Synergistic effects of common schizophrenia risk variants

Nadine Schrodé1,2,12, Seok-Man Ho3,4,12, Kazuhiko Yamamuro5,6, Amanda Dobbyn1,2, Laura Huckins1,2,7, Marliette R. Matos4, Esther Cheng4, P. J. Michael Deans1,2, Erin Flaherty4,8, Natalie Barretto4, Aaron Topol5, Khaled Algenan9, Sonya Abadali4, James Gregory10, Emily Hoelzli10, Hemali Phatnani10, Vineeta Singh11, Deeptha Girish11, Bruce Aronow11, Robert Mccullumsmith9, Gabriel E. Hoffman1,2,7, Eli A. Stahl1,2,7, Hirofumi Morishita5,6, Pamela Sklar1,2,3,5,6,8,13 and Kristen J. Brennand1,2,3,5,6,7,8*

The mechanisms by which common risk variants of small effect interact to contribute to complex genetic disorders are unclear. Here, we apply a genetic approach, using isogenic human induced pluripotent stem cells, to evaluate the effects of schizophrenia (SZ)-associated common variants predicted to function as SZ expression quantitative trait loci (eQTLs). By integrating CRISPR-mediated gene editing, activation and repression technologies to study one putative SZ eQTL (FURIN rs4702) and four top-ranked SZ eQTL genes (FURIN, SNAP91, TSNARE1 and CLCN3), our platform resolves pre- and postsynaptic neuronal deficits, recapitulates genotype-dependent gene expression differences and identifies convergence downstream of SZ eQTL gene perturbations. Our observations highlight the cell-type-specific effects of common variants and demonstrate a synergistic effect between SZ eQTL genes that converges on synaptic function. We propose that the links between rare and common variants implicated in psychiatric disease risk constitute a potentially generalizable phenomenon occurring more widely in complex genetic disorders.

Genome-wide association studies (GWAS) of SNPs continue to identify loci (143 and growing12) that are significantly associated with risk for SZ. These common variant risk loci are enriched for genes expressed in pyramidal excitatory neurons (and a subset of GABAergic interneurons)3,4, particularly in synaptic pathways2.

By integrating GWAS and postmortem brain eQTL (cis-eQTL) studies, common variants that affect disease risk through regulation of gene expression can be identified. Recent estimates calculate that approximately 45.5% of SZ GWAS loci have brain eQTLs5. While not necessarily true across all complex genetic disorders (for example, inflammatory bowel disease7), extensive evidence suggests that in SZ at least, the eQTL signal drives the GWAS signal12. What is unclear is the precise neuronal cell type(s) where common variants act, and how they impact neuronal function. Because current case-control human induced pluripotent stem cell (hiPSC) designs are underpowered to discover the impact of common variants8, it is necessary to use isogenic strategies to evaluate the impact of these genetic variants, alone and in combination, on gene expression and synaptic function across a variety of human neural cell types.

In this study, we apply a functional validation pipeline for common variants that incorporates leading genomic, hiPSC-based and CRISPR-based approaches. For prioritized variants12, we apply CRISPR editing to achieve allelic conversion when only one putative causal SNP is predicted at a SZ-associated locus, or CRISPR activation/interference (CRISPRa/i) to manipulate endogenous gene expression at loci containing several linked SNPs. We demonstrate that single and combinatorial isogenic comparisons can reveal the molecular and functional effects of common variants associated with SZ risk in a cell-type-specific manner. This work identifies a synergy between SZ eQTL genes, supporting the hypothesis that common and rare SZ-associated variants occur in the same genes and/or converge on the same pathways—an unexpected finding that may apply more broadly across complex genetic disorders.

Results

Prioritization of FURIN rs4702 for CRISPR editing and SNAP91/TSNARE1 for CRISPRa/i. Common variant SNPs and genes were prioritized from the expanding list of SZ-associated loci based on: (1) strong evidence of genetic regulation of expression12; (2) fine-mapping quantification of the number of putative causal SNPs at each locus12; (3) RNA levels in hiPSC-derived neural cells (neural progenitor cells (NPCs), neurons and astrocytes); and (4) capacity for CRISPRa/i-mediated manipulation (Fig. 1 and Supplementary Fig. 1).

Of the 108 SZ GWAS loci identified by the Psychiatric Genomics Consortium (PGC) SZ GWAS1,2, 19 harbored colocalized SZ GWAS and cis-eQTL (SNP–gene pairs within 1 Mb of a gene) signals in

1Department of Genetics and Genomics, Icahn School of Medicine at Mount Sinai, New York, NY, USA. 2Icahn Institute of Genomics and Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA. 3Department of Stem Cell and Regenerative Biology, Icahn School of Medicine at Mount Sinai, New York, NY, USA. 4Graduate School of Biomedical Science, Icahn School of Medicine at Mount Sinai, New York, NY, USA. 5Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA. 6Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA. 7Pamela Sklar Division of Psychiatric Genomics, Icahn School of Medicine at Mount Sinai, New York, NY, USA. 8Department of Neuroscience, Institute in the College of Medicine & Life Sciences, The University of Toledo, Toledo, OH, USA. 9Center for Genomics of Neurodegenerative Disease, New York Genome Center, New York, NY, USA. 10UC Department of Pediatrics Cincinnati Children’s Hospital Medical Center, Cincinnati, OH, USA. 11These authors contributed equally: Nadine Schrodé, Seok-Man Ho. 12Deceased: Pamela Sklar *e-mail: kristen.brennand@mssm.edu
the CommonMind Consortium (CMC) postmortem brain RNA-sequencing (RNA-seq) analysis. Of these, five were predicted to involve only a single protein-coding gene: furin, paired basic amino acid cleaving enzyme (FURIN, downregulated in risk allele individuals); t-SNARE domain containing 1 (TSNARE1, upregulated); contactin 4 (CNTN4, upregulated); chloride voltage-gated channel 3 (CLCN3, upregulated); and synaptosome associated protein 91 (SNAP91, upregulated). For FURIN only, the most significant GWAS SNP (rs4702, NC_000015.10:g.90883330G>A) was also the most significant eQTL SNP (Fig. 1b); fine-mapping analysis identified a single putative cis-eQTL (rs4702, probability = 0.94) (Fig. 1b). Independent evidence suggested that FURIN rs4702 is an active cis-eQTL in NEUROG2 (also known as NGN2) excitatory neurons.

We considered postmortem and hiPSC-based neural expression patterns of these five SZ eQTL genes. Postmortem brain findings were inconsistent: bulk immunopanned adult postmortem expression of FURIN was neuronal but enriched in astrocytes, SNAP91 and CNTN4 were neuron-specific, TSNARE1 was very low and CLCN3 was pan-neural, whereas cortical expression in the Genotype-Tissue Expression (GTEx) Project was robust for all but CNTN4 (Supplementary Fig. 1a), and single-cell RNA-seq from the PsychENCODE Consortium showed comparable expression of CNTN4 in all neural cell types (Supplementary Fig. 1b). Therefore, we further prioritized these five genes based on our RNA-seq analysis of control hiPSC-derived NPCs, 6-week forebrain neurons, 21-d NGN2 excitatory neurons and primary astrocytes from three independent donors, which found reliable expression of four: SNAP91, TSNARE1, CLCN3 and FURIN (Supplementary Fig. 1c).

To facilitate prioritization of other neuropsychiatric GWAS genes in hiPSC-based platforms, we make available a resource by which neural cell-type-specific expression of any gene can be easily cross-referenced in our hiPSC CRISPRa/i datasets as well as hiPSC-NPC and hiPSC-neuron RNA-seq datasets (https://schroden.shinyapps.io/BrennandLab-ExpressionApp-limited/).

CRISPR activation (CRISPRa) and interference (CRISPRi) was applied to upregulate and/or downregulate endogenous gene expression in the NGN2 excitatory neurons from two donors (Supplementary Table 1) (refs. 19,20). Three candidate guide RNAs for each gene were screened in 6-week forebrain neurons and primary astrocytes, which were significant for three SZ GW AS genes: SNAP91, CNTN4, and FURIN (Supplementary Fig. 1c). CRISPRa/i resulted in smaller perturbations in neurons generally decreasing over neuronal maturation (Supplementary Fig. 1d–f). CRISPRa/i showed decreased efficacy specifically with CRISPRi, consistent with the strong ‘active’ neuron-specific H3K27ac peaks across the entire gene (Supplementary Fig. 1g).

For most SZ GWAS genes, the number and/or distance between eQTLs in linkage disequilibrium was impractical for allelic conversion. By considering SZ GWAS, fine-mapping, brain cis-eQTL,
CRISPR-mediated single-bp edit
Seed on 96-well plate
Isolate DNA

Differentiate
Plate cells at low single-cell density
Find well with highest G allele enrich

RFLP and Sanger sequence clones

CRISPR-mediated single-bp edit
Seed on 96-well plate
Isolate DNA

D7 NGN2 excitatory neurons
D21 NGN2 excitatory neurons

ASCL1/DLX2 GABAergic neurons
hiPSC NPCs
NFIB astrocytes

D7 NGN2 neurons + miR negative Ctrl
D7 NGN2 neurons + miR338 inhibitor

Average neurite length (µm)
400
300
200
100
0.001

Average burst duration (s)
7.5
6.0
4.5
3.0
1.5
0.004

Average firing rate (Hz)
8.0
6.0
4.0
2.0
0.063

rs4702 genotype
AA
GG

rs4702 genotype
AA
GG

rs4702 genotype
AA
GG

rs4702 genotype
AA
GG

Fig. 2 | CRISPR editing demonstrates the cell-type-specific impact of rs4702 on FURIN expression and resulting neural phenotypes. a, Schematic for CRISPR-mediated single-base-pair edit, pooled screening and validation strategy. b, \log_2(FURIN expression) in rs4702 AA and GG D7 (left, \(n = 14\)) and D21 (right, \(n = 13\)) NGN2 excitatory neurons. qPCR was performed on four edited clones and two unedited clones of the same donor line in four independent experiments. c, \log_2(FURIN expression) in rs4702 AA and GG hiPSC-derived inhibitory ASCL1/DLX2 GABAergic neurons (left, \(n = 6\) each), hiPSC NPCs (center, \(n = 2\) each) and hiPSC-derived NFIB astrocytes (right, \(n = 14\) each) through qPCR. d, \log_2(FURIN expression) in rs4702 AA and GG D7 NGN2 excitatory neurons, following transfection with either a control miRNA (left, \(n = 8\) each) or miR338 inhibitor (right, \(n = 8\) each). P values were calculated using a one-sided t-test, a linear mixed-effects model controlled for variation between qPCR runs; \(n\) refers to biologically independent samples. e, Average neurite length of rs4702 AA and GG D18–D25 NGN2 excitatory neurons, measured (semiautomated) from confocal images of GFP-labeled neurons. f, g, Average burst duration (f) and mean firing rate (g) in rs4702 AA and GG D18–D25 NGN2 excitatory neurons (left) and at various time points from 15 to 35 d in vitro (right). \(n = 44\) biologically independent samples in each experiment. P values calculated using two-sided t-test; line fitted through locally weighted smoothing (locally estimated scatterplot smoothing). Shaded areas (95% confidence interval), box plots (quartiles), whiskers (largest/smallest observation within hinge ±1.5x interquartile range) and overlaid scatter plots (means of replicates from individual clones) are shown.

cell-type-specific hiPSC neural expression and CRISPRa/i efficacies in NGN2 excitatory neurons, we selected FURIN rs4702 for CRISPR editing and SNAP91/TSNARE1/CLCN3 for CRISPRa/i-based functional evaluation.

Impact of CRISPR-mediated allelic conversion of FURIN rs4702. CRISPR-based allelic conversion was used to test the impact of FURIN rs4702 through isogenic comparisons (Fig. 2 and Supplementary Fig. 2). gRNAs were designed to the target region that minimized potential off-target effects and to favor homozygous genomic editing. hiPSCs were nucleofected with a single-stranded DNA repair template and Cas9 vector (catalog no. 62988; Addgene) with gRNA, followed by a 24-h puromycin selection. To isolate low-frequency edits, a pooled TaqMan droplet digital PCR (ddPCR) screen (Fig. 2a and Supplementary Fig. 2a) was performed to enrich for cells with the desired allele; edited cells were identified via restriction fragment length polymorphism screening and validated by Sanger sequencing (Supplementary Fig. 2b). Although our targeting strategy incorporated a single point mutation in the protospacer adjacent motif (PAM) to avoid recursive cutting through...
Fig. 3 | CRISPRa/i perturbations of the synaptic genes SNAP91 and TSNARE1 lead to transcriptomic changes and synaptic phenotypes. a. Heatmap of SNAP91 and TSNARE1 CRISPRa/i clustering in NG2 excitatory neurons based on –log10(P) and the regression coefficient of GSEA. Gene set enrichment tests were performed using a competitive test (accounting for inter-gene correlation) across 698 curated neural gene sets, summarized in 8 categories (denoted by y axis color annotations). b. Word cloud analysis of enriched gene sets and summary categories from a for SNAP91 and TSNARE1 shows the most frequently occurring gene set/category words. Font size denotes frequency, which was corrected by subtraction of the respective total word frequency in all used gene sets. c. Representative confocal microscopic images of immunostained against presynaptic SYP (green) and dendritic MAP2 (blue) at D24. Scale bar, 5 μm (left). Normalized SYP+ puncta counts (top) or size (bottom) of neurons with control (scramble gRNA, CRISPRa/i, n = 104 each), or altered expression of SNAP91 (CRISPRa/i, n = 106 each) from two cell lines (C1, C2). P values were calculated using two-sided t-tests. n refers to independent images from three independent experiments. Box plots (quartiles), whiskers (largest/smallest observation within hinge ±1.5x interquartile range) and overlaid scatter plots (means of replicates from individual clones) are shown. d. Schematic of electrophysiological strategy to evaluate presynaptic and postsynaptic effects of CRISPRa/i clustering. e. sEPSC frequency and amplitude of fluorescently labeled D28–31 NG2 excitatory neurons with SNAP91 CRISPRa (left, n = 111 individual neurons) or SNAP91 CRISPRi (right, n = 128 individual neurons). Data collected across three independent experiments. P values were determined by multiple comparison tests using two-way ANOVA. Plots show scatter graphs with the mean and s.e.m. for each group. KRAB, Krüppel associated box; VPR, VP64–p65–Rta. f. Summary schematic of sEPSC frequency and amplitude of fluorescently labeled NG2 excitatory neurons from e. Orange, SNAP91 CRISPRi; brown, control; purple, SNAP91 CRISPRa neurons. g. Heatmap representation of observed synaptic phenotypes for SNAP91 and TSNARE1 CRISPRa/i based on –log10(P) and regression coefficients from c and e.

Cas9, unexpectedly, the final edited lines did not show evidence of PAM edits. We achieved seamless editing of FURIN rs4702 from AA to GG in four clonal lines from one control donor (NSB3182; Supplementary Table 1 and Supplementary Fig. 2b). The top three predicted on- and off-target cutting sites were confirmed for all hiPSCs by Sanger sequencing (Supplementary Fig. 2c).
Allelic conversion from rs4702 AA to GG resulted in decreased FURIN messenger RNA levels in 7-d NGN2 excitatory neurons by quantitative PCR (qPCR; GG, $P = 1.7 	imes 10^{-3}$ in D7 and $P = 0.61$ in D21 neurons) (Fig. 2b). rs4702 is predicted to be an eQTL for several genes beyond FURIN, both by the CMC and GTEx. In our edited NGN2 excitatory neurons, SV2B showed significantly increased expression in rs4702 GG cells, while PRC1 and RCCD1 were repressed with moderate significance and FES was unchanged (Supplementary Fig. 2d).

To evaluate cell-type-specific eQTL effects, rs4702 AA and GG hiPSCs were induced or differentiated into ASCL1/DLX2 GABAergic neurons, NFIB astrocytes, and hiPSC NPCs. Although we observed no significant differences in FURIN expression between AA and GS ASCL1/DLX2 GABAergic neurons ($P = 0.38$) and NFIB astrocytes ($P = 0.56$), unexpectedly, we observed increased FURIN in GG hiPSC NPCs (GG, $P = 1.4 	imes 10^{-2}$). Moreover, consistent with reduced neural migration following repression of FURIN in hiPSC NPCs (Supplementary Fig. 2f), rs4702 GG hiPSC NPCs (which showed increased FURIN expression) showed increased neural migration relative to isogenic AA controls (Supplementary Fig. 2e). These results highlight the complex and cell-type-specific nature of eQTL biology (Fig. 2c).

Cell-type-specific, eQTL-regulated gene expression may differ in a more physiologically relevant environment, whereby heterogeneous cell types engage in both cell- and non-cell-autonomous gene regulation. To test this, we generated both three-dimensional human cortical spheroids comprising predominantly glutamatergic (and some GABAergic) neurons and astrocytes, and human subpallial spheroids comprising largely GABAergic neurons and astrocytes (Supplementary Fig. 2f). FURIN expression in human cortical spheroids was somewhat repressed (trending toward significance, $P = 0.12$) but was unchanged in human subpallial spheroids ($P = 0.74$) (Supplementary Fig. 2g). Although generally consistent with our induced neurons, heterogeneous three-dimensional cultures may less sensitively resolve cell-type-specific eQTL effects.

One mechanism by which cell-type-specific eQTL-regulated gene expression can occur is posttranscriptional regulation. FURIN affects protein expression by several mechanisms. First, we found that FURIN represses the expression of several genes beyond FURIN (Fig. 2d), suggesting that regulatory activity of the rs4702 GG allele is dependent on miR338 expression levels. rs4702 GG and AA neurons grown on unlabeled control (nuclease dead Cas9 (dCas9)-expressing) control substrate showed significantly increased expression of several genes beyond FURIN (rs4702 GG, $P = 1.04 	imes 10^{-2}$) (Fig. 2f, top). All CRISPRa/i perturbations of SNAP91 (1.11-fold up, $P = 1.6 	imes 10^{-2}$) (Fig. 2c, top). All CRISPRa/i perturbations of SNAP91 and TSNARE1 reduced SYNPuncta size (CRISPRa, SNAP91 1.12-fold down, $P = 0.17$; CRISPRa SNAP91 1.04-fold down, $P = 0.16$; TSNARE1 1.09-fold down, $P = 1.2 	imes 10^{-4}$) (Fig. 2c, bottom). To further test the potential of this approach for high-throughput analyses, we employed an automated image analysis approach, which pointed to a reduction in synaptic puncta number and size following SNAP91 CRISPRa and TSNARE1 CRISPRa, confirming the trends we observed in our manual analysis (Supplementary Fig. 3g). Despite changes in SYNPuncta number, we did not observe significant changes in population-wide neuronal activity by MEA (Supplementary Fig. 3h). Although we had expected to prioritize electrophysiological studies through MEA analysis, we proceeded instead with intracellular recording, guided by our genomic analyses, which linked increased SNAP91 expression (the SZ-relevant direction) to previously unknown postsynaptic effects (Fig. 3b).

To specify consistent presynaptic inputs, electrophysiology experiments were conducted on sparse reciprocally labeled isogenic neurons grown on unlabeled control (nuclease dead Cas9 (dCas9)-effector + scrambled gRNA) or SNAP91 CRISPRa/i (dCas9-effector + SNAP91 gRNA) NGN2 excitatory neuronal lawns (Fig. 3d). Comparisons of control (mCherry) and SNAP91 CRISPRa/i (green fluorescent protein (GFP)) neurons on a control lawn tested effects on the postsynaptic neuron, whereas contrasting rare same-colored isogenic neurons across different lawns (either control or CRISPRa/i) queried the impacts on the presynaptic neuron (Contrasting Supplementary Fig. 3e). Strikingly, we observed reciprocal changes in spontaneous synaptogenesis in SNAP91-expressing and SNAP91 CRISPRa/i-expressing neurons with a miR338 inhibitor, but not a negative control, putting CRISPRa/i (Supplementary Fig. 3a-d, Supplementary Dataset 1 and Supplementary Table 2) to elucidate impacts on the global transcriptome and explore the possibility of convergent downstream effects. Gene set enrichment analysis (GSEA) was performed across a collection of 698 manually curated gene sets with a neural theme (subdivided into 8 categories) (Supplementary Dataset 2). We applied a competitive gene set test to evaluate enrichment of genes that were not necessarily genome-wide significant, allowing us to capture even subtle changes in expression, and identified sets of genes for which the distribution of the t-statistics differed from expectation. The results were clustered hierarchically by significance and regression coefficient of the enriched genes (Fig. 3a). Word cloud analysis of enriched gene sets (false discovery rate (FDR) < 0.1%) for SNAP91 and TSNARE1 revealed similar motifs (often containing the words ‘abnormal’, ‘neuron’, ‘morphology’, ‘brain’, or ‘development’), suggesting that perturbation of these genes might impact various neural processes (Fig. 3b).

Rodent studies suggest a role for SNAP91 in presynaptic function; similarly, TSNARE1 is predicted to function in presynaptic vesicle exocytosis. We evaluated isogenic pairs of CRISPRa/i-manipulated NGN2 excitatory neurons, cocultured with human fetal astrocytes to improve synaptic maturation (Fig. 3c–g and Supplementary Fig. 3–h). Through immunocytochemistry, CRISPRa of SNAP91 (1.2-fold down, $P = 1.1 	imes 10^{-4}$) and TSNARE1 (1.2-fold down, $P = 1.7 	imes 10^{-4}$) showed significant reduction in synaptophysin (SYNP) puncta numbers across two control donors, whereas CRISPRa increased SYNPuncta for SNAP91 (1.11-fold up, $P = 1.8 	imes 10^{-2}$) and decreased for TSNARE1 (1.13-fold down, $P = 1.6 	imes 10^{-2}$) (Fig. 3c, top). All CRISPRa/i perturbations of SNAP91 and TSNARE1 reduced SYNPuncta size (CRISPRa, SNAP91 1.12-fold down, $P = 3.1 	imes 10^{-4}$; TSNARE1 1.04-fold down, $P = 0.17$; CRISPRa SNAP91 1.04-fold down, $P = 0.16$; TSNARE1 1.09-fold down, $P = 1.2 	imes 10^{-4}$) (Fig. 3c, bottom). To further test the potential of this approach for high-throughput analyses, we employed an automated image analysis approach, which pointed to a reduction in synaptic puncta number and size following SNAP91 CRISPRa and TSNARE1 CRISPRa, confirming the trends we observed in our manual analysis (Supplementary Fig. 3g). Despite changes in SYNPuncta number, we did not observe significant changes in population-wide neuronal activity by MEA (Supplementary Fig. 3h).

Although we had expected to prioritize electrophysiological studies through MEA analysis, we proceeded instead with intracellular recording, guided by our genomic analyses, which linked increased SNAP91 expression (the SZ-relevant direction) to previously unknown postsynaptic effects (Fig. 3b).
excitatory postsynaptic currents (sEPSCs) when perturbing SNAP91 in neurons cultured on the control lawn, indicating effects on the postsynaptic neuron. Frequency (1.9-fold up, $P = 1 \times 10^{-10}$) and amplitude (1.1-fold up, $P = 6.2 \times 10^{-5}$) both increased following SNAP91 CRISPRa, whereas frequency decreased following SNAP91 CRISPRi (3.3-fold down, $P = 1 \times 10^{-5}$) (Fig. 3e,f). Perturbations of SNAP91 in presynaptic neurons in either direction reduced sEPSC magnitude ($P = 1 \times 10^{-4}$; Supplementary Table 2). Combinatorial perturbation of SNAP91 CRISPRa: 3.2-fold down, $P = 1 \times 10^{-4}$, SNAP91 CRISPRi: 2-fold down, $P = 1 \times 10^{-4}$ (Fig. 3e,f). Interestingly, perturbation of SNAP91 did not produce genome-wide significant differentially expressed genes (DEGs; Supplementary Fig. 3a), suggesting that the observed synaptic phenotypes are a consequence of multiple subtle changes in gene expression (Fig. 3a) or might reflect disruptions at the protein level. Altogether, in addition to the effects in presynaptic neurons, our findings hint at a second role for SNAP91 in postsynaptic receptor stabilization and/or retrograde signaling.

After summarizing all synaptic phenotypes in a heatmap (Fig. 3f), rather than observing a correlation between samples with SZ-relevant (upregulated) perturbations in SNAP91 and TSNARE1, we unexpectedly found a striking congruence between SNAP91 CRISPRi and TSNARE1 CRISPRs (Fig. 3g). This was apparent in the gene set enrichment patterns (Fig. 3b) and confirmed in the correlation of the SNAP91 CRISPRi and TSNARE1 CRISPR t-statistics (Fig. 3a and Supplementary Fig. 3f), suggesting that although not necessarily relevant to the study of SZ, unexpected convergent effects between SNAP91 and TSNARE1 may occur.

Overall, these findings highlight the complexity of the possible consequences of SZ-associated eQTLs, whereby subtle changes in target gene expression in either direction can converge on synaptic level phenotypes.

**Combinatorial perturbation of SZ eQTL genes improves correlation with differential expression in neuropsychiatric disorders.**

We computationally perturbed four SZ eQTL genes in a better approximation of the polygenic nature of SZ (Fig. 4 and Supplementary Fig. 4), in the direction predicted for disease risk, using CRISPRa to upregulate SNAP91, TSNAREI and CLCN4 and RNA interference (RNAi) to repress FURIN (Supplementary Fig. 4a–c, Supplementary Dataset 1 and Supplementary Table 2). Combinatorial perturbation led to differential expression of 1,261 genes (FDR < 5%, 665 up, 596 down) and impacted coregulated downstream genes and proteins. Competitive GSEA, consistent with our phenotypic analyses, found strong enrichment of synaptic gene sets following individual SNAP91 and TSNARE1 CRISPRa, while CLCN3 CRISPRa and FURIN RNAi showed more moderate outcomes; the combinatorial perturbation in fact yielded negative correlation with synaptic gene sets (Fig. 4a). The 1,261 combinatorial perturbation DEGs formed a protein network of 1,151 nodes with highly significant protein–protein interaction enrichment in the STRING database (http://string-db.org) ($P < 1 \times 10^{-14}$) (Fig. 4b); weighted gene coexpression network analysis revealed 21 coexpression modules (Fig. 4c). Two modules were highly significantly associated with the combinatorially perturbed sample signature (ME-lightgreen and ME-darkgreen) (Fig. 4d), showing significant enrichment (FDR < 5%) in gene sets commonly associated with psychiatric disorders, such as miR137, common and rare SZ-variant genes, differential expression in bipolar disorder and protein–protein interaction networks and fragile X mental retardation protein-regulated genes in autism spectrum disorder (ASD), but also, interestingly, SZ-relevant drug classes (that is, antipsychotic, nortriptyline and dopaminergic) (Fig. 4e).

To test for convergence of our SZ-eQTL perturbations with differential expression in an hiPSC neural dataset generated from control and childhood onset SZ (childhood-onset schizophrenia (COS)) cases as well as a postmortem SZ analysis (CMC)\(^2\), which originally identified our common variant target genes, we calculated the Spearman correlation of their t-statistics (Fig. 4f). All but one individual gene perturbation (CLCN3) actually correlated negatively with both datasets. However, remarkably, the combinatorial perturbation displayed highly significant positive correlation with the differential expression in SZ from both the postmortem CMC and hiPSC-based COS cohort (Fig. 4f). Very similar results were observed when comparing the t-statistics with those from multiple postmortem brain comparisons of bipolar disorder, ASD and major depressive disorder (NIMH HBCC and UCLA datasets), but not alcohol dependence (Fig. 4g). We also compared our differential expression t-statistics to genetically regulated gene expression associations of SZ (predicXcan\(^2\)). Although we did not see a significant correlation between overall sets of summary statistics ($P > 0.05$), genes that were nominally significant in our combinatorial analysis were more likely to also be nominally significant in the dorsolateral prefrontal cortex (DLPFC) genetically regulated gene expression analysis (binomial test; $P < 7 \times 10^{-5}$) (Supplementary Table 3).

Our observations suggest considerable downstream effects specific to a combinatorial perturbation of SZ eQTL genes that go beyond what would be expected from an additive effect of individually perturbed genes. These synergistic effects emphasize the importance of considering the polygenic nature of SZ and other neuropsychiatric disorders, where a combination of variants contributes to disease.

**Synergistic effects beyond the additive impact of individual SZ eQTL genes are enriched in synaptic as well as common and rare SZ-variant genes.**

In an effort to explore the nature of this synergistic effect, we modeled the additive effect of differential expression in singly perturbed SZ eQTL genes computationally (Fig. 5a), which led to similar numbers of expected DEGs as found in the combinatorial approach (641 up, 604 down versus 665 up, 596 down). Figure 5b
illustrates this for three representative genes, which exemplify the different synergistic effects that can be found when comparing additive model and combinatorial perturbation. Interestingly, the correlation between the nominally significant synergistic differentially expressed gene t-statistics and DLPFC genetically regulated gene expression z-scores was significant (P = 0.0097; Supplementary Table 3).
Fig. 5 | The synergistic effects of SZ eQTL genes converge on synaptic function and common and rare variant signatures. a, Schematic of differential expression analysis. Individual gene modifications, the implementation of the expected additive model based on the latter and the measured combinatorial perturbation allowing for the detection of synergistic effects through comparison with the additive model. b, log2(fold changes) of three representative genes (CRMP1, FMN1, DLX1) in individual SZ eQTL gene perturbations, their computed additive model and their combinatorial perturbation, illustrating different possible synergistic effects (negative, positive and none, respectively). c, Hierarchical clustering of the t-statistics for the additive model and the combinatorial perturbation. The color gradient represents t-statistic values. d, Differential expression log2(fold changes) of SNAP91, TSNARE1, CLCN3 and FURIN in the additive model and the combinatorial perturbation. e, Pie chart showing the percentages of genes that exhibit similar or more moderate differential expression (beige) following combinatorial perturbation in comparison with the expected additive model, as well as genes that are more downregulated (red) and more upregulated (blue). f, Hierarchical clustering of the differential expression log2(fold changes) of ‘more downregulated’ and ‘more upregulated’ genes, in the additive model versus the combinatorial perturbation. FMN1, as seen in b, is part of the ‘more upregulated’ category. g,h, Overrepresentation analysis, using a hypergeometric test, of 698 curated gene sets and those ‘more downregulated’ and ‘more upregulated’ genes with significant synergistic differential expression (FDR <10%, n (more downregulated) = 36 genes, n (more upregulated) = 132 genes), ranked by adjusted significance.

We identified those genes that showed larger changes following combinatorial perturbation than predicted by our additive model (Fig. 5c–f and Supplementary Fig. 4d,e). For both the additive model and the combinatorial perturbation, we performed hierarchical clustering of the t-statistics for each contrast (Fig. 5c). The differences between the predicted and observed cumulative effects (including
inverse differential expression of some genes) was not explained by unequal gene perturbation magnitudes of SNAP91 (fold change = 2.9 versus 2.2), TSNARE1 (fold change = 3.1 versus 2.9), CLCN3 (fold change = 1.5 versus 1.7) or FURIN (fold change = 0.8 versus 0.86) between the individual and combinatorial perturbations (Fig. 5d; for the CMC information, see Supplementary Table 4).

To examine the synergistic effects in more detail, we grouped genes based on differential expression between the additive model and the combinatorial perturbation (Fig. 5e). Genes were classified as 'more' differentially expressed in the combinatorial perturbation than predicted if their logarithmic fold change differed by at least 0.3 (a conservative estimation of the maximum s.d. in all samples) (Fig. 5f). Most genes (82%) were altered approximately as predicted or less (Supplementary Fig. 4d,e), while 7% (1,430 genes) were more downregulated and 11% (2,107 genes) more upregulated than expected (Fig. 5e,f). As might be expected, the latter were mainly overrepresented in weighted gene coexpression network analysis modules that showed correlation with the transcriptional profile in the combinatorial perturbation samples and vice versa (Supplementary Fig. 4f and Fig. 4d). Furthermore, testing for overrepresentation in our curated neural gene sets showed that genes more downregulated than expected from the additive model were significantly enriched for pre- and postsynaptic gene sets (Fig. 5g), particularly those involved in the secretion of glutamate and other neurotransmitters, synaptic vesicle trafficking and a postsynaptic glutamate receptor pathway (Fig. 5h). Genes more upregulated than expected in the additive model correlated with disorder signatures (Fig. 5g), particularly genes harboring rare copy number variations or nonsynonymous de novo mutations associated with SZ, as well as SZ GWAS genes (Fig. 5h). This latter result is striking, since it links common and rare variants more broadly associated to psychiatric disease.

In this study, we show that combinatorial perturbation of four SZ eQTL genes, in an approximation of the presumed transcriptional alterations they cause in SZ, results in a synergistic effect that culminates in the downstream alteration of both rare and common risk variant genes.

Discussion

We have demonstrated human-specific functional validation of a putative causal SNP (FURIN rs4702) by CRISPR editing and multi-SNP candidate genes (SNAP91, TSNARE1 and CLCN3) via single and combinatorial CRISPRa/i manipulations. CRISPR editing of a single noncoding SNP altered neuronal expression of the cis-gene target and impacted a variety of neuronal phenotypes, while CRISPRa/i effected small changes in target genes that resulted in convergent downstream transcriptomic differences capturing the effects observed in the postmortem brain. Our isogenic hiPSC-based strategies manipulated common variant loci and genes in human neurons and suggest that synergy between risk variants may impact SZ risk.

CRISPR editing at FURIN rs4702 led to significant transcriptomic and cellular effects when altering even a single noncoding SNP, but also identified a surprisingly large cell-type-specific effect not detected by postmortem studies. Located in the 3′-UTR of the FURIN gene and within the binding site of miR338, the regulatory activity of FURIN rs4702 is dependent on miR338 expression levels, which in itself is sufficient to mediate the thalamocortical disruptions observed in SZ-associated 22q11.2 deletion syndrome (14). CRISPRa/i studies further revealed the impact of SZ eQTL genes on neuronal branching, synaptic puncta density and synaptic activity. Although genetic, genomic and proteomic studies previously implicated both pre- and postsynaptic processes in SZ, our electrophysiological analyses demonstrated that even a single SZ eQTL gene (SNAP91) can impact both pre- and postsynaptic neuronal function.

Our study integrated existing PGC GWAS, fine-mapping and cis-eQTL analyses to prioritize the SNPs and genes tested in the study. Advances in analytic strategies, such as new multi-SNP cis-eQTL analyses (that is, colocalization (14), PrediXcan3) coupled with their application to larger PGC SZ GWAS genetic and genomic datasets continue to expand and refine the list of SZ-associated common variants (Supplementary Table 5) and genes (Supplementary Table 6) suitable for functional validation. The release of the unpublished PGC3 SZ GWAS (65,205 cases and 87,919 controls) and larger postmortem RNA-seq datasets will further resolve the list of putative causal variants and genes linked to SZ. Moreover, as more types of QTL studies become available, this will inform new avenues for functional validation; for example, a SNAP91 splice QTL is in high linkage disequilibrium with a SZ risk index SNP, suggesting that differential splicing may influence expression differences in SZ (9).

This work underscores the technical difficulties in adapting CRISPR-based systems as scalable platforms to test SZ eQTL genes. First, not all genes (notably FURIN) proved equally amenable to CRISPRa/i, particularly in mature neurons, reinforcing that all gRNAs must be independently validated in each neural cell type and each donor (14). Second, although MEA is a widely used, convenient and scalable method to record population-wide neuronal activity (15), in this study it did not detect phenotypes identified by electrophysiology (a similar phenomenon has been reported in refs. (16,17)). Therefore, we recommend screening synaptic characteristics across multiple assays whenever possible. Although labor-intensive, patch clamp electrophysiology remains the gold standard for studying synaptic function. Long-term, recent advances in automated patch clamping technology may yet improve the scalability of this approach for screening large numbers of SZ eQTL genes, alone and in combination. Toward this, advanced engineering of gRNA structure and gRNA expression systems will improve the efficiency of multiplexed gene regulation.

We previously hypothesized that, relative to postmortem analyses, isogenic experiments would show comparable effect sizes but decreased s.d. as a reflection of reduced biological and technical variation; instead, we report that the cis-eQTL effect sizes observed through isogenic comparisons of FURIN rs4702 were substantially larger than observed in postmortem brain analyses (14). This suggests that the cell-type-specific effects may have been diluted in postmortem cis-eQTL studies of brain homogenates. As CRISPR-based SNP edits are repeated across larger numbers of donors (particularly those with extreme polygenic risk scores), we predict that observed cis-eQTLs should remain relatively consistent between individuals, meaning that all individuals will have a similar (expected) cis-effect size from the CRISPR allelic conversion. There is currently no specific reason to reject this null model and suspect widespread epistasis, although empirical studies may prove otherwise. Nonetheless, we speculate that downstream transcriptomic and cellular (that is, synaptic) effects may vary as a result of interactions with preexisting donor-specific risk alleles, a hypothesis that may prove true more broadly across complex genetic disorders. Moreover, manipulating SZ eQTL genes across donors with high and low polygenic risk scores might make it possible to better distinguish the additive, epistatic and omnigenic models of inheritance.

Convergence between the various risk variants linked to psychiatric diseases, including SZ, ASD and bipolar disorder has long been hypothesized (14,16–18) but is yet unproven. We observed a striking convergence and synergy downstream even of our small sampling of four SZ eQTL genes. First, CRISPRa/i-based gene expression analyses revealed a surprising degree of overlap between the downstream differential expression patterns and neuronal phenotypes of two SZ eQTL genes, SNAP91 and TSNARE1. Second, combinatorial perturbation of four SZ eQTL genes in the SZ-relevant direction (upregulated: SNAP91, TSNARE1 and CLCN3; downregulated: FURIN) revealed negative synergy converging on synaptic function and pos-
itive synergy linking the rare and common variant genes implicated in psychiatric disease risk. These observations agree with a highly complex, polygenic etiology of SZ and other psychiatric disorders\(^{18,22}\) and are interesting to consider in the context of additive\(^{27,30}\), epistatic\(^{28}\) and omnigenic\(^{26,31}\) models of inheritance. For example, it is tempting to speculate whether the synergy we observed is consistent with individual risk being additive on a liability scale, or if our evidence instead suggests that SNPAP1, with reciprocal effects on synaptic function and seemingly lacking downstream transcriptional targets, might represent a core gene for SZ as conceptualized in the omnigenic model. High-throughput CRISPR-based perturbation methods\(^{1,25}\) will help to infer gene regulatory networks, resolve coregulation of core disease genes and/or identify the existence of peripheral master genes in psychiatric disease. Overall, our observations, coupled with findings that high polygenic risk increases disease liability in carriers of rare mutations\(^{26,32}\), suggest cumulative effects between rare and common risk variants. Our hope is that hiPSC-based models will illuminate the synergistic impact of common variants on cellular and molecular phenotypes, leading us toward precision psychiatry\(^{33}\) and more personalized medicine\(^{34}\).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0497-5.

Received: 29 July 2018; Accepted: 13 August 2019; Published online: 23 September 2019

References

1. Ripke, S. et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427 (2014).
2. Pardinas, A. F. et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nature Genet. 50, 381–389 (2018).
3. Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nature Genet. 50, 621–629 (2018).
4. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. Nature Genet. 50, 825–833 (2018).
5. Jaffe, A. E. et al. Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. Nat. Neurosci. 21, 1117–1123 (2018).
6. Hall, L. S. et al. A transcriptome wide association study implicates specific peripheral master genes in psychiatric disease. Overall, our observation instead suggests that SNPAP1, with reciprocal effects on synaptic function and seemingly lacking downstream transcriptional targets, might represent a core gene for SZ as conceptualized in the omnigenic model. High-throughput CRISPR-based perturbation methods\(^{1,25}\) will help to infer gene regulatory networks, resolve coregulation of core disease genes and/or identify the existence of peripheral master genes in psychiatric disease. Overall, our observations, coupled with findings that high polygenic risk increases disease liability in carriers of rare mutations\(^{26,32}\), suggest cumulative effects between rare and common risk variants. Our hope is that hiPSC-based models will illuminate the synergistic impact of common variants on cellular and molecular phenotypes, leading us toward precision psychiatry\(^{33}\) and more personalized medicine\(^{34}\).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0497-5.

Received: 29 July 2018; Accepted: 13 August 2019; Published online: 23 September 2019

References

1. Ripke, S. et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427 (2014).
2. Pardinas, A. F. et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nature Genet. 50, 381–389 (2018).
3. Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nature Genet. 50, 621–629 (2018).
4. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. Nature Genet. 50, 825–833 (2018).
5. Jaffe, A. E. et al. Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. Nat. Neurosci. 21, 1117–1123 (2018).
6. Hall, L. S. et al. A transcriptome wide association study implicates specific peripheral master genes in psychiatric disease. Overall, our observation instead suggests that SNPAP1, with reciprocal effects on synaptic function and seemingly lacking downstream transcriptional targets, might represent a core gene for SZ as conceptualized in the omnigenic model. High-throughput CRISPR-based perturbation methods\(^{1,25}\) will help to infer gene regulatory networks, resolve coregulation of core disease genes and/or identify the existence of peripheral master genes in psychiatric disease. Overall, our observations, coupled with findings that high polygenic risk increases disease liability in carriers of rare mutations\(^{26,32}\), suggest cumulative effects between rare and common risk variants. Our hope is that hiPSC-based models will illuminate the synergistic impact of common variants on cellular and molecular phenotypes, leading us toward precision psychiatry\(^{33}\) and more personalized medicine\(^{34}\).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0497-5.

Received: 29 July 2018; Accepted: 13 August 2019; Published online: 23 September 2019

References

1. Ripke, S. et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427 (2014).
2. Pardinas, A. F. et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nature Genet. 50, 381–389 (2018).
3. Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. Nature Genet. 50, 621–629 (2018).
4. Skene, N. G. et al. Genetic identification of brain cell types underlying schizophrenia. Nature Genet. 50, 825–833 (2018).
5. Jaffe, A. E. et al. Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. Nat. Neurosci. 21, 1117–1123 (2018).
6. Hall, L. S. et al. A transcriptome wide association study implicates specific peripheral master genes in psychiatric disease. Overall, our observation instead suggests that SNPAP1, with reciprocal effects on synaptic function and seemingly lacking downstream transcriptional targets, might represent a core gene for SZ as conceptualized in the omnigenic model. High-throughput CRISPR-based perturbation methods\(^{1,25}\) will help to infer gene regulatory networks, resolve coregulation of core disease genes and/or identify the existence of peripheral master genes in psychiatric disease. Overall, our observations, coupled with findings that high polygenic risk increases disease liability in carriers of rare mutations\(^{26,32}\), suggest cumulative effects between rare and common risk variants. Our hope is that hiPSC-based models will illuminate the synergistic impact of common variants on cellular and molecular phenotypes, leading us toward precision psychiatry\(^{33}\) and more personalized medicine\(^{34}\).

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of code and data availability and associated accession codes are available at https://doi.org/10.1038/s41588-019-0497-5.
50. Purcell, S. M. et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature* **506**, 185–190 (2014).
51. Sanders, S. J. et al. Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. *Neuron* **87**, 1215–1233 (2015).
52. O’Dushlaine, C. et al. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat. Neurosci.* **18**, 199–209 (2015).
53. Balloz, S. & Gillis, J. Strength of functional signature correlates with effect size in autism. *Genome Med.* **9**, 64 (2017).
54. Jia, P., Chen, X., Fanous, A. H. & Zhao, Z. Convergent roles of de novo mutations and common variants in schizophrenia in tissue-specific and spatiotemporal co-expression network. *Transl. Psychiatry* **8**, 105 (2018).
55. Anttila, V. et al. Analysis of shared heritability in common disorders of the brain. *Science* **360**, eaap8757 (2018).
56. Gandal, M. J. et al. Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359**, 693–697 (2018).
57. Purcell, S. M. et al. Recovery of trait heritability from whole genome sequence data. Preprint at bioRxiv https://doi.org/10.1101/588020 (2019).
58. Mimitou, E. P. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **17**, 1153 (2020).
59. Purcell, S. M. et al. A polygenic burden of rare disruptive mutations in neuropsychiatric disorders. *Nat. Neurosci.* **21**, 1153 (2018).
60. O’Dushlaine, C. et al. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat. Neurosci.* **18**, 199–209 (2015).
61. Purcell, S. M. et al. Discovery and functional dissection of a risk locus for schizophrenia at 6q14. *Science* **359**, 693–697 (2018).
62. Zhang, J. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **17**, 1153 (2018).
63. Boer, A. J. et al. Coupled single-cell CRISPR screening and epigenomic profiling reveals causal gene regulatory networks. *Cell* **176**, 361–376.e17 (2019).
64. Mimitou, E. P. et al. Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019).
65. Weiner, D. J. et al. Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat. Genet.* **49**, 978–985 (2017).
66. Tansey, K. E. et al. Common alleles contribute to schizophrenia in CNV carriers. *Mol. Psychiatry* **21**, 1153 (2016).
67. McMahon, F. J. & Insel, T. R. Pharmacogenomics and personalized medicine in neuropsychiatry. *Neuron* **74**, 773–776 (2012).

**Acknowledgements**
This manuscript is dedicated to Pamela Sklar, a scientist, mentor and friend who advanced the field of psychiatric genetics. This work was partially supported by NIH grants (nos. R56 MH101454 to K.J.B., E.S. and H.M., R01 MH106656 to K.J.B., U19 MH104172 and U19 MH107367 to B.A., and R01 MH109897 to P.S. and K.J.B.), the New York Stem Cell Foundation (to K.J.B.) and project ALS (grant no. 2017-03 to J.G. and H.P.). We thank the Neuroscience and Stem Cell Cores at ISMMS. This work was supported in part through the computational resources and staff expertise provided by the Scientific Computing unit at ISMMS. P. O’Reilly provided thoughtful feedback on the manuscript. The GTEx Project was supported by the Common Fund of the Office of the Director of the NIH and by National Cancer Institute, National Human Genome Research Institute, National Heart, Lung, and Blood Institute, National Institute on Drug Abuse, National Institute of Mental Health and National Institute of Neurological Disorders and Stroke. The data used for the analyses described in this manuscript were obtained from the GTEx portal on 4 January 2019.

**Author contributions**
N.S., S.-M.H., P.S. and K.J.B. contributed to the experimental design. N.S. conducted all the CRISPR editing experiments, assisted by M.R.-M., A.T. and S.A. S.-M.H. completed all the CRISPRa/i experiments. K.Y. and H.M. conducted and analyzed all the CRISPRa/i electrophysiological experiments. N.S., S.-M.H. and M.R.-M. conducted all the MEA experiments. E.C. generated all the CRISPR-edited organoids. P.S. conducted all the transcriptomic analyses, with critical advice from G.H. and E.F. All confocal imaging and semiautomated synaptic analyses were conducted by S.-M.H., M.R.-M. and N.B. V.S., D.G. and B.A. conducted all the automated high-content imaging analyses. K.A. and R.M. conducted the kinome analysis. K.J.B., N.S. and S.-M.H. wrote the manuscript.

**Competing interests**
The authors declare no competing interests.

**Additional information**
Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0497-5.
Correspondence and requests for materials should be addressed to K.J.B.
Reprints and permissions information is available at www.nature.com/reprints.
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019
For the TaqMan ddPCR mix, 1x ddPCR Supermix for Probes; catalog no. 1863024; Bio-Rad Laboratories) and 1x TaqMan SNP Genotyping Assay (rs4702 A/G, H2O to 25 μl; catalog no. 4351379; Thermo Fisher Scientific) were added to 180 ng DNA of each pooled sample as well as H2O, rs4702 AA and rs4702 GG controls. ddPCR was performed and analyzed with the QX100 Droplet Digital PCR System (catalog no. 186-3001; Bio-Rad Laboratories) according to the manufacturer’s instructions. For the pooled row sample showing the highest G allele concentration, ddPCR was performed for each well as described earlier. The resulting well with the highest G allele concentration was noted and the same well was found on the duplicated 96-well plate still in culture. Cells in that well were then seeded onto a new 96-well plate using EDTA to allow for nonhomogenous distribution. This 96-well plate in return was duplicated, DNA was isolated and the pooled TaqMan ddPCR performed as described earlier. Enrichments of 40–60% can be expected by round 2. If the enrichment is too low, a third round can be performed, although enrichments often reach close to 100% and therefore lack a heterogeneity between clones.

Restriction fragment polymorphism screening and Sanger sequencing. At 60% confluence DNA was isolated from one plate as described earlier and 2 μl of each well was used for amplification of a 600-bp region around the target site in a 20 μl reaction as stated earlier. Restriction enzyme digestion was performed by adding digestion mix (3 μl 10x NEBuffer 3.1 (New England BioLabs), 6.7 μl H2O and 0.3 μl SfiI restriction enzyme (catalog no. R0172; New England BioLabs) each) directly to PCR products and incubating for 1.5 h at 37 °C. SfiI digested DNA samples (1 μl) were run on an agarose gel and the amplicon was identified. Sanger sequencing was performed as described previously.

CRISPRa/i manipulation of endogenous gene expression. hiPSC NPCs were generated as described previously23, CRISPRa/i gRNA design and cloning, antibiotic selection of dCas9-VP and dCas9-KRAB hiPSC NPCs, gRNA lentiviral transduction and NGN2 induction24 were conducted as described previously25. More details about CRISPRa/i manipulation of endogenous gene expression can be found in the Supplementary Note.

Molecular and synaptic phenotyping. Quantitative reverse-transcription PCR. RT–qPCR was performed as described previously26. Cells were collected with TRizol and total RNA extraction was carried out according to the manufacturer’s instructions. Quantitative transcript analysis was performed using a QuantStudio 7 (Thermo Fisher Scientific). Relative expression was determined using the 2^(-ΔΔCT) method (both Thermo Fisher Scientific). The total RNA template (25 ng per reaction) was added to the PCR mix, including the primers (Supplementary Table 7). qPCR conditions were as follows: 48 °C for 15 min; 95 °C for 10 min followed by 45 cycles (95°C for 15 s, 60 °C for 60 s). All qPCR data were collected from at least three independent biological replicates of one experiment. Data analyses were performed using Prism 6 (GraphPad Software).

Immunostaining and microscopy. NGN2 excitatory neurons (23 or 24 days old) (on coverslips) were washed with ice-cold PBS and fixed with 4% paraformaldehyde/sucrose PBS solution at pH 7.4 for 20 min at room temperature. Then, fixative solution was replaced with permeabilizing solution (ice-cold 0.1% Triton X-100 PBS supplemented with 5% donkey serum), followed by incubation on ice for 20 min and another 20 min at room temperature. After washing with PBS 3 times, NGN2 excitatory neurons were incubated with blocking solution (0.1% Tween 20, 5% donkey serum in PBS) for 1 h at room temperature. The blocking solution was aspirated and replaced with the same solution with primary antibodies: anti-GFP-rabbit (1:500; catalog no. ab6556; Abcam); MAP2-mouse (1:500; catalog no. MA5-12826; Thermo Fisher Scientific); synaptophysin 1-rabbit (1:500; catalog no. 1001 002; Sino Biological Systems); synapsin (SYN) 2-rabbit (1:500; catalog no. 106 006; Sino Biological Systems); Ctip2-chicken (1:500; catalog no. ab18465; Abcam); SATB2-mouse (1:5; catalog no. ab15302; Abcam); GABA-a-rabbit (1:500; catalog no. A2052; Sigma-Aldrich); and GAD67-mouse (1:1,000; catalog no. MAB5406; Millipore Biotech), and human cortical spiny (1:500; catalog no. A2052; Sigma-Aldrich), and subpalial spiny (1:500; catalog no. MAB5406; Millipore Biotech) neurons were incubated for 24 h at 4 °C. Neurons were then incubated with secondary antibodies: Alexa Fluor 488 donkey anti-rabbit, (1:500; catalog no. 711-545-152; Jackson ImmunoResearch); Alexa Fluor 568 goat anti-chicken (1:500; catalog no. A-11041; Thermo Fisher Scientific); Alexa Fluor 647 donkey anti-mouse immunoglobulin G (H+L) (1:500; catalog no. 715-605-150; Jackson ImmunoResearch).
Immunoresearch), prepared in blocking solution for 2 h at room temperature, followed by washing with PBS 3 times.

A 20 μl aliquot of Aqua-Poly/Mount mounting solution (catalog no. 18060–20; Polysciences) was added to the overlying coverslip and air-dried for 2 d at ambient temperature. Neurons on the coverslips were imaged with a Zeiss LSM 780 confocal microscope. For presynaptic immunocytochemistry imaging, images were acquired (five images each from three biological replicates per condition and cell line in three different experimental sets) using a confocal microscope (LSM 780) with 10× ocular lens and 63× objective lens. After uniform thresholding of all SYP, SN1 1/2 and microtubule-associated protein 2 (MAP2) images, SYP, SN1 and SYN 1/2 puncta number or their size were measured. Total SYP and SN1 1/2 puncta number per image were divided by that image's respective MAP2 antibody-positive area to calculate SYP and SN1 1/2 puncta counts normalized to MAP2 antibody levels. The average puncta number of SYP and SN1 1/2 per image was used for statistical analysis. Data from 51–54 images from 3 independent experiments were analyzed with Prism 6.

For neurite tracing, GFP-labeled NGN2 excitatory neurons (D5), cultured as described earlier, were seeded at very low density (50–100 cells per coverslip) on a lawn of same-age unlabeled NGN2 excitatory neurons (5.0–6.0×10⁴ cells per coverslip). Cultures were delivered overnight in 10× volume of culture media. To maximize GFP neurites from overgrowth and tangling, thus easing imaging and analysis. Immunostaining was performed against GFP. For neurite tracing imaging, GFP NGN2 excitatory neurons were captured at a magnification of ×20 using a confocal microscope (LSM 780). Nine images containing one neuron each were captured per coverslip per condition (two technical replicates of two independent cell lines in two independent experiments). Images were stitched using the National Institutes of Health (NIH) open-source image processing software ImageJ [v.2.0.0-rc-69/1.52n, at 5% overlap and transformed as maximum projection. Using the ImageJ plug-in 'simple neurite tracer', each neuron's cell soma was identified and neurites were traced from primary neurites outwards. Tracings were exported for each individual neurite for analysis. Data from 15–20 cells per condition and 3 independent experiments were analyzed using 1-way ANOVA followed by Dunnett's multiple comparison post test.

Neurosphere assay: Control rs4702 AA and edited rs4702 GG NPC lines were dissociated using accutase and cultured in low-adherence plates for 48 h to generate neurospheres. Neurospheres were manually picked and replated into individual wells of a Matrigel-coated 96-well plate. Images of neurospheres were taken 1 h post-plating using a brightfield microscope with a 4× objective (numerical aperture 0.1). Forty-eight hours after plating neurospheres were fixed in 4% formaldehyde, stained using DAPI and imaged using an epifluorescence microscope with a 4× objective (numerical aperture 0.1). Images of each neurosphere were traced at each time point and the average radial migration was calculated using ImageJ by subtracting the average radius of each neurosphere 1 h post-plating from the radius 48 h post-plating. Twenty-eight neurospheres taken from three separate passages were analyzed for each experimental condition. Results were analyzed with an unpaired Student's t-test.

Serine/threonine-specific protein kinase activity profiling. Profiling of serine/threonine-specific protein kinase activity was performed with the PhosMaStation12 microarray (PamGene) and STK (4-well) PamChips containing 144 consensus phosphopeptide sequences (three of which were internal controls) per well, immobilized on porous ceramic membranes. Each PamChip well was blocked with 2% bovine serum albumin before 2 μM ATP and fluorescein isothiocyanate-labeled anti-phospho serine/threonine antibodies (PamGene, 157μM AT and fluorescein isothiocyanate-labeled anti-phospho serine/threonine antibodies (PamGene) were added in each well. The homogenized samples containing the active kinases and assay mix were pumped through the wells to facilitate interaction between kinases in the sample and the immobilized substrate immobilized on the chip. The degree of phosphorylation per well was measured in real time using the Evolve v.2.0.8 (PamGene) kinetic assay. The main drivers of gene expression variance were determined using the main effect panel. The limma v.3.36.1 function was used to compare the weights for heteroscedasticity adjustment by estimating the mean variance trend for log, counts. Linear models were fitted to the expression values of each gene using the limFit function. Supplementary Fig. 4b summarizes the contrasts used for the combinatorial perturbation experiment. Empirical Bayesian moderation was applied using the eBayes function to obtain more precise comparisons testing for two-way ANOVA (Sidak's post hoc test for comparison of all sample means). For presynaptic imaging, MEA and whole-cell patch clamp electrophysiology, statistical significance was tested using multiple comparisons testing for two-way ANOVA (Sidak's post hoc test for comparison of all sample means).

RNA-seq analysis. RNA-seq libraries were prepared using the Kapa Total RNA library prep kit. Paired-end sequencing reads (125 bp) were generated on an Illumina HiSeq 2500 platform (coverage per reads: 40 M), aligned to hg19 using STAR aligner v.2.5.2a; uniquely mapping reads were counted with featureCounts. The main drivers of gene expression variance were determined using the main effect panel. The limma v.3.36.1 function was used to compare the weights for heteroscedasticity adjustment by estimating the mean variance trend for log, counts. Linear models were fitted to the expression values of each gene using the limFit function. Supplementary Fig. 4b summarizes the contrasts used for the combinatorial perturbation experiment. Empirical Bayesian moderation was applied using the eBayes function to obtain more precise comparisons testing for two-way ANOVA (Sidak's post hoc test for comparison of all sample means). For presynaptic imaging, MEA and whole-cell patch clamp electrophysiology, statistical significance was tested using multiple comparisons testing for two-way ANOVA (Sidak's post hoc test for comparison of all sample means).

D3, pHAs were split as 17 K cells per well in a Matrigel-coated 48 W MEA plate (catalog no. M768-4MEA-48W; Axion Biosystems) and maintained as stated earlier. At D7, NGN2 excitatory neurons were detached, spun down and seeded on the coverslip in the well plate coated with deionized water to minimize the evaporation of marginal wells during long-term culture. Half volume of neuronal medium (supplemented with 2% FCS) was replaced with fresh medium including 2 μM cytarabine from D9 until the end of MEA recording. Electrical activity of neurons was recorded at 37°C once every week from D15 (approximately D15, D23, D29, D33 and D41). On each recording day, the plate was loaded into the Axion Axioptro MEA reader (Axion Biosystems). Recording was performed with AxoS v.2.4. The batch mode/statistic compiler tool was run after the final recording. Quantitative analysis of the recording was performed as a Microsoft Excel v.14.5.6 spreadsheet. Data from 6–12 biological replicates were analyzed using Prism 6 or R.

Whole-cell patch clamp electrophysiology. NGN2 excitatory neurons (23–24 days old), grown on PHA, with SNAP91 CRISPRa/i were prepared as described earlier. Especially on D6, a very low titer of EF1-t-dTomato or EF1α-enhanced GFP (eGFP) lentivirus was added to a well of either dCas9-effector NGN2 excitatory neurons with scramble GRNA or SNAP91 gRNA. At D7, unlabeled dCas9-effector NGN2 excitatory neurons with scramble gRNA or SNAP91 gRNA (grown separately from the labeled NGN2 excitatory neurons stated earlier) were split on PHA culture as 4.5–6.0×10⁴ cells per coverslip, serving as neuronal lawns. Additionally, NGN2 excitatory neurons labeled with dTomato or eGFP were seeded into each neuronal lawn 24 h later. Feeding the cultures was described earlier. Peak patch clamp recordings were performed from these neurons at D29–31 using borosilicate glass electrodes (3–5 MΩ). Whole-cell voltage clamp recordings were obtained with the internal solution containing: 130 mM K-glutamate; 6 mM MgCl₂; 4 mM NaCl; 10 mM HEPES Na; 0.2 mM EGTA K; 0.3 mM GTP; 2 mM Mg-ATP; 10 mM d-glucose. The pH and osmolality of the internal solution were close to physiological conditions (pH 7.3–7.4, 300–305 mOsm; pH 7.3–7.4, 300–305 mOsm). EPSCs were recorded in voltage clamp mode to monitor the spontaneous EPSC frequency and amplitude. Spontaneous EPSCs were recorded in BrainPhys Neuronal Medium (STEMCELL Technologies) at 36–50 μA. Data were low-pass filtered at 10 kHz and acquired at 10 kHz using Multichannel 700B Microelectrode Amplifier (Molecular Devices) and pClamp 10 (Molecular Devices). Series and membrane resistance was continuously monitored and recordings were discarded when these measurements changed by >20%. Recordings where series resistance exceeded 25 MΩ were rejected. Detection and analysis of EPSCs were performed using Mini Analysis v.6.0.7 (Synaptosoft). For current-clamp recordings, series resistance was monitored and canceled using a bridge circuit, and pipette capacitance was compensated. Voltage signals were low-pass filtered at 10 kHz. The baseline membrane potential was maintained near −70 mV with a current injection (5 pA). We recorded membrane potential responses to hyperpolarizing and depolarizing current pulses (500 ms in duration); then, we examined the action potential and subthreshold membrane properties using Signal v.4 (Cambridge Electronic Design).

Data analysis. Data from all phenotypic assays were first organized in a Microsoft Excel spreadsheet and analyzed with Prism 6 or R. For qPCR data analysis and synaptic imaging analysis, values are expressed as the mean ± s.e.m. Statistical significance was tested using either a one-sided Student's t-test or one-way ANOVA (Sidak's post hoc test for comparison of all sample means). For presynaptic imaging, MEA and whole-cell patch clamp electrophysiology, statistical significance was tested using multiple comparisons testing for two-way ANOVA (Sidak's post hoc test for comparison of all sample means).
The expected additive effect was modeled through the addition of the individual coefficient comparisons: (TSNARE1—Ctrl) + (SNAP91—Ctrl) + (CLCN3—Ctrl) + (FURIN—Ctrl). The synergistic effect was modeled by subtracting the additive effect from the combinatorial perturbation comparison: (multiplexed genes—multiplexedCtrls)−((TSNARE1—Ctrl) + (SNAP91—Ctrl) + (CLCN3—Ctrl) + (FURIN—Ctrl)). We categorized all genes by the direction of their change in both models and their log₂(fold change) in the synergistic model (Supplementary Fig. 4e and Supplementary Dataset 3). log₂(fold change) s.d. were calculated for all samples and never exceeded 0.3. Genes were grouped into ‘positive synergy’ if their fold change was greater than 0.3 and ‘negative synergy’ if smaller than −0.3. If the corresponding additive model log₂(fold change) showed the same or no direction, the gene was classified as ‘more’ differentially expressed in the combinatorial perturbation than predicted.

Integration with external datasets and hierarchical clustering to confirm neural cell identity was performed as described previously10. Counts data was log₂-transformed (reads per kilobase of transcript per million mapped reads) followed by principal component analysis and plotting in R. RNA-seq datasets were obtained from GTEx (www.gtexportal.org), CMC (www.synapse.org/CMC), BrainSpan (brainspan.org) and Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) (Supplementary Fig. 3c,d). More details about the RNA-seq analysis can be found in the Supplementary Note.

Oversight. All hiPSC research was conducted under the oversight of the institutional review board and Embryonic Stem Cell Research Overview committees at Icahn School of Medicine at Mount Sinai (ISMMS).

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

Data availability
All source donor hiPSCs have already been deposited at the Rutgers University Cell and DNA Repository (study 160; http://www.nimhstemcells.org/); CRISPR-edited hiPSCs are in the process of being submitted in advance of publication. The RNA-seq data are available at www.synapse.org/#!Synapse:syn20502314. Additionally, we make available the following resource by which neural cell-type-specific expression of any gene can be easily cross-referenced in our hiPSC datasets as well as case–control postmortem, hiPSC NPC and hiPSC-neuron RNA-seq datasets (https://schroden.shinyapps.io/BrennandLab-ExpressionApp-limited/). Owing to constraints reflecting the original consents, which are restricted to the study of neuropsychiatric disease only, the raw RNA-seq data will be made available by the authors upon reasonable request and institutional review board approval.

Code availability
Code is available at https://github.com/nadschro/SZvariant-synergy.

References
67. He, X. et al. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. Am. J. Hum. Genet. 92, 667–680 (2013).
68. Benner, C. et al. FINEMAP: efficient variable selection using summary data from genome-wide association studies. Bioinformatics 32, 1493–1501 (2016).
69. Brennand, K. J. et al. Modelling schizophrenia using human induced pluripotent stem cells. Nature 473, 221–225 (2011).
70. Szklarczyk, D. et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 43, D447–D452 (2015).
71. Hoffman, G. E. & Schadt, E. E. variancePartition: interpreting drivers of variation in complex gene expression studies. BMC Bioinformatics 17, 483 (2016).
72. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. Bioinformatics 27, 1739–1740 (2011).
73. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47 (2015).
74. Wu, D. & Smyth, G. K. Camera: a competitive gene set test accounting for inter-gene correlation. Nucleic Acids Res. 40, e133 (2012).
Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

| Data collection | Evolve (PamGene) |
|-----------------|------------------|
|                 | AxIS 2.4         |
|                 | pClamp 10 (Molecular Devices) |
|                 | MiniAnalysis (Synaptosoft) |
|                 | Signal 4 (Cambridge Electronic Design) |

| Data analysis | Whole transcriptome RNAseq data was processed using publicly available code from Hoffman et. al 2017 Nature Communications and other published R packages. Code will be made available on github.com/nadschro/SZvariant-synergy.
|---------------|------------------------------------------------|
|               | R version 3.5.0 (2018-04-23) |
|               | R packages: |
|               | WGCNA_1.67 |
|               | synapseClient_1.18-0 |
|               | gridExtra_2.3 |
|               | WebGestaltR_0.4.0 |
|               | biomaRt_2.38.0 |
|               | gProfileR_0.6.7 |
|               | rlist_0.4.6.1 |
|               | VennDiagram_1.6.20 |
|               | futile.logger_1.4.3 |
|               | ggpubr_0.2 |
|               | magrittr_1.5 |
|               | lme4_1.1-21 |
|               | Matrix_1.2-17 |
|               | doParallel_1.0.14 |
|               | iterators_1.0.10 |
For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All source donor hiPSCs have already been deposited at the Rutgers University Cell and DNA Repository (study 160; http://www.nimhstemcells.org/); CRISPR-edited hiPSCs are in the process of being submitted in advance of publication. Code is available at github.com/nadschro/SZvariant-synergy. RNA-Seq data are available at www.synapse.org/#!Synapse:syn20502314; additionally, we make available the following resource by which neural cell-type-specific expression of any gene can be easily cross-referenced in our hiPSC datasets as well as case/control post-mortem and hiPSC-NPC & hiPSC-neuron RNA-Seq datasets (https://schroden.shinyapps.io/BrennandLab-ExpressionApp-limited/). Owing to constraints reflecting the original consents, which are restricted to the study of neuropsychiatric disease only, the raw RNA-Seq data will be made available by the authors upon reasonable request and IRB approval.

Field-specific reporting

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

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

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

Life sciences study design

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

Sample size
Sample size for RNA-seq was > 30, which allows an adequate observation to take the benefits of the Central limit Theorem, i.e. normally distributed data.

Data exclusions
No data were excluded.
Replication

hiPSCs, hiPSC-NPCs and hiPSC-neurons underwent multiple differentiations and multiple wells for each phenotypic assay and RT-qPCR. Wholetranscriptome RNA-seq was performed on 2 replicates per donor and cell type. Four independent clones were analyzed following CRISPR editing.

All attempts at replication were successful.

Randomization

All donor samples were allocated into both, control and experimental groups and analyzed in parallel and blinded to control for covariates.

Blinding

All phenotypic assays were carried out blinded.

---

### Reporting for specific materials, systems and methods

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

| Materials & experimental systems | Methods |
|----------------------------------|---------|
| n/a                              | n/a     |
| - [ ] Antibodies                 | - [ ] Involved in the study |
| - [ ] Eukaryotic cell lines      | - [ ] ChIP-seq |
| - [X] Palaeontology              | - [ ] Flow cytometry |
| - [ ] Animals and other organisms| - [ ] MRI-based neuroimaging |
| - [ ] Human research participants|         |
| - [ ] Clinical data              |         |

### Antibodies

**Antibodies used**

- Rabbit-anti-GFP, Abcam ab6556, 1:500;
- Mouse-anti-MAP2, Thermo Fisher Scientific MAG-12826, 1:500;
- Rabbit-anti-SYNAPTOPHYSIN1, Synaptic Systems 101 002, 1:500;
- Chicken-anti-SYNAPSIN1/2, Synaptic Systems 106 006, 1:500;
- Chicken-anti-CTIP2, Abcam, ab18465, 1:500;
- Mouse-anti-SATB2, Abcam, satba4b10, 1:5;
- Rabbit-anti-GABA, Sigma A2052, 1:500;
- Mouse-anti-GAD67, Millipore, MAB5406, 1:1000;
- Alexa 488 Donkey-anti-Rabbit, Jackson immunoResearch 711-545-152, 1:500;
- Alexa 568 Goat-anti-Chicken, Thermo Fisher Scientific A-11041, 1:500;
- Alexa 647 Donkey-anti-mouse, Jackson immunoResearch 715-605-150, 1:500;

**Validation**

All antibodies were previously validated (in Brennand et al 2011, Brennand et al 2015, and Ho et al 2017) for immunocytochemistry in human cells.

### Eukaryotic cell lines

**Policy information about cell lines**

**Cell line source(s)**

- Validated control hiPSCs for CRISPR-editing and CRISPRa/i were selected from a previously reported case/control hiPSC cohort of childhood onset SZ (COS) (Hoffman et al, Nat Comm 2017). The following controls were used for CRISPR-editing (hiPSCs: NSB3182-1 (female, for deletion), NSB3182-3 (female, for rs4702 G knock-in)) and CRISPRa/i (hiPSC NPCs NSB553-51-1 (male), NSB2607-1-4 (male), NSB690-2-1 (male)).

- Commercial cell lines:
  - Human astrocytes for co-culture: ScienCell, #1800
  - HEK293T cells for virus generation: Verma Lab (https://jvi.asm.org/content/73/1/576)
  - MEFs for hiPSC feeder culture: Millipore #PMEF-CF

**Authentication**

Source fibroblasts were re-genotyped using PsychChip and exome sequencing. hiPSCs were reprogrammed via sendai viral reprogramming (Life Technologies) and validated by karyotyping, gene expression and protein immunocytochemistry. All other cell lines were purchased and not re-validated.

**Mycoplasma contamination**

Cells were tested for mycoplasma monthly with all test being negative.

**Commonly misidentified lines**

(See ICLAC register)

- none
### Human research participants

Policy information about studies involving human research participants.

| Population characteristics | No subjects were consented or recruited by the stem cell investigators. Only de-identified fibroblast samples, with limited clinical information (noted in Table 1) were provided by the team at NIMH. |
|----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Recruitment                | Subjects were recruited as part of a longitudinal study of childhood onset schizophrenia, since closed, conducted at the NIMH and led by Judith Rapoport.                                                                 |
| Ethics oversight           | Informed consent was obtained from all fibroblast donors at the National Institute of Mental Health, under the review of the Internal Review Board of the NIMH. All hiPSC work was reviewed by the Internal Review Board of the Icahn School of Medicine at Mount Sinai. This work was also reviewed by the Embryonic Stem Cell Research Oversight Committee at the Icahn School of Medicine at Mount Sinai. |

Note that full information on the approval of the study protocol must also be provided in the manuscript.