersected with cell type-specific differentially expressed genes (DEGs) from a single-cell genomics study of post-mortem cortical brain tissues from ASD patients [7]. We used genes expressed in each cell type as the background to perform the hypergeometic test for overlap between each bait gene list and a list of ASD DEGs in each cell type.

Western blot

For Western blot analyses, 10μL of HiUGE-iBioID purified samples were subjected to SDS-PAGE. After transferring to nitrocellulose membranes, the blot was probed for HA-epitope (rabbit anti-HA, Cell Signaling #3724, 1:1000; or rat anti-HA, Sigma #12158167001, 1:1000), or PSD95 (mouse anti-PSD95, Abcam #ab2723, 1:1000) at 4°C overnight. Equal amounts of input protein from mouse brain lysate were also subjected to SDS-PAGE, and the blot was probed for GAPDH (rabbit anti-GAPDH, Abcam #ab9485 or Cell Signaling #2118, 1:1000) at 4°C overnight. Matching IRDye secondary antibodies (LI-COR, 1:10,000) were incubated with the blot for 1 hour at RT, and the immunosignal was detected using Odyssey FC imager (LI-COR).

Immunocytochemistry

Cells were fixed with 4% PFA and 4% sucrose on DIV 14, then blocked with blocking buffer (Abcam #ab126587, 1:10 in PBS with 0.3% Triton-X). Immunocytochemistry was performed by incubating with primary antibody overnight at 4°C, and with secondary antibody for 1 hour at RT. Antibodies and dilutions consisted of: mouse anti-V5-epitope (ThermoFisher #R960-25, 1:500), guinea pig anti-Homer1 (Synaptic Systems #160004, 1:2000), goat anti-mouse Alexa Fluor Plus 594 (ThermoFisher #A32742, 1:1000), and goat anti-guinea pig Alexa Fluor 488 (ThermoFisher #A11073, 1:1000). Cells were counterstained with 4′,6-diamidino-2-phenylindole (DAPI), mounted with FluorSave reagent (Millipore Sigma #345789), cover-slipped, and imaged on Zeiss Axio Imager M2 with Apotome module.

Co-immunoprecipitation (co-IP)

Expression vectors of Myc-DDK-tagged human ORF clones of ANKS1B and SYNGAP1 were purchased from Origene (#RC211877, #RC229432). V5-epitope or GFP-tagged mutants were cloned into the same expression backbone. Truncations of Syngap1 consist of the following: N-Trunc: Δ a.a. 2-361; C1-Trunc: Δ a.a.730-1343; C2-Trunc: Δ a.a. 848-1343; C3-Trunc: Δ a.a. 1181-1343. Mutation of Syngap1 (SYNGAP1-c.2214_2217del) was introduced by site-directed mutagenesis (QuikChange Lightning, Agilent #210518). HEK293T cells were co-transfected with expression vectors using Lipofectamin 3000 (ThermoFisher). Four days following transfection, cells were lysed for Nanobody Trap experiments following the manufacturer's protocol (Chromotek #gta-20, #gtma-20, or #v5tma-20). The bound proteins were eluted by boiling in 2X Western sample buffer and subjected to SDS-PAGE (immunoprecipitated IP fraction). The membrane was probed for Myc-epitope (Santa Cruz #sc-40, 1:250), V5-epitope (Cell Signaling #13202S, 1:1000), or GFP (Cell Signaling #2956S, 1:2000). The lysate was also subjected to SDS-PAGE (input fraction). Here, the membrane was probed for Myc-epitope (Santa Cruz #sc-40, 1:250) and GAPDH (Abcam #ab9485 or Cell Signaling #2118, 1:1000). Matching IRDye secondary antibodies (LI-COR, 1:10,000) were used to visualize immuno-signals on an Odyssey FC imager.

AAV / Cas9-mediated expression disruption

To disrupt Syngap1 and Anks1b expression, the following genomic sequences were targeted by gRNAs using AAV: Syngap1: acggactcggtctcagcccatgg; Anks1b: attgtcccactgtttggacaggg. AAVs (PHP.eB serotype) were applied to H11-Cas9 primary neuronal cultures, with empty gRNA virus serving as negative control. Effective disruption of Syngap1 and Anks1b expression was confirmed by Western blot (Syngap1: Sigma #SAB2501893 or ThermoFisher #PA5-58362, 1:1000; Anks1b: ThermoFisher #PA5-98554, 1:1000).

Lentiviral-mediated Scn2a CRISPRa

The all-in-one gRNA-Cas9 expression plasmid used for CRISPRa was generated by modifying the hUBC-dSpCas9-2xVP64-T2A-BSD plasmid (Addgene #162333) to remove the T2A-BSD selection marker and include a U6-gRNA scaffold. The non-targeting control gRNA and the gRNAs targeting the *Scn2a* promoter were selected from the Caprano Mouse CRISPR Activation Pooled Library [88] (No_current_500 (non-targeting scramble): tttttagacctaattcgcgc; Scn2a_gRNA1: cagcgattccacttgtggcc; Scn2a_gRNA2: gttgaatgttgctttgccaa; and Scn2a_gRNA3: aattacagcgattccacttg). Individual gRNAs were purchased as oligonucleotides (Integrated DNA Technologies) and cloned into the gRNA expression plasmids using BsmBI sites.

HEK293T cells were acquired from the American Tissue Collection Center (ATCC). The cells were maintained in DMEM high glucose supplemented with 10% FBS and 1x GlutaMAX Supplement (Gibco, 35050061) and cultured at 37°C with 5% CO2. Lentivirus was produced as previously described [89]: 16 hours before transfection, $7x10^6$ cells were plated in 12ml of transfection media (Opti-MEM I Reduced Serum Medium (Gibco #31985070), 1x GlutaMAX Supplement, 5% FBS, 1mM sodium pyruvate (Gibco #11360070), 1x MEM NEAA (Gibco #11140050)) in a 10cm plate. On the day of transfection, 6mL of transfection media were removed, and the cells were transfected with Lipofectamine 3000 (Invitrogen #L3000008) and 5.4μg pMD2.G (Addgene #12259), 9.9μg psPAX2 (Addgene #12260), and 12μg of the expression vector, according to the manufacturer's instructions. The medium was exchanged with transfection media 6hr after transfection, and the viral supernatant was harvested 24 and 48 hrs after this medium change. The viral supernatant was pooled and passed through a PVDF 0.45 μm filter (Millipore #SLHV033RB), and concentrated to 50x in 1xPBS using Lenti-X Concentrator (Clontech #631232) in accordance with the manufacturer's protocol. The lentivirus was titered by qPCR as previously described [90]. The concentrated lentivirus was snap-frozen and stored at -80°C as single-use aliquots.

To assess the effect of CRISPRa transcriptional activation, cultured neurons were treated with Scn2a-CRISPRa or non-targeting control lentiviral vectors at DIV0. On DIV11, the cDNA was prepared using Cells-to-cDNA kit (ThermoFisher #AM1723) following the manufacturer's instructions. Predesigned KiCqStart SYBR green primers (Sigma #KSPQ12012) targeting mouse Scn2a and Actb were used for qPCR experiments with PowerUp SYBR green master mix (ThermoFisher #A25742). Specific on-target amplifications were confirmed by gel electrophoresis and TOPO sequencing of the PCR products (ThermoFisher #450030). The gRNA2 was selected for the experiments in this study based upon its superior ability to upregulate Scn2a over gRNA1 and gRNA3 in cultured neurons.

AAV-mediated expression of SCN1B and FGF12

To express additional SCN1B or FGF12, cDNAs (Origene #RC209565, #RC215868) were cloned into an AAV-expression vector downstream of the hSyn promotor. A non-targeting AAV-Flex-GFP vector (from Dr. Il Hwan Kim) was used as a negative control. Cultured neurons were treated with these AAV vectors (PHP.eB serotype) at DIV0 and lysed for Western blot analysis on DIV11. Effective expression was confirmed by Western blot (SCN1B: Cell Signaling #13950S, 1:1000; FGF12: Proteintech #13784-1-AP, 1:1000).

Microelectrode Array (MEA)

48-well MEA plates (Axion Biosystems #M768-KAP-48 or #M768-tMEA-48W) were coated with 1 mg/mL poly-L-lysine in borate buffer (pH 8.5). Forebrain tissue was rapidly isolated from P0-1 mice, dissociated with papain, and spotted at a density of 150,000 cells per the inner growth area of each well. For experiments that required genotyping, brain tissue was temporarily stored in Hibernate A solution (ThermoFisher #A1247501) at 4°C following the manufacturer's instructions until ready for dissociation and plating. Viral treatments were applied at the day of plating. Recordings were conducted on a Maestro MEA system (Axion Biosystems) on DIV 8, 11, and 14 for 10 min for each session, after 10 min of acclimation to the recording chamber (37°C, 5% CO_2 environment). After each recording, a half-change of growth media (Neurobasal A supplemented with B27, GlutaMax, and 10 μg/mL Gentamicin) was performed. Recording data were analyzed using the Axion Neural Metric Tool. Single electrode bursts were defined as a minimum of 5 spikes, separated by an inter-spike interval (ISI) of no more than 100 msec. Network bursts were defined as a minimum of 10 spikes, separated by an ISI of no more than 100 msec with at least 25% of the electrodes active. Statistical analyses (one-way ANOVA followed by pair-wise Tukey HSD *post-hoc* tests) were performed using JMP Pro 15 (SAS).

Behavioral Tests

Elevated zero maze: Anxiety-like behaviors were assessed in the elevated zero maze as described under 40-60 lux illumination [91]. Mice were housed overnight in the test room and tested individually the next day. Animals were placed into the closed area of the maze and provided 5 min of free exploration. Videos were scored by trained observers blinded to the genotype and sex of the animals using the Noldus Observer XT15 program for the percent time in the open areas and the distance traveled.

Hole-board test for repetitive behaviors: Mice were housed in the test room overnight and the next day were examined in the hole-board test as described [91]. Individual mice were placed into a 42 x 42 x 30 cm open field (Omnitech Electronics) and allowed free exploration of the apparatus for 10 min under 180 lux illumination. The hole-board apparatus consisted of a white Plexiglas floor with 16 equally spaced holes (3 cm in diameter) arranged in 4 rows. Head-dips into the holes were filmed and the videos were scored with the TopScan program (CleverSys) for the numbers of head-dips and the location of each head-dip.

Self-grooming: Individual mice were housed in the test room overnight and the next day were habituated to clean home-cages for 10 min prior to testing [92]. Mice were filmed for 10 min to collect self-induced grooming. The videos were converted subsequently for analysis using Noldus MediaRecorder2. Grooming behavior was scored using TopScan software (CleverSys) for the duration of self-grooming.

Ultrasonic vocalizations (USVs): Male WT and *Scn2a*$^{+/R102Q}$ mice were housed individually 1 week before testing. Initially, males were primed twice with soiled bedding from breeder C57BL/6J cages for 2 days. Subsequently, the males were placed on clean bedding for 1 day, then on soiled bedding for another 2 days, and finally were returned overnight to soiled bedding. Males are tested the next day. They were acclimated to the recording chambers for 2 min and then they were introduced to an unfamiliar C57BL/6J female (12-16 weeks of age) for 8 min. Ultrasonic calls were recorded over the entire 10 min test as waveform audio files. The data were analyzed by scorers blinded to the genotypes of the mice using Avisoft SASLab Pro for the numbers of calls, call duration, and USV frequencies.

Statistics: The data were presented as means and standard errors of the mean (SEMs). The zero maze, hole-board, and self-grooming data were analyzed by independent-samples t-tests and the USV data were analyzed by repeated-measures ANOVA followed by Bonferroni corrected pair-wise comparisons. A $p<0.05$ was considered significant. All statistical analyses were performed with IBM SPSS Statistics 28 programs (IBM, Chicago, IL) and the data were graphed using GraphPad Prism.

Supplementary Figures and Legends

Fig. S1. HiUGE labeling of 14 high-risk ASD proteins.

(**A**) Representative images showing wide-spread TurboID-mediated biotinylation across the brain following labeling Tubb3 with HiUGE. (**B-O**) Representative images showing correct localization (arrowheads) of 14 high-risk ASD proteins labeled with a highly antigenic "spaghetti monster" fluorescent protein (smFP), with boxed regions enlarged in (**Q**) to show colocalization with a synaptic marker Homer1. (**P**) Representative image of the negative control.

Fig. S2. Western blot of HiUGE-iBioID purified samples following streptavidin pulldown.

(**A**) Schematic illustration of enriching biotinylated proteins by streptavidin pulldown. (**B-F**) Western blot images showing detection of TurboID-HA fusion proteins at the expected molecular masses for 14 high-risk ASD proteins, (**B-C**) detection of an expected synaptic interactor (PSD95) is also confirmed. (**G**) Western blot image showing detection of soluble TurboID-HA (as a survey for background) by multiplexed insertion of IRES-TurboID-HA donor at the 3'-UTRs of Syngap1, Scn2a, and Hnrnpu.

Fig. S3. Overlap of HiUGE-iBioID interactomes with published datasets.

(**A**) Significance of interactome overlap with differentially expressed genes (DEGs) found in ASD patients across excitatory neurons, interneurons, and non-neuronal cell populations. (**B**) Significance of interactome overlap with ASD risk genes identified by Fu and colleagues, 2022. Thresholds for statistical significance were delineated on the graphs.

Fig. S4. Structure-function analysis of SYNGAP1-ANKS1B interaction.

(**A**) Schematic illustration of assessing the ANKS1B interaction with SYNGAP1 truncations using human cDNA constructs expressed in HEK293T cells. (**B**) Co-immunoprecipitation results showing the interaction with ANKS1B is ablated in C1- and C2- SYNGAP1 truncations while retained in C3- and N- truncations.

Fig. S5. Additional data of the Scn2a$^{+/R102Q}$ phenotypic rescue experiment.

(**A**) Quantitative PCR screening of three different CRISPRa gRNAs targeting Scn2a *in vitro*. The gRNA2 shows the best performance in upregulating Scn2a expression, and is used for subsequent experiments (n = 3 wells). Specificity of the qPCR assay is validated by sequencing the amplicon. (**B**) Additional neural metrics from the MEA recording. One-way ANOVA followed by *post-hoc* Tukey HSD tests (n= 48 wells). (**C**) mRNA expression levels of Scn2a, Scn1b, and Fgf12 in cultured Scn2a$^{+/R102Q}$ mutant neurons are comparable to that of WT (n = 4 wells). *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; n.s.: non-significant.

Fig. S6-S19. Individual interactome networks of 14 ASD-risk proteins.

Interactome networks associated with 14 ASD-risk proteins are shown in each Figure. Blue lines denote STRING interactions and red lines signify identified HiUGE-iBioID interactions. Annotations denote exemplary significant gene ontology (GO) terms associated with the protein clusters segregated by MCL. Unless otherwise specified, the Biological Process pathway database was used. CC: Cellular Component pathway database. Charts of GO results of the overall bait interactomes using the Molecular Function (MF) pathway database are shown as well, below each network plot.

Table S1. Proteomic results and statistics of HiUGE-iBioID experiments.

Proteomic results and statistics of HiUGE-iBioID experiments are shown. Batch-1: Anks1b, Syngap1, Shank2, Nckap1; Batch-2: Nbea, Ctnnb1, Lrrc4c, Iqsec2; Batch-3: Ank3, Scn2a, Scn8a; Batch-4: Arhgef9, Hnrnpu; Batch-5: Shank3; Batch-6: Syngap1-Trunc; and Batch-7: Soluble TurboID.

Table S2. HiUGE-iBioID interactomes and their overlaps with ASD gene lists.

Filtered gene lists of the detected interactomes associated with 14 high-risk ASD targets are shown. The genes that overlap with the SFARI database and Fu et al, 2022 TADA-ASD gene list are shown in a separate tab.

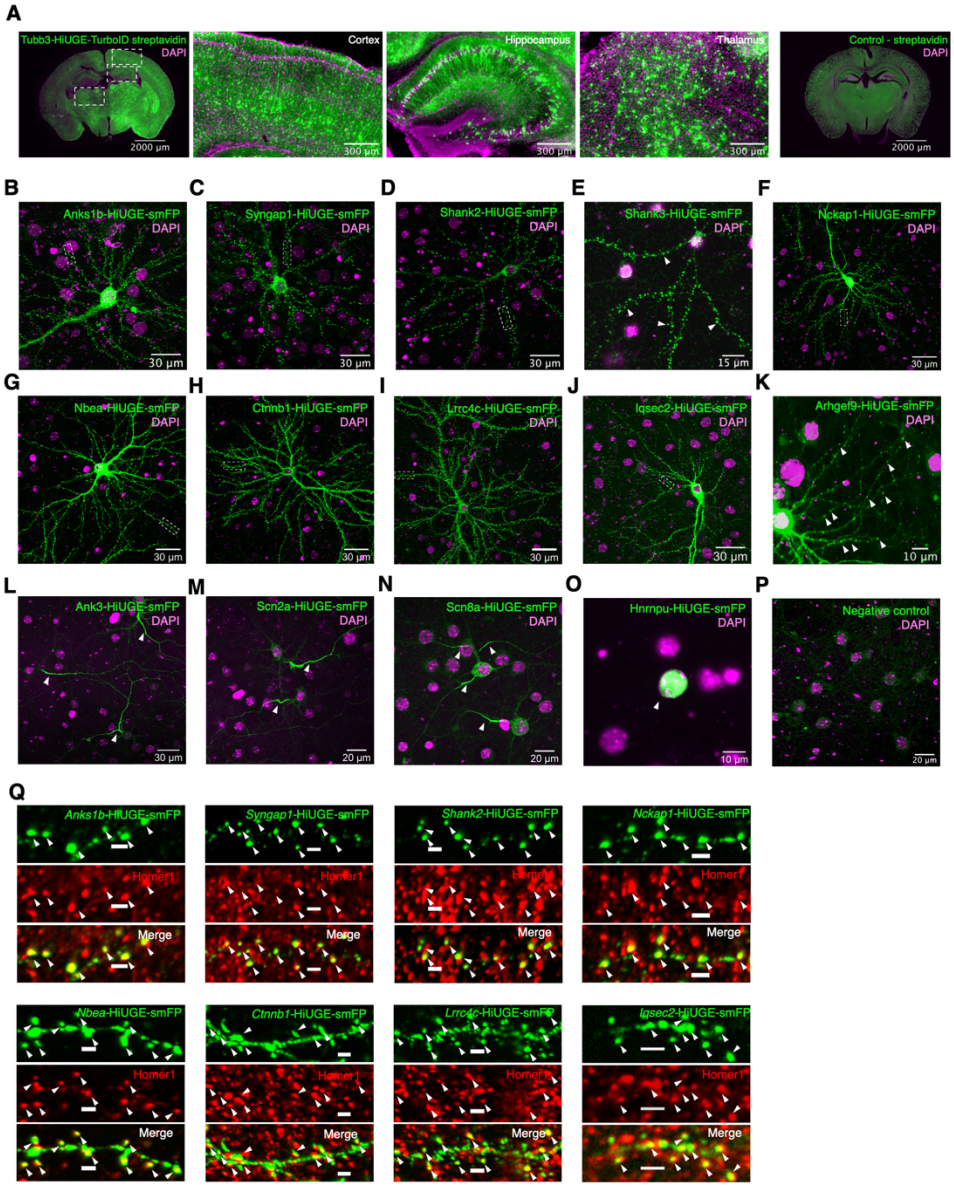
Table S3. Proteomic results and statistics of spatial co-perturbation experiments.

Proteomic results and statistics of spatial co-perturbation experiments are shown, including comparisons of Syngap1-Het synaptosome proteome with WT in the cortex and striatum, and comparison of Scn2a$^{+/R102Q}$ LOPIT-DC fraction-5 proteome with WT.

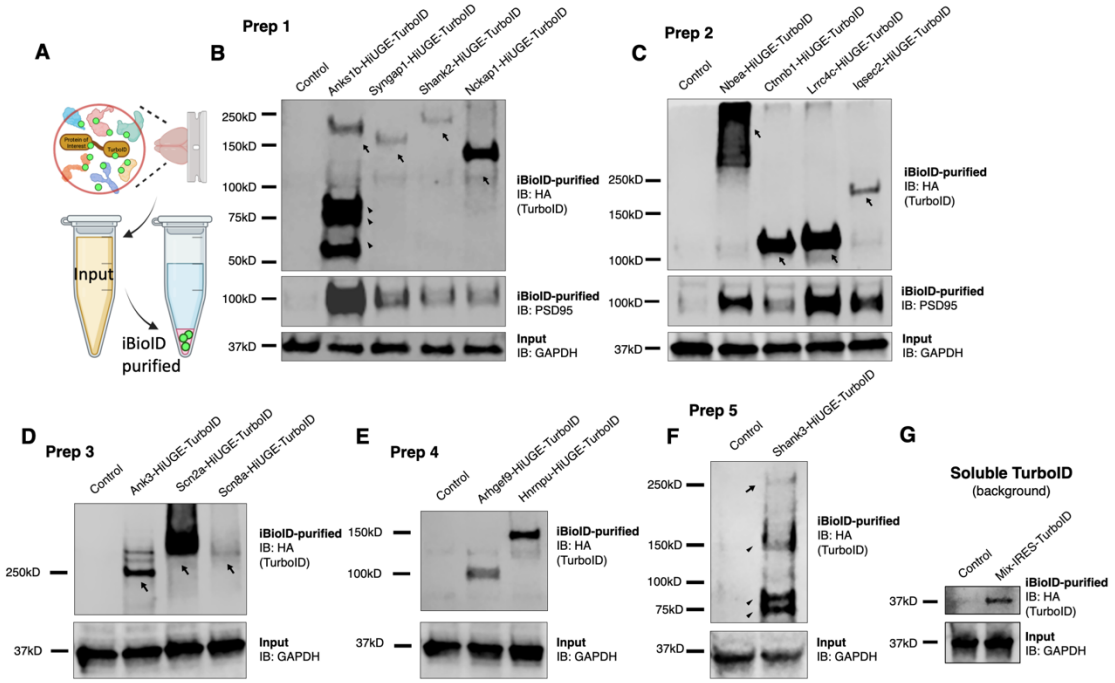
Table S4. Statistics of behavioral characterization of Scn2a$^{+/R102Q}$ mice.

Metrics and statistic summaries for the elevated zero maze test, hole-board test, self-grooming test, and USV social interaction test are reported. For USV test, no statistically significant difference was detected in the pre-social (baseline) responses: # of calls: WT (8.0 ± 1.6), Scn2a$^{+/R102Q}$ (5.9 ± 2.8); Call duration (ms, excluding non-vocal trials): WT (7.3 ± 0.6), Scn2a$^{+/R102Q}$ (6.7 ± 0.6); Call frequency (kHz, excluding non-vocal trials): WT (32.7 ± 1.4), Scn2a$^{+/R102Q}$ (33.3 ± 1.6), Means ± SEM, $p > 0.05$, independent samples t-tests, two-tailed.

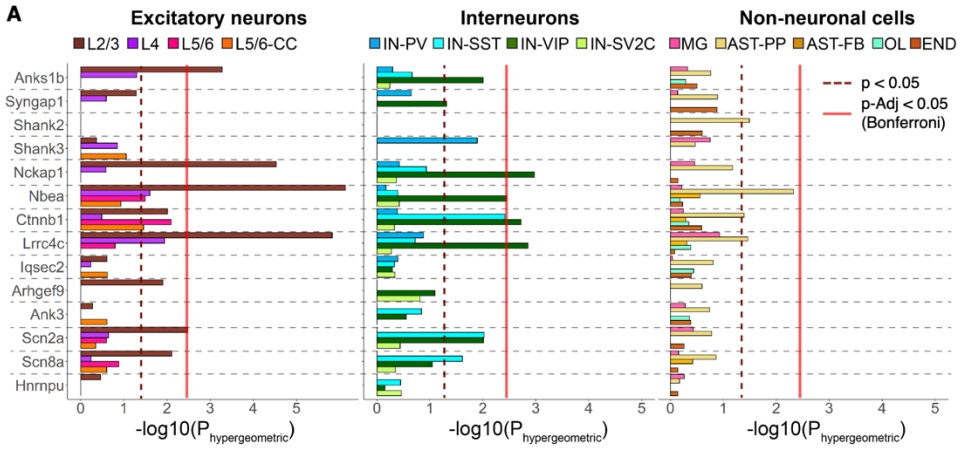
Supplementary Fig. S1

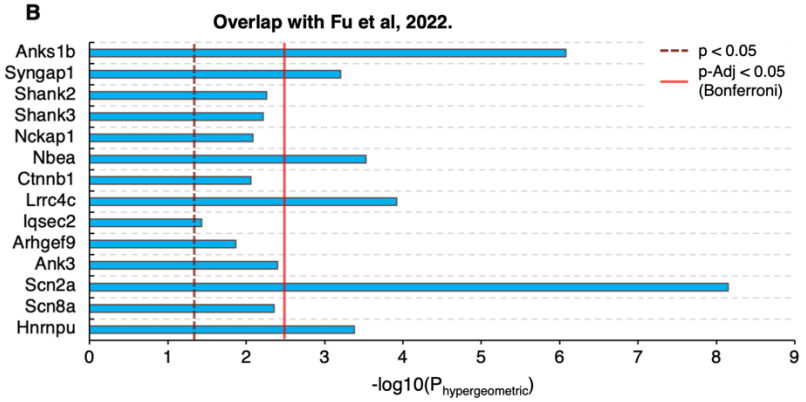


Supplementary Fig. S2

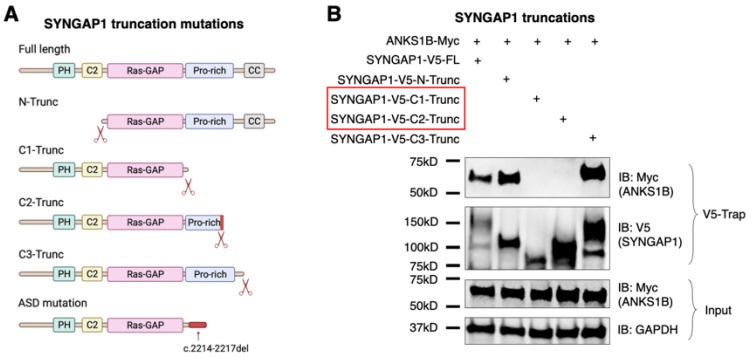


Supplementary Fig. S3

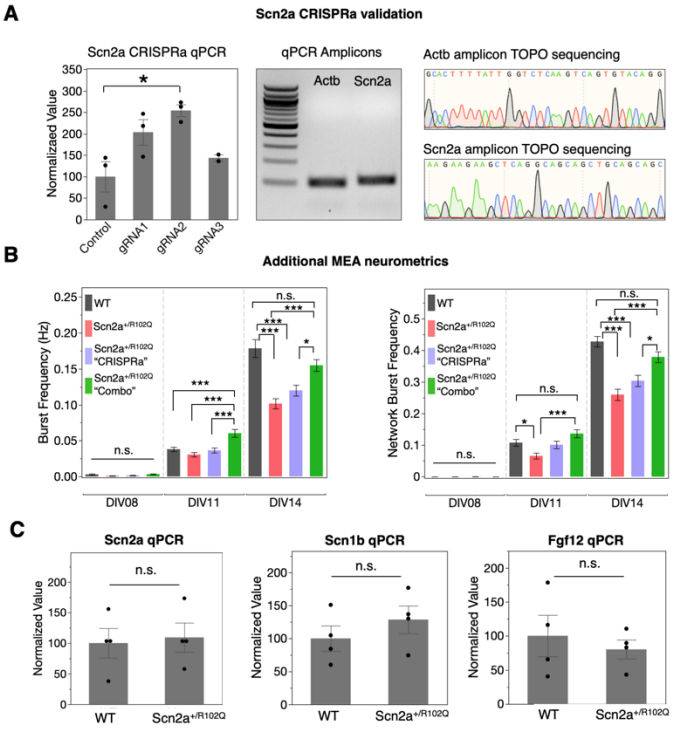




Supplementary Fig. S4

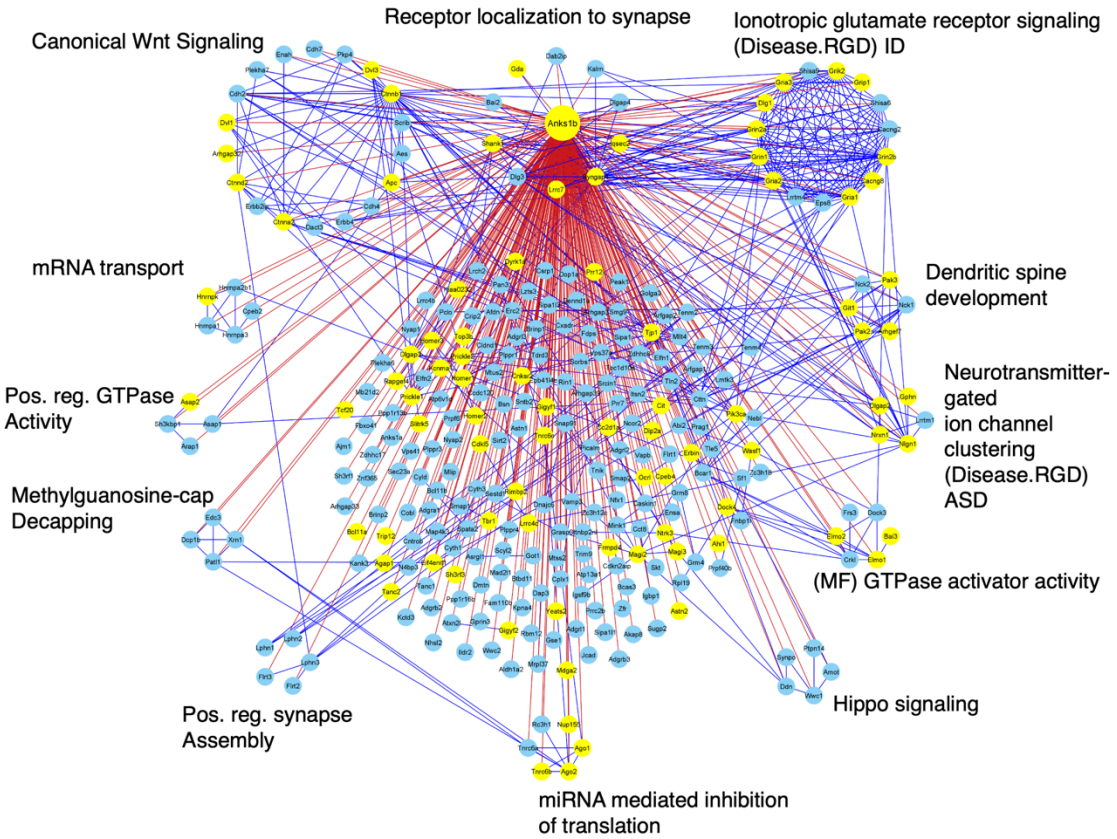


Supplementary Fig. S5

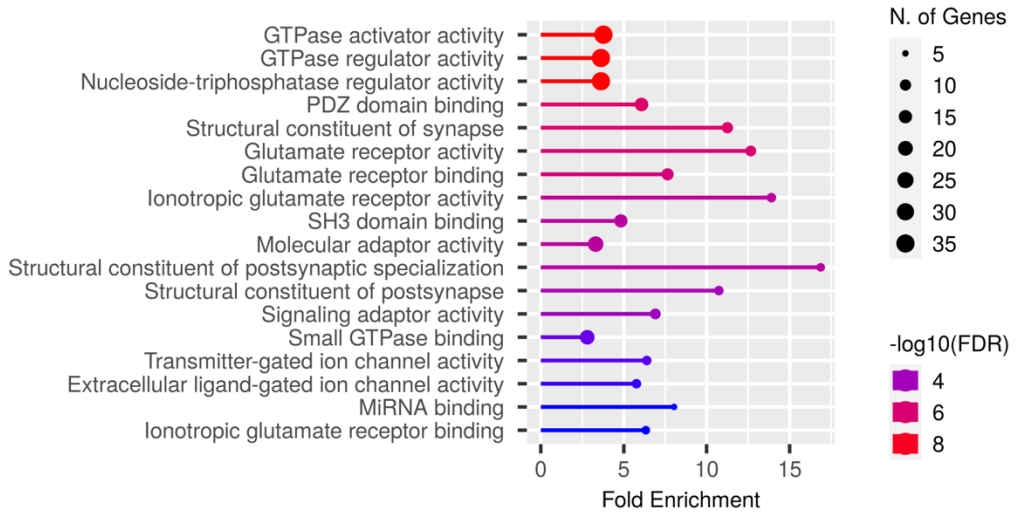


Supplementary Fig. S6

Anks1b interactome

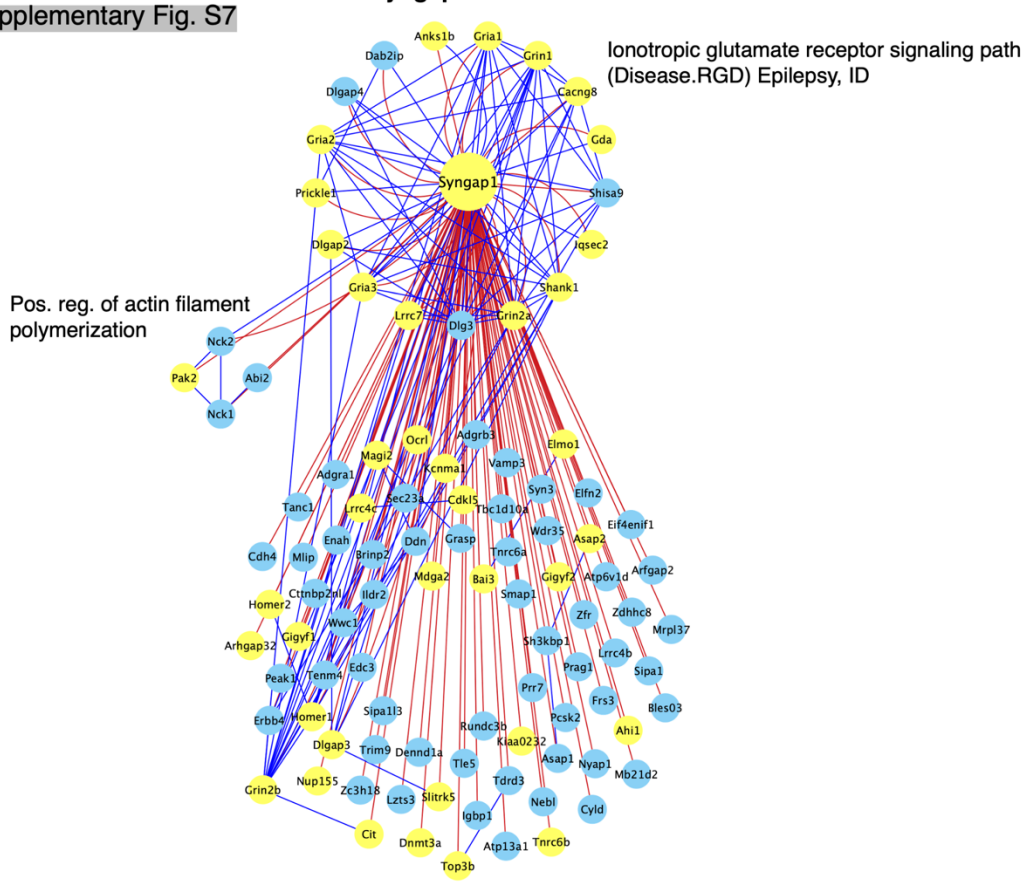


GO : Molecular Function

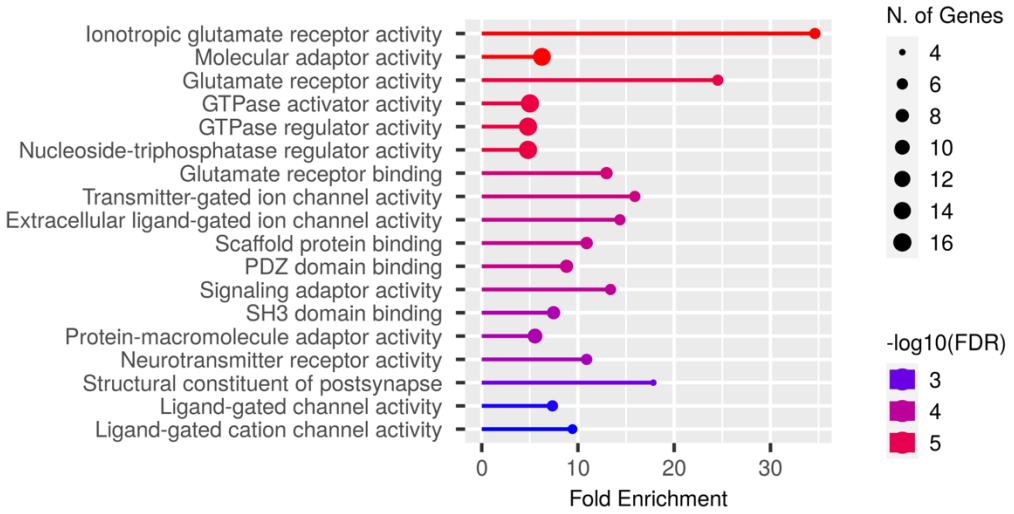


Supplementary Fig. S7

Syngap1 interactome

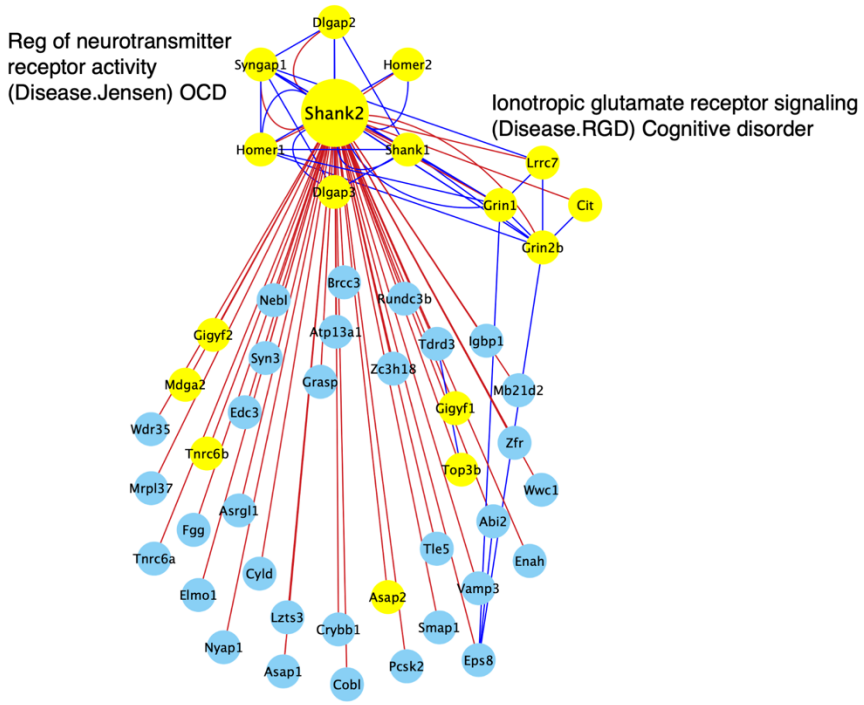


GO : Molecular Function

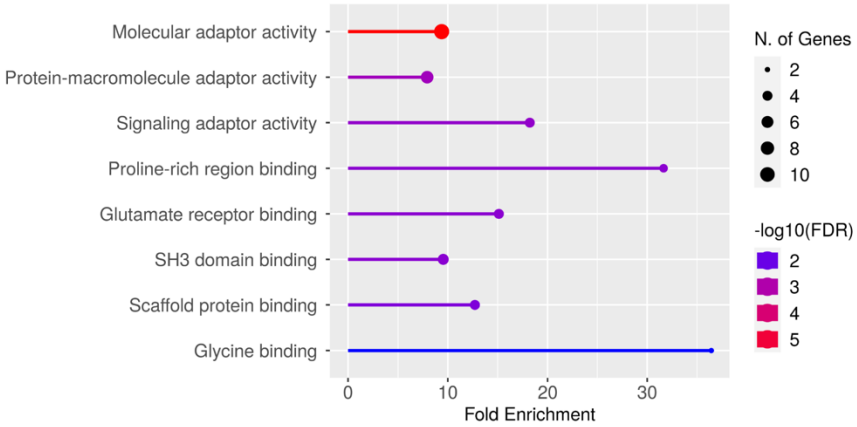


Supplementary Fig. S8

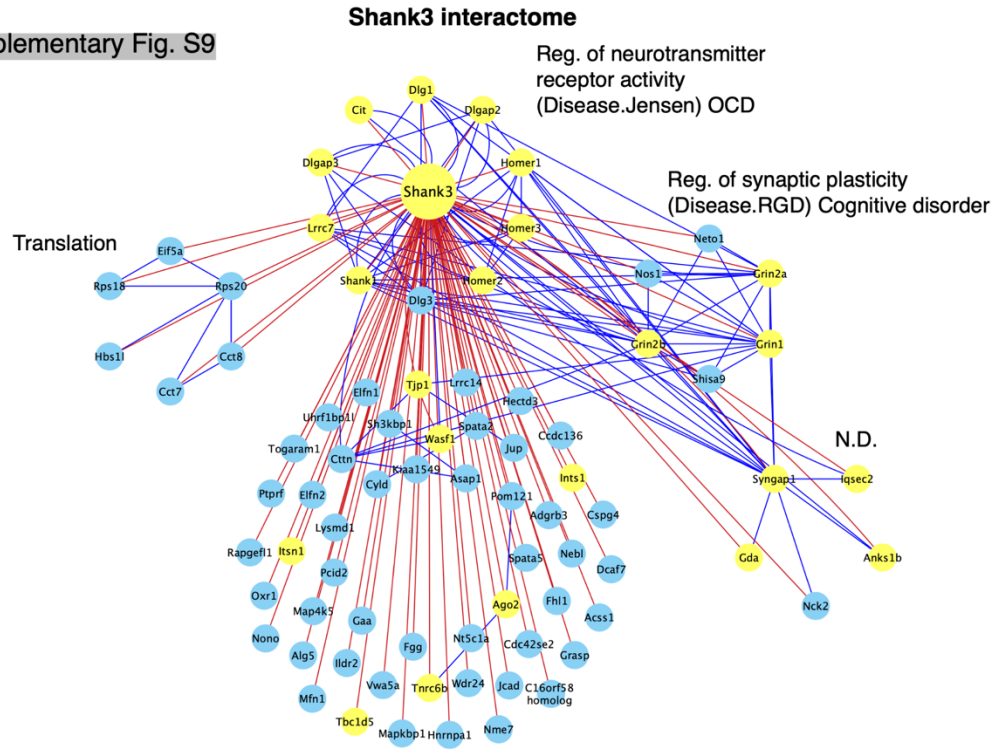
Shank2 interactome

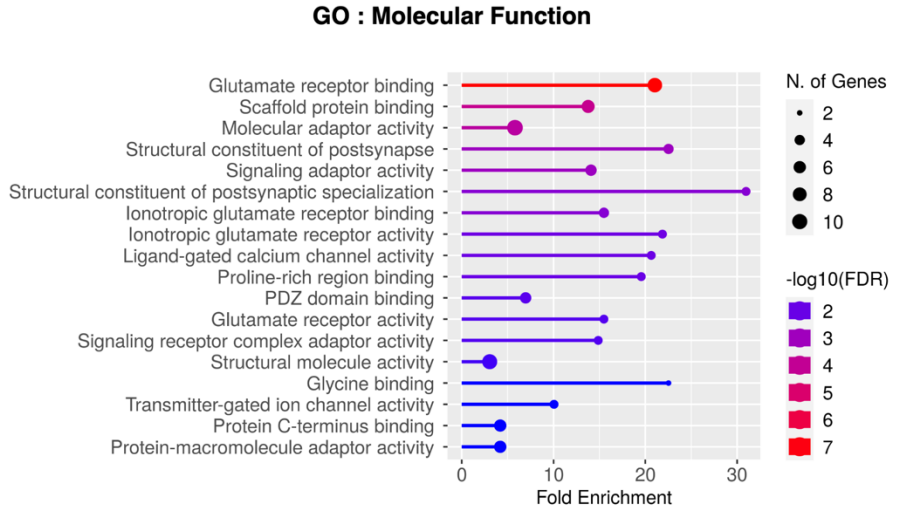


GO : Molecular Function



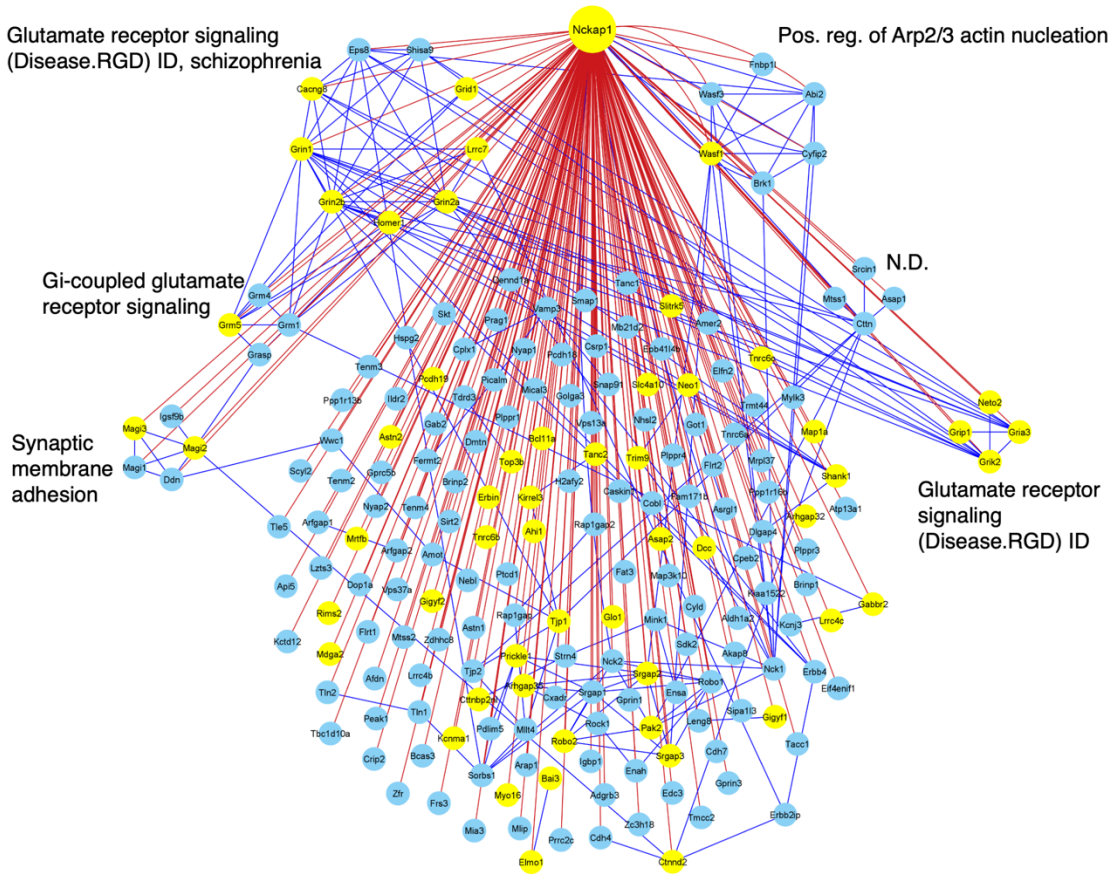
Supplementary Fig. S9



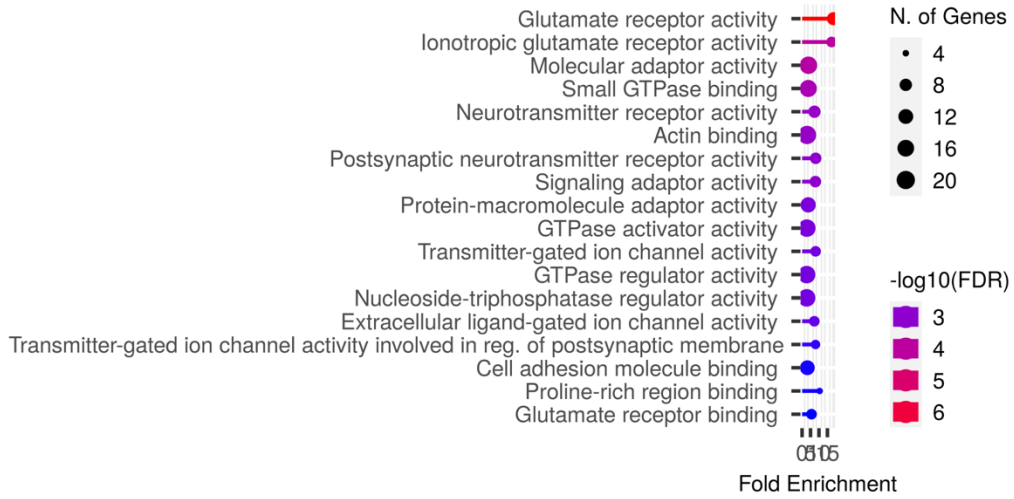


Supplementary Fig. S10

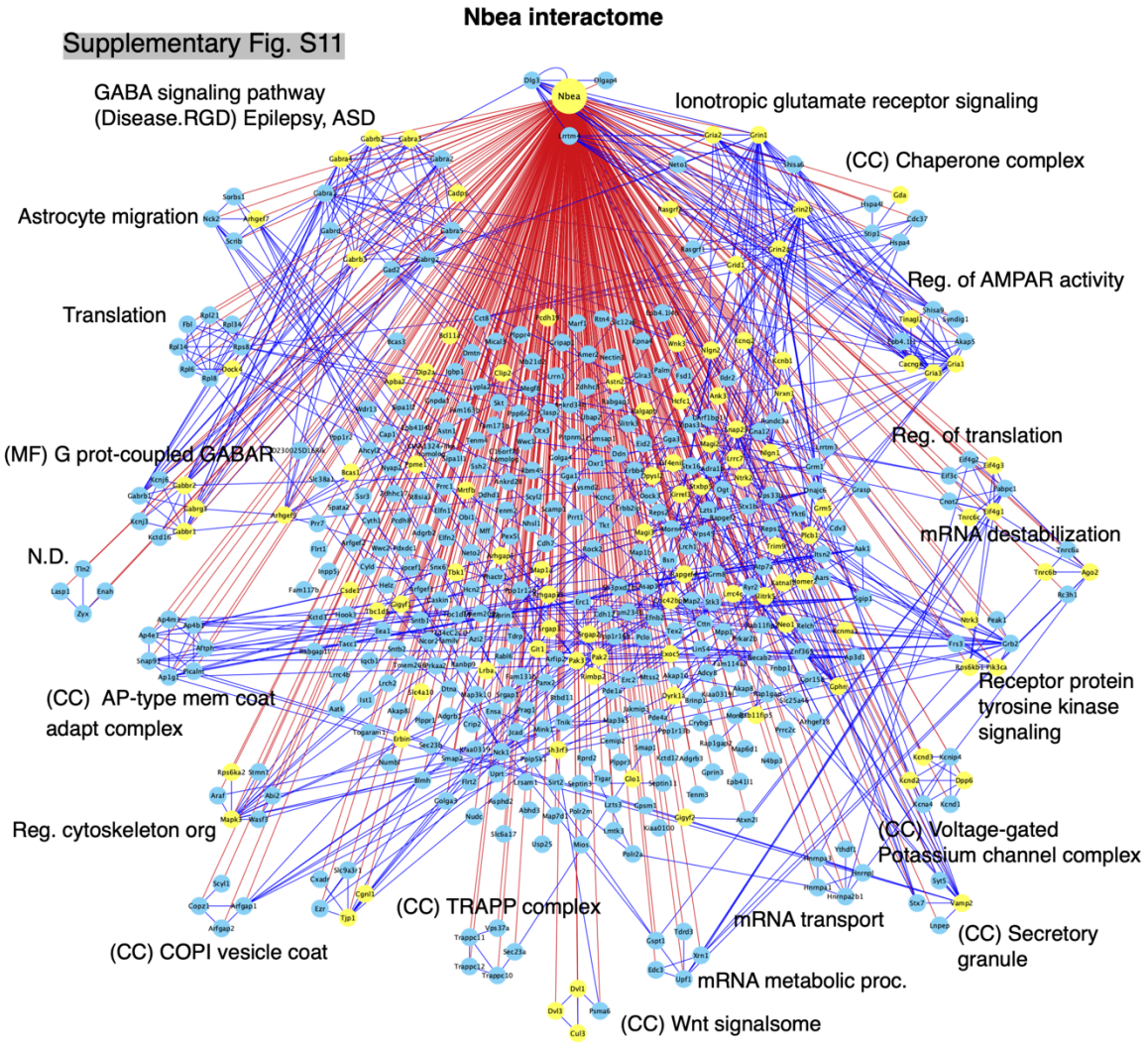
Nckap1 interactome

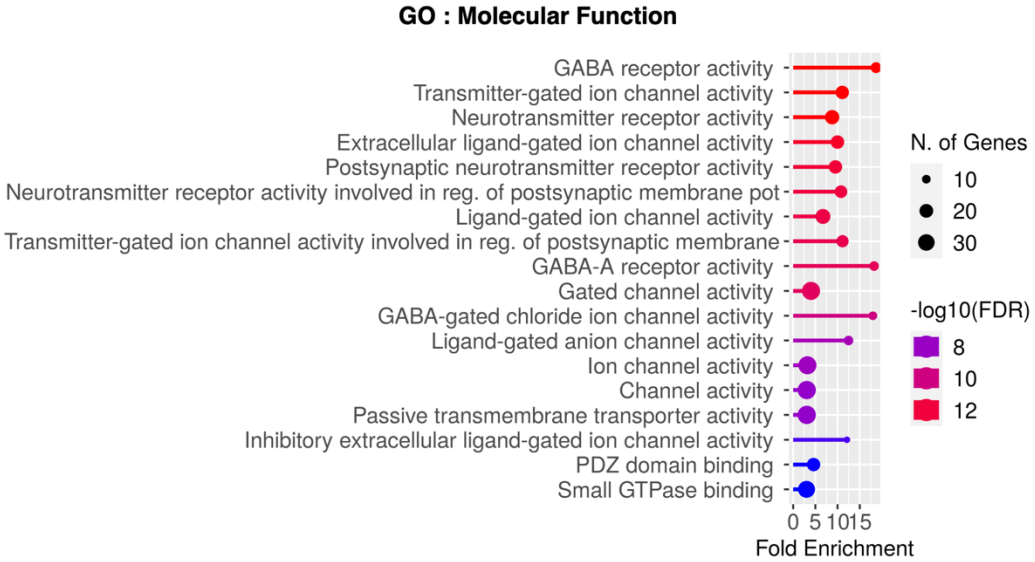


GO : Molecular Function

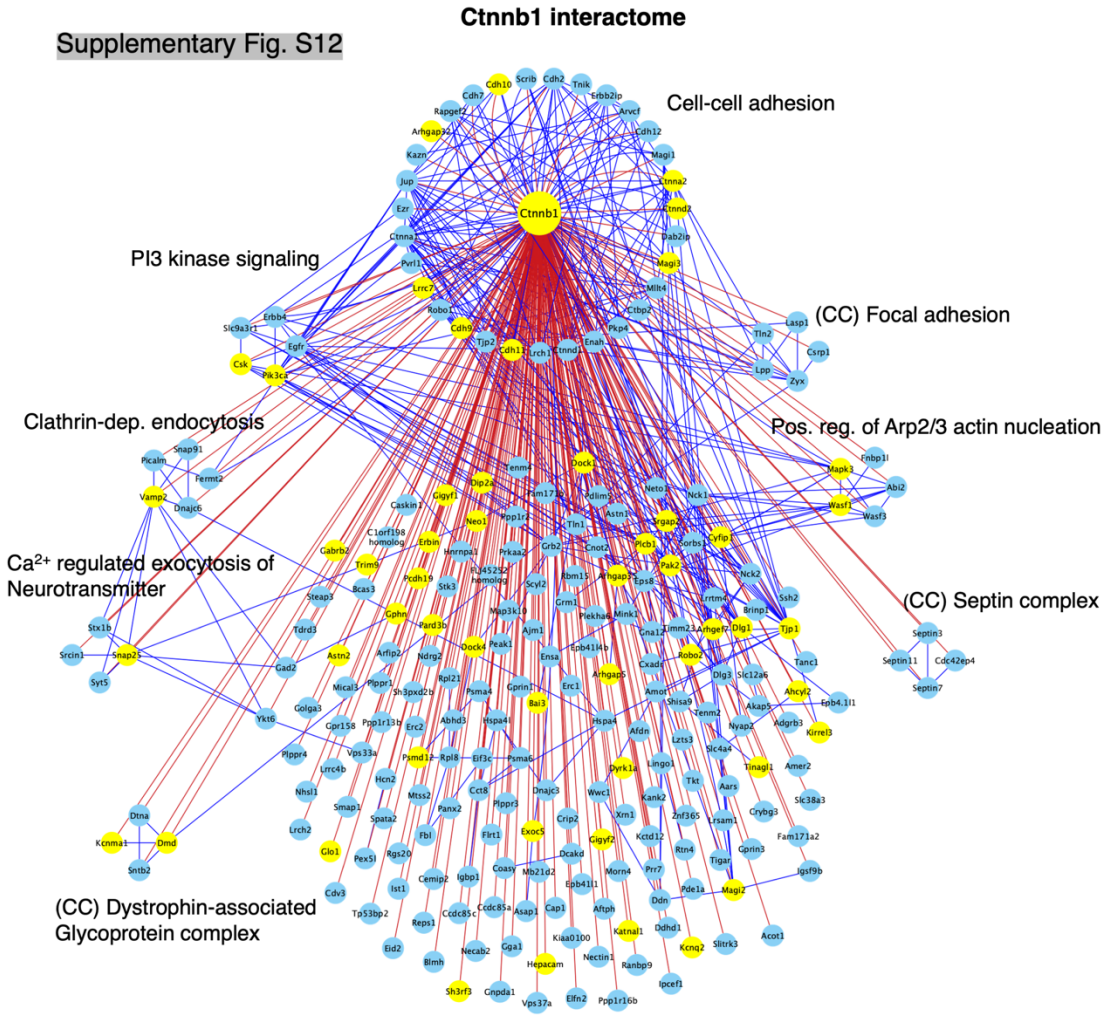


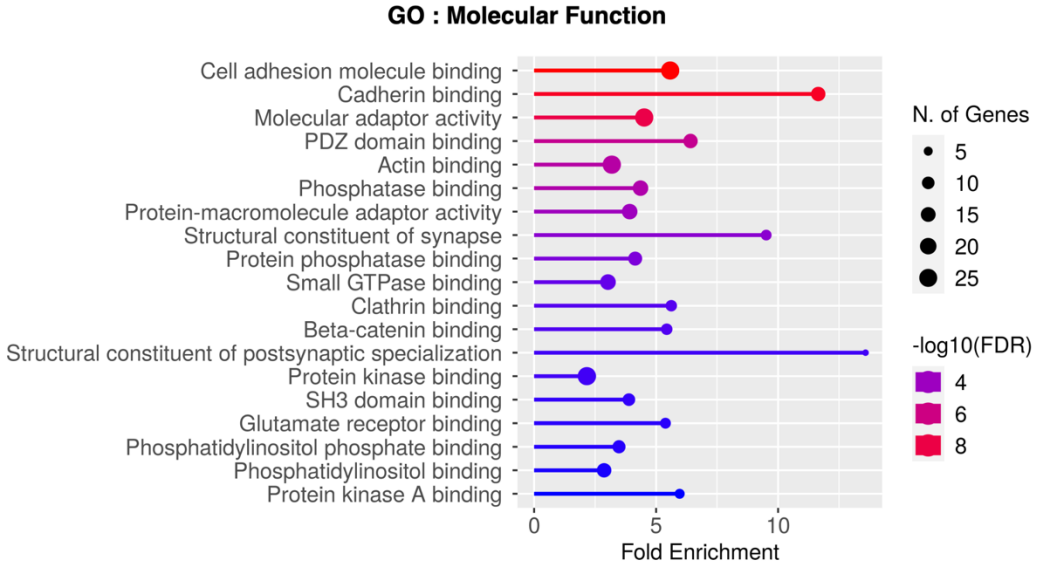
Supplementary Fig. S11



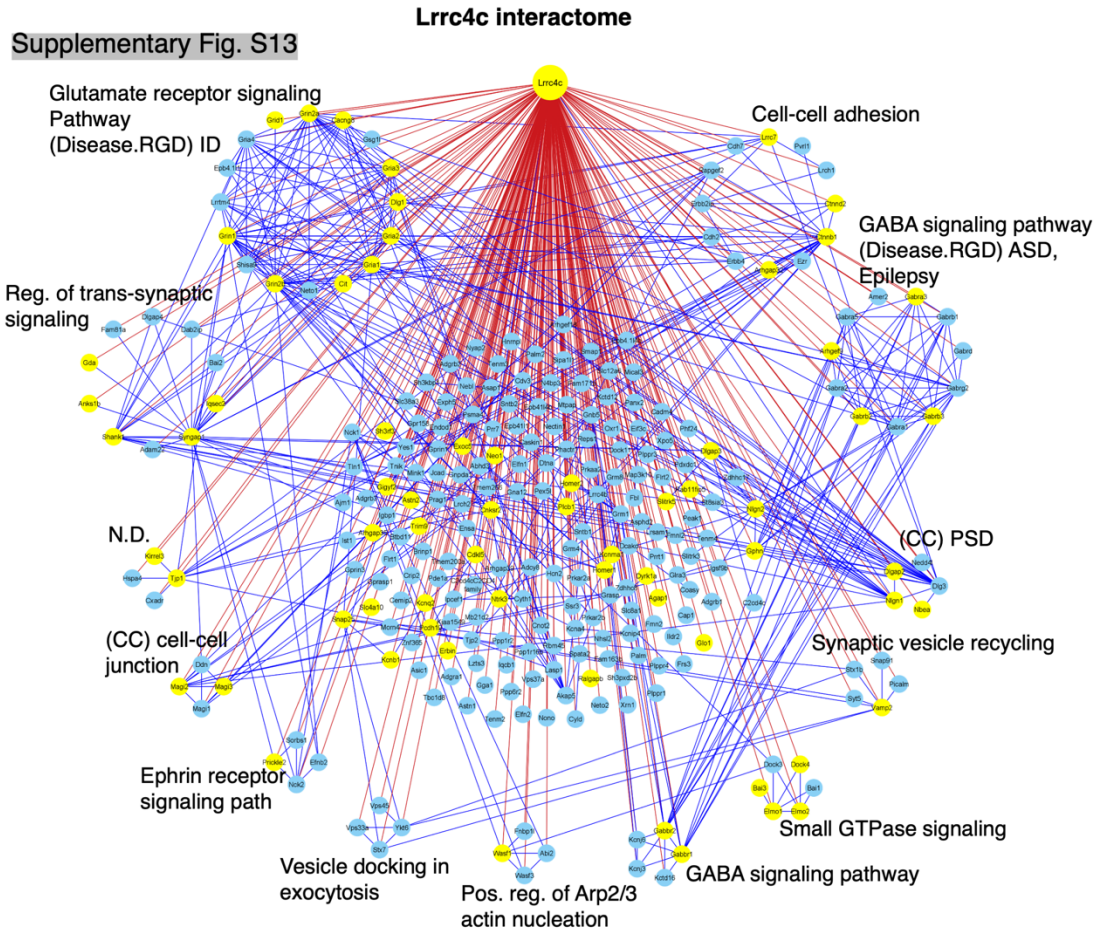


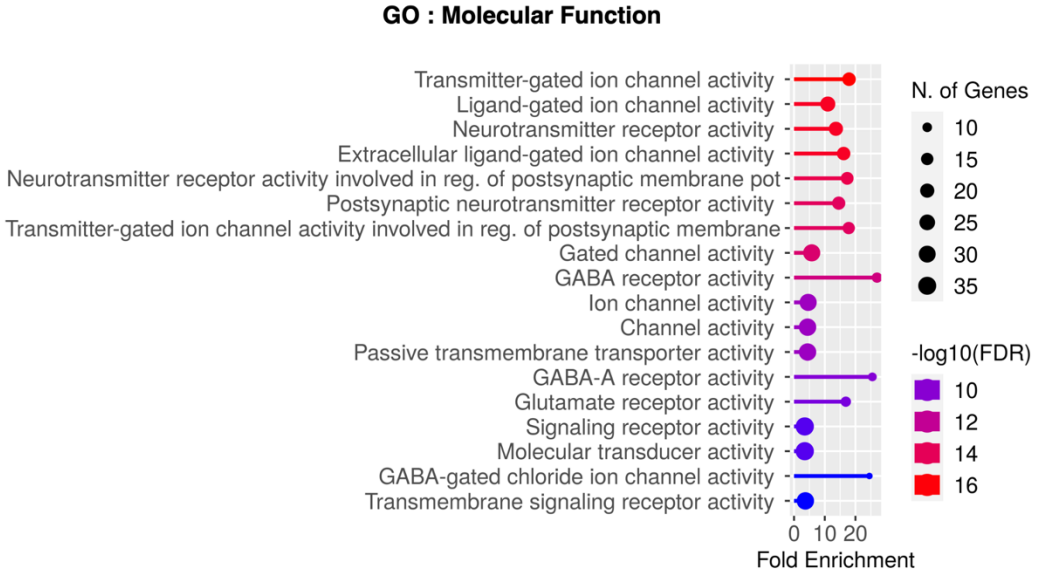
Supplementary Fig. S12





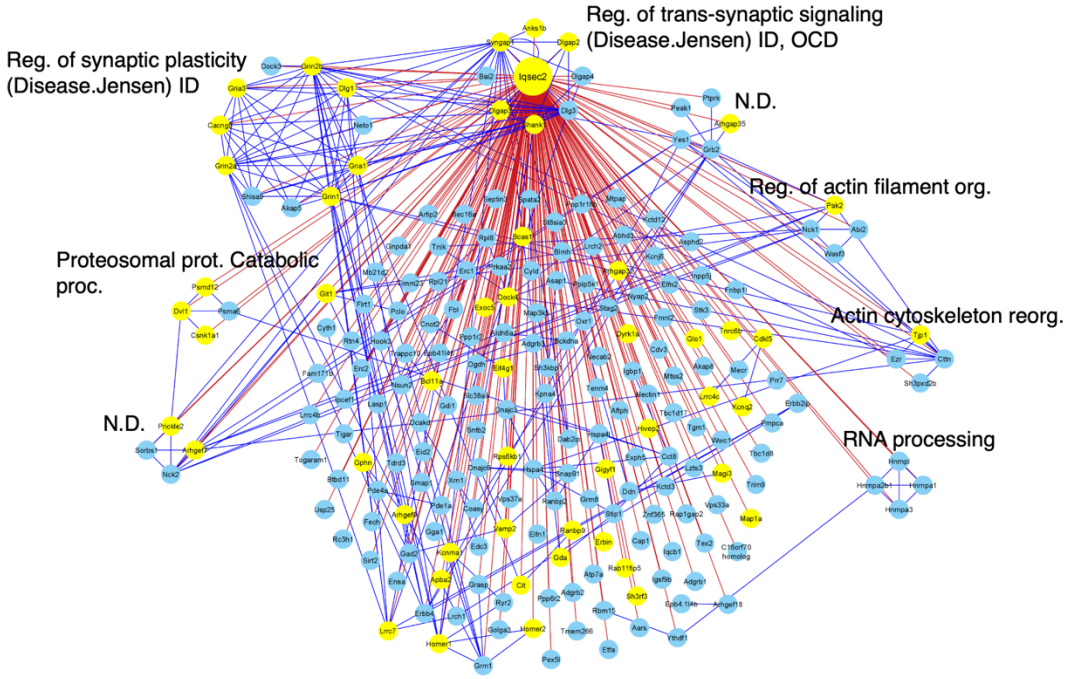
Supplementary Fig. S13



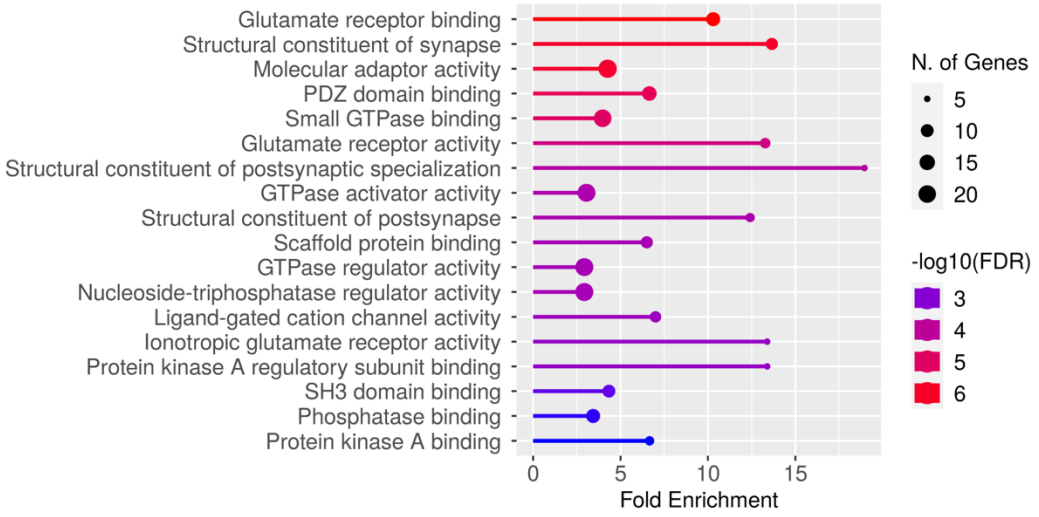


Supplementary Fig. S14

Iqsec2 interactome

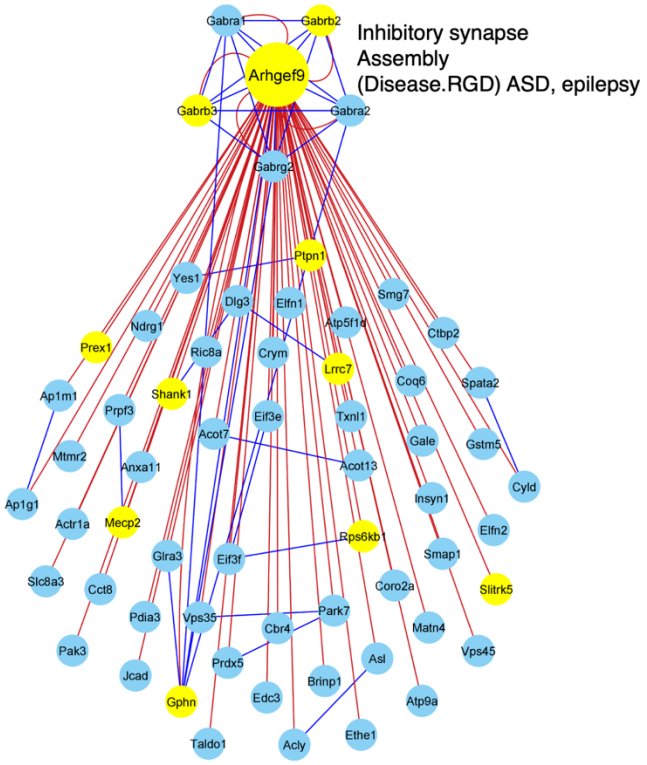


GO : Molecular Function

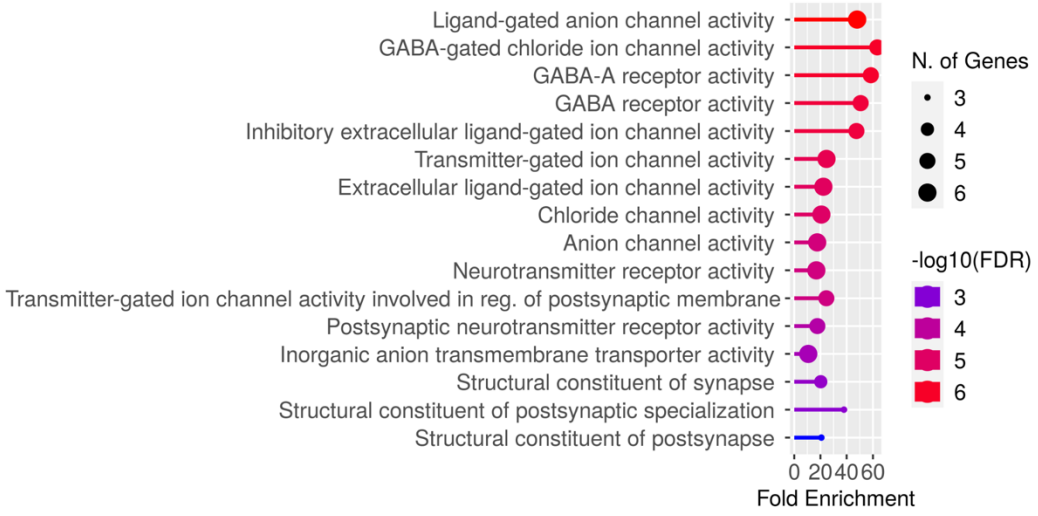


Supplementary Fig. S15

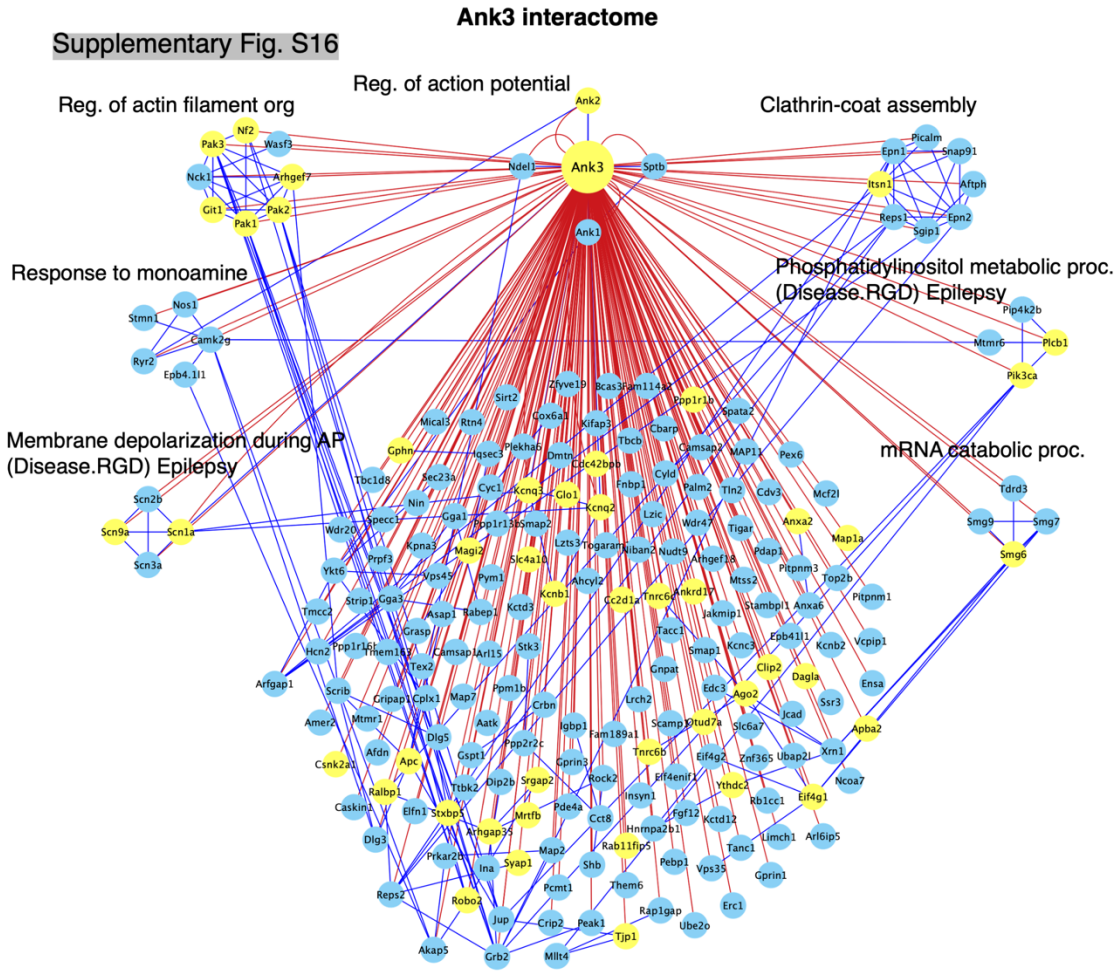
Arhgef9 interactome

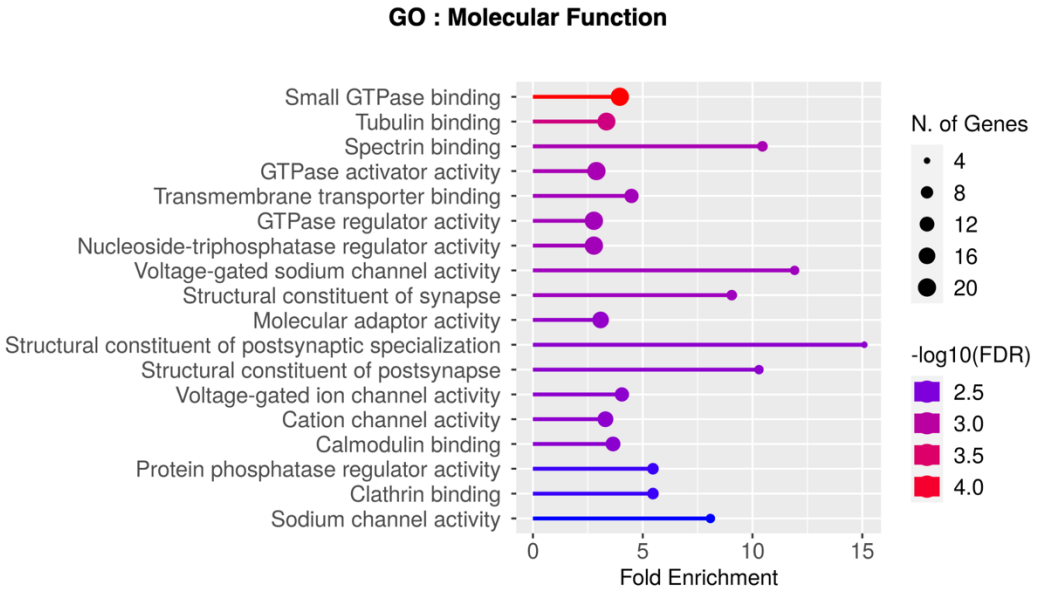


GO : Molecular Function



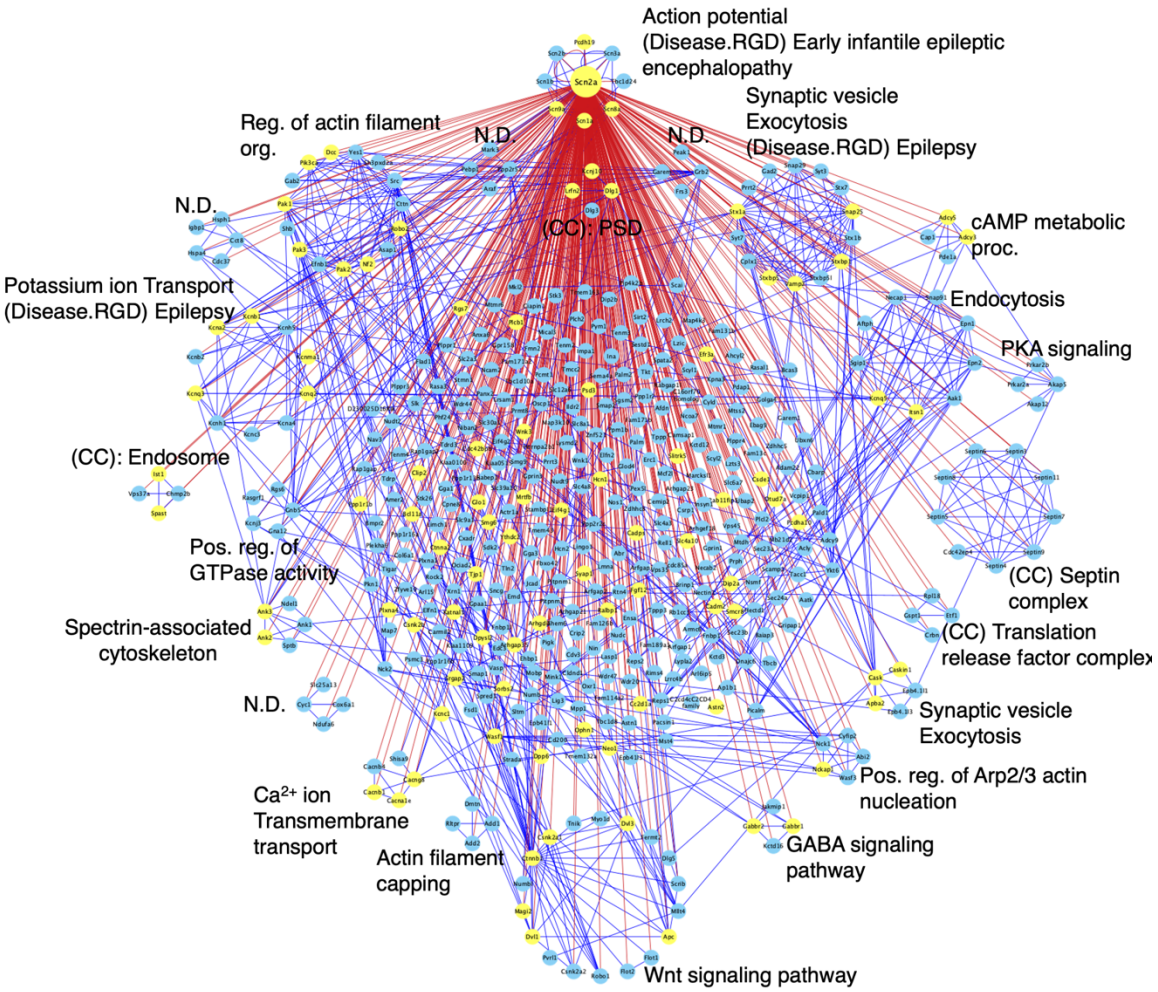
Supplementary Fig. S16



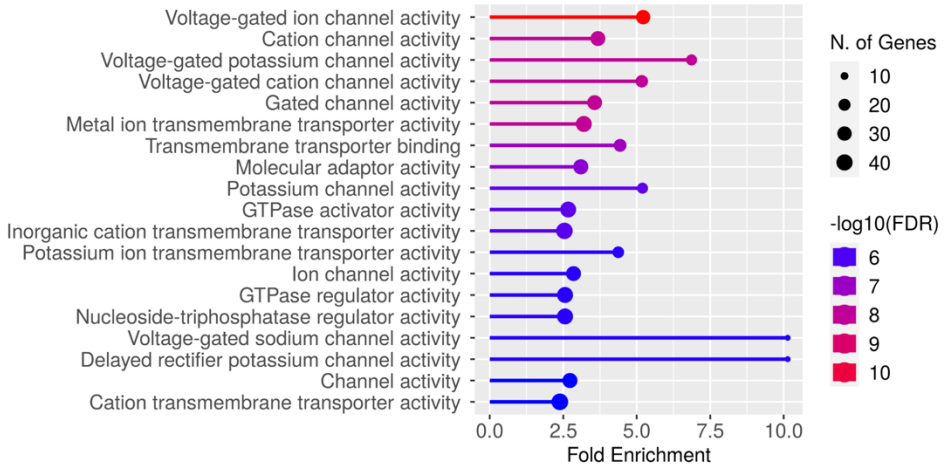


Supplementary Fig. S17

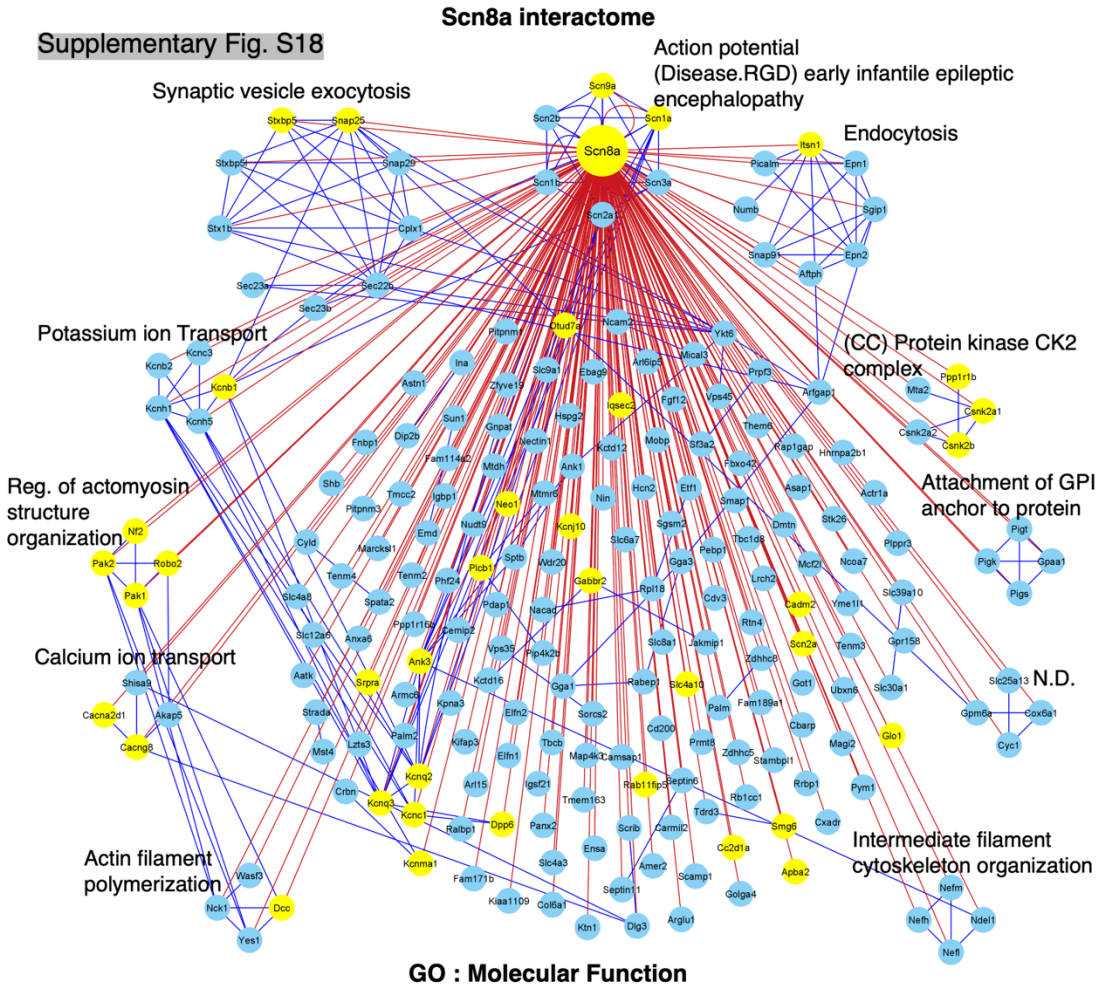
Scn2a interactome

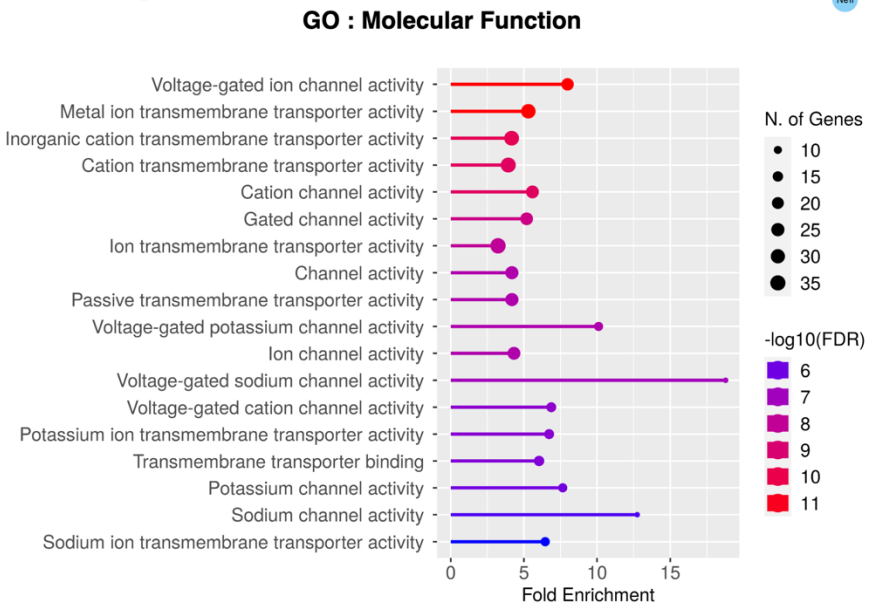
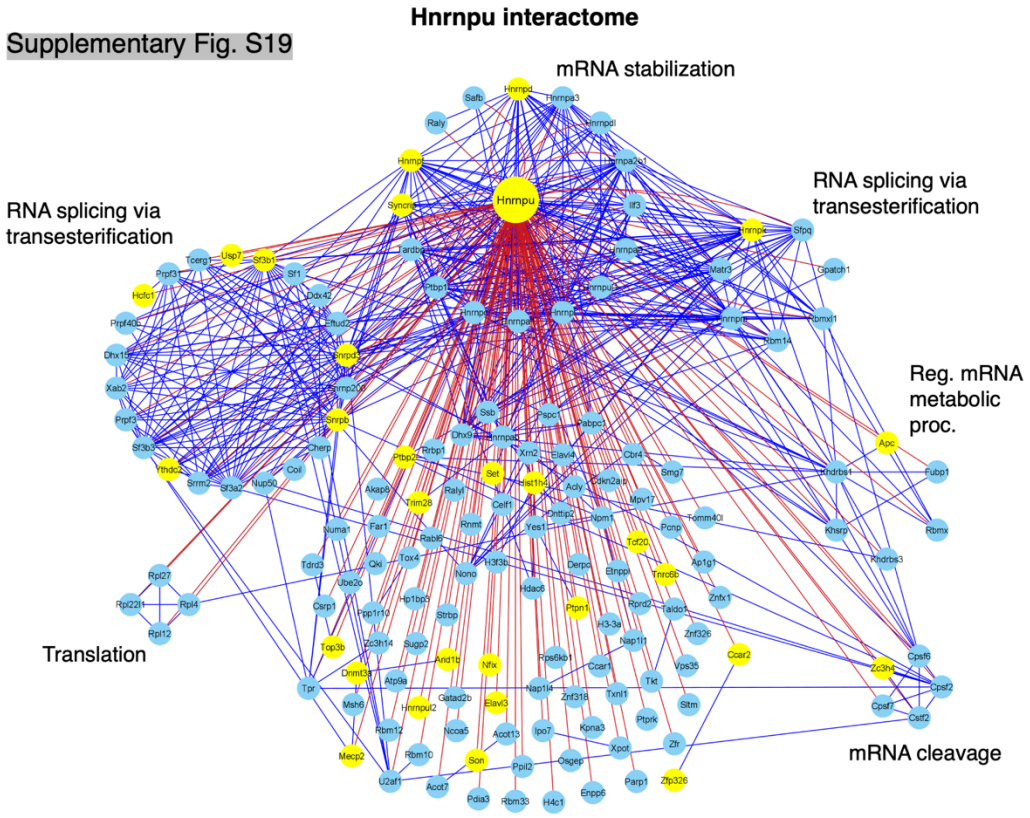
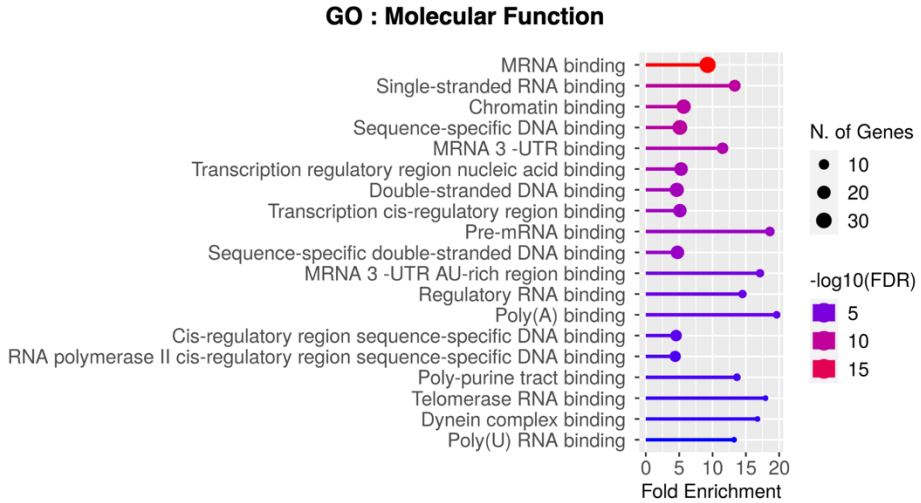


GO : Molecular Function



Supplementary Fig. S18





Supplementary Fig. S19