

# Proximity analysis of native proteomes reveals phenotypic modifiers in a mouse model of autism and related neurodevelopmental conditions

Received: 20 February 2024

Accepted: 23 July 2024

Published online: 09 August 2024

 Check for updates

Yudong Gao<sup>1,2</sup>, Daichi Shonai<sup>3</sup>, Matthew Trn<sup>1</sup>, Jieqing Zhao<sup>4</sup>, Erik J. Soderblom<sup>1,5</sup>, S. Alexandra Garcia-Moreno<sup>6</sup>, Charles A. Gersbach <sup>1,6,7,8</sup>, William C. Wetsel <sup>1,9,10,11</sup>, Geraldine Dawson <sup>9</sup>, Dmitry Velmeshev<sup>10</sup>, Yong-hui Jiang <sup>12,13,14</sup>, Laura G. Sloofman <sup>15,16</sup>, Joseph D. Buxbaum <sup>15,16,17,18</sup> & Scott H. Soderling <sup>1,10</sup> 

One of the main drivers of autism spectrum disorder is risk alleles within hundreds of genes, which may interact within shared but unknown protein complexes. Here we develop a scalable genome-editing-mediated approach to target 14 high-confidence autism risk genes within the mouse brain for proximity-based endogenous proteomics, achieving the identification of high-specificity spatial proteomes. The resulting native proximity proteomes are enriched for human genes dysregulated in the brain of autistic individuals, and reveal proximity interactions between proteins from high-confidence risk genes with those of lower-confidence that may provide new avenues to prioritize genetic risk. Importantly, the datasets are enriched for shared cellular functions and genetic interactions that may underlie the condition. We test this notion by spatial proteomics and CRISPR-based regulation of expression in two autism models, demonstrating functional interactions that modulate mechanisms of their dysregulation. Together, these results reveal native proteome networks in vivo relevant to autism, providing new inroads for understanding and manipulating the cellular drivers underpinning its etiology.

Autism spectrum disorder (hereinafter “autism”) is a neurodevelopmental condition associated with social communication difficulties and restricted and repetitive behaviors. Autism presents with significant clinical heterogeneity and complex genetic etiology<sup>1</sup>. Decades of research have identified and curated an evolving list of gene mutations associated with autism risk, many of which converge on pathways mediating synaptic/axonal functions and gene regulation<sup>2–7</sup>, and exhibit cell-type specific expression patterns across the brain<sup>8</sup>. Until recently, evidence of shared biology across these risk genes has

been inferred primarily from RNA-level gene expression results<sup>9–11</sup>, or interpreted from protein interaction analyzes<sup>12–15</sup>, although these are often derived from non-neuronal cells. Thus, there is a prevailing notion in the literature that molecular convergence in autism may be optimally reflected at the protein level in the brain.

One hallmark of neurons is their distinctive subcellular compartmentation - such as the synapse and axonal initial segment (AIS) - that are pivotal for neurotransmission<sup>16,17</sup>. In addition, gene expression regulators, many of which reside within the nucleus, also orchestrate

key milestones of neurodevelopment<sup>6</sup>. It is not surprising that autism risk converges on genes encoding proteins associated with these subcellular compartments<sup>18–22</sup>. Identifying the autism-associated proteome architecture of these compartments could define how seemingly diverse genetic mutations are functionally connected, thereby providing a roadmap to reveal a converging autism etiology. Current efforts to dissociate protein complexes and protein-protein interactions (PPIs) rely upon techniques such as cellular fractionation, immunoaffinity purification, cross-linking, and proximity-based proteomic methods, including biotin-identification (BioID)<sup>23–30</sup>. However, due to inherent specificity constraints of antibody-dependent immunoprecipitation or recombinant exogenous promoter-based strategies that express proteins at non-physiologic levels in non-native cell types, access to the native proximity proteomes of endogenously expressed autism risk proteins within brain tissue remains a significant challenge. Given these limitations, high-fidelity proteomes organized around autism risk proteins in the brain are clearly needed for better mechanistic insights.

Previously, we published a two-vector CRISPR/Cas9 approach, Homology independent Universal Genome Engineering (HiUGE), which enables rapid endogenous neuronal gene knock-in for protein modification *in vivo*<sup>31</sup>. Here, we have 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<sup>32</sup>, for scalable proximity proteomics of endogenous protein complexes. The approach is flexible also from a protein engineering perspective, as the proteins can be modified by placing TurboID within terminal coding exons or internally within introns using splicing sequences<sup>33</sup>. Importantly, because this strategy uses AAV transduction of readily available Cas9-transgenic mice, it obviates the significant time and cost burden of the traditional method of generating *in vivo* fusions of endogenous proteins by producing transgenic mice via germline transmission. Coupling this approach with *in vivo* BioID (iBioID)<sup>23</sup>, we have targeted 14 genetic drivers of autism that are mapped to the synapse, AIS, or the nucleus. Out of these 14 targets, 13 are SFARI gene score 1, one is gene score 2, also noted as syndromic, and 7 are from the autism risk gene lists of Satterstrom et al.<sup>3</sup>, or Fu et al.<sup>7</sup>. With this approach, we have modified the endogenous proteins with TurboID and then unraveled their native proximal proteomes directly from brain tissue. We have identified 1252 proteins within these 14 proximity proteomes and 3264 proximity PPIs associated with these autism targets. Amongst them, 16% are proteins encoded by mouse orthologs of SFARI genes<sup>2</sup>, 8% overlap with differentially expressed genes (DEGs) found in brain tissue of autistic individuals<sup>3</sup>, and 65% of the PPIs are not reported in STRING queries<sup>34,35</sup>. Notably, the proximity proteomes contain products of many newly discovered autism risk genes identified in recent studies<sup>7,36–38</sup> and reveal shared biological processes among autism proteins that may be predictive of interactions influencing autism phenotypes.

We have tested this notion by identifying intersections between HiUGE-iBioID proximity proteomes and co-perturbed proteins in two autism mouse models (Syngap1 synaptopathy and Scn2a channelopathy). In the Syngap1 model, we find that its binding with Anks1b is disrupted by an autism-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 communication in mice. The mutation also downregulates a key Scn2a modulatory protein cluster discovered in its proximity proteome, and results in aberrant attenuation of neural activity. Strikingly, re-expression of this cluster rescues this autism-associated electrophysiological impairment.

Together, our results establish a scalable platform to engineer endogenous proteins and map native proximity proteomes that are associated with genetic risks for neurodevelopmental conditions such

as autism. Our findings also reveal an intersectional approach to prioritize candidates based on proteomic co-perturbation. These data support a protein-centric model to reveal mechanisms of autism and related neurodevelopmental disorders and potential mitigation approaches.

## Results

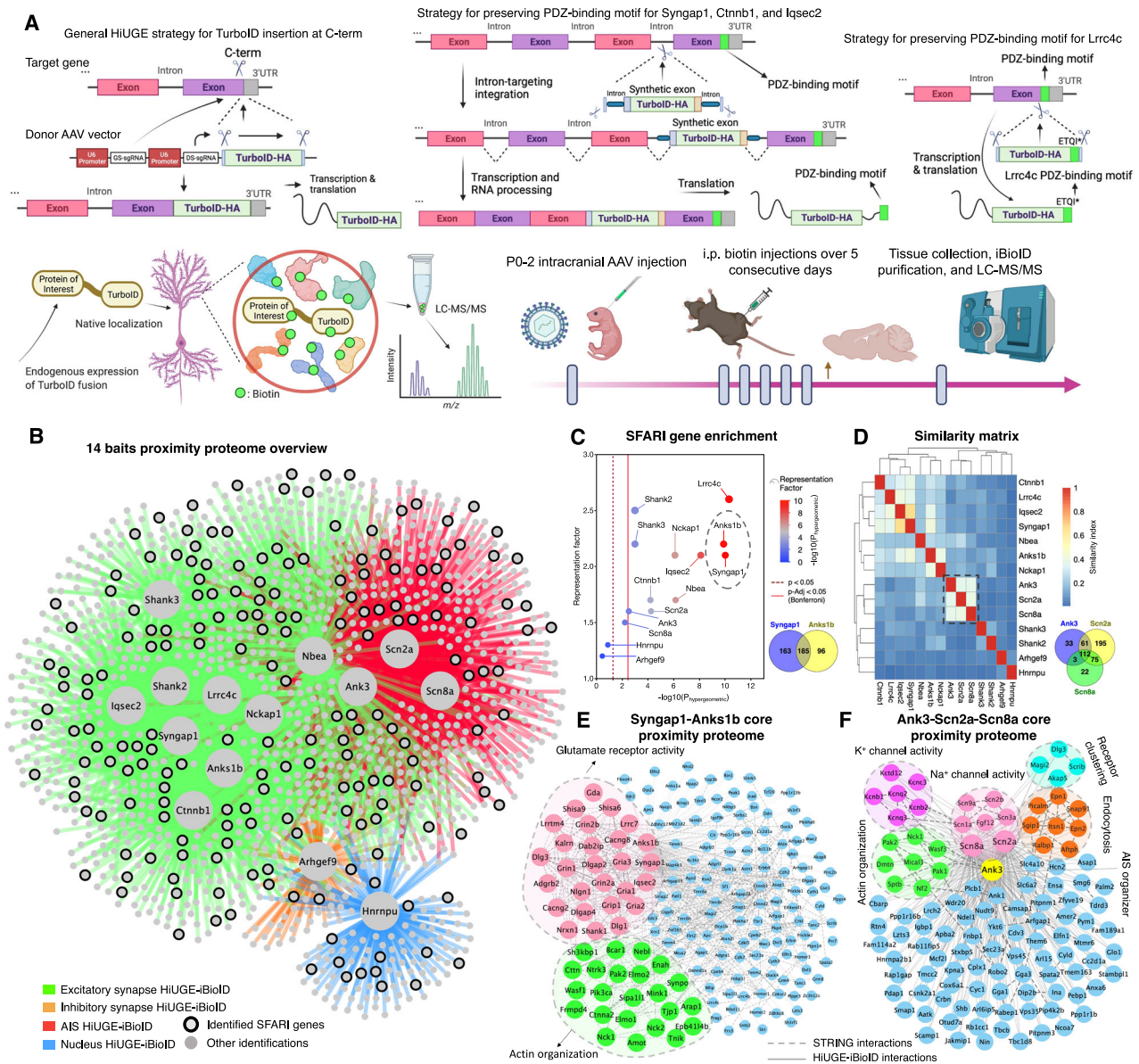
### Endogenous labeling of autism risk proteins using HiUGE

Previously we have used overexpression of bait proteins fused to various biotin ligases to discover proximity proteomes from brain tissue<sup>23,39</sup>. 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<sup>40</sup>. 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 an approach for iBioID experiments with *in-frame* TurboID fusions introduced into endogenous gene 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 autism is driven by mutations in a large number of risk genes whose products may interact with each other in unknown functional ways, we next targeted 14 high-confidence autism risk genes from the SFARI gene list (*Anks1b*, *Syngap1*, *Shank2*, *Shank3*, *Nckap1*, *Nbea*, *Ctnnb1*, *Lrrc4c*, *Iqsec2*, *Arhgef9*, *Ank3*, *Scn2a*, *Scn8a*, and *Hnrnpu*) that are expressed in neuronal compartments of the synapse, AIS, and nucleus<sup>2,41,42</sup>. Note, 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. Importantly, for targets with C-terminal (C-term) PDZ-binding motifs (*Syngap1*, *Ctnnb1*, *Iqsec2*, and *Lrrc4c*), intron-targeting<sup>33</sup> or custom donor strategies were used to preserve these critical protein interaction sites (Fig. 1A, S28). First, to confirm the proper localization of HiUGE-labeled targets, a highly immunogenic spaghetti monster (smFP) tag<sup>43</sup>, 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 nuclear compartments (Fig. S1L–O). We further validated that the HiUGE-labeling was colocalized with the immunofluorescence of specific antibodies, demonstrating the localization of these proteins was not affected by the tag fusion (Fig. S2). 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 then biotinylated surrounding proteins by supplying biotin via intraperitoneal (i.p.) injections over 5 consecutive days starting at -P21. Western blot analyzes of the purified streptavidin-precipitations from forebrain lysates collected at -P26 detected epitope-tagged baits at the expected molecular masses, confirming correct TurboID fusion protein expression (Fig. S3A–H).

### HiUGE-iBioID reveals endogenous proximity proteomes associated with autism risk proteins of diverse subcellular compartments

LC-MS/MS analysis of the 14 HiUGE-iBioID samples detected a total of 1252 proteins that were enriched in the bait proteomes, with expected interactions faithfully captured (Fig. 1B, S11–24, Supplementary Data S1.2). Importantly, 65% of the interactions detected were absent from STRING queries (using a generous stringency interaction score of 0.15<sup>34,35</sup>) - likely a reflection of prior studies being conducted in non-



**Fig. 1 | HiUGE-iBioID reveals endogenous proximity proteomes of 14 autism risk proteins. A** Schematic illustration of HiUGE-iBioID and its workflow. GS-sgRNA: gene-specific-gRNA. DS-sgRNA: donor-specific-gRNA. Strategies to preserve C-term PDZ-binding motifs of Syngap1, Ctnnb1, Iqsec2, and Lrrc4c are detailed. **B** Overview of 14 proximity proteomes that segregates according to expected bait functions. **C** Enrichment analysis of overlapping SFARI genes using

hypergeometric probability. The solid red line denotes Bonferroni adjusted p-value at 0.05. **D** Proximity proteome clustering based on a similarity matrix. **E** Core proximity proteome between Syngap1 and Anks1b that show highly significant overlaps with SFARI genes. **F** Core proximity proteome amongst Ank3, Scn2a, and Scn8a that are clustered based on similarity. Modules of proteins were isolated by MCL clustering or GO analysis.

neuronal cell types or methods other than proximity proteomics. Gene ontology (GO) analyzes 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); thereby demonstrating high fidelity identification of local proximity proteomes (Fig. S11–24, S27). These networks were 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. S21–23, Supplementary Data S2) echoed a previous study that suggested their roles in regulating cytoskeletal stability and polarized trafficking at the AIS<sup>24</sup>. Reciprocal analyzes also supported the reproducibility of the proximity proteomes, showing that the baits frequently cross-identified

each other (Fig. S4A, Supplementary Data S2). Additional HiUGE-iBioID of proteomic “hubs” (e.g., Homer1 and Wasf1, which were commonly associated with many baits) demonstrated they could reversely detect the majority of baits in validation experiments (Fig. S4B–H). In addition, comparisons showed that HiUGE-iBioID proximity proteomes significantly overlap with, but also differ from proteomic detections derived from independent antibody immunoprecipitations we performed for Anks1b and Scn2a (Fig. S4I, J). Interactions specific to HiUGE-iBioID compared to these antibody pull-downs were expected since BioID is based on covalent labeling of nearby proteins and thus does not require transient or weak interactions to survive the pull-down and washing steps of immunoaffinity purification. Of note, proteomic detections unique to HiUGE-iBioID conformed to the expected top molecular function GO terms of Anks1b and Scn2a, while detections unique to immunoprecipitations did not (Fig. S4K, L).

We also sought to determine if the proximity proteomes showed a significant overlap with differentially expressed genes (DEGs) from specific cell types in autistic individuals by cross-referencing a recent human tissue single-cell genomics study<sup>8</sup>. HiUGE-iBioID networks showed the highest level of overlap with autism DEGs in cortical layer 2/3 (L2/3) excitatory neurons (Fig. S5A), consistent with the study suggesting L2/3 neurons are significantly affected in autism<sup>8</sup>. Notably, most networks were also significantly enriched for other autism risk genes (SFARI, Satterstrom et al.<sup>3</sup>, and Fu et al.<sup>7</sup>; Fig. 1C, S5B), demonstrating the convergence of autism genetic susceptibilities at the protein interaction level in previously unknown ways. We also performed enrichment analyses between these networks and an autism genome-wide association meta-analysis (GWAS) study<sup>44</sup>; however, the result was largely nonsignificant or marginal, likely due to limited power (Supplementary Data S7). It has also been noted that larger autism cohorts<sup>38</sup>, or new approaches, are needed to determine the potential genome-wide significance of a large number of moderate-risk genes. Importantly, the HiUGE-iBioID networks of high-confidence autism risk proteins formed physical communities with other candidate proteins of moderate confidence (Supplementary Data S2, overlap tab). These genes represent a resource of moderate autism candidates that likely should be prioritized in future studies of genetic contributions due to their possible functional interactions with high-confidence autism risk genes. Furthermore, we noted the proximity proteomes contained numerous potentially druggable targets, including ~80 kinases and phosphatases, 18 of which are encoded by SFARI gene mouse orthologs: *Camk2a*, *Camk2b*, *Cask*, *Cdc42bpb*, *Cdkl5*, *Csnk1e*, *Csnk2a1*, *Dyrk1a*, *Mapk3*, *Ntrk2*, *Ntrk3*, *Oclri*, *Pak1*, *Pak2*, *Ppp3ca*, *Rps6ka2*, *Taok2*, *Wnk3*.

Finally, similarity clustering of the bait proximity proteomes revealed subgroups that largely segregated according to their expected cellular compartments (Fig. 1D). Networks of the overlapping proximity proteomes between two baits with highly significant SFARI gene enrichment (*Syngap1* and *Anks1b*) and three similarity-clustered AIS baits (*Ank3*, *Scn2a*, and *Scn8a*) revealed shared pathways that emphasized glutamate receptor and voltage-gated channel activities, respectively, and common modules that were involved in actin organization in both networks (Fig. 1E,F). These proximity proteomes may contain proteins that function together with the autism-associated baits, modulation of which may reveal the complex genetic risks of autism or serve as potential inroads for normalizing relevant phenotypes.

### **Syngap1 mutations lead to reshaping of its synaptic proximity proteome and loss of Anks1b binding**

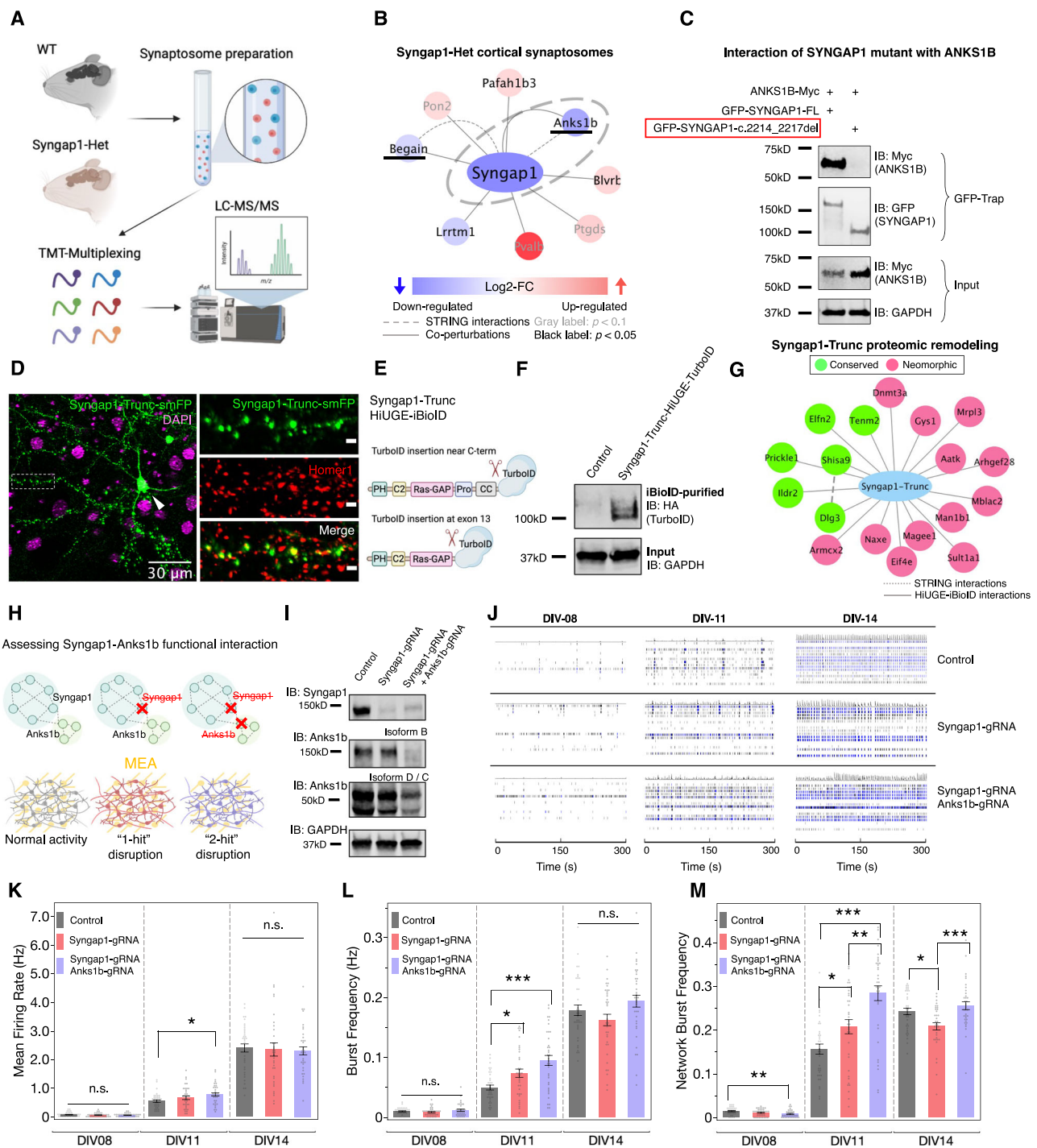
We hypothesized that proximity proteomic discovery could inform functional interactions relevant to the phenotypes underlying autism. To test this proposition, we first focused on the intersection between *Syngap1* and *Anks1b*, as their proteomes show a high degree of overlap and exhibit significant SFARI gene enrichment (Fig. 1C). Markov cluster algorithm (MCL)<sup>45</sup> analysis of their overlapping proximity proteomes identified a large PPI community (27 proteins, including *Syngap1* and *Anks1b*), that was significantly enriched for pathways of “autistic disorder” and “glutamate receptor activity”, indicating this shared cluster may regulate excitatory synaptic transmission in autism. Although molecular interaction between *Syngap1*-*Anks1b* has been reported previously<sup>27,46–48</sup>, an understanding of their functional interaction remains limited. Prior studies have indicated that *Anks1b* and *Syngap1* both regulate synaptic activity and plasticity through NMDA-type glutamate receptors<sup>49,50</sup>, they are dispersed from the PSD in response to synaptic activity in a CaMKII-dependent manner<sup>51–53</sup>, and are associated with similar autism-like phenotypes in haploinsufficiency models<sup>54,55</sup>. Hence, we sought to test whether *Syngap1* and *Anks1b* functionally interact in driving neuronal phenotypes.

We first analyzed the synaptic proteomes in wildtype (WT) and *Syngap1* heterozygous (*Syngap1*-Het) mice<sup>56–58</sup> (Fig. 2A) to determine if *Anks1b* was influenced by haploinsufficiency of *Syngap1*. Fractionation steps were performed to isolate the synaptosomes<sup>26,59,60</sup> from the cerebral 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, Supplementary Data S4), supporting the notion of a very close functional interaction between *Syngap1* and *Anks1b* not only in typical physiology, but also in autism-associated synaptopathy.

As truncating mutations of *Syngap1* confer significant genetic risk for autism<sup>50,55</sup>, 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-term 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. S6). A pathological mutation found in a human patient (SYNGAP1-c.2214\_2217del<sup>61</sup>), which resulted in a frame-shift and premature truncation in this region, also abolished the SYNGAP1-ANKS1B interaction (Fig. 2C). A previous report showed that a 2 a.a. substitution in the C-term PDZ-binding motif of *Syngap1* impaired several synaptic PPIs, but not *Anks1b*<sup>62</sup>. As an orthogonal validation of the putative *Anks1b* interaction region identified above, we sought to test if a more disruptive truncation of *Syngap1* that ablated this region could lead to a deeper remodeling of its proximity proteome, including a loss of the *Anks1b* interaction in vivo. Endogenously expressed smFP-labeled *Syngap1* truncated at exon 13 (*Syngap1*:1-744-smFP) retained synaptic localization in cultured neurons, although mis-localization in the soma was detected as well (Fig. 2D). We then targeted this locus to generate TurboID fusion to the truncated *Syngap1* protein. 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, Supplementary Data S1). Thus, we confirmed that the region identified by structure-function analysis in vitro is also crucial for the *Syngap1*-*Anks1b* interaction in vivo.

### **Depletion of Anks1b exacerbates electrophysiological abnormalities associated with Syngap1 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* loss-of-function (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 DIV 11, but not at DIV 8 or DIV 14 (Fig. 2J–M). This result phenocopied previous reports demonstrating that *Syngap1*-LOF atypically accelerated synaptic and network activities during the synaptogenesis period<sup>63,64</sup>. Interestingly, further depletion of *Anks1b* exacerbated the abnormally 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 the development of neuronal activity. Based on the phenotype and common biological pathways found in the overlapping proximity proteomes, this effect may be due to the altered developmental trajectory of the glutamate receptor module found in both networks.



**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. Proteins that overlap with the Syngap1 proximity proteome are underlined. **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). Scale bar in the enlarged view represents 2  $\mu$ m. **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** Proximity proteomic network showing conserved and neomorphic interactions in Syngap1 truncation. Note that interaction with Anks1b is no longer detected. **H** Schematics assessing phenotypes of the Syngap1-Anks1b functional interaction. **I** Western blot confirming disruption of Syngap1 and Anks1b expression. **C**, **D**, **I** Representative experiments are based on three replicates with similar results. **J** Representative raster plots of neural activities at DIV-08, 11, and 14. **K–M** Neurometrics 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). Plots are mean  $\pm$  SEM.

## HiUGE-iBioID and spatial co-perturbation proteomics reveal targets that can restore spontaneous activity of neurons harboring a patient-derived Scn2a<sup>+R102Q</sup> mutation

Another critical rationale for interrogating endogenous proximity proteomes is the prospect of identifying proteins functioning with the autism-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 autism<sup>3,65–67</sup>. In recent years, whole exome sequencing (WES) from autistic individuals identified a missense mutation, SCN2A-p.R102Q, that was previously reported<sup>68,69</sup> but uncharacterized. Clinical observations of one patient with this mutation included meeting DSM-5 diagnostic criteria for autism spectrum disorder (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 autism, we generated a Scn2a point-mutant heterozygous mouse model (Scn2a<sup>+R102Q</sup>) (Fig. 3A). Behavioral testing of the Scn2a<sup>+R102Q</sup> mice found that compared to WT littermates, the mutants presented with hyperactivity and reduced anxiety in the elevated zero maze, excessive repetitive behaviors in the self-grooming and the hole-board tests, and reduced ultrasonic vocalizations (USV) during social interactions (Fig. 3B). Social behavioral tests were also conducted with the resident-intruder and social dyadic assays where C3H/HeJ males served as social partners of WT and Scn2a<sup>+R102Q</sup> males. In both the resident-intruder and the dyadic tests, the social behaviors of the mutants appear to be relatively similar to WT, although abnormal numbers of social events and reduced withdrawal were noted (Fig. S7). Hence, the primary behavioral deficits of Scn2a<sup>+R102Q</sup> mice lie within social communication and repetitive behaviors. It is known that mutations in Scn2a can result in either infantile epileptic encephalopathy (IEE) or autism, depending upon the nature of the mutational effect (gain- or loss-of-function)<sup>70</sup>. Cortical neurons cultured from Scn2a<sup>+R102Q</sup> mice exhibited a significant attenuation in neural firing activity (~40% decrease at DIV 14;  $p < 0.001$ , Fig. 3D), consistent with a role in autism but not IEE. Next, we analyzed the spatial proteome of these mutants versus WT mice using fractionation steps of whole brain tissue adapted from the Localization of Organelle Proteins by Isotope Tagging after Differential ultraCentrifugation (LOPIT-DC) procedure<sup>30</sup>, as we have published previously<sup>71</sup>. Interestingly, we identified a downregulation of a protein cluster associated with voltage-gated sodium channel (VGSC) activity in Scn2a<sup>+R102Q</sup> mice. This VGSC modulatory cluster included candidates intersecting with the Scn2a HiUGE-iBioID proximity proteome: an auxiliary  $\beta$  subunit, Scn1b<sup>72</sup>, and an intracellular VGSC modulator, Fgf12<sup>73</sup> (Fig. 3C; Supplementary Data S4).

Based on the finding of the downregulated VGSC modulatory cluster, a rescue strategy to restore the electrophysiological deficits in Scn2a<sup>+R102Q</sup> neurons was devised. First, lentiviral-mediated CRISPR-activation (CRISPRa) was used to upregulate endogenous Scn2a expression (Fig. S8A). A similar strategy has recently been used to phenotypically rescue Scn2a haploinsufficiency<sup>74</sup>; however, it 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<sup>+R102Q</sup> neurons showed normal transcript levels (Fig. S8B), suggesting the reduced protein levels in the mutant were 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 potentially could be diluted by the mutant allele. Indeed, CRISPRa treatment only partially rescued the Scn2a<sup>+R102Q</sup> phenotype (Fig. 3D, purple bars), suggesting this approach

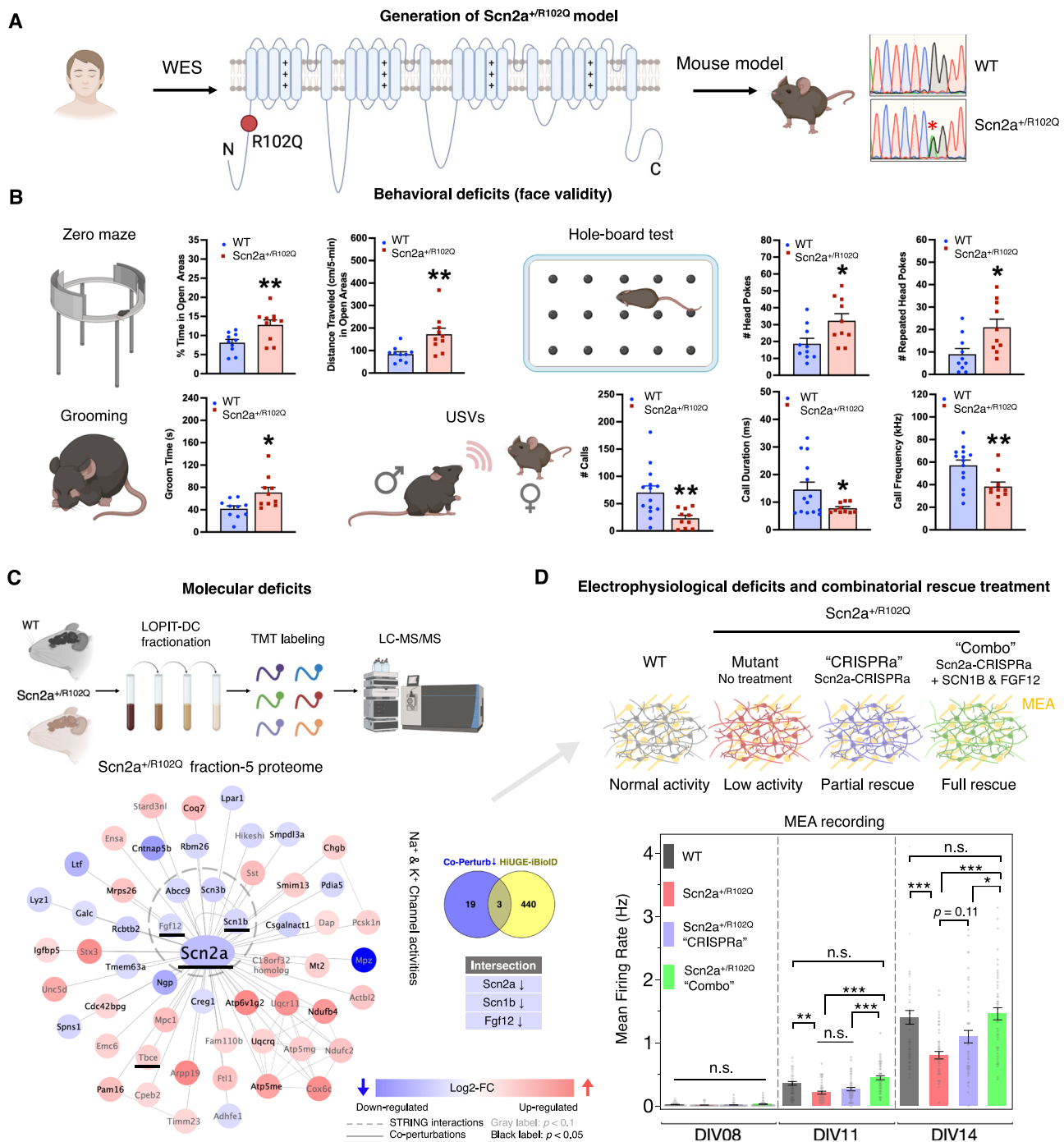
is unlikely to work in the context of missense mutations. A treatment strategy was next tested by supplying either additional SCN1B or FGF12 via AAV-mediated expression to augment these downregulated intersecting proteins identified in the VGSC modulatory cluster. MEA analysis revealed that neither expressing SCN1B or FGF12 alone, or when added together, was sufficient to rescue the phenotype (Fig. S8C). Finally, a combinatory (“combo”) approach was tested, combining the upregulation of all three of the depleted VGSC cluster proteins. This combo approach resulted in a full restoration of neural firing activity metrics at DIV 14 (Fig. 3D, S8D), suggesting the potential dominant-negative effect of the mutation in Scn2a-CRISPRa was overcome by the increased expression of SCN1B and FGF12; thereby confirming the hypothesis that this modulatory cluster is crucial for the phenotypic rescue of Scn2a<sup>+R102Q</sup>. Collectively, these results strongly indicate that intersectional proteomics between HiUGE-iBioID and spatial co-perturbation can be an informative approach to discover novel approaches for the functional rescue of abnormal phenotypes and potential therapeutic targets.

## Discussion

Here we report a strategy combining the advantages of HiUGE and iBioID to resolve native proximity proteomes associated with 14 autism-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 autism genetics and proteomics. The findings also emphasize an effective proteomic-driven systems-biology approach to discover molecular mechanisms and potential treatment targets.

Compared to immunoprecipitation or recombinant BioID expression methods *in vitro*, HiUGE-iBioID is expected to have 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 proximity proteomes are covalently marked as they exist *in vivo* and thus, the proteins can be purified subsequently under stringent conditions without the need to optimize samples for weak or transient interactions. This feature is especially important for protein complexes organized by transmembrane baits, which must be extracted from the lipid bi-layer under conditions that are unfavorable for maintaining many PPIs. Fourth, the selection of the bait protein is not limited by viral packaging capacity or availability of high-specificity and validated antibodies.

HiUGE-iBioID is a technique that, from a biochemical perspective, combines the strengths of endogenous protein-based purification from tissue that antibodies afford with the advantages of covalent marking of proximity proteomes by BioID. Nevertheless, an essential question is how its proximity proteomes compare with analogous studies using immunoprecipitation or recombinant BioID expression. To address this important issue, we performed a direct comparison of HiUGE-iBioID to antibody immunoprecipitations of Anks1b and Scn2a from mouse brain tissue (Fig. S4 I, J). While there was a significant overlap in the proteomes between HiUGE-iBioID and each immunoprecipitation, the differences between them were even greater (Fig. S4 K, L). For example, immunoprecipitation of the voltage-gated sodium channel, Scn2a, identified the expected proteins that overlap with HiUGE-iBioID, including clusters of Na<sup>+</sup> / K<sup>+</sup> channels and Fgf12, consistent with previous reports<sup>28,75</sup>. However, GO analysis of the fraction that did not overlap with HiUGE-iBioID yielded top-ranked molecular function terms of “structural constituent of the ribosome” and multiple RNA processes. In contrast, the fraction unique to HiUGE-iBioID were enriched for the expected top GO terms of voltage-gated ion channel functions. Similar conclusions were noted from the comparison of HiUGE-iBioID to the recent publication of



**Fig. 3 | Intersectional proteomics reveal hidden molecular mechanism of a patient-derived *Scn2a* mutation.** **A** Generation of a mouse model based on a clinically identified *Scn2a* missense mutation (R102Q) in autistic individuals. WES: whole exome sequencing. **B** Behavioral face-validity of *Scn2a*<sup>+/R102Q</sup> mutants was assessed by the zero maze as the percent time and distance traveled in the open areas; the hole-board test as numbers of head pokes and repeated head-pokes; the self-grooming test; and the ultrasonic vocalizations (USVs) as numbers of calls, call durations, and call frequencies during social interaction. No differences were detected in the metrics of pre-social (baseline) responses in the USV tests. \**p* < 0.05, \*\**p* < 0.01, WT vs. *Scn2a*<sup>+/R102Q</sup> mice; independent samples t-tests, two-tailed (*n* = 10 or 14 mice per group). Statistics are summarized in Supplementary Data S6. Plots

are mean ± SEM. **C** Spatial proteomics reveals co-perturbations in *Scn2a*<sup>+/R102Q</sup> mutants, two-tailed heteroscedastic t-test on log<sub>2</sub>-transformed data. MCL analysis discovered a key cluster associated with voltage-gated channel activity that is downregulated, including three targets that intersect with the *Scn2a* HiUGE-iBioID proximity proteome (underlined). **D** *Scn2a*<sup>+/R102Q</sup> 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). Plots are mean ± SEM.

immunoprecipitation-based proteomes of Syngap1, Shank3, *Scn2a*, and *Cttnb1*<sup>28</sup> (Fig. S9). We also compared recently published data of recombinant bait-BioID expression in cultured mouse neurons to HiUGE-iBioID for the baits Syngap1, Shank3, and *Lrrc4c* that were

shared between both studies<sup>27</sup> (Fig. S10). While the BioID and HiUGE-iBioID proximity proteomes had more in common with each other than the analogous immunoprecipitation comparisons, there were again more differences between them than there were similarities.

While HiUGE-iBioID specific proximity proteomes for each had the top GO term “Glutamate receptor binding” for all three synaptic baits, the recombinant BioID specific data were enriched for top GO terms such as “Tubulin binding”, “RNA binding”, “Heat shock protein binding”, and “Acidic amino acid transmembrane transporter activity”. While further comparisons are needed to confirm the above observations, the available data suggest that endogenous proximity proteomics using HiUGE-iBioID is a valuable addition to the existing methods.

Consistent with the observation that HiUGE-iBioID performs well, each proximity proteome revealed the expected and additional biological insights following further analysis with MCL to partition biologically-relevant communities (Fig. S11–24). For example, neurobeachin (Nbea) (Fig. S16) is known to regulate the surface levels of ionotropic GABA and glutamate receptors as well serve as an A-Kinase (PKA) Anchoring Protein (AKAP)<sup>76–79</sup>. 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 “autism spectrum disorder” and “epilepsy”, both to which Nbea is implicated<sup>80,81</sup>. Interestingly, GABA and glutamate surface levels are thought to be modulated by distinct pathways influenced by Nbea<sup>82</sup>. Consistent with this idea, clusters enriched for “AP-type membrane coat adapter 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”, although additional experimental analyzes will be needed to test these new proteomics-derived hypotheses.

In addition to specificity advantages, HiUGE-iBioID is easily scalable in terms of time and resources when compared to the alternative of the traditional transgenic animal approaches to generate endogenous protein fusions *in vivo*. We have found that forebrain tissue collected from as few as two mice is sufficient for one biological replicate. Therefore, a typical mouse litter (~8 pups) is sufficient to analyze a bait proximity proteome in biological triplicates. We anticipate this method can be optimized even further for larger-scale applications, where costs can be limiting. These approaches could include multiplexed mass-spectrometry techniques such as isobaric labeling<sup>83,84</sup>, or new advancements in mass spectrometry scan rates that reduce instrument time.

We also recognize a few limitations of our method. 1) The proximity proteomes reported here are based on an enrichment with the bait over negative control, thus it is possible some proteins are present yet not identified as significantly enriched<sup>85</sup>. We recognize that not all previously reported interactions were found to be enriched in our experiments, including *Dlgap1* and *Dlgap4* for the Shank2 and Shank3 baits<sup>86</sup>. 2) Since the fusion proteins are expressed at an endogenous level, efficient biotinylation may be challenging for some low-expression proteins. 3) Common to all protein fusion strategies, it is unrealistic to guarantee that the tag does not alter any interaction. The ability to generate fusion proteins by inserting TurboID directly into coding exons, or by splicing TurboID in from introns for proteins that cannot be altered at their N- or C-termini, provides flexible gRNA selection as well as protein engineering design. 4) Off-target integrations can occur in cells using CRISPR approaches, which is why it is important to verify the bait protein is detected in the resulting proteomics analysis. In our prior publication on the HiUGE approach, we experimentally determined the off-target rate is low and largely results in integration in non-coding regions. As the integration of TurboID into each cell in non-dividing neurons is an independent effect, it is likely that rare off-target integrations that occurred in-frame with a coding

region would make limited contribution to the total detected proximity proteome. 5) HiUGE-iBioID relies on sparse editing of cells within brain tissue. While we have demonstrated that the efficiency is sufficient for biochemical approaches such as proximity proteomics, one should keep in mind that some cell types may edit better than others, presenting potential bias in the data. 6) Unlike transgenic mouse production, HiUGE editing can result in cells that are either heterozygous or homozygous tagged in the same tissue, potentially contributing to functional heterogeneity in the edited cells. Furthermore, germline modifications can leverage Cre lines to specify cell types, while the currently presented HiUGE-iBioID approach does not. We note, however, that cell-type specificity for HiUGE-iBioID could be achieved by using Cre-dependent expression of Cas9. We envision additional future developments to our approach will include adapting *in silico* structure prediction tools (e.g. AlphaFold<sup>87</sup>, ESMFold<sup>88</sup>, and RoseTTAFold<sup>89</sup>) to minimize potential disturbances by structure-based optimization of TurboID fusions. Additionally, rapid advancements in the sensitivity of mass-spectrometry<sup>90–92</sup> promise to enable further HiUGE-iBioID analysis, even for low-level proteins.

A crucial finding from the current HiUGE-iBioID application is that diverse autism genes physically form protein interaction networks with each other and can be co-perturbed in genetic models of autism. This result confirms that divergent genetic mutations can converge at the protein level to drive autism neurobiology, providing a working paradigm to prioritize co-regulated candidates in physiology and atypical brain development. Furthermore, we demonstrate that high-confidence autism genes interact with other lower-confidence autism genes at the proteome level. Indeed, three proteins detected in the proximity proteome dataset (*Its1n1*, associated with *Scn2a*, *Scn8a*, *Ank3*, and *Shank3* baits; *Nav3*, associated with *Scn2a* bait; and *Hnrnpul2*, associated with *Hnrnpu* bait) were discovered subsequently as autism risk genes of genome-wide significance<sup>38</sup> during the preparation of this manuscript. These results emphasize the possibility that additional genes – whose significance in autism are as yet unknown – exist in the dataset. Thus, HiUGE-iBioID may be useful to prioritize lower-confidence autism genes, which could either play a role in regulating core autism drivers or serve as targets for pharmacological developments. In addition, we expect the endogenous proximity proteomics data will stimulate a new impetus for predictive modeling of autism genetics<sup>93,94</sup> and cell-type specific analyzes harnessing single-cell proteomics<sup>95–97</sup>.

Informed by the proximity proteomes and proteomic co-perturbation results, we focused on a tightly co-regulated pair, *Syngap1* and *Anks1b*<sup>27,46–48</sup>, and verified the functional significance of their interaction. We identified a putative internal region on *Syngap1* that is crucial for interaction with *Anks1b*, and confirmed that *Syngap1* truncation led to remodeling of its proximity proteome, including a loss of interaction with *Anks1b*. The data support the notion that remodeling of synaptic protein complexes, such as previously reported PDZ-associated alterations<sup>62</sup> and the disruption of specific interactions like *Syngap1*-*Anks1b*, could be possible mechanisms relevant to *Syngap1* synaptopathy beyond nonsense-mediated decay<sup>50</sup>. Further, we identified several neomorphic proteomic interactions associated with the *Syngap1* truncation (Fig. 2G). The mechanisms of whether or how these abnormal interactions might contribute to *Syngap1*-associated phenotypes remain to be carefully assessed, and cannot be easily extrapolated to patients. The additive effects of *Syngap1* and *Anks1b* deficits seen in the aberrant neural activity indicate that the downregulation of *Anks1b* in *Syngap1*-Het mice is not a compensatory effect, but rather it may contribute to a mechanism that alters neuronal development. The disordered region of *Syngap1*, identified as a critical domain mediating binding with *Anks1b*, contains proline-rich stretches, a poly-histidine motif, and is phosphorylated at multiple sites<sup>98,99</sup>. This architecture suggests their interaction could be under activity-dependent kinase modulation. Accordingly, within the *Anks1b*-



Syngap1 core proximity proteome, there are 12 kinases / phosphatases, 4 of which are likely autism risk genes themselves (Pak2, Cdk15, Ntrk3, Dyrk1a)<sup>2</sup>. Thus, an intriguing speculation is that these gene products may play a role in regulating synaptogenesis via acting on the Syngap1/Anks1b complex, and contributing to the activity-dependent shuttling of Syngap1/Anks1b during plasticity.

A notable finding of this study is the discovery of a mechanism 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 proximity proteome with co-perturbation spatial-proteome in a mouse model harboring this mutation. A qPCR analysis revealed that the abundances of Scn2a, Scn1b, and Fgf12 mRNAs in cultured Scn2a<sup>+R102Q</sup> neurons are comparable to the WT (Fig. S8B), 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-term intracellular tail of Scn2a affects VGSC assembly remains to be determined. Similarly, the effects of these perturbations need to be assessed further as some alterations may be benign with respect to phenotypes.

Both  $\beta$ -subunits and FGF family proteins play critical roles in maintaining neuronal excitability through regulating VGSC kinetics<sup>72,100</sup>. Thus, their downregulation in Scn2a<sup>+R102Q</sup> is likely a contributing factor to VGSC channelopathy. These results also explain why Scn2a-CRISPRa treatment alone is insufficient to rescue the electrophysiological deficits, especially since the treatment does not differentiate the functional allele from the point-mutant allele<sup>74</sup>. We further show that supplying additional SCN1B and FGF12 without Scn2a-CRISPRa does not rescue the deficits either (Fig. S8C). Thus, it appears that upregulation of all three components is required to provide the molecular environment necessary 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 should 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<sup>65,101</sup>, future studies are needed to further dissect their unique subcellular effects.

Together, our results show that HiUGE-iBioID provides a CRISPR/proximity proteomics method to reveal native proteomes in vivo. Combined with co-perturbation proteomics, our intersectional approach offers a generalizable strategy to identify and prioritize candidates for discovering new biology and potential therapeutic targets. Thus, future work could adopt a similar approach to investigate genetic co-perturbations and PPIs in other models. We expect that the framework developed in this study will encourage further research of the native proximity proteomes in the brain, enhancing our understanding of proteome organization in various aspects of cellular neurobiology and disease.

## 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<sup>36</sup>, was a gift from Dr. Gavin Rumbaugh. The Scn2a<sup>+R102Q</sup> mouse model of the human c.305 G > A (p.R102Q) mutation<sup>68,69</sup> was created by the Duke Transgenic and Knockout Mouse Shared Resource. The Scn2a<sup>+R102Q</sup> mice were generated using a heterozygous breeding scheme and genotyping was performed by sequencing the amplicon using the following primer set: Scn2a-s, acagacatggcggaaacatgag; and Scn2a-as, agcaggagaggaaagaaagaagc. C3H/HeJ males (#000659; Jackson Laboratory, Bar Harbor,

ME) served as social partners in the resident-intruder and social dyadic tests. 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. Mice were group-housed in the Duke University's Division of Laboratory Animal Resources facility, with an ambient temperature of 72 ± 2 °F, humidity of 30–70%, and light cycle of 12 h on/off.

### Single-vector HiUGE TurboID knock-in

HiUGE plasmids were constructed based on the previously described method<sup>31</sup>. 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 promoter, 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<sup>102</sup>, 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: cggcgggtatcagaaaaatcgtgg, Shank2: aacagctgctggacagataagg, Shank3: cgtgcctcagggcagctggatgg, Nckap1: gcatttctcagcaacacataagg, Nbea: ggcatatgatgatcagaacagg, Arhgef9: cattctggcaaaacttcagtagg, Ank3: gaagaagaaatccggaacgtgg, Scn2a: ggacaaggggaaagatatcaggg, Scn8a: tcagggagtcgaagtgtcagtagg, Hnrnpu: tggagtcagcattatcaccaagg, Homer1: tttagctcattccagtagcttgg, and Wasfl: gtctgatgaagtagactggctgg. To protect the PDZ-binding motifs of Syngap1, Cttnb1, and Iqsec2, an intron-targeting strategy<sup>33</sup> was used, where the donor was flanked by intron / exon boundary sequences from an obligatory intron to enable internal TurboID insertion. The following intronic sequences were targeted: Syngap1: acttattgagacgcttcgctgggg, Cttnb1: aacaggctccagatgcatgggg, and Iqsec2: agggccaactccaatagggagg. To insert TurboID while protecting the C-term PDZ-binding motif of Lrrc4c that has only one coding exon, a custom donor was used where the coding sequence of a.a. ETQI was appended to the C-term of the TurboID donor, thus preserving the native motif. The genomic target for Lrrc4c: gaggctcattcggatcaataacgg. Exon 13 of Syngap1 was targeted for Syngap1 truncation and disruption: acggactcggctcagccatgg. To endogenously express soluble TurboID as a survey for background detection, C-term or 3'-UTR sequences were targeted with a stop codon - internal ribosome entry site (IRES) - TurboID-HA donor. The genomic target sequences were: Syngap1: aggaggtctgtgacgctgggtgg, Scn2a: agtttggcatagacctctgagg, Hnrnpu: aaacagtcgactcttctggaagg, Cttnb1: ttataagcttctactactaaagg, Iqsec2: tactggggagcaggatagcttgg, Homer1: tttagctcattccagtagcttgg, and Wasfl: gtctgatgaagtagactggctgg.

### AAV preparation

AAV was prepared following previously described methods<sup>23,31</sup>. Briefly, concentrated AAV virus was produced in HEK293T cells grown in six 15 cm dishes by triple-transfection with 15  $\mu$ g HiUGE vector, 30  $\mu$ g pADdeltaF6, and 15  $\mu$ g serotype plasmid (pUCmini-iCAP-PHP.eB, a gift from Viviana Gradinaru<sup>103</sup>, 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 0.4  $\mu$ g HiUGE vector, 0.8  $\mu$ g pADdeltaF6, and 0.4  $\mu$ g serotype 2/1 plasmids. Three days following transfection, the virus-containing medium was filtered through Costar Spin-X columns (Sigma #8162) as previously described<sup>31</sup>. Briefly, for each preparation, 0.5 mL AAV-containing supernatant was filtered through the Spin-X column and temporarily stored at 4 °C until ready to use.

### AAV injection and biotin injection for HiUGE-iBioID

Neonatal (P0-2) H11-Cas9 mice were anesthetized by hypothermia and injected intracranially with the purified AAV (2  $\mu$ L per hemisphere,

PHP.eB serotype, > 10<sup>10</sup> GC/μL titer). Donor vector backbone AAV (empty gRNA) was used as negative control. Approximately 3 weeks after injection, mice received daily intraperitoneal injections (i.p.) of biotin (50 mg/kg) over 5 consecutive days. Mice were deeply anesthetized with isoflurane and euthanized by decapitation. 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 × g for 30 min at 4 °C. The supernatant lysate was desalted with Zebra columns 7K MWCO (ThermoFisher #89894 or #89892). The flow-through was combined with 150 μL magnetic Streptavidin 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, 1 M KCl once, 0.1 M Na<sub>2</sub>CO<sub>3</sub> once, 2 M urea in 10 mM Tris-HCl once, and RIPA buffer 2 times. Biotinylated proteins were eluted by boiling the beads in 90 μL 2× sample buffer, supplemented with 2.5 mM biotin, and used for downstream LC-MS/MS and Western blot analyzes.

### 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 dithiothreitol for 30 min at 80 °C, alkylated with 20 mM iodoacetamide for 30 min at room temperature (RT), then supplemented with a final concentration of 1.2% phosphoric acid and 723 μL of S-Trap (Protifi) binding buffer (90% MeOH/100 mM 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% acetonitrile (ACN)/0.2% FA. All samples were 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/ACN), 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% ACN with 0.1% formic acid at a flow rate of 400 nanoliters/minute (nL/min) with a column temperature of 55 °C. 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 msec. The total cycle time for each CV was 0.66 s, with total cycle times of 2 sec between like full MS scans. A 20 s dynamic exclusion was employed to increase depth of coverage. The total analysis cycle time for each injection was approximately 2 h.

Following UPLC-MS/MS analyzes, 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 Infernyx 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; 2 ppm precursor and 0.8 Da production 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 filtration, 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 log<sub>2</sub>-transformed and analyzed using the PolySTest online tool<sup>104</sup>. Proteomic detection was defined as proteins identified by at least 2 peptides in LC-MS/MS. Protein abundances were considered significantly enriched if they met FDR < 0.05 (PolySTest), and fold change ≥ 2 (Log<sub>2</sub>-FC ≥ 1) compared to controls. The enriched gene lists were filtered against known experimental and omnipresent biological contaminants, and soluble TurboID background detections. Note that this high-stringency filter removed a few well-known interactors such as Dlg4, Shank2, and Shank3 from the dataset of several baits.

### Synaptosomal preparation and proteomic analysis of Syngap1-Het mice

Synaptosomal preparation was performed with four adult Syngap1-Het mice and four WT controls. Briefly, mice were deeply anesthetized with isoflurane and euthanized by decapitation. The rapidly isolated brain tissue was sliced to 1 mm 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 1000 × g to remove cell debris and nuclei. The supernatant was further centrifuged at 12,000 × g 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 × g 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<sup>+R102Q</sup> mice

LOPIT-DC fractionation was performed with three adult Scn2a<sup>+R102Q</sup> mice and three WT controls, following a previously described method<sup>30,71,105</sup>. Briefly, mice were deeply anesthetized with isoflurane and euthanized by decapitation. The 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, pH 7.4), supplemented

with cOmplete protease inhibitor cocktail (Sigma #11873580001). The tissue was homogenized for 15 passes in a 2 mL 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  $\mu$ m 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  $\times$  g. Fraction-5, determined to be enriched with Scn2a in a pilot experiment<sup>71</sup>, was used for tandem mass tag (TMT)-multiplexed proteomic analysis.

### TMT-multiplexed quantitative LC-MS/MS analysis

Samples were supplemented with 100  $\mu$ L of 8 M urea and probe sonicated. Protein concentrations were determined via Bradford Assay and ranged from 1.8 to 2.3 mg/mL. Samples were normalized to 120  $\mu$ g 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  $\mu$ L of 20% SDS, reduced with 10 mM dithiothreitol for 45 min at 32 °C, alkylated with 20 mM iodoacetamide for 45 min at RT, then supplemented with a final concentration of 1.2% phosphoric acid and 70  $\mu$ L of S-Trap (Protifi) binding buffer (90% MeOH/100 mM TEAB). Proteins were trapped on the S-Trap micro cartridge, digested using 100 ng/ $\mu$ L 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  $\mu$ L 200 mM triethylammonium bicarbonate, pH 8.0 (TEAB). 40  $\mu$ L 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  $\mu$ L 100% ACN and was added to 75  $\mu$ g of each sample. Samples were incubated for 1 h at RT. After 1-hour reaction, 8  $\mu$ L of 5% hydroxylamine was added and incubated for 15 min at RT to quench the reaction. Samples were combined then lyophilized to dryness.

For offline fractionation, samples were resuspended in 300  $\mu$ L 0.1% formic acid. 400  $\mu$ g was fractionated into 48 unique high pH reversed-phase fractions using pH 9.0 20 mM ammonium formate as mobile phase A and neat ACN as mobile phase B. The column used was a 2.1 mm  $\times$  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. Forty-eight 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  $\mu$ L 1%TFA/2% ACN prior to LC-MS analysis.

Quantitative LC-MS/MS was performed on 2  $\mu$ L (1  $\mu$ g) 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  $\times$  180  $\mu$ m trapping column (5  $\mu$ L/min at 99.9/0.1 v/v water/ACN), after which the analytical separation was performed using a 1.8  $\mu$ m Acquity HSS T3 C18 75  $\mu$ m  $\times$  250 mm column (Waters Corp.) with a 90-min linear gradient of 5 to 30% ACN with 0.1% formic acid at a flow rate of 400 nanoliters/minute (nL/min) with a column temperature of 55 °C. 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 to 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 msec. The total cycle time for each CV was 1 s, with total cycle times of 3 s between like full MS scans. A 45 s dynamic exclusion was employed to increase depth of coverage. The total analysis cycle time for each sample injection was approximately 2 h.

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 (deamidation), 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 percent 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)  $\geq 1.2$  (abs(Log2-FC)  $\geq 0.263$ ). Those with  $p$ -value < 0.1 and abs(fold change)  $\geq 1.2$  (abs(Log2-FC)  $\geq 0.263$ ) were considered indicative candidates.

### Protein network visualization

Protein networks were constructed using Cytoscape<sup>106</sup> 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>)<sup>34,35</sup> 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 and gene ontology (GO) analysis were used to detect protein communities within each proximity proteome, with additional adjustments made based on known protein functions.

### Gene set enrichment analyzes

Gene ontology (GO) enrichment of proximity proteomes was searched against a custom statistical domain of all identified brain proteins (9686 unique proteins, Supplementary Data S5) from cumulative mouse brain proteomic studies in our lab ( $n=107$ ), using ShinyGO<sup>107</sup>. 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. SynGO function analyzes<sup>108</sup> were also conducted with the synaptic bait proteomes versus the nucleus Hnnpu proteome as control. 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](http://www.nemates.org/MA/progs/overlap_stats.html)). In addition, we mapped autism risk genes identified at FDR < 0.1 (102 genes) in Satterstrom et al.<sup>3</sup> and FDR < 0.05 (185 genes) in Fu et al.<sup>7</sup> 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](https://github.com:lauragails/gene_baiting.git)). Bonferroni adjustments were performed and thresholds for statistical significance were delineated on the graphs. We also performed an enrichment analysis between our bait proteomes and the Grove et al. 2019 autism genome-wide association meta-analysis study<sup>44</sup> with MAGMA, as described<sup>109</sup>. Briefly, MAGMA consists of three steps, which were all run using default settings. SNPs were first mapped to genes in the annotation step, then a *p*-value was computed for each gene. Finally, competitive gene-set analysis tests were performed. Mouse genes in each proximity network were mapped to human entrez gene IDs when possible, using HGNC gene nomenclature (<https://www.genenames.org/>, Downloaded 06/06/2023).

For cell-type specific representations, the list of genes for each bait were intersected with mouse orthologs of cell type specific differentially expressed genes (DEGs) from a single-cell genomics study of post-mortem cortical brain tissues from autistic individuals<sup>8</sup>. We used genes expressed in each cell type as the background to perform the hypergeometric test for overlap between each bait gene list and a list of autism DEGs in each cell type.

To compare HiUGE-iBioID datasets against immunoprecipitation results from brain lysate, or with recently published autism proteomic interaction datasets<sup>27,28</sup>, GO analyzes were conducted using input gene lists unique to each set with ShinyGO. Bait self-identifications were excluded from analyzes across studies. Gene names were converted to mouse orthologs if applicable. For analyzes that lack a consensus statistical domain, mouse genome was used as a non-biased background.

### Western blot

For Western blot analyzes, 10 $\mu$ L 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 #11867423001, 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 hr at RT, and the immunosignal was detected using Odyssey FC imager (LI-COR).

### Immunocytochemistry and immunohistochemistry

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 hr at RT. Antibodies and dilutions consisted of: mouse anti-V5-epitope (ThermoFisher #R960-25, 1:500), rabbit anti-V5-epitope (Cell Signaling #13202 S, 1:1000), guinea pig anti-Homer1 (Synaptic Systems #160004, 1:2000), rabbit anti-Syngap1 (ThermoFisher #PA5-58362, 1:1000), rabbit anti-Shank2 (Cell signaling #12218 S, 1:1000), mouse anti-Ank3 (ThermoFisher #33-8800, 1:1000), mouse anti-Scn2a (Antibodiesinc #75-024, 1:1000), goat anti-mouse Alexa Fluor Plus 594

(ThermoFisher #A32742, 1:1000), goat anti-guinea pig Alexa Fluor 488 (ThermoFisher #A11073, 1:1000), and goat anti-rabbit Alexa Fluor 488 (ThermoFisher #A11008, 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 or Zeiss LSM 710 confocal microscope. For immunohistochemistry, brains were collected after intracardiac perfusion with ice cold saline and 4% PFA, cryoprotected in 30% sucrose solution, and sectioned at 40  $\mu$ m. Sections were blocked with blocking buffer and incubated with Alexa Fluor 594 or 568-conjugated streptavidin (ThermoFisher #S32356 or #S11226, 1:1000) overnight at 4 °C. After washes, sections were counterstained with DAPI and cover-slipped for imaging on Zeiss Axio Imager M2 with Apotome module or Zeiss LSM 710 confocal microscope. Images were processed using Fiji<sup>110</sup>.

### Immunoprecipitation

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 consisted of the following: N-Trunc:  $\Delta$  a.a. 2-361; C1-Trunc:  $\Delta$  a.a.730-1343; C2-Trunc:  $\Delta$  a.a. 848-1343; C3-Trunc:  $\Delta$  a.a. 1181-1343. Mutation of Syngap1 (SYNGAP1-c.2214\_2217del) was introduced by site-directed mutagenesis (Quik-Change 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 #13202 S, 1:1000), or GFP (Cell Signaling #2956 S, 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.

For immunoprecipitation from brain lysate, mice were deeply anesthetized with isoflurane and euthanized by decapitation. The forebrain tissues were homogenized in T-PER protein extraction reagent (ThermoFisher #78510), cleared by centrifugation following the manufacturer's instruction, and equally divided into replicates. The following antibodies were added to the lysate at a final concentration of 10  $\mu$ g/mL: mouse anti-Anks1b (Santa Cruz #sc-376610), mouse anti-Scn2a (Antibodiesinc #75-024), or mouse control IgG (Vector labs #I-2000-1) and incubated overnight on a rotator. The next day, 25 $\mu$ L Pierce Protein A/G magnetic beads (ThermoFisher #88802) were added to each immunoprecipitation replicates and incubated for 1 hr under RT. Beads were washed in wash buffer containing 150 mM NaCl and boiled in 2X sample buffer to elute the precipitated complexes for label-free LC-MS/MS analysis. Peptide signals were subjected to trimmed-mean normalization and summed on the protein level. Single-peptide detections and contaminants such as mouse IgG were excluded from the analysis. Enrichment was defined as Log<sub>2</sub>-FC  $\geq$  1 and *p*-value < 0.05 using the Duke Proteomics and Metabolomics Shared Resource (DPMSR) Proteome Discoverer Data Visualization Tool.

### AAV / Cas9-mediated expression disruption

To disrupt Syngap1 and Anks1b expression, the following genomic sequences were targeted by gRNAs using AAV: Syngap1: acggactc ggtctcagccatgg; Anks1b: attgtcccactgtttggacaggg. AAVs (PHP.eB serotype) were applied to H1I-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<sup>111</sup> (No\_current\_500 (non-targeting scramble): ttttagaccttaattcgccg; Scn2a\_gRNA1: cagcggattcacttggtggcc; Scn2a\_gRNA2: gttgaatgttgctttgcca; and Scn2a\_gRNA3: aattacagcgattcacttg). 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 Type Culture Collection (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% CO<sub>2</sub>. Lentivirus was produced as previously described<sup>112</sup>. Briefly, 16 h before transfection, 7 × 10<sup>6</sup> cells were plated in 12 ml of transfection media (Opti-MEM I Reduced Serum Medium (Gibco #31985070), 1x GlutaMAX Supplement, 5% FBS, 1 mM sodium pyruvate (Gibco #11360070), 1x MEM NEAA (Gibco #11140050) in a 10 cm plate. On the day of transfection, 6 mL of transfection media were removed, and the cells were transfected with Lipofectamine 3000 (Invitrogen #L3000008) and 5.4 μg pMD2.G (Addgene #12259), 9.9 μg psPAX2 (Addgene #12260), and 12 μg of the expression vector, according to the manufacturer's instructions. The medium was exchanged with transfection media 6 h after transfection, and the viral supernatant was harvested 24 and 48 h after this medium change. The viral supernatant was pooled and passed through a PVDF 0.45 μm filter (Millipore #SLHV033RB), and concentrated to 50× in 1×PBS using Lenti-X Concentrator (Clontech #631232) in accordance with the manufacturer's protocol. The lentivirus was titered by qPCR as previously described<sup>113</sup>. Briefly, DNA was extracted from cells infected with the lentivirus following a 3-day incubation period. The multiplicity of infection (MOI) was determined by qPCR. 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 DIV 11, 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 promoter. 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 DIV 11. 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).

P0-1 mice were euthanized by decapitation. Forebrain tissue was rapidly isolated, 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<sub>2</sub> 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 interspike 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 (SAS).

### Behavioral tests

**Elevated zero maze.** Anxiety-like behaviors were assessed in the elevated zero maze as described under 40–60 lux illumination<sup>114</sup>. Mice were housed overnight in the test room and tested individually the next day. Animals were placed into a 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 (Leesburg, VA) 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<sup>114</sup>. Individual mice were placed into a 4; 2 × 42 × 30 cm open field (Omnitech Electronics, Columbus, OH) and given 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, Reston, VA) 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<sup>115</sup>. Mice were filmed for 10 min for self-induced grooming. Subsequently, the videos were converted for analysis using Noldus MediaRecorder2. Grooming behavior was scored using TopScan software (CleverSys, Reston, VA) for the duration of self-grooming.

**Ultrasonic vocalizations (USVs).** Male WT and Scn2a<sup>+/-R102Q</sup> mice were housed individually 1 week before testing. Initially, males were primed twice with soiled bedding from estrous C57BL/6J females for 2 days. Subsequently, males were placed on clean bedding for 1 day, then on soiled bedding for another 2 days, and finally were returned to clean bedding overnight. Males were tested the next day. They were acclimated to the recording chambers for 2 min and then 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 (Glienicke/Nordbahn, Germany) for the numbers of calls, call duration, and USV frequencies.

**Social behavior tests.** Social behavior in male WT and Scn2a<sup>+/-R102Q</sup> mice was assessed in the resident-intruder and social dyadic

assays<sup>116</sup>. The Scn2a males were housed individually on the same bedding for 14 days prior to testing; the partner C3H/HeJ mice (Jackson Laboratory, Bar Harbor, ME) were group-housed at 3–4 mice/cage. Both social tests were conducted with C3H/HeJ males matched for body weight with the Scn2a males. Note, the C3H/HeJ mice provide high amounts of social investigation without initiating unprovoked agonistic or attack behaviors<sup>117</sup>. Social testing began within two hr after the lights had extinguished and was conducted under red light. Prior to testing, the chow, water bottle, and metal frame holding the water bottle were removed. The filter was removed from the top of the cage so behaviors could be filmed and the mouse could not escape. Tests were terminated if attacks by either Scn2a or C3H/HeJ males occurred for more than 1 min or if a mouse was injured. All behaviors were tested in the Social Stereo-Scan apparatus (Cleversys) where behaviors can be filmed simultaneously from the top and sides of the test chamber. All videos were scored using Noldus Observer (version 15) by trained personnel blinded to the genotypes of the mice.

Mice were tested first in the resident-intruder assay. Here, a C3H/HeJ male was introduced into the home-cage of the WT or Scn2a<sup>+R102Q</sup> male. Animals were allowed to interact for 5 min, after which the C3H/HeJ male was removed to its home-cage. Following this testing, the cages housing the individual Scn2a mice were replaced with new cages and bedding. Seven days later dyadic testing began where a WT or Scn2a<sup>+R102Q</sup> male was paired with an unfamiliar C3H/HeJ male in a novel clear Plexiglas chamber (32 × 20 × 20 cm). Mice were placed at opposite ends of the chamber divided in half with a white polyfoam partition. Following a 5 min habituation, the partition was removed and the mice were permitted to interact for 5 min. All behaviors in both social tests were calculated as a rate per 5 min [(numbers of incidences/total test time in sec) × 300 s] and were analyzed over four response domains<sup>116</sup>. Mild-social investigation consisted of approaching, sniffing or nosing (ano-genital region), climbing onto the side or back, and/or grooming the face of the partner. Reactivity referred to boxing, holding, kicking, startling, and/or jumping in the presence of the partner. Agonistic behaviors denoted tail rattling to the partner and/or feinting, chasing, lunging, biting, climb-grooming, and/or attacking the partner. Withdrawal behaviors (no acknowledgement of partner) included walking-away, turning-away (without leaving proximity of the partner), or digging in the bedding when the partner contacts or attempts to interact.

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

### Reporting summary

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

### Data 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). Key plasmids from this study have also been deposited to Addgene. The proteomic and MEA data generated in this study are provided in the Supplementary Information / Source Data file. The proteomic data have also been deposited in the MassIVE database under accession code MSV000095141. Source data are provided with this paper.

### Code availability

Requests for custom computer code used in this study should be directed to and will be fulfilled by the Corresponding Author, Dr. Scott Soderling (scott.soderling@duke.edu). Code used in this study is also available in a public Github repository ([git@github.com:lauragails/gene\\_baiting.git](https://github.com:lauragails/gene_baiting.git)).

### References

- Lord, C. et al. Autism spectrum disorder. *Nat. Rev. Dis. Prim.* **6**, 5 (2020).
- Banerjee-Basu, S. & Packer, A. SFARI Gene: an evolving database for the autism research community. *Dis. Model Mech.* **3**, 133–135 (2010).
- 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 (2020).
- Wang, T. et al. Large-scale targeted sequencing identifies risk genes for neurodevelopmental disorders. *Nat. Commun.* **11**, 4932 (2020).
- Zoghbi, H. Y. Postnatal neurodevelopmental disorders: meeting at the synapse? *Science* **302**, 826–830 (2003).
- Ayhan, F. & Konopka, G. Regulatory genes and pathways disrupted in autism spectrum disorders. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **89**, 57–64 (2019).
- Fu, J. M. et al. Rare coding variation provides insight into the genetic architecture and phenotypic context of autism. *Nat Genet.* **54**, 1320–1331 (2022).
- Velmeshev, D. et al. Single-cell genomics identifies cell type-specific molecular changes in autism. *Science* **364**, 685–689 (2019).
- 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 (2013).
- Voineagu, I. et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* **474**, 380–384 (2011).
- 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 (2019).
- O’Roak, B. J. et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**, 246–250 (2012).
- Li, T. et al. A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat. Methods* **14**, 61–64 (2017).
- Neale, B. M. et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012).
- Sakai, Y. et al. Protein interactome reveals converging molecular pathways among autism disorders. *Sci. Transl. Med.* **3**, 86ra49 (2011).
- 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 (2019).
- Terenzio, M., Schiavo, G. & Fainzilber, M. Compartmentalized signaling in neurons: from cell biology to neuroscience. *Neuron* **96**, 667–679 (2017).
- De Rubeis, S. et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**, 209–215 (2014).
- Santini, E. & Klann, E. Reciprocal signaling between translational control pathways and synaptic proteins in autism spectrum disorders. *Sci. Signal* **7**, re10 (2014).
- Grant, S. G. Synaptopathies: diseases of the synaptome. *Curr. Opin. Neurobiol.* **22**, 522–529 (2012).
- Fujitani, M., Otani, Y. & Miyajima, H. Pathophysiological roles of abnormal axon initial segments in neurodevelopmental disorders. *Cells* **10**, 2110 (2021).

22. 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 (2020).
23. Uezu, A. et al. Identification of an elaborate complex mediating postsynaptic inhibition. *Science* **353**, 1123–1129 (2016).
24. Hamdan, H. et al. Mapping axon initial segment structure and function by multiplexed proximity biotinylation. *Nat. Commun.* **11**, 100 (2020).
25. Bayes, A. et al. Characterization of the proteome, diseases and evolution of the human postsynaptic density. *Nat. Neurosci.* **14**, 19–21 (2011).
26. Distler, U. et al. In-depth protein profiling of the postsynaptic density from mouse hippocampus using data-independent acquisition proteomics. *Proteomics* **14**, 2607–2613 (2014).
27. Murtaza, N. et al. Neuron-specific protein network mapping of autism risk genes identifies shared biological mechanisms and disease-relevant pathologies. *Cell Rep.* **41**, 111678 (2022).
28. Pintacuda, G. et al. Protein interaction studies in human induced neurons indicate convergent biology underlying autism spectrum disorders. *Cell Genom.* **3**, 100250 (2023).
29. Piersimoni, L., Kastritis, P. L., Arlt, C. & Sinz, A. Cross-linking mass spectrometry for investigating protein conformations and protein-protein interactions - a method for all seasons. *Chem. Rev.* **122**, 7500–7531 (2022).
30. Geladaki, A. et al. Combining LOPIT with differential ultracentrifugation for high-resolution spatial proteomics. *Nat. Commun.* **10**, 331 (2019).
31. Gao, Y. et al. Plug-and-play protein modification using homology-independent universal genome engineering. *Neuron* **103**, 583–597 e588 (2019).
32. Branon, T. C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
33. Zhong, H. et al. High-fidelity, efficient, and reversible labeling of endogenous proteins using CRISPR-based designer exon insertion. *Elife* **10**, e64911 (2021).
34. 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 (2021).
35. 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 (2000).
36. Beyreli, I., Karakahya, O. & Cicek, A. E. DeepND: deep multitask learning of gene risk for comorbid neurodevelopmental disorders. *Patterns* **3**, 100524 (2022).
37. Seiffert, S. et al. Modulating effects of FGF12 variants on Na(V)1.2 and Na(V)1.6 being associated with developmental and epileptic encephalopathy and Autism spectrum disorder: A case series. *EBioMedicine* **83**, 104234 (2022).
38. Zhou, X. et al. Integrating de novo and inherited variants in 42,607 autism cases identifies mutations in new moderate-risk genes. *Nat Genet.* **54**, 1305–1319 (2022).
39. O’Neil, S. D. et al. Action potential-coupled Rho GTPase signaling drives presynaptic plasticity. *Elife* **10**, e63756 (2021).
40. 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*. <https://doi.org/10.1101/2021.11.19.469212> (2022).
41. S. C. E. a. & Consortium, S.SPARK: a US cohort of 50,000 families to accelerate autism research. *Neuron* **97**, 488–493 (2018).
42. 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 (2012).
43. Viswanathan, S. et al. High-performance probes for light and electron microscopy. *Nat. Methods* **12**, 568–576 (2015).
44. Grove, J. et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet* **51**, 431–444 (2019).
45. 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 (2002).
46. Li, J. et al. Long-term potentiation modulates synaptic phosphorylation networks and reshapes the structure of the postsynaptic interactome. *Sci. Signal* **9**, rs8 (2016).
47. Li, J. et al. Spatiotemporal profile of postsynaptic interactomes integrates components of complex brain disorders. *Nat. Neurosci.* **20**, 1150–1161 (2017).
48. Wilkinson, B., Li, J. & Coba, M. P. Synaptic GAP and GEF complexes cluster proteins essential for GTP signaling. *Sci. Rep.* **7**, 5272 (2017).
49. Tindi, J. O. et al. ANKS1B gene product AIDA-1 controls hippocampal synaptic transmission by regulating GluN2B subunit localization. *J. Neurosci.* **35**, 8986–8996 (2015).
50. Gamache, T. R., Araki, Y. & Huganir, R. L. Twenty years of SynGAP research: from synapses to cognition. *J. Neurosci.* **40**, 1596–1605 (2020).
51. 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 (2016).
52. 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 (2011).
53. 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 (2015).
54. Carbonell, A. U. et al. Haploinsufficiency in the ANKS1B gene encoding AIDA-1 leads to a neurodevelopmental syndrome. *Nat. Commun.* **10**, 3529 (2019).
55. 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 (2013).
56. Kim, J. H., Lee, H.-K., Takamiya, K. & Huganir, R. L. The role of synaptic GTPase-activating protein in neuronal development and synaptic plasticity. *J. Neurosci.* **23**, 1119–1124 (2003).
57. Nakajima, R. et al. Comprehensive behavioral analysis of heterozygous Syngap1 knockout mice. *Neuropsychopharmacol. Rep.* **39**, 223–237 (2019).
58. Komiya, N. H. et al. SynGAP regulates ERK/MAPK signaling, synaptic plasticity, and learning in the complex with postsynaptic density 95 and NMDA receptor. *J. Neurosci.* **22**, 9721–9732 (2002).
59. 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 (1980).
60. Schmeisser, M. J. et al. Autistic-like behaviours and hyperactivity in mice lacking ProSAP1/Shank2. *Nature* **486**, 256–260 (2012).
61. Mignot, C. et al. Genetic and neurodevelopmental spectrum of SYNGAP1-associated intellectual disability and epilepsy. *J. Med Genet* **53**, 511–522 (2016).
62. Kilinc, M. et al. Endogenous Syngap1 alpha splice forms promote cognitive function and seizure protection. *Elife* **11**, e75707 (2022).
63. Llamas, N. et al. SYNGAP1 controls the maturation of dendrites, synaptic function, and network activity in developing human neurons. *J. Neurosci.* **40**, 7980–7994 (2020).
64. Clement, J. P. et al. Pathogenic SYNGAP1 mutations impair cognitive development by disrupting maturation of dendritic spine synapses. *Cell* **151**, 709–723 (2012).

65. 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 (2019).
66. 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 (2017).
67. Sanders, S. J. et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–241 (2012).
68. Wang, T. et al. De novo genic mutations among a Chinese autism spectrum disorder cohort. *Nat. Commun.* **7**, 13316 (2016).
69. Callaghan, D. B. et al. Whole genome sequencing and variant discovery in the ASPIRE autism spectrum disorder cohort. *Clin. Genet* **96**, 199–206 (2019).
70. Berecki, G. et al. Functional correlates of clinical phenotype and severity in recurrent *SCN2A* variants. *Commun. Biol.* **5**, 515 (2022).
71. Courtland, J. L. et al. Genetic disruption of *WASHC4* drives endolysosomal dysfunction and cognitive-movement impairments in mice and humans. *Elife* **10**, e61590 (2021).
72. Brackenbury, W. J. & Isom, L. L. Na Channel beta subunits: over-achievers of the ion channel family. *Front. Pharm.* **2**, 53 (2011).
73. Turner, C. A., Watson, S. J. & Akil, H. The fibroblast growth factor family: neuromodulation of affective behavior. *Neuron* **76**, 160–174 (2012).
74. Tamura, S. et al. CRISPR activation rescues abnormalities in *SCN2A* haploinsufficiency-associated autism spectrum disorder. *bioRxiv*. <https://doi.org/10.1101/2022.03.30.486483> (2022).
75. Wildburger, N. C. et al. Quantitative proteomics reveals protein-protein interactions with fibroblast growth factor 12 as a component of the voltage-gated sodium channel 1.2 (*nav1.2*) macromolecular complex in Mammalian brain. *Mol. Cell Proteom.* **14**, 1288–1300 (2015).
76. Repetto, D. et al. Molecular dissection of neurobeachin function at excitatory synapses. *Front. Synaptic Neurosci.* **10**, 28 (2018).
77. 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 (2018).
78. Wang, X. et al. Neurobeachin: a protein kinase A-Anchoring, beige/Chediak-Higashi protein homolog implicated in neuronal membrane traffic. *J. Neurosci.* **20**, 8551–8565 (2000).
79. Nair, R. et al. Neurobeachin regulates neurotransmitter receptor trafficking to synapses. *J. Cell Biol.* **200**, 61–80 (2013).
80. Mulhern, M. S. et al. NBEA: developmental disease gene with early generalized epilepsy phenotypes. *Ann. Neurol.* **84**, 788–795 (2018).
81. Castermans, D. et al. The neurobeachin gene is disrupted by a translocation in a patient with idiopathic autism. *J. Med. Genet.* **40**, 352–356 (2003).
82. Farzana, F. et al. Neurobeachin regulates glutamate- and GABA-receptor targeting to synapses via distinct pathways. *Mol. Neurobiol.* **53**, 2112–2123 (2016).
83. 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 (2003).
84. Bachor, R., Waliczek, M., Stefanowicz, P. & Szewczuk, Z. Trends in the design of new isobaric labeling reagents for quantitative proteomics. *Molecules* **24**, 701 (2019).
85. Takano, T. et al. Chemo-genetic discovery of astrocytic control of inhibition in vivo. *Nature* **588**, 296–302 (2020).
86. Stark, C. et al. BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34**, D535–D539 (2006).
87. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
88. Lin, Z. et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023).
89. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
90. Slavov, N. Unpicking the proteome in single cells. *Science* **367**, 512–513 (2020).
91. Liang, Y. et al. Fully automated sample processing and analysis workflow for low-input proteome profiling. *Anal. Chem.* **93**, 1658–1666 (2021).
92. Rosenberger, F. A., Thielert, M. & Mann, M. Making single-cell proteomics biologically relevant. *Nat. Methods* **20**, 320–323 (2023).
93. 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 (2020).
94. Brueggeman, L., Koomar, T. & Michaelson, J. J. Forecasting risk gene discovery in autism with machine learning and genome-scale data. *Sci. Rep.* **10**, 4569 (2020).
95. Schoof, E. M. et al. Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat. Commun.* **12**, 3341 (2021).
96. Brunner, A. D. et al. Ultra-high sensitivity mass spectrometry quantifies single-cell proteome changes upon perturbation. *Mol. Syst. Biol.* **18**, e10798 (2022).
97. Vistain, L. F. & Tay, S. Single-cell proteomics. *Trends Biochem. Sci.* **46**, 661–672 (2021).
98. Walkup, W. G. T. et al. Phosphorylation of synaptic GTPase-activating protein (*synGAP*) by  $Ca^{2+}$ /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 (2015).
99. Walkup, W. G. et al. A model for regulation by *SynGAP*-alpha1 of binding of synaptic proteins to PDZ-domain ‘Slots’ in the post-synaptic density. *Elife* **5**, e16813 (2016).
100. Goldfarb, M. et al. Fibroblast growth factor homologous factors control neuronal excitability through modulation of voltage-gated sodium channels. *Neuron* **55**, 449–463 (2007).
101. 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 (2017).
102. Concordet, J. P. & Haeussler, M. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res.* **46**, W242–W245 (2018).
103. Chan, K. Y. et al. Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat. Neurosci.* **20**, 1172–1179 (2017).
104. Schwammler, 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 Proteom.* **19**, 1396–1408 (2020).
105. 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 116. <https://doi.org/10.1002/0471142301.ns0116s42> (2008).
106. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
107. Ge, S. X., Jung, D. & Yao, R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics* **36**, 2628–2629 (2020).
108. Koopmans, F. et al. SynGO: an evidence-based, expert-curated knowledge base for the synapse. *Neuron* **103**, 217–234 e214 (2019).



109. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol.* **11**, e1004219 (2015).
  110. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
  111. Sanson, K. R. et al. Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nat. Commun.* **9**, 5416 (2018).
  112. Schmidt, R. et al. CRISPR activation and interference screens decode stimulation responses in primary human T cells. *Science* **375**, eabj4008 (2022).
  113. Gordon, M. G. et al. lentiMPRA and MPRAflow for high-throughput functional characterization of gene regulatory elements. *Nat. Protoc.* **15**, 2387–2412 (2020).
  114. Wang, X. et al. Synaptic dysfunction and abnormal behaviors in mice lacking major isoforms of Shank3. *Hum. Mol. Genet* **20**, 3093–3108 (2011).
  115. Pappas, A. L. et al. Deficiency of Shank2 causes mania-like behavior that responds to mood stabilizers. *JCI Insight* **2**, e92052 (2017).
  116. Rodriguiz, R. M., Chu, R., Caron, M. G. & Wetsel, W. C. Aberrant responses in social interaction of dopamine transporter knockout mice. *Behav. Brain Res* **148**, 185–198 (2004).
  117. Nehrenberg, D. L. et al. An anxiety-like phenotype in mice selectively bred for aggression. *Behav. Brain Res* **201**, 179–191 (2009).
  118. Hulsen, T., de Vlieg, J. & Alkema, W. BioVenn - a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488 (2008).
- performed gene-set analyzes. C.A.G., J.D.B., and S.H.S. supervised the collaborative study. Y.G. and S.H.S. wrote the original manuscript. All authors contributed to manuscript editing and revision.

## Competing interests

S.H.S. and Y.G. have a patent application related to the HiUGE technology (16/968,904). The intellectual property was licensed to CasTag Biosciences. S.H.S. is a founder of CasTag Biosciences; Duke University as an institution holds equity in CasTag Biosciences. C.A.G. 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. G.D. 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. G.D. 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. J.D.B. is a consultant for Bridgebio. W.C.W. is a consultant for Onsero Therapeutics. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-51037-x>.

**Correspondence** and requests for materials should be addressed to Scott H. Soderling.

**Peer review information** *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024

## Acknowledgements

Research reported in this publication was supported by a donation to the Duke Center for Autism and Brain Development, NIH R01MH11684 (SHS), UM1HG012053 (CAG), R01MH125236 (CAG), RM1HG011123 (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. We would like to thank Drs. Tyler Bradshaw and Jamie Courtland for performing synaptosome and fractionation proteomics, Dr. Ramona M. Rodriguiz, Mr. Christopher R. Means, Mr. Nathan Franklin, and Ms. Sarah Page Steffens for conducting the behavioral experiments, and Dr. Rodriguiz for helping with the behavioral statistical analyzes. 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. Figure 1A, Fig. 2A, E, H, Fig. 3A–D, Fig. S3A, Fig. S6A, created with BioRender.com, released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license. Proportional Venn diagrams were made using BioVenn<sup>118</sup>.

## Author contributions

Y.G. and S.H.S. conceived the study. Y.G., M.T., D.S. and E.J.S. performed HiUGE-iBioID experiments and analyzes. SAGM designed CRISPRa constructs and J.Z. performed validation experiments. Y.G. and M.T. performed MEA experiments. W.C.W. designed and managed the behavioral experiments and statistical analyzes. G.D. and Y.J. contributed to the generation of Scn2a<sup>+/R102Q</sup> model. D.V. and L.G.S.

<sup>1</sup>Department of Cell Biology, Duke University School of Medicine, Durham, NC, USA. <sup>2</sup>Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA. <sup>3</sup>Department of Molecular Genetics and Microbiology, Duke University School of Medicine, Durham, NC, USA. <sup>4</sup>Department of Biology, Duke University, Durham, NC, USA. <sup>5</sup>Proteomics and Metabolomics Shared Resource, Duke University School of Medicine, Durham, NC, USA. <sup>6</sup>Department of Biomedical Engineering, Duke University, Durham, NC, USA. <sup>7</sup>Center for Advanced Genomic Technologies, Duke University, Durham, NC, USA. <sup>8</sup>Department of Surgery, Duke University Medical Center, Durham, NC, USA. <sup>9</sup>Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine,

Durham, NC, USA. <sup>10</sup>Department of Neurobiology, Duke University School of Medicine, Durham, NC, USA. <sup>11</sup>Mouse Behavioral and Neuroendocrine Analysis Core Facility, Duke University School of Medicine, Durham, NC, USA. <sup>12</sup>Department of Genetics, Yale University School of Medicine, New Haven, Connecticut, USA. <sup>13</sup>Department of Neuroscience, Yale University School of Medicine, New Haven, Connecticut, USA. <sup>14</sup>Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut, USA. <sup>15</sup>Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount Sinai, New York, NY, USA. <sup>16</sup>Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. <sup>17</sup>The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. <sup>18</sup>Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ✉e-mail: [Scott.Soderling@duke.edu](mailto:Scott.Soderling@duke.edu)