Transcriptomic immaturity inducible by neural hyperexcitation is shared by multiple neuropsychiatric disorders

村野　友幸

博士（医学）

総合研究大学院大学
生命科学研究科
生理科学専攻

平成３０（２０１８）年度

Transcriptomic immaturity inducible by neural hyperexcitation is shared by multiple neuropsychiatric disorders

Tomoyuki, Murano
SOKENDAI (The Graduate University for Advanced Studies)
School of Life Science
Department of Physiological Sciences

Table of Contents

Introduction

Neuropsychiatric disorders—such as schizophrenia, bipolar disorder, major depressive disorder, and autism spectrum disorder—are common, with over a third of the population in most countries being diagnosed with at least one such disorder at some point in their life[1]. Almost all neuropsychiatric disorders are currently classified mainly on the basis of clinical signs and symptoms. However, there is evidence that patients with different clinical diagnoses share similar biological features, such as genetic mutations, molecular expression, and brain activity[2–6]. Recently, psychiatry has undergone a tectonic shift to incorporate the concepts of modern biology. There have been recent attempts to reclassify psychiatric disorders according to biological domains (e.g., genes, neural circuits, behavior), such as through the Research Domain Criteria (RDoC) initiative[7]. Therefore, identifying appropriate biomarkers that can be used for transdiagnostic assessment of neuropsychiatric disorders is essential for improving the classification of these diseases and understanding their biological basis.

Using coexpression network analysis, a recent study revealed that cross-disorder gene expression overlaps could be used to characterize five major neuropsychiatric disorders[8]. Some of these overlapping gene groups were well characterized biologically by Gene Ontology enrichment or cell-type specificity, but the biological properties of other gene groups were rather unclear. Thus, nonbiased coexpression network analyses do not necessarily detect modules that extract the biological features of neuropsychiatric disorders. Thus, in order to improve the characterization of neuropsychiatric disorders, it might be helpful to detect modules of coexpressed genes and conduct gene expression analysis based on the findings derived from studies on animal models of neuropsychiatric disorders.

To date, our group have screened more than 180 strains of genetically engineered mice using a large-scale, and comprehensive battery of behavioral tests, and we have identified several strains with abnormal behaviors related to neuropsychiatric disorders such as schizophrenia, bipolar disorder, and intellectual disability[9]. We discovered common endophenotypes in the brains of multiple strains of these genetically engineered mice with behavioral abnormalities. We termed one such endophenotype in the hippocampus of adult mice the immature dentate gyrus (iDG) phenotype[10–13]. In this phenotype, the molecular and electrophysiological properties of adult DG neurons in the genetically engineered mice were similar to those of immature DG neurons in typically developing infants. For example, the expression of calbindin, a marker of maturity in DG neurons, was decreased, and the expression of calretinin, a marker of immaturity, was increased[10–15]. Molecular changes similar to some of those found in mice with iDG have been observed in the postmortem brains of patients with

schizophrenia[16], bipolar disorder[16], and epilepsy[17–19]. Furthermore, there is growing evidence that changes in molecular markers of pseudoimmaturity are also present in other brain areas of patients with schizophrenia[20-28], bipolar disorder[26], autism[26], and alcoholism[29]. Therefore, we proposed that pseudoimmaturity of the brain could potentially be a useful transdiagnostic biomarker[9].

Pseudoimmaturity of the brain can be induced in adulthood. Previously, we found that chronic fluoxetine treatment reversed the maturation status of DG neurons in adult wild-type mice, a phenomenon that we termed dematuration[30,31]. Likewise, recent studies suggest that several maturation-related genes and electrophysiological properties in the DG of wild-type adult mice assume an immature-like status after treatment with pilocarpine or electroconvulsive stimulation[16,32]. As mentioned above, an iDG-like phenotype has been found in patients with epilepsy[17–19]. Therefore, we hypothesized that the neural hyperexcitation may induce pseudoimmaturity of the brain in adulthood; however, this hypothesis has not been tested in human samples.

Some studies suggest that hyperexcitation of neurons may underlie abnormalities related to certain types of neuropsychiatric disorders. Individuals with epilepsy are at increased risk of developing schizophrenia, and vice versa[33,34]; additionally, patients with epilepsy can display psychotic symptoms that resemble those found in patients with schizophrenia[35]. Excitatory/inhibitory imbalances have been proposed to be involved in the pathogenesis and pathophysiology of schizophrenia[36–39]. Hyperactive action-potential firing has also been observed in hippocampal granule-cell-like neurons derived from induced pluripotent stem cells (iPSCs) of patients with bipolar disorder[40]. Recent studies have suggested that human patients with Alzheimer's disease and temporal lobe epilepsy may harbor common underlying disease mechanisms[17,41–43]. Considering these findings, I hypothesized that the immature-like gene expression patterns induced by neural hyperexcitation may overlap with the abnormal gene expression patterns in the brains of patients with neuropsychiatric disorders and the related animal models. If this is the case, I hypothesized that this overlap can be used to perform transdiagnostic characterization of neuropsychiatric disorders.

To test this hypothesis, I first performed a meta-analysis of microarray datasets, comparing the changes in gene expression in the rat DG after seizure induction with the differences in gene expression in infant mice versus adult mice. To assess consistency across species, I also conducted a similar comparison using the human fetal hippocampus. The overlap between gene sets was estimated using the Running Fisher test[44], which is a nonparametric statistical method. The fold change and direction of change (up- or downregulation) for each gene between the two conditions were used to define the ranked gene signatures. For

comparison across different arrays, orthologs were used for each organisms. To determine the similarities between two datasets, I used a combination of rank-based enrichment statistics and ontology-based meta-analysis. This method enables us to statistically assess the pairwise correlations between any two datasets, including datasets from different species and organs[28,29,45,46]. The gene expression patterns in the rat DG after seizure induction significantly overlapped with those specific to the immature mouse DG and with those specific to the early-stage human fetal hippocampus. From the set of overlapping genes, I defined two groups: maturity-marker genes and immaturity-marker genes that are inducible by neural hyperexcitation. I assessed the expression patterns of these two groups of maturation-related genes in 87 public gene expression datasets derived from the postmortem brains of patients with various neuropsychiatric disorders and from neural cells derived from patient iPSCs. I further analyzed 12 datasets from the brains of related animal models. Through this analysis, I characterized the expression patterns of maturation-related genes that are inducible by neural hyperexcitation across different disease conditions and animal models of these diseases.

Results

Hyperexcitation induces immature-like gene expression

To examine the developmental changes in gene expression patterns in the rodent DG, I created a microarray dataset from postnatal days 8, 11, 14, 17, 21, 25, and 29 (GSE113727) and compared it with a dataset from 33-week-old adult mice (GSE42778)[12]. Within the entire mouse DG dataset, the largest overlap for changes in gene expression after pilocarpine injection was for the comparison between day 8 infant and 33-week-old adult mice (Supplementary Figure 1a). I included the dataset from postnatal day 8 infant mice for subsequent analysis. The expression levels of 6552 genes were increased in the DG of infant mice compared with adult mice, whereas the expression levels of 8637 genes were decreased (absolute fold change > 1.2 and t-test $P < 0.05$). Next, I assessed the changes in gene expression induced by neural hyperexcitation in a rodent model. I obtained publicly available microarray datasets from the DG of adult rats after seizures induced by injection of pilocarpine (GSE47752)[47]. The expression levels of 7073 genes were significantly changed in the DG of epileptic-seizure rats 1 day after pilocarpine injection compared with rats treated with saline (absolute fold change > 1.2, $P < 0.05$).

To investigate whether the neuronal hyperexcitation datasets contain immature-like gene expression patterns, I assessed the overlap between the set of genes with altered expression in immature mice and the set of genes with altered expression in adult seizure-model rats using the Running Fisher algorithm on the BaseSpace platform to determine the significance of the overlap (see Supplementary Methods for details). I found a striking degree of similarity: 2807 genes showed changes in expression in both datasets (overlap $P = 3.8 \times 10^{-11}$) (Figure 1a). Among these 2807 genes, I named the 726 genes whose expression levels decreased in both datasets hyperexcitation-induced maturity-related genes (hiM genes (mouse): green bar in Figure 1a) and the 938 genes whose expression levels increased in both datasets hyperexcitation-induced immaturity-related genes (hiI genes (mouse): red bar in Figure 1a). Comprehensive lists of hiM and hiI genes are in Supplementary Data 1. The overlap for genes with positively correlated expression (red and green bars) was larger than the overlap for genes with negatively correlated expression (light and dark yellow bars), indicating that the directions of expressional change in the two datasets are more often similar than they are different. These results suggest that neuronal hyperexcitation induces a pattern of immature-like gene expression in the adult DG.

Next, I compared the changes in expression during development in the human fetal hippocampus with those in rats after seizure induction to assess consistency across species. I

obtained publicly available microarray datasets for the human fetal hippocampus during development (GSE25219)[48]. Within the entire fetal hippocampal dataset, the largest overlap for changes in gene expression after pilocarpine injection was for the comparison between 8- to 9-week fetuses and 19- to 23-week fetuses (Supplementary Figure 1b). Again, I found a striking degree of similarity: 2043 genes showed changes in expression in both datasets (overlap $P = 1.8 \times 10^{-12}$) (Figure 1b). Among these 2043 genes, I termed the 579 genes whose expression decreased in both datasets hiM genes (human) (green bar in Figure 1b) and the 716 genes whose expression increased in both datasets hiI genes (human) (red bar in Figure 1b). The overlaps for genes with positively correlated expression (red and green bars) were larger than the overlaps for genes with negatively correlated expression (light and dark yellow bars), suggesting that, similar to the results in mice, the gene expression changes in the rat DG after seizure induction are comparable to the reverse of the changes that occur as the human hippocampus develops.

hiM/hiI genes exhibit different biological properties

To characterize the biological features associated with the hiM and hiI gene groups in mice and humans, I conducted pathway enrichment analyses in BaseSpace. The 20 biogroups that had the largest overlap with the hiM and hiI genes are listed in Tables 1. Among mouse hiM genes, 4 out of the top 20 biogroups are associated with synapse and channel activity (e.g., transmission of nerve impulse, synapse, and synaptic transmission) (Table 1); among human hiM genes, 6 out of the top 20 biogroups are also associated with synapse and channel activity (e.g., transmission of nerve impulse, synaptic transmission, axon, and synapse) (Table 1). Among the mouse hiI genes, 4 out of the top 20 biogroups were associated with the nucleus (e.g., genes involved in the cell cycle and genes involved in DNA replication) (Table 2). Among the human hiI genes, 15 out of the top 20 biogroups were associated with the nucleus (e.g., genes involved in the cell cycle, chromosomes, and response to DNA damage stimulus) (Table 2). Notably, there is little overlap in the top 20 biogroups for the hiM and hiI genes (Table 2). Thus, the biogroups related to the hiM and hiI genes are likely to be functionally different. I further characterized the biological significances of hiM and hiI genes by comparing them with coexpression gene modules obtained from mouse DG development datasets and conducting pathway analyses of those modules. For the detail of the results, see Supplementary Figure 4.

I also compared datasets from the DG of typically developing infants with datasets from the rat DG at three different timepoints after seizure induction by injection of pilocarpine or kainate (day 1, day 3, and day 10) and performed principal component analysis on the

changes in mouse hiM and hiI genes at different timepoints (Supplementary Figure 2a, 2b; Supplementary Notes). The time-course of changes in the mouse hiM genes after seizure induction was different from the time-course of changes in the mouse hiI genes. In addition, I conducted a spatial pattern analysis of the mouse hiM and hiI genes, which indicated that their protein products have slightly different patterns of subcellular localization (Supplementary Figure 2c; Supplementary Notes). The mouse hiM genes tend to be strongly expressed at the plasma membrane, with expression changes stabilizing by the third day after seizure induction. In contrast, the hiI genes tend to be expressed in the nucleus, and changes in expression after seizure induction are slower to stabilize (Supplementary Figure 2b). Together, these results indicate that the hiM and hiI genes have different spatiotemporal patterns of changes in expression.

Gene expression analyses of patient samples by hiM/hiI genes

Next, I investigated whether and to what extent the expression changes in maturation-related genes induced by hyperexcitation overlap with gene expression patterns in various neuropsychiatric disorders. As above, I evaluated similarities between the changes in gene expression patterns in different groups using overlap *P*-values calculated by the Running Fisher algorithm (Figure 2a). Similarity indexes for each comparison were defined as the −log of the overlap *P*-values with hiM or hiI genes, denoted by hiM-index or hiI-index, respectively. High values of the hiM-/hiI-index indicate that there is a large overlap between the dataset analyzed and hiM/hiI genes. I obtained the hiM-/hiI-indexes for the datasets from human patients and plotted them in two-dimensional (2-D) space to show the extent of overlap between datasets and hiM/hiI genes (Figure 2a).

I initially performed this 2-D analysis on a dataset containing the expression profile of the prefrontal cortex in the postmortem brains of patients with schizophrenia (schizophrenia dataset #1: details in Supplementary Data 2) (Figure 2b). The expression of 1744 genes differed between patients and healthy controls (significance level of 0.05). The numbers of hiM and hiI genes with altered expression in schizophrenia dataset #1 were 87 and 76, respectively, and the overlap *P*-values were 2.4×10^{-10} and 6.9×10^{-7}, respectively. The hiM-index and hiI-index for this dataset were 9.62 ($= -\log(2.4\times10^{-10})$) and 6.16 ($= -\log(6.9\times10^{-7})$). This result corresponds to a point in 2-D space (Figure 2b). Points that fall below the unity line (dashed line) indicate datasets in which changes in hiM genes are dominant, whereas points above the unity line indicate datasets with dominant changes in hiI genes. The angle from the unity line indicates the degree of hiM or hiI dominance. The same analysis was performed for other schizophrenia datasets (schizophrenia datasets #2–#16),

including those obtained from different areas of the postmortem brain and from cultured neurons derived from the iPSCs of patients. Scatter plots of the results from the schizophrenia datasets are shown in Figure 3a. Thirteen out of sixteen points were below the unity line. Most of the schizophrenia datasets exhibited hiM-index-dominant patterns, showing high hiM-index values and low hiI-index values (Figure 3a).

I extended the same analysis to 87 disease datasets from seven other neuropsychiatric diseases (amyotrophic lateral sclerosis (ALS), Alzheimer's disease (AD), autism spectrum disorder (ASD), Parkinson's disease (PD), bipolar disorder (BPD), Huntington's disease (HD), and major depressive disorder (MDD); Supplementary Data 2). The results from each dataset are shown in Figure 3b-3h. The overall distribution patterns of each disease are shown in Figure 3i. The ALS datasets tended to show higher hiI-index values than hiM-index values, indicating an hiI-index-dominant pattern (Figure 3b). The AD datasets showed different patterns in the hiM-/hiI-index depending on the type of sample; datasets from the postmortem brains of patients with AD tended to show high values only for the hiM-index, and datasets from patient iPSCs tended to show high values only for the hiI-index (Figure 3c). Datasets from ASD did not show any dominant patterns in either the hiM-index or the hiI-index (Figure 3d). Most datasets from patients with PD, BPD, HD, and MDD did not show pronounced values for the hiM-index or the hiI-index (Figure 3e, 3f, 3g, 3h). I applied same analyses on the microarray expression datasets from a recent report by Gandal and colleagues, in which confounding factors are controlled relatively well. Datasets from postmortem brains of patients with schizophrenia and alcohol abuse disorder showed high values for the hiM-index, and those from patients with ASD showed high values for the hiI-index (Supplementary Figure 3).

Thus, the 2-D analysis revealed that some neuropsychiatric diseases have characteristic patterns in the hiM-/hiI-indexes; for example, most datasets from patients with schizophrenia exhibited a higher hiM-index than hiI-index, whereas ALS datasets showed an hiI-index-dominant pattern (Figure 3i). Meanwhile, different diseases sometimes shared similar changes in the hiM- or hiI-index; for example, some of the schizophrenia, ASD, and AD datasets shared a high hiM-index, and some of the ALS and AD datasets shared a high hiI-index. The other four diseases—PD, BPD, HD, and MDD—did not feature pronounced changes in the hiM-/hiI-indexes, suggesting that these diseases may not share an endophenotype of pseudoimmaturity inducible by neural hyperexcitation. These results raise the possibility that there are patterns of gene expression perturbations that are shared and distinct across these neuropsychiatric disorders.

Genetic and environmental risks induce pseudoimmaturity

Previous studies suggest that many genetic risk factors and environmental factors, such as seizure, hypoxia, and infection, contribute to the development of neuropsychiatric disorders[2,49,50]. I next applied the 2-D analysis technique to datasets from genetic animal models of disorders and from animals that had experienced risk events.

First, I obtained publicly available datasets for mice that had experienced putative risk events for schizophrenia, bipolar disorders, and Alzheimer's disease, including seizure (#1: GSE49030, #2: GSE4236)[51,52], ischemia (#1: GSE32529, #2: GSE35338)[53,54], and infection (mimicked by CpG; GSE32529)[53,54]. All the studies used here included datasets for different time points after the risk event; hence, I was able to examine the time-course of changes in the hiM- and hiI-indexes to reveal the short- and long-term effects of risk events on the expression patterns of hiM/hiI genes. The results showed that datasets from the hippocampus of mice treated with kainite, a seizure-inducing drug, exhibited time-course changes in the hiM- and hiI-indexes; the hiM-index tended to be dominant in the early stage after seizure induction, and the hiI-index became more dominant in the late stages (Figure 4a: seizure #1). The results from other datasets on seizure, ischemia, and infection showed time-course pattern changes in the hiM- and hiI-indexes that roughly matched those observed in seizure dataset #1, being relatively hiM-index-dominant in the early stage and then relatively hiI-index-dominant in the later stage (Figure 4a: seizure, ischemia, and CpG infection). These results indicate that different types of putative risk events for neuropsychiatric disorders induce roughly similar time-course changes in the expression of maturation-related genes induced by neural hyperexcitation.

Next, I obtained datasets from animal models with a genetic risk of a neurodegenerative disease: mice with transgenic expression of a G93A mutant form of human SOD1, as a model of ALS (#1: GSE46298, #2: GSE18597)[55,56]; transgenic mice with mutant human amyloid precursor protein (APP) and presenilin1 (PSEN1) genes, which cause familial Alzheimer's disease (#1: GSE64398, #2: GSE64398)[57]; and Df16(A) heterozygous mice carrying a chromosome 16 deletion syntenic to human 22q11.2 microdeletions, as a model of schizophrenia (GSE29767)[58]. I also acquired datasets from Schnurri-2 (Shn-2) knockout mice as a model of schizophrenia[12] and intellectual disability[14,59,60] and from mice with heterozygous knockout of the α-isoform of calcium/calmodulin-dependent protein kinase II (α-CaMKII$^{+/-}$)[10,13] as a model of bipolar disorder[74]. These mice show an array of behavioral abnormalities, including locomotor hyperactivity and severe deficits in working memory. The expression profiles of several maturation marker genes in exhibit immature-like patterns in these mice[10,12]. The threshold current of granule cells to induce spikes was low in these mice,

indicating that the granule cells of these mice are highly excitable[10,12]. I performed 2-D analysis on these datasets and evaluated the changes in the hiM-/hiI-indexes of these model mice. Both datasets from transgenic mice with the SOD1(G93A) mutation exhibited an hiI-index higher than the hiM-index in the later stages of disease progression (Figure 4b). These hiI-index-dominant patterns were also observed in the results derived from human patients with ALS (Figure 3b). In the mice with mutant human APP and PSEN1, both the hiM- and hiI-indexes increased in the dataset from the hippocampus, and only the hiI-index increased in the dataset from the cortex during the course of disease progression (Figure 4c). These patterns are neutral or hiI-index dominant and partially mimic the results from human patients with Alzheimer's disease (Figure 3c). The Df16(A) heterozygous mice and α-CaMKII$^{+/-}$ mice showed hiM-index-dominant patterns, similar to results from human patients with schizophrenia (Figure 4d). Shn-2 KO mice showed high values for both the hiM- and hiI-indexes (Figure 4d). Thus, the results from the 2-D analysis of animal models are to some extent consistent with the results from human patients, indicating that these model mice have patterns of pseudoimmaturity induced by neural hyperexcitation that are similar to those in human patients. Potential confounding factors, such as types of samples, brain areas, and age, were less numerous in the animal-model analyses than in the analyses of samples from human patients, establishing the causal links between experimental manipulations corresponding to disease conditions and pseudoimmaturity inducible by hyperexcitation.

Discussion

In this study, I demonstrated that neural hyperexcitation induces changes in the pattern of gene expression in the DG that are similar to the patterns in the immature hippocampus of typically developing human fetuses. From the pool of genes, I identified two groups of genes, and found that these are shared by multiple neuropsychiatric disorders, such as schizophrenia, Alzheimer disorders, and ALS.

Many of the datasets from patients with schizophrenia and from the postmortem brains of patients with Alzheimer's disease exhibited hiM-index-dominant pattern changes. The hiM genes include genes encoding a GABA receptor, voltage-dependent calcium channel, glutamate receptor, and voltage-dependent sodium channel (Supplementary Data 1). These genes have been implicated in the pathological changes in the brains of patients with schizophrenia and Alzheimer's disease[61–64]. Thus, many of the synaptic genes that changed in the brains of patients with schizophrenia or Alzheimer's disease could be genes whose expression increases during maturation and decreases with neural hyperexcitation. Although reductions in the expression of some synaptic genes in these disorders are well documented, to our knowledge, my results are the first to raise the possibility that neuronal hyperexcitation may also induce reductions in such synaptic molecules. Most of the datasets from patients with ALS and Alzheimer's disease exhibited hiI-index-dominant patterns. The hiI genes include genes encoding DNA methyltransferase, cyclin D1, cyclin-dependent kinase 1, integrin beta 3 binding protein, and tumor protein p53 (Supplementary Data 1). These genes are known to be important in chromosomal modification and DNA repair, and abnormal functioning of these systems has been observed in patients with ALS and Alzheimer's disease[65–69]. Thus, some of the genes that are considered to be important in the development of these disorders are immaturity-related genes, whose expression decreases during maturation and can be increased by neural hyperexcitation. As for the datasets from patients with PD, BPD, HD, and MDD, most of them did not overlap with either hiM or hiI genes, suggesting that there might not be major pathological changes in transcriptomic pseudoimmaturity inducible by neural hyperexcitation in the datasets of these four diseases. Thus, I was able to characterize the gene expression patterns in disease datasets of each disease category using the hiM- and hiI-indexes.

My study has some limitations. First, the number of available datasets was limited. All the datasets except the one for mouse development were obtained from the BaseSpace Correlation Engine. On this platform, vast numbers (over 21,000) of complex biological and clinical datasets are available. Although I used all the gene expression dataset hits from my keyword query to avoid sampling bias, the number of datasets was still small, from 8 datasets

for ALS to 16 datasets for schizophrenia. Further accumulation of the studies will improve the reliability of my results. Another limitation is that the datasets used in this study are from different types of samples, including various areas within the central nervous system, such as the hippocampus, prefrontal cortex, striatum, and spinal cord. The gene expression abnormalities in patients could differ depending on the brain area[48]. I also used datasets from cultured neurons differentiated from the iPSCs of human subjects, and it is controversial whether the pattern of gene expression in these neurons is comparable to that of neurons in the patients' brains[70,71]. It is also possible that the altered gene expression in the postmortem brains is due to the effects of medication rather than pathological changes from the disease itself[72]. I compared the expression patterns of the genes derived from rodents and humans. There could be critical differences among different species, which might have caused the analysis to miss important molecular regulatory elements shared among those neuropsychiatric disorders. Other conditions that were not controlled in this study and can be potential confounding factors include the age of subjects at death, the storage conditions of the samples, the genetic background of animals, and the animal housing conditions. For these reasons, I need to be careful in interpreting the results of the analyses. It should be noted, however, that despite the variety of sample types used, I was able to identify some shared and distinct patterns of gene expression, which is mostly due to the advantages of utilizing the Running Fisher test; this test, using a rank-based nonparametric algorithm, can evaluate similarity between datasets from different species or organs[44]. Moreover, I applied 2-D analysis to the animal models, in which potential confounding factors such as species, sample types, and brain areas are well controlled. The results from these models were comparable to those from human samples, supporting the idea that pseudoimmaturity inducible by hyperexcitation is a feature shared by multiple neuropsychiatric disorders. It should be noted, however, that a gene expression pattern resembling those induced by neural hyperexcitation may not be induced by neural hyperexcitation *per se* but by other factors such as inflammation, hypoxia, and infection. Another potential advantage of using the Running Fisher test is that this method is applicable to the dataset of gene expression profiles from an individual patient or control. In the future, it might be interesting to apply this method to the characterization and classification of gene expression patterns in the brains of individual patients and in neurons derived from patients' iPSCs.

Recent attempts such as the RDoC initiative have tried to reclassify psychiatric disorders according to biological domains (e.g., genes, neural circuits, behavior)[7]. While Gandal *et al*. conducted nonbiased coexpression analyses[8], this study utilized gene groups derived from the findings based on the studies of animal models of neuropsychiatric disorders.

Characterization by these gene groups enabled us to extract novel biological features of some neuropsychiatric disorders that are related to pseudoimmaturity inducible by neural hyperexcitation. Detecting such domains that extract the biological features of each neuropsychiatric disorder will move this diagnostic framework forward, from criteria based on signs and symptoms to those including biological dimensions.

In conclusion, the biological domain of pseudoimmaturity inducible by neural hyperexcitation is a common endophenotype among several neuropsychiatric disorders. Future studies are needed to find translational indices that correspond to these features and can be applicable to human patients for better diagnosis of these neuropsychiatric disorders. My findings here may promote the development of biomarkers, leading to a better diagnosis of neuropsychiatric disorders.

Methods

Microarray experiments to examine mouse DG development

Wild-type mouse DGs were sampled at postnatal days 8, 11, 14, 17, 21, 25, 29 (C57BL/6J × BALB/cA background; male, n = 5)[73], and microarray experiments were performed with Mouse Genome 430 2.0 Array (Affymetrix, Santa Clara, CA) as previously described[10]. RNA was isolated by using the TRIzol method (Invitrogen, Carlsbad, CA) from the hippocampus of mice, followed by purification, using RNeasy columns (Qiagen, Valencia, CA). Double-stranded cDNA was synthesized from the total RNA, and invitro transcription reaction was then performed on biotin-labeled RNA that was made from the cDNA. Labeled RNA was hybridized with Mouse Genome 430 2.0 Array (Affymetrix, Santa Clara, CA) containing 45101 probe sets, and washed according to the manufacturer's recommendations. The hybridized probe array was then stained with streptavidin-conjugated phycoerythrin, and each GeneChip was scanned by an Affymetrix GeneChip Scanner 3000 (GCS3000). GeneChip analysis was performed with Microarray Analysis Suite version 5.0. All of the gene represented on the GeneChip were globally normalized. Genes with a P-value < 0.05 (without correction for multiple testing) and an absolute fold change > 1.2 were included in the differentially expressed gene datasets. The microarray data, including those used in this study, were deposited in the Gene Expression Omnibus (GEO) database under accession number GSE113727. I also obtained a dataset for 33-week-old wild-type mice, which our group previously reported (C57BL/6J × BALB/cA background) (GSE42778)[12]. I integrated these two datasets into one to construct the dataset for the development of the wild-type mouse DG used in this study (P8 versus adult, fold change > 1.2, P < 0.05). All animal experiments were approved by the Institutional Animal Care and Use Committee of Fujita Health University, based on the Law for the Humane Treatment and Management of Animals and the Standards Relating to the Care and Management of Laboratory Animals and Relief of Pain. Every effort was made to minimize the number of animals used.

Data collection and processing

Except for the mouse DG developmental dataset mentioned above, the 99 gene expression datasets used in this study were obtained from publicly available databases (listed in Supplementary Data 2). I used all the hits from my keyword query for gene expression studies to avoid sampling bias. If a single study included several different datasets, I chose the one that I considered the most comprehensive (e.g., if the study includes datasets from male, female, and all subjects, I used datasets from all subjects). All gene expression datasets were

analyzed with the BaseSpace Correlation Engine (formerly known as NextBio) (https://japan.ussc.informatics.illumina.com/c/nextbio.nb; Illumina, Cupertino, CA), a database of biomedical experiments. BaseSpace is a repository of analyzed gene expression datasets that allows researchers to search expression profiles and other results[44]. The datasets registered in BaseSpace undergo several preprocessing, quality control, and organization stages. Quality control ensures the integrity of the samples and datasets and includes evaluations of pre- and postnormalization boxplots, missing value counts, and *P*-value histograms (after statistical testing) with false-discovery rate analysis to establish whether the number of significantly altered genes is larger than that expected by chance.

Genes with a P-value < 0.05 (without correction for multiple testing) and an absolute fold change > 1.2 were included in the differentially expressed gene datasets. This sensitivity threshold is typically the lowest used with commercial microarray platforms and the default criterion in BaseSpace analyses[44]. Correction for multiple testing was omitted to minimize false negatives at this stage. All data from the Affymetrix GeneChip series were downloaded from the NCBI GEO database. Affymetrix Expression Console software (specifically, the robust multiarray average algorithm) was used to preprocess the data.

I used the expression values (on a log base 2 scale) to calculate the fold changes and *P*-values between two conditions (infants–adults and patients–healthy controls). To determine the fold changes, I began with the expression values of the probes/genes in the test data sets and divided them by those of the control data sets. If the fold change was < 1.0, these values were converted to the negative reciprocal, or −1/(fold change). Genes with an absolute fold change > 1.2 and a *t*-test *P*-value < 0.05 were imported into BaseSpace Correlation Engine according to the instructions provided by the manufacturer. The rank order of these genes was determined by their absolute fold change. All statistical analyses were performed in BaseSpace, and the similarities between any two datasets were evaluated as overlap *P*-values using the Running Fisher algorithm[44]. The Bonferroni correction was used to adjust the significance level according to the number of datasets pairs[44].

Data availability statement

The microarray and RNAseq data that support the findings of this study have been available in the BaseSpace Correlation Engine (https://japan.ussc.informatics.illumina.com/c/nextbio.nb; Illumina, Cupertino, CA) and Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/). All detailed information is listed in Supplementary Data 2.

Computing overlap P-values of gene expression patterns in different datasets

BaseSpace can be used to compare the signatures in publicly available microarray datasets with a signature provided by the user using a "Running Fisher" algorithm, as previously described[25,29,44-46]. To enable comparison across different arrays, orthologs were identified for each pair of organisms. Ortholog identification was based on information obtained from Mouse Genome Informatics (MGI) at Jackson Lab (http://www.informatics.jax.org), HomoloGene at NCBI (http://www.ncbi.nlm.nih.gov), and Ensembl (http://www.ensembl.org). The overlap *P*-value, i.e., the direction of the correlation between two given gene signature sets (*b1*, *b2*), and the *P*-values between subsets of gene signatures are calculated as follows.

Each gene signature set was rank-ordered according to the absolute fold-change value. Upregulated and downregulated genes were denoted by positive and negative signs, respectively, to indicate directionality. A directional subset was generated for each direction, such as *b1*+, *b1*-, *b2*+, and *b2*-.

Next, all of the subset pairs were identified as *b1Di*, *b2Dj*, where *Di* and *Dj* were the available directions (+ or −) in *b1* and *b2*, respectively. The Running Fisher algorithm was applied to each subset pair. The top ranking genes in the first subset *b1Di* were collected as a group, *G*, and the second subset *b2Dj* was scanned from top to bottom in rank order to identify each rank with a gene matching a member in group *G*. At each matching rank, *K*, the scanned portion of the second subset *b2Dj,* consisted of *N* genes, and the overlap between group *G* and these *N* genes was defined as *M*. Fisher's exact test was performed at rank *K* to evaluate the statistical significance of observing *M* overlaps between a set of size *G* and a set of size *N*, where the set of size *G* comes from platform *P1,* and the set of size *N* comes from platform *P2*, given the sizes of *P1* and *P2* as well as the overlap between *P1* and *P2*. At the end of the scan, the best *P*-value was retained, and a multiple-hypothesis-testing correction factor was applied. The negative log of the multiple-testing-corrected best *P*-value ($P_{b1Di \rightarrow b2Dj}$) was a score ($S_{b1Di \rightarrow b2Dj}$) for the subset pair. Here, the subscript $b1Di \rightarrow b2Dj$ indicates that *b*1*Di* was the first subset used to define the top genes *G,* and *b*2*Dj* was the second subset that was used for the scan.

$$S_{b1Di \rightarrow b2Dj} = -\ln P_{b1Di \rightarrow b2Dj} \quad (1)$$

Next, the Running Fisher algorithm was performed in the reverse direction. The same procedure in this reverse direction produced another score ($S_{b2Dj \rightarrow b1Di}$) for the same subset pair. The two scores were averaged to represent the magnitude of the similarity between the two subsets.

$$S_{b1Dib2Dj} = \frac{S_{b1Di \rightarrow b2Dj} + S_{b2Dj \rightarrow b1Di}}{2} \quad (2)$$

The *P* value ($P_{b1Dib2Dj}$) between *b1Di* and *b2Dj* was calculated using the following equation:

$$P_{b1Dib2Dj} = exp\ (\ -S_{b1Dib2Dj}\ ) \quad (3)$$

A positive sign was assigned to pairwise correlation scores (S_{b1+b2+} and S_{b1-b2-}) for a subset pair of the same direction (*b1+b2+*, *b1-b2-*), and a negative sign was assigned to pairwise correlation scores (S_{b1+b2-} and S_{b1-b2+}) for a subset pair of opposite directions (*b1+b2-*, *b1-b2+*). Then, the overall score (S_{b1b2}) between *b1* and *b2* was calculated from the correlation scores (S_{b1+b2+}, S_{b1-b2-}, S_{b1+b2-}, and S_{b1-b2+}) of subset pairs using the following equation:

$$S_{b1b2} = \frac{S_{b1+b1+} + S_{b1-b2-}}{2} - \frac{S_{b1+b1-} + S_{b1-b2+}}{2} \quad (4)$$

The sign of S_{b1b2} reflected whether the two signatures were positively or negatively correlated. The overall *P*-value (P_{b1b2}) between *b1* and *b2* was calculated using the following equation:

$$P_{b1b2} = exp\ (\ -|S_{b1b2}|\ ) \quad (5)$$

This overall *P*-value is referred to as the "overlap *P*-value" between two gene expression patterns in this paper.

Prediction of the subcellular of proteins coded by genes

To predict the subcellular localizations of proteins coded by genes in each dataset, COMPARTMENTS (http://compartments.jensenlab.org)[75] was used. COMPARTMENTS is a web resource that integrates evidence on protein subcellular localization from manually curated literature, high-throughput screens, automatic text mining, and sequence-based prediction methods. For each gene queried, it provides a score reflecting the localization to multiple cellular compartments (e.g., the plasma membrane, nucleus, cytosol, and so on) based on aggregating data from prediction algorithms (e.g., PSORT and YLoc).

The WoLF PSORT program[76], a sequence-based protein localization predictor, was used for the prediction of protein subcellular localization in this study. PSORT predicts subcellular localization based on various sequence-derived features such as sorting signals, binding domains, and amino acid composition. All PSORT scores were sorted by the number of stars (Spsort) assigned to a sequence-based prediction, from 0 to 3 (e.g., Rreb1: nucleus = 3, cytosol = 2, and other = 0; Fam107a: nucleus = 2, cytosol = 2, and other = 0). Detailed information about PSORT and Spsort score have been described previously[6,7]. In this study, the score of each gene was determined by the highest Spsort score among the subcellular

compartment(s). If there were two or more compartments sharing the highest score, the score of each gene was determined by dividing the total number by the number of compartments with the highest score (e.g., Rreb1: nucleus = 1; Fam107a: nucleus = 0.5, cytosol = 0.5). The subcellular distribution patterns of hiM/hiI genes are expressed as the proportion of the integrated scores of the top 50 genes.

Principal component analysis

Principal component analysis (PCA) was performed to reveal the relationship between datasets of marker genes. PCA was performed for datasets composed of hiM/hiI genes at three different timepoints: day 1, day 3, and day 10 after treatment. From the comprehensive gene lists, the 500 genes with the largest fold changes were used for PCA. The two primary components in the results of PCA, which are denoted as principal component 1 (PC1) and PC2, correspond to the x-/y-axes on the graphs. For data processing and PCA, I used R for Mac OS X (The R Foundation for Statistical Computing).

Supplementary Notes

To evaluate the time-course of changes in immature-like gene expression patterns after seizure induction, I compared datasets from the DG of typically developing infants with datasets from rat DG at three different timepoints after seizure induction by injection of pilocarpine or kainite (day 1, day 3, and day 10). As shown in Supplementary Figure 2a, the overlap *P*-values for the hiM genes were smaller than those for the hiI genes on the first day after seizure induction in both the pilocarpine and kainate datasets. Interestingly, the overlap *P*-values for the hiI genes were smaller than those for the hiM genes on day 3. These differences in the overlap *P*-values of the hiM/hiI genes with the number of days after seizure induction suggest that the expression levels of maturity-/immaturity-related genes induced by neural hyperexcitation change over time; the expression changes of hiM genes in the DG occur earlier than those of hiI genes after seizure induction. The similarity of the results for pilocarpine and kainite also suggests that the induction of immature-like gene expression is not dependent on the specific seizure-inducing drugs but on the induced neural hyperexcitation itself.

I performed principal component analysis (PCA) to visualize the relationships between hiM/hiI genes at different timepoints after seizure induction. For these analyses, I generated integrated datasets comprised of hiM or hiI genes at three different timepoints after seizure induction by pilocarpine or kainite (pilocarpine hiM gene sets, pilocarpine hiI gene sets, kainite hiM gene sets, and kainite hiI gene sets) and performed PCA for each dataset. Datasets at every timepoint after seizure induction and those of untreated control were plotted in a two-dimensional space with coordinates corresponding to principal components (PC) 1 and 2 (Supplementary Figure 2b). The PC1/PC2 coordinates of the hiM genes are different for day 1, day 3, and the untreated control, but those for day 3 and day 10 are close to each other (Supplementary Figure 2b), indicating that the expression changes in hiM genes after seizure induction become stable by day 3. However, the PC1/PC2 coordinates for hiI genes differed from each other at all timepoints after seizure (Supplementary Figure 2b), indicating that the expression changes in hiI genes continue after day 3 until at least day 10. These results also suggest that the changes in expression of the hiM genes follow a different temporal pattern than the changes in expression of hiI genes.

I conducted a bioinformatic estimation of the subcellular distributions of the protein products of hiM/hiI genes. To predict their subcellular distribution, I used the open-source tool COMPARTMENTS, which provides information on the estimated subcellular localization of genes of interest (see Methods). The subcellular distributions of hiM genes in the plasma membrane/nucleus were 28%/16% (in the pilocarpine dataset) and 26%/15% (kainite), and

those of hiI genes were 16%/20% (pilocarpine) and 16%/23% (kainate) (Supplementary Figure 2c). The proportion of expression in the plasma membrane was higher in the hiM gene group than in the hiI gene group, while that in the nucleus was higher in the hiI gene group than in the hiM gene group. These results indicate that the protein products of hiM/hiI genes tend to have spatially distinct localizations.

I applied my analysis on the microarray expression datasets generated by Gandal and colleagues, in which confounding factors are largely controlled (Supplementary Figure 3). The hiM-index and hiI-index of datasets from patients with schizophrenia were 7.85 and 2.46, respectively, indicating hiM-index-dominant pattern. This result indicates that schizophrenia datasets used in Gandal's study also show transcriptomic pseudoimmaturity inducible by neural hyperexcitation. The hiM-index and hiI-index of datasets from bipolar disorder patients were 2.19 and 1.48, respectively, and those of major depressive disorder were both 0. These indexes are smaller than those of schizophrenia, which is mostly consistent with my results in Figure 3. The hiM-index and hiI-index of autism were -6.74 and 28.66, respectively. This hiI-index-dominant pattern is different from my results, as shown in Figure 3d and 3i. Although the pattern in autism is apparently different from mine, datasets from autism show significant overlap with pseudoimmaturity inducible by hyperexcitation, regarding hiI-index. I have added analysis of the datasets from alcohol abuse disorder and inflammatory bowel disease. Their hiM-/hiI-indexes were 25.85/-5.92 and -7.03/9.31, respectively.

To further evaluate biological significances of hiM and hiI genes, I performed weighted gene coexpression network analysis (WGCNA), which is a common method in systems biology for describing the correlation patterns among genes across microarray samples based on the k-means algorithms. I applied this method to the datasets from mouse DG development (postnatal day 8, 11, 14, 17, 21, 25, and 29 mouse infants and 33-week-old adult mice), and extracted 5 gene modules (turquoise, blue, brown, yellow, and grey) (Supplementary Figure 4a). I compared hiM and hiI gene groups with these 5 gene modules and evaluated their similarities by Running Fisher test (Supplementary Figure 4b). The hiM and hiI genes showed partially different patterns of overlap with these 5 modules of coexpressed genes during DG development. While hiM genes significantly overlap with turquoise, blue, and yellow modules, hiI genes significantly overlap with only turquoise module (Supplementary Figure 4b). Additionally, I performed pathway analyses on these 5 modules in BaseSpace (Supplementary Figure 4c). The turquoise module includes biogroups associated with the nucleus (e.g., "cell division", "Genes involved in Cell Cycle", and "mitosis"), and blue module includes biogroups associated with synapse (e.g., "dendritic spine" and "synapse"). Similar biogroups were found in results of pathway analysis in Table 1 and

Table 2 (for example, biogroups such as “synaptic transmission” and “synapse” were involved in the results of pathway analysis for hiM genes; biogroups such as “Genes involved in Cell Cycle”, “mitosis”, and “cell division” were involved in the results of pathway analysis for hiI genes).

All information on the datasets and gene lists used in this study are provide as Supplementary Data 1 to 12.

References

1. Cross-national comparisons of the prevalences and correlates of mental disorders. WHO International Consortium in Psychiatric Epidemiology. *Bull World Health Organ* **78,** 413–426 (2000).

2. Cardno, A. G. & Owen, M. J. Genetic Relationships Between Schizophrenia, Bipolar Disorder, and Schizoaffective Disorder. *Schizophr Bull* **40,** 504–515 (2014).

3. Hall, J., Trent, S., Thomas, K. L., O'Donovan, M. C. & Owen, M. J. Genetic Risk for Schizophrenia: Convergence on Synaptic Pathways Involved in Plasticity. *Biological Psychiatry* **77,** 52–58 (2015).

4. Forero, D. A. *et al.* A network of synaptic genes associated with schizophrenia and bipolar disorder. *Schizophrenia Research* **172,** 68–74 (2016).

5. Douaud, G. *et al.* A common brain network links development, aging, and vulnerability to disease. *PNAS* **111,** 17648–17653 (2014).

6. Argyelan, M. *et al.* Resting-State fMRI Connectivity Impairment in Schizophrenia and Bipolar Disorder. *Schizophr Bull* **40,** 100–110 (2014).

7. Insel, T. R. & Cuthbert, B. N. Brain disorders? Precisely. *Science* **348,** 499–500 (2015).

8. Gandal, M. J. *et al.* Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science* **359,** 693–697 (2018).

9. Hagihara, H., Takao, K., Walton, N. M., Matsumoto, M. & Miyakawa, T. Immature Dentate Gyrus: An Endophenotype of Neuropsychiatric Disorders. *Neural Plasticity* **2013,** (2013).

10. Yamasaki, N. *et al.* Alpha-CaMKII deficiency causes immature dentate gyrus, a novel candidate endophenotype of psychiatric disorders. *Molecular Brain* **1,** 6 (2008).

11. Ohira, K. *et al.* Synaptosomal-associated protein 25 mutation induces immaturity of the dentate granule cells of adult mice. *Molecular Brain* **6,** 12 (2013).

12. Takao, K. *et al.* Deficiency of Schnurri-2, an MHC Enhancer Binding Protein, Induces Mild Chronic Inflammation in the Brain and Confers Molecular, Neuronal, and Behavioral Phenotypes Related to Schizophrenia. *Neuropsychopharmacology* **38,** 1409–1425 (2013).

13. Hagihara, H. *et al.* Circadian Gene Circuitry Predicts Hyperactive Behavior in a Mood Disorder Mouse Model. *Cell Reports* **14,** 2784–2796 (2016).

14. Nakao, A. *et al.* Immature morphological properties in subcellular-scale structures in the dentate gyrus of Schnurri-2 knockout mice: a model for schizophrenia and intellectual disability. *Molecular Brain* **10,** 60 (2017).

15. Hagihara, H., Fujita, M., Umemori, J., Hashimoto, M. & Miyakawa, T. Immature-like molecular expression patterns in the hippocampus of a mouse model of dementia with Lewy body-linked mutant β-synuclein. *Mol Brain*, **11** 38 (2018).

16. Shin, R. *et al.* The immature dentate gyrus represents a shared phenotype of mouse models of epilepsy and psychiatric disease. *Bipolar Disord* **15,** 405–421 (2013).

17. You, J. C. *et al.* Epigenetic suppression of hippocampal calbindin-D28k by ΔFosB drives seizure-related cognitive deficits. *Nature Medicine* **23,** 1377 (2017).

18. Maglóczky, Z., Halász, P., Vajda, J., Czirják, S. & Freund, T. F. Loss of Calbindin-D28K immunoreactivity from dentate granule cells in human temporal lobe epilepsy. *Neuroscience* **76,** 377–385 (1997).

19. Karádi, K. *et al.* Correlation between calbindin expression in granule cells of the resected hippocampal dentate gyrus and verbal memory in temporal lobe epilepsy. *Epilepsy & Behavior* **25,** 110–119 (2012).

20. Zhang, Z. J. & Reynolds, G. P. A selective decrease in the relative density of parvalbumin-immunoreactive neurons in the hippocampus in schizophrenia. *Schizophrenia Research* **55,** 1–10 (2002).

21. Hashimoto, T. *et al.* Gene Expression Deficits in a Subclass of GABA Neurons in the Prefrontal Cortex of Subjects with Schizophrenia. *J. Neurosci.* **23,** 6315–6326 (2003).

22. Fung, S. J. *et al.* Expression of Interneuron Markers in the Dorsolateral Prefrontal Cortex of the Developing Human and in Schizophrenia. *Am. J. Psychiatry* **167,** 1479–1488 (2010).

23. Pantazopoulos, H., Woo, T.-U. W., Lim, M. P., Lange, N. & Berretta, S. Extracellular Matrix-Glial Abnormalities in the Amygdala and Entorhinal Cortex of Subjects Diagnosed With Schizophrenia. *Arch Gen Psychiatry* **67,** 155–166 (2010).

24. Berretta, S. Extracellular matrix abnormalities in schizophrenia. *Neuropharmacology* **62,** 1584–1597 (2012).

25. Hagihara, H., Ohira, K., Takao, K. & Miyakawa, T. Transcriptomic evidence for immaturity of the prefrontal cortex in patients with schizophrenia. *Molecular Brain* **7,** 41 (2014).

26. Gandal, M. J., Nesbitt, A. M., McCurdy, R. M. & Alter, M. D. Measuring the Maturity of the Fast-Spiking Interneuron Transcriptional Program in Autism,

Schizophrenia, and Bipolar Disorder. *PLOS ONE* **7,** e41215 (2012).

27. Mellios, N. *et al.* Molecular Determinants of Dysregulated GABAergic Gene Expression in the Prefrontal Cortex of Subjects with Schizophrenia. *Biological Psychiatry* **65,** 1006–1014 (2009).

28. Hashimoto, T. *et al.* Alterations in GABA-related transcriptome in the dorsolateral prefrontal cortex of subjects with schizophrenia. *Mol. Psychiatry* **13,** 147–161 (2008).

29. Murano, T., Koshimizu, H., Hagihara, H. & Miyakawa, T. Transcriptomic immaturity of the hippocampus and prefrontal cortex in patients with alcoholism. *Scientific Reports* **7,** srep44531 (2017).

30. Kobayashi, K. *et al.* Reversal of hippocampal neuronal maturation by serotonergic antidepressants. *PNAS* **107,** 8434–8439 (2010).

31. Ohira, K., Takeuchi, R., Iwanaga, T. & Miyakawa, T. Chronic fluoxetine treatment reduces parvalbumin expression and perineuronal netsin gamma-aminobutyric acidergic interneurons of the frontal cortex in adultmice. *Molecular Brain* **6,** 43 (2013).

32. Imoto, Y., Segi-Nishida, E., Suzuki, H. & Kobayashi, K. Rapid and stable changes in maturation-related phenotypes of the adult hippocampal neurons by electroconvulsive treatment. *Molecular Brain* **10,** 8 (2017).

33. Chang, Y.-T. *et al.* Bidirectional relation between schizophrenia and epilepsy: A population-based retrospective cohort study. *Epilepsia* **52,** 2036–2042 (2011).

34. Hyde, T. M. & Weinberger, D. R. Seizures and schizophrenia. *Schizophr Bull* **23,** 611–622 (1997).

35. Cascella, N. G., Schretlen, D. J. & Sawa, A. SCHIZOPHRENIA AND EPILEPSY: IS THERE A SHARED SUSCEPTIBILITY? *Neuroscience research* **63,** 227 (2009).

36. Coyle, J. T. Glutamate and Schizophrenia: Beyond the Dopamine Hypothesis. *Cell Mol Neurobiol* **26,** 363–382 (2006).

37. Mouri, A., Noda, Y., Enomoto, T. & Nabeshima, T. Phencyclidine animal models of schizophrenia: Approaches from abnormality of glutamatergic neurotransmission and neurodevelopment. *Neurochemistry International* **51,** 173–184 (2007).

38. Kehrer, C., Maziashvili, N., Dugladze, T. & Gloveli, T. Altered Excitatory-Inhibitory Balance in the NMDA-Hypofunction Model of Schizophrenia. *Front Mol Neurosci* **1,** (2008).

39. O'Donnell, C., Gonçalves, J. T., Portera-Cailliau, C. & Sejnowski, T. J. Beyond excitation/inhibition imbalance in multidimensional models of neural circuit changes

in brain disorders. *eLife* **6,** e26724 (2017)

40. Mertens, J. *et al.* Differential responses to lithium in hyperexcitable neurons from patients with bipolar disorder. *Nature* **527,** 95–99 (2015).

41. Vossel, K. A. *et al.* Incidence and impact of subclinical epileptiform activity in Alzheimer's disease. *Annals of Neurology* **80,** 858–870 (2016)

42. Lam, A. D. *et al.* Silent Hippocampal Seizures and Spikes Identified by Foramen Ovale Electrodes in Alzheimer's Disease. *Nat Med* **23,** 678–680 (2017).

43. Vossel, K. A., Tartaglia, M. C., Nygaard, H. B., Zeman, A. Z. & Miller, B. L. Epileptic activity in Alzheimer's disease: causes and clinical relevance. *The Lancet Neurology* **16,** 311–322 (2017).

44. Kupershmidt, I. *et al.* Ontology-Based Meta-Analysis of Global Collections of High-Throughput Public Data. *PLOS ONE* **5,** e13066 (2010).

45. Takao, K. & Miyakawa, T. Genomic responses in mouse models greatly mimic human inflammatory diseases. *PNAS* **112,** 1167–1172 (2015).

46. Ryan, S. D. *et al.* Isogenic Human iPSC Parkinson's Model Shows Nitrosative Stress-Induced Dysfunction in MEF2-PGC1α Transcription. *Cell* **155,** 1351–1364 (2013).

47. Dingledine, R. *et al.* Data Descriptor: Transcriptional profile of hippocampal dentate granule cells in four rat epilepsy models. *Scientific Data* **4,** (2017).

48. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478,** 483–489 (2011).

49. van Os, J., Kenis, G. & Rutten, B. P. F. The environment and schizophrenia. *Nature* **468,** 203–212 (2010).

50. DeMichele-Sweet, M. a. A. *et al.* Genetic risk for schizophrenia and psychosis in Alzheimer disease. *Mol. Psychiatry* **23,** 963-972 (2017).

51. Wilson, D. N. *et al.* Microarray analysis of postictal transcriptional regulation of neuropeptides. *J Mol Neurosci* **25,** 285 (2005).

52. Hermey, G. *et al.* Genome-Wide Profiling of the Activity-Dependent Hippocampal Transcriptome. *PLOS ONE* **8,** e76903 (2013).

53. Vartanian, K. B. *et al.* LPS preconditioning redirects TLR signaling following stroke: TRIF-IRF3 plays a seminal role in mediating tolerance to ischemic injury. *Journal of Neuroinflammation* **8,** 140 (2011).

54. Stevens, S. L. *et al.* Multiple Preconditioning Paradigms Converge on Interferon Regulatory Factor-Dependent Signaling to Promote Tolerance to Ischemic Brain Injury. *J. Neurosci.* **31,** 8456–8463 (2011).

55. Nardo, G. *et al.* Transcriptomic indices of fast and slow disease progression in two mouse models of amyotrophic lateral sclerosis. *Brain* **136,** 3305–3332 (2013).

56. Lerman, B. J. *et al.* Deletion of galectin-3 exacerbates microglial activation and accelerates disease progression and demise in a SOD1 G93A mouse model of amyotrophic lateral sclerosis. *Brain and Behavior* **2,** 563–575 (2012).

57. Matarin, M. *et al.* A Genome-wide Gene-Expression Analysis and Database in Transgenic Mice during Development of Amyloid or Tau Pathology. *Cell Reports* **10,** 633–644 (2015).

58. Xu, B., Hsu, P.-K., Stark, K. L., Karayiorgou, M. & Gogos, J. A. Derepression of a Neuronal Inhibitor due to miRNA Dysregulation in a Schizophrenia-Related Microdeletion. *Cell* **152,** 262–275 (2013).

59. Srivastava, S. *et al.* Loss-of-function variants in *HIVEP2* are a cause of intellectual disability. *European Journal of Human Genetics* **24,** 556–561 (2016).

60. Steinfeld, H. *et al.* Mutations in HIVEP2 are associated with developmental delay, intellectual disability, and dysmorphic features. *Neurogenetics* **17,** 159–164 (2016).

61. Harrison, P. J. The hippocampus in schizophrenia: a review of the neuropathological evidence and its pathophysiological implications. *Psychopharmacology* **174,** 151–162 (2004).

62. Glausier, J. R. & Lewis, D. A. Dendritic spine pathology in schizophrenia. *Neuroscience* **251,** 90–107 (2013).

63. Fatemi, S. H., Folsom, T. D. & Thuras, P. D. GABAA and GABAB receptor dysregulation in superior frontal cortex of subjects with schizophrenia and bipolar disorder. *Synapse* **71,** (2017).

64. Lisman, J. E. *et al.* Circuit-based framework for understanding neurotransmitter and risk gene interactions in schizophrenia. *Trends Neurosci.* **31,** 234–242 (2008).

65. Chestnut, B. A. *et al.* Epigenetic Regulation of Motor Neuron Cell Death through DNA Methylation. *J. Neurosci.* **31,** 16619–16636 (2011).

66. Martin, L. J. & Wong, M. Aberrant Regulation of DNA Methylation in Amyotrophic Lateral Sclerosis: A New Target of Disease Mechanisms. *Neurotherapeutics* **10,** 722–733 (2013).

67. Nguyen, M. D. *et al.* Cell Cycle Regulators in the Neuronal Death Pathway of Amyotrophic Lateral Sclerosis Caused by Mutant Superoxide Dismutase 1. *J.*

Neurosci. **23,** 2131–2140 (2003).

68. Dorszewska, J. *et al.* Mutations of TP53 Gene and Oxidative Stress in Alzheimer's Disease Patients. *Advances in Alzheimer's Disease* **03,** 24 (2014).

69. Moskalev, A. A. *et al.* Gadd45 proteins: Relevance to aging, longevity and age-related pathologies. *Ageing Research Reviews* **11,** 51–66 (2012).

70. Brennand, K. J. *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* **473,** 221–225 (2011).

71. Wen, Z. *et al.* Synaptic dysregulation in a human iPS cell model of mental disorders. *Nature* **515,** 414–418 (2014).

72. Chan, M. K. *et al.* Evidence for disease and antipsychotic medication effects in post-mortem brain from schizophrenia patients. *Mol. Psychiatry* **16,** 1189–1202 (2011).

73. Hagihara, H., Toyama, K., Yamasaki, N. & Miyakawa, T. Dissection of hippocampal dentate gyrus from adult mouse. *J. Vis. Exp.* **33,** 1543 (2009).

74. Butler, G. *et al.* Limbic epilepsy in transgenic mice carrying a Ca2+/calmodulin-dependent kinase II alpha-subunit mutation. *PNAS* **92,** 6852–6855 (1995).

75. Binder, J. X. *et al.* COMPARTMENTS: unification and visualization of protein subcellular localization evidence. *Database (Oxford)* **2014,** (2014).

76. Horton, P. *et al.* WoLF PSORT: protein localization predictor. *Nucleic Acids Res* **35,** W585–W587 (2007).

Competing interests

The authors declare no competing non-financial interests but the following competing financial interests: Tsuyoshi Miyakawa received research grants from Astellas Pharma Inc, and Katsunori Tajinda and Mitsuyuki Matsumoto are employees of Astellas Pharma Inc.

Acknowledgments

This work was supported by JSPS Grants-in-Aid for Scientific Research on Innovative Areas, No. 25116526, 15H01297 (“Microendophenotype”), and 16H06462 (“Scrap & Build”); JSPS KAKENHI Grant No. 25242078; and a grant from Astellas Pharma Inc.

I thank Hideo Hagihara, Hisatsugu Koshimizu, Satoko Hattori, Shoji Hirotaka, Nao Hirata, Yoshihiro Takamiya, Mika Tanaka, Keizo Takao, Katsunori Tajinda, Mitsuyuki Matsumoto, Yoko Kagami, Harumi Mitsuya, Aki Miyakawa, Makoto Tominaga, Tsuyoshi Miyakawa, and other members of the Miyakawa laboratory for their support. I would like to express the greatest appreciation to my family, Ikuyo Kunihiro, Tadayuki Kunihiro, Aya Murano, Shigeko Murano, and Ichiro Murano, for their moral support and warm encouragements.

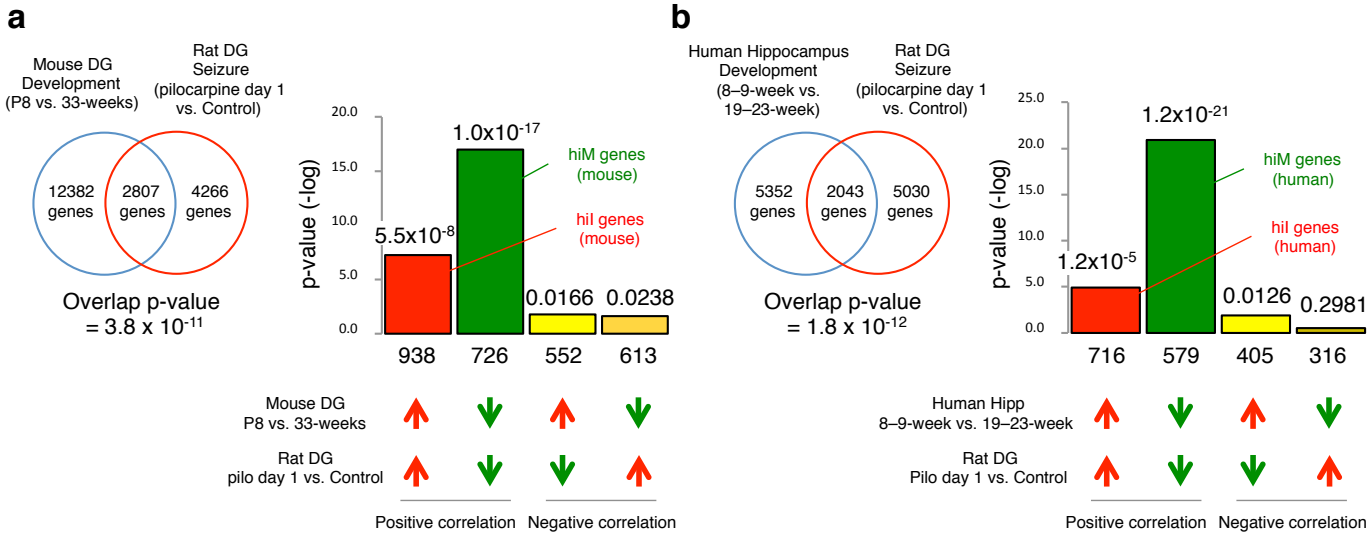


Figure 1. Neural hyperexcitation induces immature-like gene expression patterns in mouse and human

The patterns of changes in gene expression in the rat DG 1 day after pilocarpine treatment compared with developmental changes in the mouse DG and human hippocampus. Venn diagrams illustrating the overlap in genome-wide gene expression changes between the rat DG after seizure induction (GSE47752) and the DG of typically developing mouse infants (GSE113727; P8 infants compared with 33-week-old adults) (a) or the hippocampus of typically developing human fetuses (GSE25219: 19- to 23-week fetuses compared with 8- to 9-week fetuses) (b). Bar graphs illustrate the −log of the overlap *P*-values for genes upregulated (red arrows) or downregulated (green arrows) by each condition. The Bonferroni correction was used to adjust the significance level according to the number of dataset pairs (see the Methods section and Supplementary Methods). Genes that were downregulated in both conditions were defined as mouse hiM genes (green bar in (a)), and genes that were upregulated in both conditions were defined as mouse hiI genes (red bar in (a)). Similarly, human hiM genes and human hiI genes were defined as the groups of genes with positive correlations between the conditions of development and seizures (b).

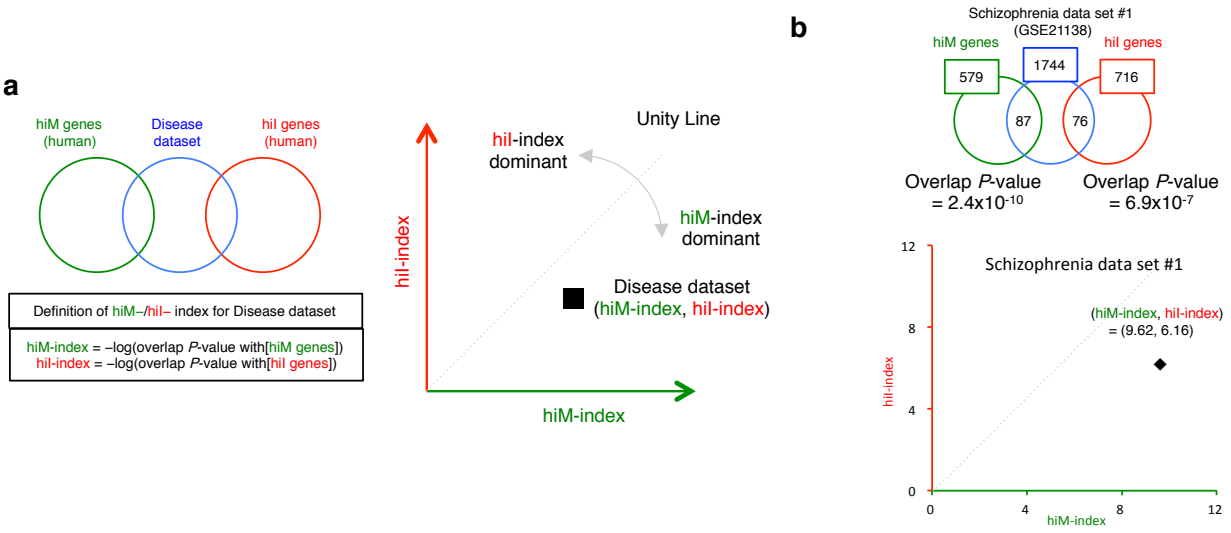


Figure 2. Overview of the two-dimensional (2-D) analysis of disease datasets

(a) Genes with expression changes in the disease datasets are compared with the hiM and hiI gene groups. The hiM- and hiI-indexes were defined as the −log of the overlap *P*-values with the hiM and hiI genes, respectively. The gene expression patterns of the disease datasets are plotted in two-dimensional coordinates, in which the x- and y-axes are defined by the hiM- and hiI-indexes. Each dataset is characterized as hiM- or hiI-dominant by the ratio of the hiM-index to the hiI-index, and the degree of hiM- or hiI-dominance is evaluated by deviation from the unity line. The distance of each dataset from the origin shows the degree of overlap with hiM-/hiI-genes. (b) 2-D analysis applied to a dataset of postmortem brains (prefrontal cortex) from patients with schizophrenia (schizophrenia dataset #1). The expression levels of 1744 genes are significantly changed in this disease dataset. Of these, 87 and 76 genes overlap with the hiM/hiI genes. The overlap *P*-values between the disease dataset and the hiM/hiI genes are 2.4×10^{-10} and 6.9×10^{-7}. The hiM- and hiI-indexes for the disease dataset are therefore 9.62 and 6.16, indicating that this dataset is hiM-dominant. The results of the 2-D analysis for this dataset are plotted in the two-dimensional coordinates defined by the hiM- and hiI-indexes.

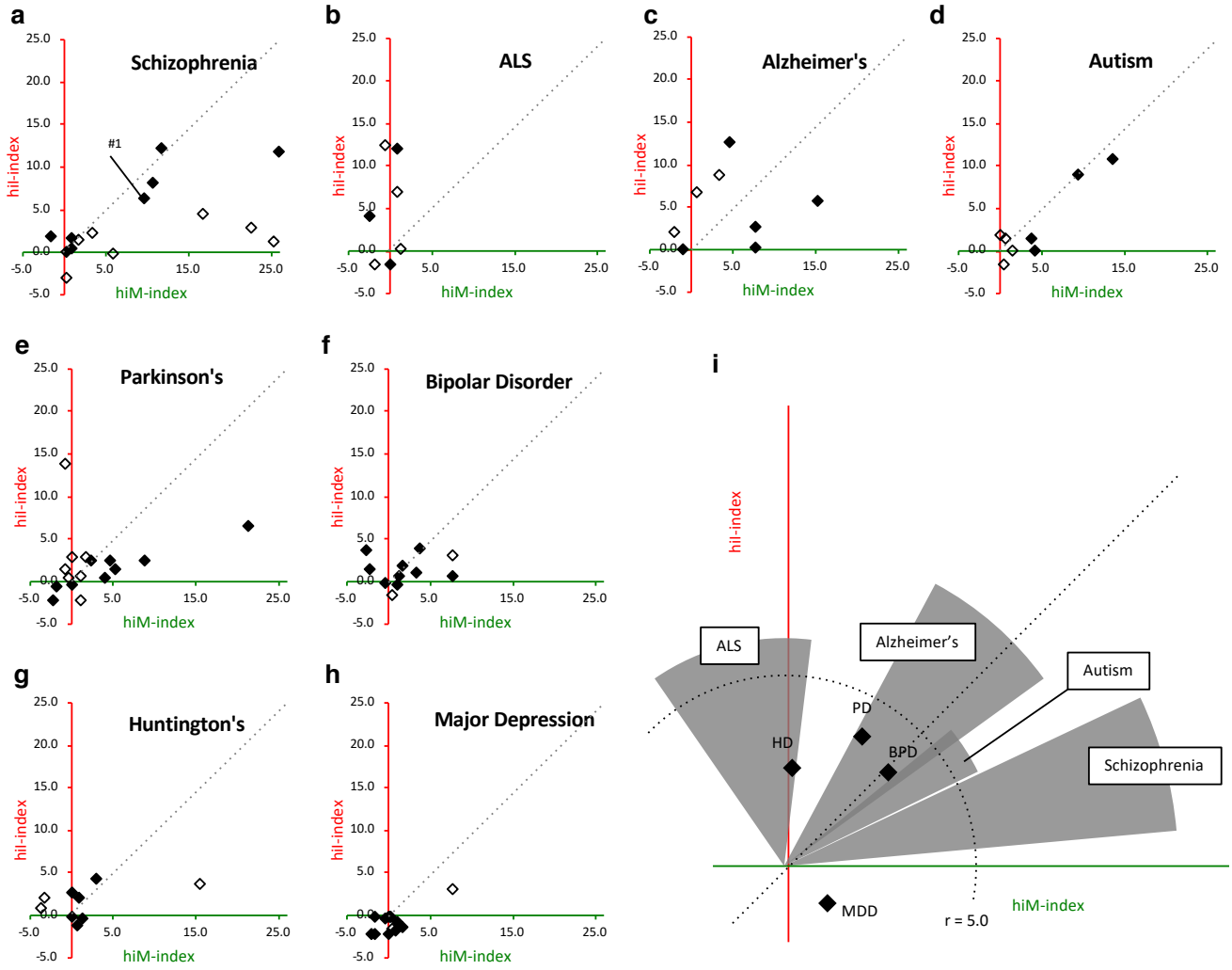


Figure 3. Two-dimensional (2-D) analysis for disease datasets from various neuropsychiatric disorders

(a-h) Results of the 2-D analysis of datasets for schizophrenia (a), ALS (b), Alzheimer's disease (c), autism (d), Parkinson's disease (e), bipolar disorder (f), Huntington's disease (g), and major depressive disorder (h). Each point corresponds to the results of one independent study. Filled points indicate datasets from the postmortem brain or spinal cord (ALS) of patients, and open points indicate those from cultured neural cells from patient iPSCs. (i) The distribution patterns of the hiM- and hiI-indexes for all diseases analyzed. The extent of the changes in hiM-/hiI-indexes is assessed by the average distance of all datasets for each disease from the origin. Four diseases whose average distances from the origin are over 5.0 are shown as circular sectors, and the others are shown as points. The radii of the circular sectors indicate the average distance of all datasets in each disease from the origin, and the central angles of the circular sectors are the average deviation ± SEM from the unity line. Each point indicates the average distance from the origin and average deviation from unity line.

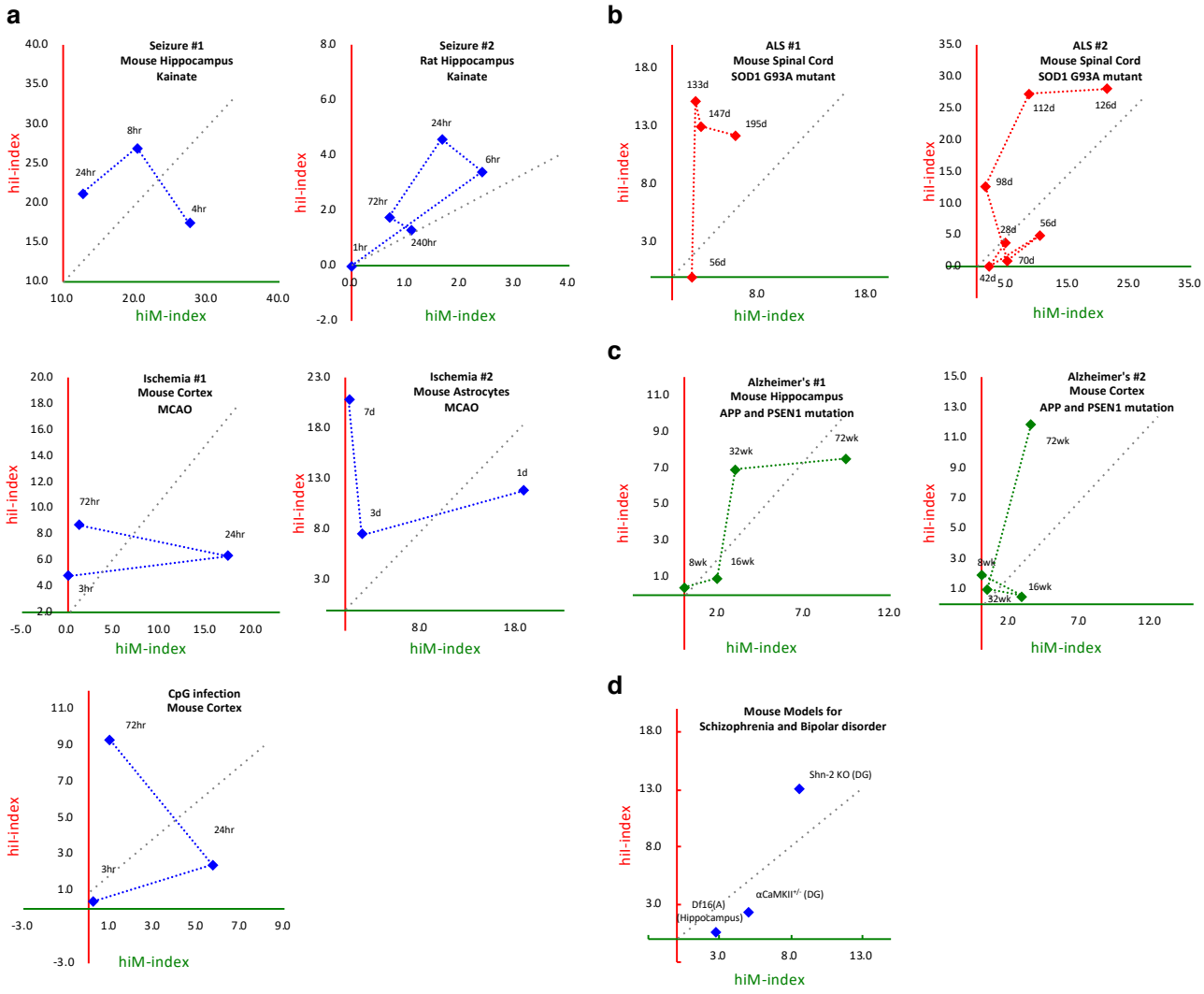


Figure 4. Time-dependent changes in the hiM- and hiI-indexes in animals subjected to various putative risk events for neuropsychiatric disorders and in genetic mouse models of schizophrenia, bipolar disorder, ALS, and Alzheimer's disease

(a) Pattern of changes in the hiM- and hiI-indexes in the mouse and rat hippocampus after treatment with kainite (seizure #1 (GSE1831) and seizure #2 (GSE4236)), in mouse cortex and astrocytes after middle cerebral artery occlusion (MCAO; ischemia #1 (GSE32529), ischemia #2 (GSE35338)), and in mouse cortex after CpG infection (GSE32529). (b) Pattern of changes in the hiM- and hiI-indexes of the spinal cord of an ALS mouse model with the SOD1(G93A) mutation. (c) Pattern of changes in the hiM- and hiI-indexes of the hippocampus and cortex of an Alzheimer's disease mouse model with mutations in APP and PSEN1. (d) hiM- and hiI-indexes in mouse models of schizophrenia and bipolar disorder.

hiM genes (mouse)			
Biogroup name	direction	Common genes	p-value
multicellular organismal signaling	down	56	4.00E-23
transmission of nerve impulse	down	51	3.10E-19
axon	down	45	7.30E-19
neuron differentiation	down	58	4.70E-17
synapse	down	57	5.70E-17
neuron development	down	50	1.50E-16
neuronal cell body	down	45	5.20E-16
cell junction	down	56	3.00E-15
cell part morphogenesis	down	43	4.70E-15
neuron projection development	down	42	6.20E-15
cell projection part	down	56	6.50E-15
cell morphogenesis involved in neuron differentiation	down	35	2.90E-14
cell morphogenesis involved in differentiation	down	40	6.70E-14
cellular chemical homeostasis	down	53	2.50E-13
cell-cell signaling	down	44	2.60E-13
synaptic transmission	down	37	3.60E-13
regulation of neurological system process	down	31	1.10E-12
regulation of transmission of nerve impulse	down	30	1.40E-12
Genes involved in Neuronal System	down	31	1.70E-12
cellular ion homeostasis	down	49	1.80E-12

hiI genes (mouse)			
Biogroup name	direction	Common genes	p-value
response to wounding	up	60	1.10E-16
positive regulation of developmental process	up	66	9.30E-16
cardiovascular system development	up	61	7.70E-15
circulatory system development	up	61	7.70E-15
proteinaceous extracellular matrix	up	32	7.80E-15
Genes involved in Cell Cycle	up	47	1.00E-14
extracellular matrix	up	37	1.30E-14
Genes involved in Cell Cycle, Mitotic	up	42	6.30E-14
protein domain specific binding	up	60	7.00E-14
positive regulation of signal transduction	up	58	8.70E-14
Genes involved in DNA Replication	up	32	9.20E-14
cell cycle	up	67	3.00E-13
Genes involved in Adaptive Immune System	up	56	3.20E-13
Focal adhesion	up	31	5.60E-13
kinase binding	up	44	5.50E-12
cytoskeleton organization	up	54	5.60E-12
neuron differentiation	up	52	5.80E-12
basement membrane	up	19	7.50E-12
regulation of cell migration	up	36	1.10E-11
MAPKinase Signaling Pathway	up	20	1.40E-11

Table 1. Summary of results from the pathway analyses of mouse hiM and hiI genes

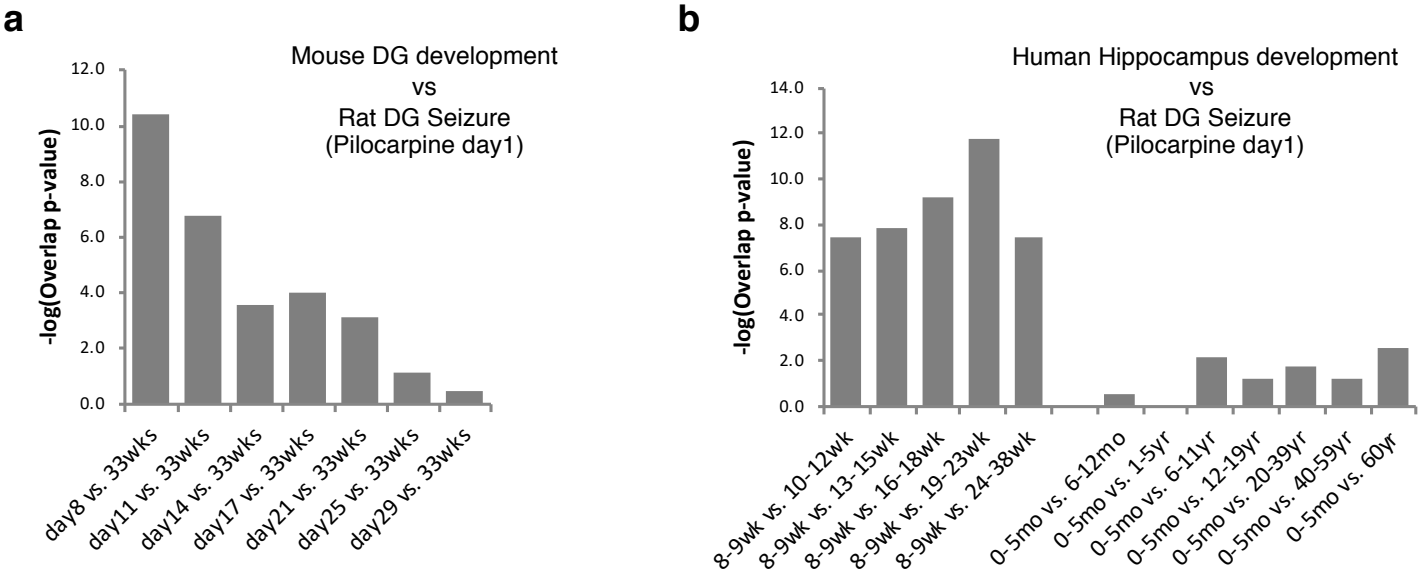
The 20 biogroups with the most significant similarities to mouse hiM and hiI genes. Green columns indicate biogroups that are related to the plasma membrane. Red columns indicate biogroups that are related to reactions in the nucleus.

hiM genes (human)			
Biogroup name	direction	Common genes	p-value
multicellular organismal signaling	up	87	4.40E-58
transmission of nerve impulse	up	83	1.40E-54
synaptic transmission	up	75	6.70E-50
neuron projection	up	70	2.60E-44
axon	up	42	2.90E-33
synapse	up	51	2.60E-32
neuron development	up	58	1.60E-30
Genes involved in Neuronal System	up	41	3.40E-30
Genes involved in Transmission across Chemical Synapses	up	32	1.50E-29
regulation of neurological system process	up	35	6.50E-29
regulation of transmission of nerve impulse	up	34	9.20E-29
cell projection part	up	53	1.80E-27
passive transmembrane transporter activity	up	44	1.80E-26
ion channel activity	up	43	2.20E-26
cell part morphogenesis	up	49	1.80E-24
behavior	up	38	2.90E-24
single-organism behavior	up	35	3.70E-24
cell morphogenesis involved in neuron differentiation	up	44	4.30E-24
neuron projection development	up	47	4.40E-24
dendrite	up	37	2.00E-23

hil genes (human)			
Biogroup name	direction	Common genes	p-value
chromosome	down	65	1.30E-32
response to DNA damage stimulus	down	60	4.80E-32
Genes involved in Cell Cycle	down	50	3.60E-31
Genes involved in Cell Cycle, Mitotic	down	43	1.20E-30
Genes involved in DNA Replication	down	34	2.00E-30
interphase	down	47	8.20E-30
interphase of mitotic cell cycle	down	46	5.10E-29
Genes involved in Mitotic M-M/G1 phases	down	29	2.60E-26
Cell cycle	down	25	7.60E-23
DNA repair	down	40	2.30E-22
wound healing	down	50	7.40E-22
response to ionizing radiation	down	21	1.20E-20
nuclear division	down	35	8.00E-20
mitosis	down	35	8.00E-20
cell division	down	38	2.20E-19
G1/S transition of mitotic cell cycle	down	26	1.70E-18
cardiovascular system development	down	46	3.20E-18
circulatory system development	down	46	3.20E-18
S phase	down	22	3.80E-18
blood coagulation	down	41	6.10E-18

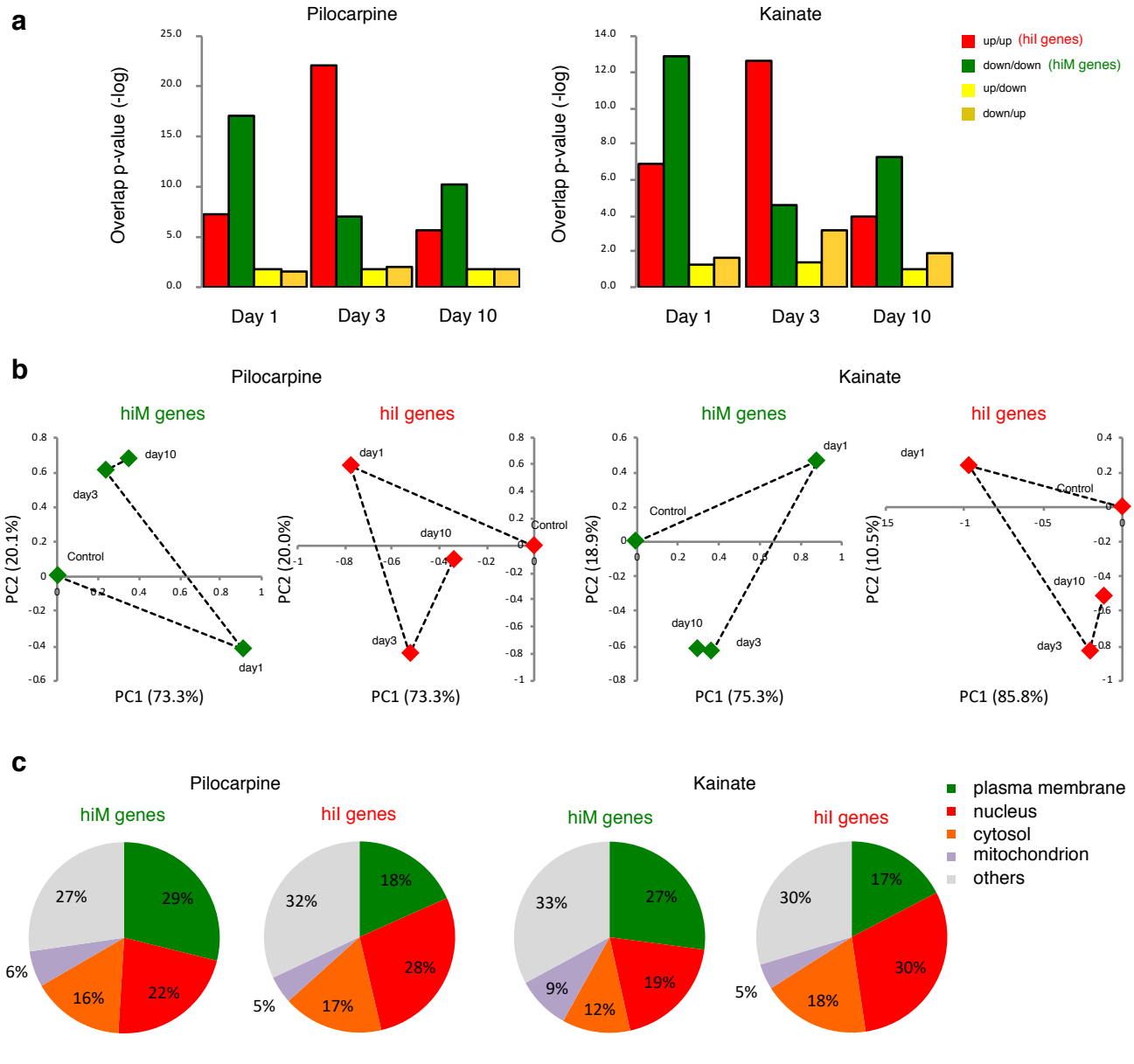
Table 2. Summary of results from the pathway analyses of human hiM and hiI genes

The 20 biogroups with the most significant similarities to human hiM and hiI genes. Green columns indicate biogroups that are related to the plasma membrane. Red columns indicate biogroups that are related to reactions in the nucleus.



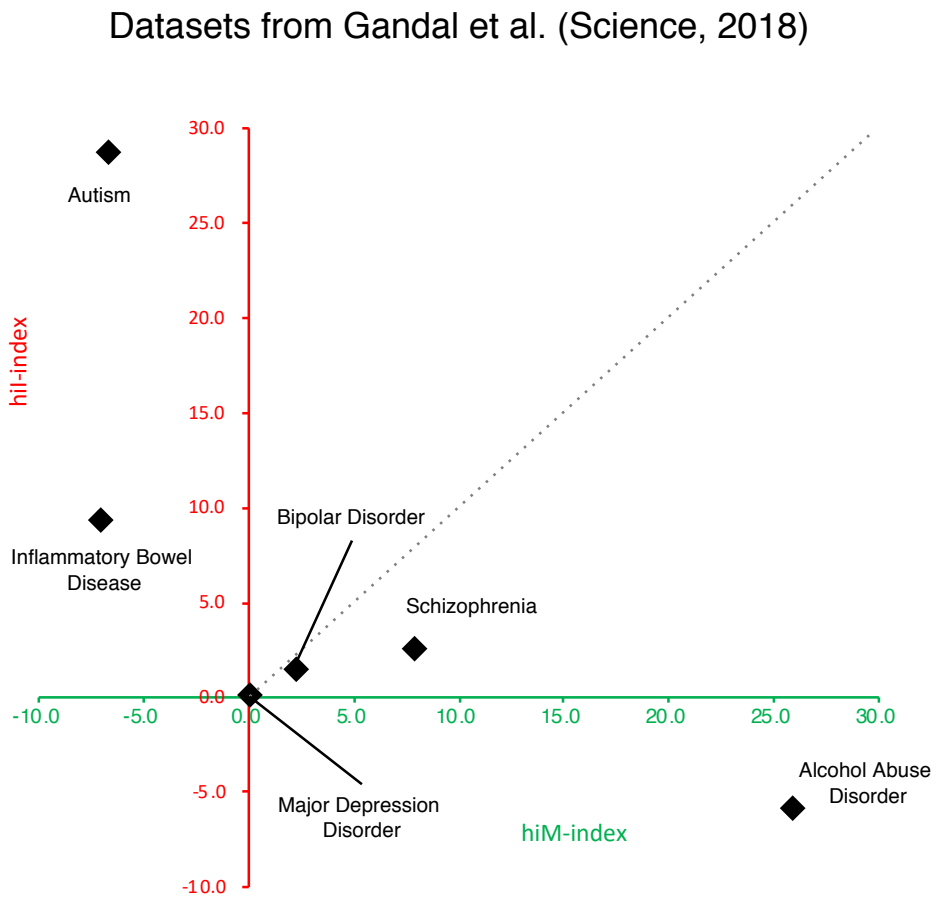
Supplementary Figure 1. The pattern of changes in gene expression in rat DG 1 day after pilocarpine treatment compared with various developmental stages in mouse DG and human hippocampus

Gene expression changes in rat DG after seizure induction (GSE47752) and the DG of developing mice at various stages (GSE113727, infants: P8, P11, P14, P17, P21, P25, P29; adults: 33 weeks) (a) or the hippocampus of typically developing human fetuses (GSE25219: 10–12-week, 13–15-week, 16–18-week, 19–23-week, and 24–38-week fetuses compared with 8–9 week fetuses, or 6–12-month, 1–5-year, 6–11-year, 12–19-year, 20–39-year, 40–59-year, and 60+-year adults compared with 0–5-month infants) (b). Bar graphs illustrate the −log of overlap *P*-values between conditions. The dataset from the DG of day 8 infant mice versus 33-week adults had the largest overlap with the dataset for rat DG 1 day after seizure induction versus controls. The dataset from the hippocampus of 19–23-week fetuses versus 8–9-week fetuses showed the largest overlap with the dataset for rat DG 1 day after seizure induction versus controls.



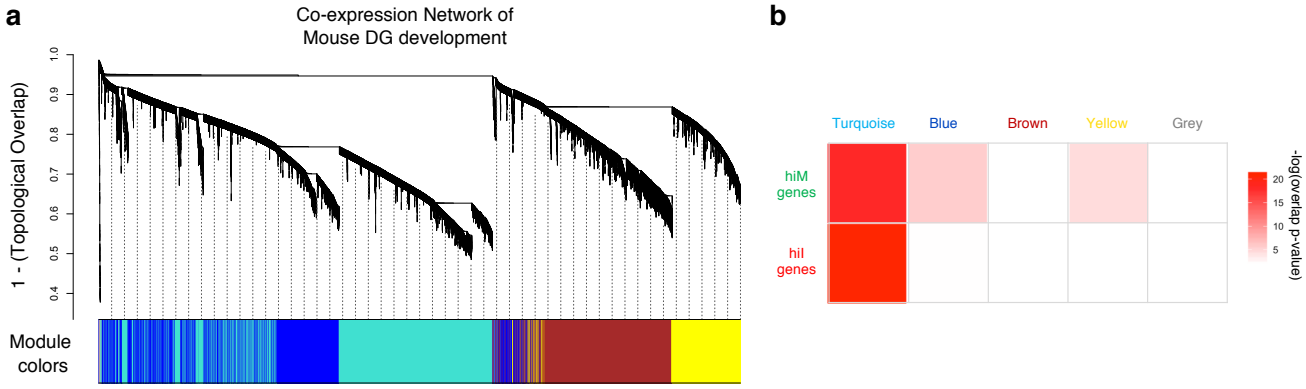
Supplementary Figure 2. Differences in the spatiotemporal patterns of the hyperexcitation-induced expression changes in hiM and hiI genes

(a) Gene expression changes in rat DG at three time points after seizure induction (pilocarpine or kainite) compared with the differences between infant and adult mouse DG. Bar graphs illustrate the overlap *P*-values for genes upregulated or down-regulated in each condition. The Bonferroni correction was used to adjust the significance level according to the number of dataset pairs. (b) The results of principal component analysis for the time-course expression pattern changes in hiM and hiI genes after seizure induction. (c) Subcellular distribution patterns of hiM and hiI genes. Pie charts show the subcellular distribution ratios of the 50 genes with the highest fold change included in each gene set.



Supplementary Figure 3. Results of 2-D analysis performed on the datasets from cerebral cortical samples of subjects with major neuropsychiatric disorders (Gandal et al., Science, 2018)

Each dot corresponds to results of microarray datasets from each disease. hiM-index-dominant pattern were found in schizophrenia and bipolar disorder/major depression disorder show small changes in both of hiM and hiI index, which is mostly consistent with our results in Figure 3. Autism show changes in hiI index but not in hiM index, which is different pattern obtained in our results. Alcohol abuse disorder and inflammatory bowel disorders, which were not assessed in Figure 3, show hiM-index-dominant and hiI-index-dominant patterns.



c

Module: Turquoise (7649 genes)

Biogroup name	common genes	p-value
cell division	201	2.60E-39
Genes involved in Cell Cycle	179	1.20E-30
mitosis	144	6.30E-30
Genes involved in Cell Cycle, Mitotic	155	3.60E-29
Genes involved in Developmental Biology	167	1.50E-23
positive regulation of cell differentiation	196	4.10E-23
Genes involved in Axon guidance	120	1.20E-22
Genes involved in DNA Replication	101	1.30E-21
brain development	161	5.40E-21
negative regulation of cellular component organization	131	5.90E-21

Module: Blue (5424 genes)

Biogroup name	common genes	p-value
GTPase regulator activity	120	3.70E-11
dendrite	127	2.10E-09
dendritic spine	62	5.00E-07
regulation of small GTPase mediated signal transduction	91	7.30E-07
synapse	149	9.20E-07
regulation of anti-apoptosis	25	1.30E-06
neuronal cell body	118	1.50E-06
Genes involved in Neuronal System	76	2.50E-06
Genes involved in Transmission across Chemical Synapses	57	3.00E-06
enzyme activator activity	92	7.50E-06

Module: Brown (4568 genes)

Biogroup name	common genes	p-value
Large Drosha complex	9	0.004
cytoplasmic sequestering of protein	7	0.0045
gene silencing	16	0.0109
transcription elongation factor complex	9	0.0178
mitotic spindle	6	0.0182
Genes involved in Transcriptional activity of SMAD2/SMAD3:SMAD4 heterotrimer	10	0.0197
Genes involved in Signaling by TGF-beta Receptor Complex	16	0.0203
vacuolar transport	11	0.0219
lysosomal transport	10	0.0358
G alpha q Pathway	9	0.0373

Module: Yellow (2454 genes)

Biogroup name	common genes	p-value
perinuclear region of cytoplasm	64	6.80E-06
sarcolemma	22	1.00E-05
Genes involved in Developmental Biology	60	5.50E-05
protein N-terminus binding	22	5.70E-05
Golgi-associated vesicle membrane	9	0.0001
mitochondrial membrane	63	0.0002
Genes involved in NGF signalling via TRKA from the plasma membrane	28	0.0003
Genes involved in Keratan sulfate/keratin metabolism	9	0.0003
cell-substrate adhesion	21	0.0003
14-3-3 protein binding	7	0.0004

Module: Grey (76 genes)

Biogroup name	common genes	p-value
cell surface receptor signaling pathway involved in heart development	2	0.0008
regulation of transcription regulatory region DNA binding	2	0.0013
Beta1 integrin cell surface interactions	3	0.0013
pattern specification process	5	0.0019
cellular lipid catabolic process	3	0.0023
negative regulation of leukocyte apoptotic process	2	0.0023
Genes involved in Nuclear signaling by ERBB4	2	0.0029
Genes involved in Tight junction interactions	2	0.0031
Genes involved in Amyloids	2	0.0036
BMP receptor signaling	2	0.004

Supplementary Figure 4. Comparison between gene modules obtained from mouse DG development datasets and hiM/hiI genes

(a) Network dendrogram from coexpression topological overlap of genes across disorders. (b) Overlap between genes in hiM/hiI genes and 5 gene modules obtained from coexpression analyses in (a). Color bars show -log of overlap *P*-value. (c) Summary of results from the pathway analyses of 5 modules.

Supplementary Data listed below will be published with Murano et al., *Commun Biol*, (2019).

Supplementary Data 1.
Gene list of hiM/hiI genes in mouse and human.

Supplementary Data 2.
Detailed information on the 99 microarray datasets used in this study.

Supplementary Data 1.
Gene lists of hiM and hiI genes in mouse and human.

Supplementary Data 2.
Detailed information on the 99 microarray datasets used in this study.

Supplementary Data 3.
The lists of hiM and hiI genes included in datasets from patients with amyotrophic lateral sclerosis (ALS).

Supplementary Data 4.
The lists of hiM and hiI genes included in datasets from patients with Alzheimer's disease (ALZ).

Supplementary Data 5.
The lists of hiM and hiI genes included in datasets from patients with autism spectrum disorder (ASD).

Supplementary Data 6.
The lists of hiM and hiI genes included in datasets from patients with bipolar disorder (BPD).

Supplementary Data 7.
The lists of hiM and hiI genes included in datasets from patients with Huntington's disease (HD).

Supplementary Data 8.
The lists of hiM and hiI genes included in datasets from patients with major depressive disorder

(MDD).

Supplementary Data 9.

The lists of hiM and hiI genes included in datasets from patients with Parkinson's disease (PD).

Supplementary Data 10.

The lists of hiM and hiI genes included in datasets from patients with schizophrenia (SCZ).

Supplementary Data 11.

The lists of hiI genes shared in all datasets in each disease categories.

Supplementary Data 12.

The lists of hiI genes shared in all datasets in each disease categories.