The functional genetic link of NLGN4X knockdown and neurodevelopment in neural stem cells

Lingling Shi1, Xiao Chang1, Peilin Zhang2, Marcelo P. Coba1,4, Wange Lu2,3 and Kai Wang1,4,5,*

1Zilkha Neurogenetic Institute, 2Eli and Edythe Broad Center for Regenerative Medicine and Stem Cell Research, 3Department of Biochemistry and Molecular Biology, 4Department of Psychiatry and 5Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA

Received March 20, 2013; Revised March 20, 2013; Accepted May 15, 2013

Genetic mutations in NLGN4X (neuroligin 4), including point mutations and copy number variants (CNVs), have been associated with susceptibility to autism spectrum disorders (ASDs). However, it is unclear how mutations in NLGN4X result in neurodevelopmental defects. Here, we used neural stem cells (NSCs) as in vitro models to explore the impacts of NLGN4X knockdown on neurodevelopment. Using two shRNAmir-based vectors targeting NLGN4X and one control shRNA-mir vector, we modulated NLGN4X expression and differentiated these NSCs into mature neurons. We monitored the neurodevelopmental process at Weeks 0, 0.5, 1, 2, 4 and 6, based on morphological analysis and whole-genome gene expression profiling. At the cellular level, in NSCs with NLGN4X knockdown, we observed increasingly delayed neuronal development and compromised neurite formation, starting from Week 2 through Week 6 post differentiation. At the molecular level, we identified multiple pathways, such as neurogenesis, neuron differentiation and muscle development, which are increasingly disturbed in cells with NLGN4X knockdown. Notably, several postsynaptic genes, including DLG4, NLGN1 and NLGN3, also have decreased expression. Based on in vitro models, NLGN4X knockdown directly impacts neurodevelopmental process during the formation of neurons and their connections. Our functional genomics study highlights the utility of NSCs models in understanding the functional roles of CNVs in affecting neurodevelopment and conferring susceptibility to neurodevelopmental diseases.

INTRODUCTION

In the past few years, genetic studies have implicated a strong connection between the synaptic complex with a range of neurodevelopmental or neuropsychiatric disorders, such as autism and schizophrenia (1–4). The key component of this complex include presynaptic neurexins, postsynaptic neuroligins (NLGNs), SHANKs, postsynaptic density protein 95 (PSD-95) and disks large-associated proteins (DLGAPs). Among them, neuroligins represent some of the most well-known postsynaptic molecules that play key roles in establishing cell adhesions with presynaptic neurexins at the synaptic junction (5–8). All neuroligins contain an N-terminal hydrophobic sequence with the characteristics of a cleaved signal peptide followed by a large esterase homology domain, a highly conserved single transmembrane region and a short cytoplasmic domain. Among the four members of the neuroligin family, NLGN4X is located in the X chromosome; it has 63–73% amino acid identity with the other neuroligins (9). The NLGN4X gene contains six exons and codes for a protein of 816 amino acids, and has a widespread gene expression pattern (9).

Multiple genetic studies have now implicated a significant role for neuroligins, especially NLGN4X, in the susceptibility to several neurodevelopmental disorders. Xp22.3 deletions that include the NLGN4X gene have long been associated with autism (10). In 2003, a study screened mutations in NLGN3 and NLGN4X in a cohort with autism spectrum disorders (ASDs), and identified de novo frameshift mutations in NLGN4X in one affected individual’s mother (11). Another frameshift mutation in NLGN4X was observed in all affected individuals in a large French family, in which 10 males had X-linked mental retardation, 2 had autism and 1 had pervasive developmental disorder (12). A few follow-up studies suggested that nonsense changes in NLGN4X may also contribute to autism susceptibility (13,14), although a study on Quebec population failed to identify NLGN4X mutations in 96 subjects affected with autism (15). Another family-based association analysis in 100 families with ASDs yielded only modest associations at

*To whom correspondence should be addressed. Tel: +1 3234423917; Fax: +1 3234422145; Email: kaiwang@usc.edu

© The Author 2013. Published by Oxford University Press. All rights reserved.
For Permissions, please email: journals.permissions@oup.com
NLGN4X and other neuroligins, suggesting that NLGN4X mutations probably represent only rare causes of autism (16). Several recent studies have also linked various NLGN4X deletions with autism susceptibility (17). Additionally, novel de novo splice variants in NLGN4X have also been observed in autistic individuals (18). In summary, mounting evidence suggested that very rare and loss-of-function mutations in NLGN4X, including deletions, are associated with susceptibility to ASDs.

Despite these prior genetic studies, it is still unclear how mutations in NLGN4X confer susceptibility to multiple related diseases, and what biological processes are compromised due to NLGN4X haploinsufficiency. These types of questions can be partially addressed in animal models by behavioral and molecular studies. In 2008, Jamain et al. established a mouse model with loss-of-function mutation in Nlg4, and found that these mice have highly selective deficits in reciprocal social interactions and communication that are reminiscent of ASDs in humans (19). Recently, the same group demonstrated that Nlg4n null mouse may serve as a construct-valid model of heritable monogenic autism (20). However, analysis on models derived from human tissues may provide complementary insights into mouse models. Recently, in vitro cellular models, such as neurons derived from humans (21) or induced pluripotent stem cells (hiPSCs) (22–25), have been increasingly recognized as important tools to gain insights into the functional roles of genetic mutations during neurodevelopment. Several models for Rett syndrome (26), schizophrenia (27) and fragile X mental retardation syndrome (28), have already been reported, implicating the potential utility of using in vitro human-based models to study functional impacts of genetic mutations on neurodevelopment. We have also recently established neural stem cell (NSC) models derived from both hiPSCs and human embryonic stem cells, and demonstrated that knockdown of NRXN1 expression in NSCs compromise neurodevelopmental pathways, based on gene expression profiling at Week 4 post differentiation (29). Other groups have also examined the utility of iPS-SCs-derived neurons to investigate synapse formation and function, and found that HEK-293T cells expressing NLGN3 and NLGN4X, but not those containing autism-associated mutations, are able to induce presynaptic differentiation in hiPSCs-derived neurons (30).

In the current study, we addressed a central hypothesis that if deletions of NLGN4X influence susceptibility to autism, then knockdown of gene expression levels for NLGN4X may influence the course of neurodevelopment in an in vitro model. Similar to a previous study (29), we used NSCs as models, and used shRNA mir to re-create NLGN4X haploinsufficiency; however, we are particularly interested in the time course relationships between gene knockdown and genome function, so we measured gene expression levels across six time points, after differentiating the NSCs into neurons. To address the concern on nonspecific knockdown, we evaluated two different knockdown vectors in the study. With this in vitro model, we attempted to observe whether there are cellular and molecular differences between NSCs with or without NLGN4X knockdown, and related these observations to prior knowledge on the potential functionality of neuroligins in synaptic development and function.

RESULTS
Normal developmental time course in neural stem cells
We differentiated NSCs into neurons during a 6-week period, and collected RNA samples at multiple time points (Weeks 0, 0.5, 1, 2, 4, 6) for whole-genome gene expression profiling. For each time point, two to three replicates were assayed. We constructed a hierarchical clustering plot on their expression levels, and found that these cells fall into two major groups (Fig. 1). The three replicates in Week 4 and two replicates in Week 6 cluster together, and are separated from those samples collected at Weeks 0, 0.5, 1 and 2. This observation is also consistent with Kanopka et al. (21), where they demonstrated that Week 4 is a critical time point during differentiation of primary normal human neuronal progenitors, when a significant number of genes associated with ASDs are either induced or repressed. Additionally, Kanopka et al. identified 4427 genes that have expression alterations during developmental course at Week 4 (21). We found that 1696 of 4427 genes from the Kanopka et al. study are also identified as differentially expressed between Weeks 0 and 4 in our study. Notably, several prominent synaptic genes, such as NRXN1, NRXN2, NLGN1, NLGN2, HOMER1, HOMER2, GRM3 and GRIA2, are found in our results, suggesting that these genes have altered gene expression during the normal developmental time course.

NLGN4X knockdown results in morphological differences in neural differentiation
To knockdown NLGN4X expression reliably in NSCs and differentiated neurons, we procured four commercially available shRNA mir vectors and tested them in HEK-293T cells. We confirmed that all of them can knockdown NLGN4X expression by a qPCR, but with varying efficiencies (Supplementary Material, Fig. S1). We selected two shRNA vectors, sh7 and sh8, in following studies, given that they generally resulted in high knockdown efficiency under the conditions of high cell confluence. We next tested these shRNA mir vectors in NSCs. After packaging control shRNA mir (shCon) or NLGN4X shRNA mir (sh7 and sh8) into lentivirus and infecting NSCs, we observed RFP signals within 1 week for >95% of all cells (Fig. 2A), suggesting efficient production of the shRNA mir. Seven days post infection, with doxycycline induction, we observed high knockdown efficiency (46–76%) in the NSCs for both sh7 and sh8 vectors, compared with NSCs with shCon, suggesting that the vectors can efficiently knockdown NLGN4X expression in a controlled manner (Fig. 2B).

We next examined whether there are developmental differences between the NSCs with or without NLGN4X knockdown, by examining cell morphology across a 6-week period after differentiating NSCs into the neuronal lineage. In the very early-stage post differentiation, the cells with NLGN4X knockdown and those without knockdown show typical NSC morphology (Fig. 3A). At Day 3, we observed signs that some neurons began to emerge in the culture dish with control shRNA mir, including neurites at the cell surface, but not in the two dishes with NLGN4X knockdown (Fig. 3B). At 7 days post differentiation, there were limited protrusion and branches of neurites in...
cells with NLGN4X knockdown, and we cannot identify obvious cell–cell connections; however, in dishes with control shRNA-mir, we observed obvious neurites forming cell–cell interactions (Fig. 3B). After 2 weeks, when we examine the controls, we can see that a large number of neurons are clearly formed, and some axons are already visible in the cells treated with control shRNAmir (Fig. 3B). At the same time, when we examined the cells with NLGN4X knockdown, we observed less formation of neurons and inter-cell connections. Additionally, we examined zoomed-in photos of individual neuronal cells at different time points, and found similar patterns of defects in neurite formation (Supplementary Material, Fig. S2). These differences suggested that NLGN4X knockdown affects cellular phenotypes during differentiation of NSCs into neurons, by delaying the neurodevelopmental process and by compromising neurite formation and the establishment of cell–cell interactions.

Alteration of gene expression due to NLGN4X knockdown during neurodevelopment

Using whole-genome gene expression data, we next compared transcriptome alterations between three types of cells (sh7, sh8 and shCon) at different time points: Weeks 0, 0.5, 1, 2, 4 and 6 (Fig. 3). For each time point and each sample, two to three replicates were used in expression profiling. We found that the NLGN4X gene expression level has decreased 20–80% at all of the time points, similar to what we observed in a qPCR in Figure 2. Hierarchical clustering showed that samples from the same time points tightly clustered with each other, suggesting that cells with or without NLGN4X knockdown have relatively small differences, compared with differences due to developmental time course (Fig. 4). Additionally, within each time point, cells treated with sh7 and sh8 tended to cluster together and were separated from cells treated with shCon, clearly indicating that NLGN4X knockdown induced similar whole-genome expression changes in the two cell populations treated with shRNAmir.

We are also interested in whether NLGN4X knockdown affects the temporal patterns of gene expression during the developmental time course, that is, whether the ‘trends’ of time-dependent gene expression are different between the cells. For this purpose, we compared the overall patterns of gene expression between cells treated with sh8 or shCon using ‘timecourse’ (31), a statistical package specifically designed for this purpose. All genes are ranked by the Hotelling’s T2 values, representing the magnitude of changes in temporal patterns over the time course. In this list, NLGN4X ranks as the 192nd most altered gene (Fig. 5). In previous analysis, we found that several synaptic genes, such as NRXN1, NRXN2, NLGN1, NLGN2, HOMER1, HOMER2, GRM3 and GRIA2, are differentially expressed between Weeks 0 and 4 in cells without knockdown; among these genes, only GRM3 appears to have clearly altered temporal patterns due to NLGN4X knockdown (Fig. 5), yet most other genes show weak trends of altered temporal patterns (ranked between 1000th and 2000th). HOMER1 is an exception that ranked at 23493rd. In comparison, two commonly used housekeeping genes, GAPDH and ACTB, ranked at 4460th and
9489th, respectively. We repeated the above analysis by comparing cells treated with sh7 and shCon, and obtained similar results. This analysis indicated that NLGN4X knockdown may have weak effects in impacting the temporal patterns of gene expression for synaptic genes during the developmental time course.

Biological pathways and networks affected by NLGN4X knockdown

To examine the biological pathways that were affected by NLGN4X knockdown, we identified genes that were differentially expressed at each time point between NSCs with or without NLGN4X knockdown. For Weeks 0, 0.5 and 1, we did not identify any genes that are differentially expressed, but an increasing number of genes have differential expression after Week 2. An enrichment analysis using the DAVID web server (32) identified multiple Gene Ontology (GO) categories that were significantly altered (Tables 1 and 2). We listed the top 10 most enriched genetic pathways at Week 6 for the down-regulated and up-regulated genes in sh8-treated NSCs, respectively, and then compared the significance of the selected pathways at different time points. The top 10 down-regulated pathways were most related to the nervous system and development, such as ‘nervous system development’ (GO:0007399, ranked as first), ‘neurogenesis’ (GO:0022008, ranked as third), ‘neuron differentiation’ (GO:0030182, ranked as fifth) and ‘organ development’ (GO:0048513, ranked as seventh), ‘cell development’ (GO:0048468, ranked as seventh), ‘tissue development’ (GO:0009888, ranked as ninth). However, two muscle-related pathways have also been altered as a result of NLGN4X knockdown, including ‘muscle organ development’ (GO:0007517, ranked as fourth) and ‘muscle contraction’ (GO:0006936, ranked as sixth). For the down-regulated pathways in Week 6, all of them were also enriched in Weeks 2 and 4, but with an obvious increase of statistical significance with time. Though less pronounced than down-regulated pathways, the upregulated list was also enriched with several development-related pathways, such as ‘organ development’, ‘tissue development’ and ‘nervous system development’. These results suggested that NLGN4X knockdown has significant impacts on other genes that are involved in the neurodevelopmental process. Similar observations were made on cells treated with sh7 (Supplementary
Material, Table S1), although it does not have as high knock-down efficiency as sh8 (Supplementary Material, Fig. S1, Fig. 2). We found that muscle-related and development-related pathways are also significantly enriched, yet the former has even stronger statistical significance. These results corroborate previous studies showing that neurexins and neuroligins play important roles in smooth muscle development and promote angiogenesis (33), in addition to their well-known roles in nervous system development.

We next examined the overlap of the differentially expressed (DE) genes at different time points (Table 3). At Week 2, 25 of 43 (58.1%) DE genes are also differentially expressed at Week 4, and 31 of 43 (72.1%) are also identified at Week 6. In addition, 522 of 774 (71.3%) DE genes at Week 4 are also identified at Week 6. However, the percentage of overlap between Week 1 and the following weeks are much smaller (3.6–22.5%), suggesting that genes differentially expressed at Week 2 are more likely to be continuously differentially expressed in later weeks. This observation is consistent with the pathway enrichment analysis (Tables 1 and 2), where the same down-regulated pathways at Week 6 also tend to have down-regulation in Weeks 2 and 4, but not Week 1.

To further evaluated the biological networks that are perturbed, we compiled a list of 60 well-known proteins involved in a postsynaptic complex (Supplementary Material, Table S2) and mapped them onto the human protein–protein interaction (PPI) network (34). We extracted their directly connected gene neighbors to construct a network, which contains 537 node and 2601 edges. Although it is difficult to discern details of the network visually, it is clear that all sub-modules of the network contain differentially expressed genes, especially down-regulated genes (Fig. 6A). Among these 60 postsynaptic proteins, 11 have differential expression, so we constructed a sub-network on these 11 genes and their first-degree interaction partners (Fig. 6B). In this sub-network, the \textit{NLGN4X} gene has two directly connected genes, \textit{DLG4} and \textit{DLGAP2}. \textit{DLG4} is significantly down-regulated, and it serves as a hub protein in this sub-network that connects with \textit{NLGN1}, \textit{NLGN2} and \textit{NLGN3}. \textit{NLGN1} and \textit{NLGN3} are significantly down-regulated as a result of \textit{NLGN4X} knockdown. We also

Figure 3. Differences in cellular phenotypes between cells treated with three shRNA vectors during differentiation of NSCs at multiple time points. (A) Stereo image on NSCs at Day 0 post differentiation, demonstrating a similar cell morphology in three culture dishes treated with sh7, sh8 and shCon, respectively. (B) RFP image on NSCs 3 days, 7 days and 14 days post differentiation. We observed increasingly clear trends that fewer axons and synaptic connections formed in differentiated neurons treated with sh7 and sh8.
Figure 4. Hierarchical clustering of gene expression levels across multiple time points for two cell populations with NLGN4X knockdown (sh7 and sh8) and one cell population without knockdown (shCon). For each time point, the average expression values from two to three replicate samples were used.
Figure 5. Alteration of temporal patterns of gene expression as a result of NLGN4X knockdown during the timecourse (Weeks 0.5, 1, 2, 4 and 6). Several synaptic genes are depicted, yet only NLGN4X and GRM3 have significantly altered temporal patterns.

Table 1. Biological pathways that were significantly down-regulated in differentiated NSCs with NLGN4X knockdown at Week 6, and their significance levels in Weeks 1, 2 and 4

<table>
<thead>
<tr>
<th>Pathway</th>
<th>W1</th>
<th>W2</th>
<th>W4</th>
<th>W6</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0007399 Nervous system development</td>
<td>Count</td>
<td>0</td>
<td>9</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>2.382E - 03</td>
<td>1.174E - 14</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>2.975E - 02</td>
<td>1.807E - 13</td>
</tr>
<tr>
<td>GO:0048513 Organ development</td>
<td>Count</td>
<td>0</td>
<td>16</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>1.327E - 06</td>
<td>1.506E - 12</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>1.681E - 05</td>
<td>2.314E - 01</td>
</tr>
<tr>
<td>GO:0022008 Neurogenesis</td>
<td>Count</td>
<td>0</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>1.074E - 02</td>
<td>4.548E - 09</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>1.278E - 01</td>
<td>6.987E - 08</td>
</tr>
<tr>
<td>GO:0007517 Muscle organ development</td>
<td>Count</td>
<td>0</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>6.772E - 06</td>
<td>2.414E - 12</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>8.575E - 05</td>
<td>3.709E - 11</td>
</tr>
<tr>
<td>GO:0030182 Neuron differentiation</td>
<td>Count</td>
<td>0</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>7.646E - 02</td>
<td>9.730E - 08</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>6.348E - 01</td>
<td>1.495E - 06</td>
</tr>
<tr>
<td>GO:0006936 Muscle contraction</td>
<td>Count</td>
<td>0</td>
<td>7</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>1.049E - 06</td>
<td>1.685E - 13</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>1.328E - 05</td>
<td>2.589E - 12</td>
</tr>
<tr>
<td>GO:0048468 Cell development</td>
<td>Count</td>
<td>0</td>
<td>6</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>1.418E - 02</td>
<td>2.708E - 07</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>1.655E - 01</td>
<td>4.160E - 06</td>
</tr>
<tr>
<td>GO:0060284 Regulation of cell development</td>
<td>Count</td>
<td>0</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>7.889E - 02</td>
<td>1.586E - 08</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>6.468E - 01</td>
<td>2.437E - 07</td>
</tr>
<tr>
<td>GO:0009888 Tissue development</td>
<td>Count</td>
<td>0</td>
<td>7</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>3.396E - 03</td>
<td>4.648E - 06</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>4.216E - 02</td>
<td>7.140E - 05</td>
</tr>
<tr>
<td>GO:0032989 Cellular component morphogenesis</td>
<td>Count</td>
<td>0</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>P-value</td>
<td>NA</td>
<td>6.051E - 02</td>
<td>2.547E - 04</td>
</tr>
<tr>
<td></td>
<td>FDR</td>
<td>NA</td>
<td>5.464E - 01</td>
<td>3.905E - 03</td>
</tr>
</tbody>
</table>

Total count of down-regulated genes: 11 36 375 842
noted that NRXN2, but not NRXN1, has down regulation; this is consistent with the previous report that NRXN1 knockdown did not affect gene expression for neuroligins (29). In summary, the sub-network demonstrated that a considerable number of important synaptic genes, especially DLG4, NLGN1, NLGN3, NRXN2, directly or indirectly interacted with NLGN4X and form a down-regulated network driven by NLGN4X knockdown. We recognize that the current protein–protein interaction information is still incomplete, but this analysis further reconfirmed our hypothesis that NLGN4X knockdown affects gene expression for a network containing synaptic genes.

**DISCUSSION**

In the current study, we investigated the functional significance of NLGN4X deletions in neurodevelopment, using NSCs as in vitro models. During differentiation of NSCs into mature neurons across a 6-week period, we found that knockdown of a single neuronal gene NLGN4X leads to morphological changes as well as transcriptome alterations. Altogether, our results suggested that NLGN4X deletion influences normal developmental course for neurons, and may alter several biological pathways including nervous system development and neuron differentiation. These cellular alterations could play a role in the molecular pathophysiology of several neurodevelopmental diseases affected by NLGN4X deletion, including autism and schizophrenia.

The ability to differentiate NSCs into neuronal lineage in a tightly controlled manner has made NSCs as attractive in vitro model systems to study complex neurodevelopmental disorders. For many neurodevelopmental disorders, animal models are generally not available, not representative enough or not ideal to use. On the other hand, the neuronal system is also difficult
to study in human models, as live neurons, especially developing neurons, are not readily available from the patients to understand the molecular pathophysiology of the diseases. Therefore, NSCs especially those derived from patients (such as NSCs generated from patient-specific fibroblasts or patient-specific iPS cells) may find more use in future genetic studies. By analyzing morphological, electrophysiological, transcriptional and functional differences of neurons derived from specific patients and control subjects, or from the same patient with and without gene knockdown, we may better understand the molecular mechanism of disease pathogenesis. Additionally, these in vitro models may also facilitate drug discovery (for example, in drug screening applications), and may facilitate pharmacogenomic studies to understand how deficiency in candidate genes impact cellular responses to pharmacological treatments.

There are several potential enhancements to the current study that we wish to elaborate here. First, we studied the neurons from the perspective of cell morphology and gene expression, but one interesting hypothesis to explore is whether there are electrophysiological consequences as a result of $NLGN4X$ knockdown. This can be accomplished by electrophysiological recording of single neurons by patch clamp techniques (35,36), which provides a high-resolution method of observing the function of individual ionic channels in a variety of normal and pathological cell types. Second, developing neurons are highly heterogeneous, and it is possible that one subtype of neuronal cells or astrocytes is more influenced by $NLGN4X$ knockdown than the other subtypes. However, by pooling all cells (including astrocytes) together in gene expression analysis, we lose the resolution to look at specific subtypes of neurons. This can be addressed by technical advancement in isolating specific subtypes of cells from culture dish, or even single-cell microarray analysis or single-cell RNA-Seq analysis (37). Third, there are multiple other ways in which genome expression and function can be modulated, including DNA methylation, nucleosome positioning, and post-transcriptional or post-translational modifications (38–40). Our analysis did not evaluate the impact of $NLGN4X$ deletions on these aspects, although the prior probability that $NLGN4X$ affects these biological processes is low, given the known functional role of $NLGN4X$. Finally, we used gene expression microarrays rather than more modern sequencing techniques to measure gene expression levels. Several studies have compared the performance of microarray versus RNA-Seq; these studies demonstrated that both have their advantages and limitations, although the biological interpretations (such as pathway enrichment) from both platforms are mostly consistent (41–43). Given the need for examine multiple time points in several cell lines each with two to three replicates, we elected to use gene expression microarrays as this is a more economical choice for our study.

In conclusion, by introducing targeted gene knockdown in NSCs and measuring whole-genome gene expression levels, in the context of time course series data, we developed an interesting approach for studying the functional significance of CNVs during neurodevelopment. The models that we established here would help confirm how candidate CNVs identified from previous CNV association studies on complex neuropsychiatric and neurodevelopmental disorders modulate disease risk, and also help uncover biological pathways and networks underlying disease susceptibility. We believe that these models can be further developed in a high-throughput fashion, whereas we simultaneously introduce multiple genetic knockdowns in 96-well plates, and follow-up the same batch of cells across multiple weeks, then assay the morphological changes and gene expression levels both in a high-throughput fashion, for systems understanding of candidate gene functions. These methods may ultimately represent high-throughput means to understanding the functional role of human DNA sequence variants in complex neurodevelopmental diseases.

Figure 6. Network of known protein–protein interactions that includes all first-degree neighbors of 60 well-known postsynaptic proteins. (A) Sub-networks formed by all proteins and neighbors (B) sub-network formed by of 11 DE proteins and their neighbors. Proteins involved in a postsynaptic complex are labeled in diamonds, yet their first-degree neighbors in the interaction network are labeled in circles. Upregulated genes are colored by red, yet down-regulated genes are colored in green.
MATERIALS AND METHODS

Neural stem cells (NSCs) induction and neuronal differentiation

We have previously generated hiPSC lines using 2.0 × 10^6 human fetal dermal fibroblasts (HDF, acquired from ATCC) transfected with 4 µg CAG.OSM-puDtk reprogramming transposon and 2 µg pCyl43 transposase plasmid through nucleofection (Amaza Nucleofector technology). Embryoid Bodies (EBs) were formed first by splitting the hiPSCs colonies into appropriate size and seeding on 6 cm ultra-low attachment dish (BD Biosciences) with ES culture media without bFGF, and changing the medium every 2 days. On Day 5 of EBs formation, we switched to N2 media (DMED/F12 with N2 supplement (Invitrogen) and 1% penicillin/streptomycin), 50% B27 media (DMED/F12 with B27 supplement (Invitrogen) and 1% penicillin/streptomycin), with 20 ng/ml bFGF) and changing the medium every 2 days. The knockdown efficiency was tested by qPCR (see below), and we selected two most efficient shRNAmir vectors for use in downstream studies.

Screening shRNAmir with high knockdown efficiency on NLGN4X

Four commercial shRNAmir targeting NLGN4X were acquired from Open Biosystems (catalog number: RHS4696-101317495, RHS4696-101318313, RHS4696-101318401, RHS4696-101324262, hereafter referred to as sh5, sh6, sh7, sh8, respectively), then expanded in LB medium with 100 µg/ml ampicillin and purified using Plasmid Maxi kit (Qiagen). The total RNAs were also subject to gene expression microarray analysis (see below) and each cell, two to three replicate samples were used in all samples. After normalization, data were transformed as the target mRNA signal relative to untreated control samples, and then plotted on GraphPad Prism 5.0 (GraphPad Software, Inc).

Lentivirus production and infection of NSCs

Human TRIPZ lentivirus inducible shRNAmir for NLGN4X plasmids stock was expanded in LB medium with 100 µg/ml ampicillin and purified using a Plasmid Maxi kit (Qiagen). The non-targeting TRIPZ lentivirus inducible shRNAmir control was made by integrating the non-silencing scrambled shRNA-mir sequence at MluI and XhoI restriction sites on a pTRIPZ vector, which does not match any known mammalian genes following the Open Biosystems shRNAmir manual. The TRIPZ vector was introduced into HEK-293T cells using polyethylenimine (Sigma) as the transfection reagent. Lentivirus was collected 48–72 h after transfection by a centrifuge at 28 000 rpm for 1.5 h at 4°C using Beckman Counter Optima L-100 XP ultracentrifuge. The lentivirus particles were resuspended in PBS and stored at −80°C. The lentivirus was titrated using HEK-293T cells following the Open Biosystems shRNAmir manual. For the infection of NSCs, we first cultured the NSCs in N2/B27 media. When the cells reached confluence, they were trypsinized and collected into a 1.5 ml Eppendorf tube with 150 µl media. The lentivirus was then added to the tube and incubated at 37°C for 1 h, then re-plated on a polyornithine and fibronectin double coated 6 cm plate with 2.5 ml media. The cells were incubated overnight and transferred to the fresh media in the morning next day. To induce the shRNAmir expression, 1 µg/µl doxycycline (Enzo Life Science) was added to the media. For long-term shRNAmir expression, doxycycline was refreshed every 2 days.

Quantitative real-time-PCR

Total RNA was extracted using RNeasy mini kit (Qiagen), in combination of RNase-free DNase (Qiagen) to remove the potential genomic DNA contamination. The cell lysate was homogenized by passing five times through a blunt 27-gauge needle. RNA concentration was determined by Nanodrop 1000 Spectrophotometer (Thermo Scientific). Reverse transcription was performed with 1.5 µg RNA using a ProtoScript M-MuLV First Strand cDNA Synthesis kit using random primers (New England Biolabs). The quantitative real-time PCR was carried on with gene-specific primers and IQ SYBR Green Supermix using Bio-Rad CFX96 system (Bio-Rad). The mRNA starting quantity was determined by the relative standard curve method using Bio-Rad CFX Manager software. The sequences of primers used were below:

- NLGN4X Forward: 5′-TGGACTCACAGGCGCCCCA-3′;
- Reverse: 5′-CGTGTGCCTGCAGCGACTCA-3′
- GAPDH Forward: 5′-CATGTTCGTACGGGTTGAA-3′;
- Reverse: 5′-AGTGATGGCCGACTGTGGT-3′

The primers for NLGN4X were designed following a previous study (44), which targets exon 7 of NLGN4X and have high specificity to avoid cross-targeting NLGN4Y. The expression of genes of interest was normalized to GAPDH in all samples. After normalization, data were transformed as the target mRNA signal relative to untreated control samples, and then plotted on GraphPad Prism 5.0 (GraphPad Software, Inc).

Whole-genome gene expression analysis

The total RNAs were also subject to gene expression microarray analysis at selected time points during neuronal differentiation. Briefly, total RNAs were isolated from cell lysate by RNeasy mini kit (Qiagen), and their qualities (OD260/280, RNA ratio, RNA integrity number) were determined by BioAnalyzer 2100 (Agilent). 250 ng of total RNA was subject to standard Illuma protocols for cRNA production, amplification, and then hybridized to Human12 Expression array. The data were imported into Illumina GenomeStudio, and quantile normalization was used to calculate gene expression values. For each time point and each cell, two to three replicate samples were used in the expression microarray analysis.
The gene expression values were loaded into the MeV software version 4.8 (45) for clustering and differential expression analysis. The HCT module was used to generate hierarchical clustering, with average linkage clustering on Euclidean distance. The DE genes across time series were identified by ‘time-course’ (31), an R package available in Bioconductor. For each individual time point, the ‘limma’ (46) R package from Bioconductor repository was used to detect the DE genes between the samples with and without knockdown.

The lists of significantly DE genes were analyzed by the DAVID web server (32) for enrichment of GO categories. Additionally, the results of gene expression (fold change and P-values) were overlaid with known protein–protein interactions from MiML database (34) in the Cytoscape software (47) for network-based analysis and visualization.

SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS
We thank members of the Wang laboratory and Lu laboratory for helpful suggestions and comments. We thank Dr Le Ma and Dan Gibson at the Zilkha Neurogenetic Institute at USC for their help and suggestions during the course of the experiments.

Conflict of Interest statement. None declared.

FUNDING
The study is supported by start-up funds from the Zilkha Neurogenetic Institute to K.W., and in part by NIH grant P30 MH08977 (PI: Levitt) and R01 HG006465 (PI: K.W.).

REFERENCES