Spatial and Temporal Mapping of De Novo Mutations in Schizophrenia to a Fetal Prefrontal Cortical Network

Suleyman Gulsuner,1,9 Tom Walsh,1,9 Amanda C. Watts,1,9 Ming K. Lee,1 Anne M. Thornton,1 Silvia Casadei,1 Caitlin Rippey,1 Hashem Shahin,1,3 Consortium on the Genetics of Schizophrenia (COGS),7,10 PAARTNERS Study Group,6,11 Vishwajit L. Nimgaonkar,4 Rodney C.P. Go,5 Robert M. Savage,6 Neal R. Swerdlow,7 Raquel E. Gur,8 David L. Braff,7 Mary-Claire King,1,* and Jon M. McClellan2

1Department of Medicine and Department of Genome Sciences
2Department of Psychiatry
University of Washington, Seattle, WA 98195, USA
3Department of Biology, Bethlehem University, Bethlehem, Palestine
4Departments of Psychiatry and Human Genetics, University of Pittsburgh, Pittsburgh, PA 15213, USA
5Department of Epidemiology
6Department of Psychiatry and Behavioral Neurobiology
University of Alabama at Birmingham, Birmingham, AL 35294, USA
7Department of Psychiatry, University of California at San Diego, La Jolla, CA 92093, USA
8Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA
9These authors contributed equally to this work
10A full list of Consortium on the Genetics of Schizophrenia (COGS) contributors may be found in the Supplemental Information
11A full list of PAARTNERS Study Group contributors may be found in the Supplemental Information
*Correspondence: mcking@uw.edu
http://dx.doi.org/10.1016/j.cell.2013.06.049

SUMMARY

Genes disrupted in schizophrenia may be revealed by de novo mutations in affected persons from otherwise healthy families. Furthermore, during normal brain development, genes are expressed in patterns specific to developmental stage and neuroanatomical structure. We identified de novo mutations in persons with schizophrenia and then mapped the responsible genes onto transcriptome profiles of normal human brain tissues from age 13 weeks gestation to adulthood. In the dorsolateral and ventrolateral prefrontal cortex during fetal development, genes harboring damaging de novo mutations in schizophrenia formed a network significantly enriched for transcriptional coexpression and protein interaction. The 50 genes in the network function in neuronal migration, synaptic transmission, signaling, transcriptional regulation, and transport. These results suggest that disruptions of fetal prefrontal cortical neurogenesis are critical to the pathophysiology of schizophrenia. These results also support the feasibility of integrating genomic and transcriptome analyses to map critical neurodevelopmental processes in time and space in the brain.

INTRODUCTION

Schizophrenia is a complex brain disorder characterized by aberrant perceptions, thought processes, and behavior. It is one of the leading causes of disability worldwide (Press, 2008) and is both highly heritable and highly genetically heterogeneous (McClellan and King, 2010). Individually rare gene-disrupting copy number variants (CNVs) contribute substantially to the disorder (Walsh et al., 2008; Xu et al., 2008; Stefansson et al., 2008; International Schizophrenia Consortium, 2008)—in particular, CNVs that disrupt genes involved with signaling, synaptic plasticity, and neurodevelopmental processes (Walsh et al., 2008; Kirov et al., 2012). These mutations are de novo or recent in origin, with de novo CNVs specifically enriched in patients with sporadic, rather than familial, illness (Xu et al., 2008; Kirov et al., 2012). Most rare gene-disrupting CNVs detected in affected persons are unique, although some recur independently at genomic hot spots, including on chromosomes 1q21.1, 3q29, 15q11.2, 15q13.1, 16p11.2, and 22q11.2 (McClellan and King, 2010) and at the neuropeptide receptor VIPR2 (Vacic et al., 2011).

Advances in genomics and neurobiology have enabled the next generation of these studies. The analysis of de novo mutations in schizophrenia can now include point mutations and small insertions and deletions (indels), as well as CNVs (Girard et al., 2011; Xu et al., 2011, 2012). In parallel, transcriptome databases of the human brain have been generated (Kang et al., 2011; Hawrylycz et al., 2012). Integrating these approaches has led to new insights into genetic aberrations in mental illness. For example, patterns of gene expression distinguishing frontal and temporal cortex have been shown to be significantly attenuated in brains of persons with autism (Voineagu et al., 2011). Furthermore, genes implicated in schizophrenia, including some harboring de novo mutations, have been shown to cluster in networks that are highly expressed in the brain, particularly during prenatal development (Xu et al., 2012; Gilman et al., 2012).
The goal of this project was to identify temporal and spatial processes of brain development critical to the neuropathogenesis of schizophrenia. Our approach was to examine the contribution of de novo mutations to schizophrenia and then to characterize functional networks in the brain of the genes harboring these mutations. The study design integrates genomics, transcriptomics, and proteomics. We first identified genes harboring de novo putatively damaging mutations in persons with schizophrenia from otherwise healthy families. We then evaluated the extent to which the proteins encoded by these genes interact and the extent to which they are transcriptionally coexpressed in different brain regions across developmental stages. The coexpression and protein interaction profiles were then used to generate a network whose interconnectedness was quantifiable by the number of connections (edges) between implicated genes. Across brain regions and developmental stages, we compared the interconnectedness of networks of genes harboring de novo mutations in probands versus networks from 10,000 simulations of genes harboring de novo mutations in unaffected persons.

RESULTS

Identification of De Novo Mutations by Whole-Exome Sequencing

De novo mutations were identified by exome sequencing of quads and trios comprised of a proband with schizophrenia, his/her unaffected parents, and, whenever available, an unaffected sibling. Families were selected for negative family history of severe mental illness other than the proband; that is, the probands were sporadic, or singleton, cases. The study design was based on quads because unaffected siblings provide ideal controls for ancestry, for nongenetic familial effects, and for sampling strategy (Sebat et al., 2007). A quad-based study design also optimizes experimental conditions because the affected and unaffected persons are sequenced in the same way and the same variant-assessment filters are applied to all subjects, blind to disease status. In each family, de novo variants were identified by comparing exome sequences of each child to his/her unaffected parents. Variants were experimentally validated as present in the child and absent from both parents, then classified as damaging or not damaging to protein function as defined in the Experimental Procedures. Exome sequencing of genomic DNA was carried out for 399 persons, including 105 probands with schizophrenia, 84 unaffected siblings, and their 210 unaffected parents (Table S1 and Figure S1 available online).

We first considered all de novo point mutations in probands and unaffected siblings, regardless of predicted effect on protein function. Of the 105 probands, 66 probands (63%) carried 96 de novo point mutations, 0.91 mutations per individual; of the 84 unaffected siblings, 39 siblings (46%) carried 66 such mutations, 0.79 mutations per individual (Table 1). The difference reflects a modest enrichment among probands in the likelihood of carrying at least one de novo mutation (X^2 = 5.10; p = 0.024). Combining silent, missense, and nonsense mutations yields 155 de novo point mutations in coding sequence, of which 68% (105/155) were nonsynonymous, which is similar to most previous studies of schizophrenia and autism (Xu et al., 2012; Iossifov et al., 2012; O’Roak et al., 2012b; Neale et al., 2012; Sanders et al., 2012) (Table S2). Mutation frequency was distributed as a Poisson distribution with mean 0.82 (Figure 1A).

The likelihood of carrying at least one de novo mutation was related to paternal age at conception (Kong et al., 2012). Among offspring of the youngest third of fathers, ages 19–28 years, 47% (33/70) carried at least one de novo mutation; among offspring of the oldest third of fathers, ages 33–45 years, 76% (50/66) carried at least one de novo mutation (X^2 = 13.96, p = 0.0009; Table S1 and Figure 1B). The relationship between paternal age and de novo mutations was the same for affected and unaffected siblings.

We next considered de novo mutations of all classes (point mutations, indels, and CNVs) predicted to damage protein function. Of the probands, 47/105 (45%) carried 57 de novo damaging mutations, 0.54 damaging mutations per individual; of the unaffected siblings, 25/84 (30%) carried 35 such mutations, 0.42 per individual (Table S3). This difference reflects a modest enrichment among probands in the likelihood of carrying at least one de novo damaging mutation (Figure 1C; odds ratio [OR] = 1.91, X^2 = 4.45; p = 0.035). Based on this sample, the proportion of schizophrenia attributable to de novo damaging mutations is 21% (see Experimental Procedures).

Schizophrenia Gene Networks

We next examined how genes harboring de novo putatively damaging mutations might reveal networks critical to development of schizophrenia. The analysis addressed three questions.

### Table 1. De Novo Mutations in Probands with Schizophrenia and Their Unaffected Siblings

<table>
<thead>
<tr>
<th>Mutation type</th>
<th>Probands (n = 105)</th>
<th>Unaffected Siblings (n = 84)</th>
<th>Total (n = 189)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Damaging^a</td>
<td>Benign</td>
<td>All de novo</td>
</tr>
<tr>
<td>Frameshift</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Nonsense</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>Missense</td>
<td>41</td>
<td>15</td>
<td>56</td>
</tr>
<tr>
<td>Silent</td>
<td>0</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>Splice</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>CNV</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>57</td>
<td>46</td>
<td>103</td>
</tr>
</tbody>
</table>

^aPredicted to be damaging based on criteria described in text.
First, do genes with de novo damaging mutations in schizophrenia form larger and/or more interconnected protein interaction networks than do genes with de novo damaging mutations in controls? Second, is there evidence for transcriptional coexpression of genes harboring these mutations in any brain regions across different stages of development? Third, combining protein interaction and transcriptome data, what networks are generated by the genes harboring de novo damaging mutations in schizophrenia, and which genes are included in these networks?

Overall, 54 different genes harbored de novo damaging point mutations, indels, or single-gene-disrupting CNVs in probands (Table S3A). Each of these genes was included in protein interaction and coexpression analyses. For de novo CNVs, genes were included if disrupted by breakpoints, but not if completely deleted or duplicated.

For comparison, we considered genes carrying de novo damaging mutations in healthy siblings from this study (Table S3B) and from four previous exome studies of schizophrenia and autism (Xu et al., 2012; Iossifov et al., 2012; O’Roak et al., 2012b; Sanders et al., 2012). From the five studies combined, 264 genes in unaffected siblings harbored de novo mutations predicted to be damaging based on the same criteria used to define de novo damaging events in our probands (Experimental Procedures and Table S4). We carried out the same analyses on 10,000 sets of 54 control genes, each randomly selected from the pool of 264 genes. We constructed a network for each set of 54 genes, yielding 10,001 networks, including the network based on case genes, and counted the numbers of nodes and edges for each.

Protein-Protein Interaction Networks of Genes with De Novo Mutations

We used the GeneMania physical interaction data set (Mostafavi et al., 2008) to examine whether genes disrupted by de novo damaging events in probands clustered in protein-protein
harboring de novo mutations in probands with schizophrenia. We then evaluated transcriptional coexpression across different brain regions at different developmental periods for genes harboring de novo mutations in probands with schizophrenia and for genes harboring de novo mutations in unaffected siblings. Using publicly available RNA sequencing (RNA-seq)-based gene expression levels from the BrainSpan Atlas of the Developing Human Brain (http://www.brainspan.org), we generated networks reflecting profiles of coexpressed genes.

Tissues from the BrainSpan Atlas were divided into four anatomic regions — frontal cortex (FC), temporal and parietal regions (TP), sensory-motor regions (SM), and subcortical regions (SC) — and three developmental stages — fetal (13–26 postconception weeks), early infancy to late childhood (4 months to 11 years), and adolescence to adulthood (13–23 years) (Table S5). The number of tissues included by the BrainSpan Atlas ranged from 5 to 6 for children and young adults to 11 to 12 for fetal tissues. For each pair of proband genes and each pair of control genes, we calculated Pearson correlation coefficients for RNA-seq expression levels across tissues within each brain region and each developmental stage. For each set of tissues, we therefore calculated correlation coefficients for (54 x 53)/2 = 2,860 gene pairs. Gene pairs with |R| ≥ 0.8 were defined as “connected” in subsequent network constructions. Network interconnectedness was measured by the number of connections, or edges (Barabási and Oltvai, 2004).

Numbers of edges in the networks of schizophrenia genes and control genes for each brain region by developmental stage are shown in Figure 3. The most striking result is that, in fetal frontal cortex, coexpression of genes harboring de novo damaging mutations in probands yields a network with significantly greater interconnectedness than networks derived from genes harboring de novo damaging mutations in controls (nominal p = 0.00006; corrected for multiple comparisons, p = 0.0007) (Figure 3 and Table S6). Connectedness was not significantly greater for case genes than for control genes at other developmental periods. There was no significant enrichment for connectedness in networks based on case genes harboring de novo benign mutations (Figure S2).

Given the significant interconnectedness of coexpressed case genes in fetal frontal cortex, we carried out the same analyses for four anatomical subregions of fetal frontal cortex: dorsolateral prefrontal cortex, ventrolateral prefrontal cortex, medial prefrontal cortex, and orbital frontal cortex. Coexpression profiles of genes mutant in schizophrenia differed most significantly from those of unaffected siblings in tissues from fetal dorsolateral prefrontal cortex (nominal p = 0.0001; corrected p = 0.0005) and fetal ventrolateral prefrontal cortex (nominal p = 0.0004; corrected p = 0.002) (Figure 4 and Table S7).

A Merged Network of Protein Interaction and Transcriptional Coexpression in the Dorsolateral and Ventrolateral Prefrontal Cortex during Fetal Development

We then merged the coexpression network derived for schizophrenia genes in fetal dorsolateral and ventrolateral prefrontal cortex with the PPI network derived for the same genes. The merged network encompasses 50 of the 54 genes harboring damaging de novo mutations in subjects with schizophrenia, with 126 connections among these genes (Figure 5). Genes in the network function in biological processes critical to neurogenesis and synaptic integrity, including axon guidance and

Figure 2. Interconnectedness of PPI Networks Based on Genes Harboring De Novo Mutations

(A and B) The 54 genes harboring de novo damaging mutations in schizophrenia cases yield a PPI network of 18 genes (A) and 23 interactions (B), indicated by the vertical dotted lines. For comparison, 10,000 sets of 54 genes were selected at random from 264 genes harboring de novo damaging mutations in unaffected siblings (controls). Each of these 10,000 gene sets also yields a PPI network. The numbers of genes and connecting links for each of these 10,000 control networks were plotted in (A) and (B), respectively. p values for enrichment of genes and connections in the case network were estimated by the proportion of control networks with more genes or interactions than the case network. The network based on cases is enriched both for number of genes (p = 0.002) and for number of connections between genes (p = 0.0005).

(C and D) Analogous networks based on genes harboring de novo benign events in cases and in controls. The networks based on genes with de novo benign events in cases do not differ from networks based on genes with de novo benign mutations in controls.
neuronal migration, signaling, cell proliferation, transport, synaptic transmission, and transcriptional regulation. Expression levels in fetal frontal cortex of individual network genes are provided in Figure 6. Expression levels of most of the genes are high early in fetal development, decline at the end of fetal development and in childhood, and gradually increase again in early adulthood. Expression of other genes either increases throughout fetal and adult development or remains at stable levels over time (Figure 6).

**DISCUSSION**

**The Prefrontal Cortex and Its Role in Schizophrenia**

Damaging de novo mutations in persons with schizophrenia converge in a network of genes coexpressed in the dorsolateral and ventrolateral prefrontal cortex during fetal development. These results support the neurodevelopmental hypothesis, proposed 25 years ago, that disruptions in fetal prefrontal cortical development are core mechanisms underlying schizophrenia (Weinberger, 1987; Marin, 2012). The specific timing and anatomic locale of the gene network and the biological functions of its component genes suggest that mutations leading to schizophrenia disrupt the orchestration of cortical neurogenesis.

Human brain evolution is distinguished by the remarkable expansion of the surface area and cytoarchitectural complexity of the neocortex, particularly in the prefrontal cortex. Healthy adult cortical function emerges from synchronized spatiotemporal neuronal migration during fetal development, which depends upon the complex regulation of cellular proliferation, signaling, and transcription pathways (Rakic, 2009). Many genes orchestrate these processes, the disruption of any one of which may result in loss of cortical integrity and neuronal homeostasis,
leading to neuropsychiatric disease (Ramocki and Zoghbi, 2008).

Prefrontal cortical networks organize input from other cortical and subcortical brain regions to plan and direct motor, cognitive, affective, and social behaviors (Kolb et al., 2012). The dorsolateral and ventrolateral prefrontal cortices are distinct neuroanatomic structures, with complementary neurocognitive functions. The dorsolateral prefrontal cortex is linked to cognitive control, which requires the coordination of attention and working memory to address changing environments and task demands. The ventrolateral prefrontal cortex is associated with reward learning and decision making, which depend upon the ability to gauge potential outcomes in the context of incoming stimuli as a guide to making choices (Gläscher et al., 2012). These executive functions endow humans with enormous capacity for problem solving and adaptation to a wide range of environments.

Several lines of evidence support the role of the prefrontal cortex in schizophrenia. Deficits in executive functions are characteristic of the illness and often predate its onset (Kalkstein et al., 2010). The prefrontal cortex matures late in brain development, with gray matter consolidation and myelination extending into early adulthood—the timing of which is consistent with ages of onset of schizophrenia (Gogtay et al., 2004). Neuroimaging studies of schizophrenia have demonstrated anatomical and functional deficits in the region (Shepherd et al., 2012). Postmortem brain studies have found a variety of cellular pathological findings in prefrontal cortical tissue (Eisenberg and Berman, 2010). Further, a substantial portion of prefrontal cortical neurons express monoaminergic receptors that are targets of antipsychotic medications (Masana et al., 2012). Our results suggest that deficits in prefrontal cortex executive functions are manifestations of schizophrenia, with full presentation occurring as key brain regions mature. Integrating genomic data with transcriptome network analyses can help pinpoint anatomical and developmental mechanisms that are disrupted in this process.

A key feature of schizophrenia appears to be brain disconnectivity—i.e., aberrant connections with and between different brain regions (Fitzsimmons et al., 2013). The prefrontal cortex is normally highly interconnected with other cortical, basal ganglia, and limbic regions. Human brain development depends upon the coordinated temporal and spatial regulation of neuronal migration in order to establish integrated cortical networks (Uhlhaas and Singer, 2012). Disruptions of individual components within developing neuronal networks may impact the functional integrity of larger dynamic brain systems (Vidal et al., 2011). Therefore, our results do not preclude the involvement of other brain regions in schizophrenia nor imply that damaging mutations have an impact only during fetal development. Genes harboring de novo damaging mutations in schizophrenia are...
expressed in multiple brain regions (and other tissues) across the lifespan. Neuroimaging studies of schizophrenia describe multiple anatomical abnormalities, including decreased total brain volumes and losses of gray matter in the anterior cingulate, frontal and temporal lobes, hippocampus, amygdala, thalamus, and insula (Shepherd et al., 2012).

Roles in Brain Development of Genes with De Novo Mutations in Schizophrenia

Genes implicated in schizophrenia and autism function in processes important to fetal brain development, including axon guidance, neuronal cell mobility, synaptic function, and chromosomal remodeling (Gilman et al., 2012; Xu et al., 2012). Most of the genes with damaging de novo mutations in probands are known to play critical roles in developing brain, including neuronal migration (ITGA3, LAMA2; Anton et al., 1999; Relucio et al., 2009), axon guidance (SLIT2, SLIT3; Zhang et al., 2012), signal transduction and cell adhesion (CELSR2; Shima et al., 2007), cellular proliferation (MKI67; Duchrow et al., 1995), transcriptional regulation of glial and neuronal development (DTX1, HIF1A, NCO2; Patten et al., 2006; Pacary et al., 2007; Jepsen et al., 2007), and neurotransmitter signaling and synaptic transmission (ADCY9, CACNA1I, GLS, SLC18A2; Martin et al., 2007; Simons et al., 2013).

Several network genes with de novo mutations in probands function in neurotransmitter pathways. ADCY9 is involved in neuronal signaling in glutamate and GABA pathways (Martin et al., 2007). SLC18A2 encodes the vesicular monoamine transporter VMAT2, which is responsible for pumping serotonin, dopamine, norepinephrine, epinephrine, and histamine into vesicles for use in synaptic transmission (Simons et al., 2013). GLS catalyzes the hydrolysis of glutamine to glutamate. Studies of GLS knockout mice suggest that dysregulation of glutaminase function could play a role in schizophrenia and that medications inhibiting glutaminase may have a therapeutic benefit for the illness (Gaetan-Salomon et al., 2012).

CACNA1I, the only gene disrupted by damaging de novo mutations in more than one patient, is a brain-specific, T type calcium channel involved in regulation of differential patterns of neuronal firing in the thalamus, striatum, nucleus accumbens, and prefrontal cortex (Chemin et al., 2002; Talley et al., 1999; Yunker et al., 2003). In rats, T type calcium channel antagonists reduced amphetamine-induced psychomotor activity and glutamate release, patterns predictive of antipsychotic effects (Uslaner et al., 2012). The therapeutic activity of antipsychotic medications, including clozapine, may stem in part from the inhibition of T type calcium channels (Choi and Rhim, 2010).

Our results provide further evidence for the genetic complexity and heterogeneity of schizophrenia. Several network genes have been implicated previously in schizophrenia and other neuropsychiatric disorders, including LAMA2 (Xu et al., 2012), DNTM1 (Costa et al., 2007), and SLC25A1 and CLTCL1, both of which are located in the chromosome 22q11.2 deletion region that confers substantial risk for schizophrenia. Expression profiles of CELSR2, HIVEP1, ITGA3, SERPINI1, and SLIT2 differed between neurons induced from stem cells of persons with schizophrenia versus controls (Brennand et al., 2011). Nine patients carried de novo damaging mutations in more than one gene (Table S8), suggesting the possibility of oligogenic influences on disease (Girirajan et al., 2012).

Figure 5. Network of Genes Harboring Damaging De Novo Mutations in Schizophrenia that Are Coexpressed in Fetal Dorsolateral and Ventrolateral Prefrontal Cortex

Merging data from networks of PPI network and of coexpression in fetal dorsolateral (DFC) and fetal ventrolateral (VFC) yields a network of 50 genes (nodes) with 126 connections (edges). Red lines represent coexpressed genes, and blue lines represent physical connections between protein products of the genes.
Causality

Given the degree of genetic heterogeneity of mental illness, establishing causal links between individual genes and schizophrenia will depend on both biological and statistical evidence. The sample size of this study both limits the power of network analyses and is not sufficient to prove causality of any one gene. Resequencing studies of candidate genes in very large cohorts may provide such statistical evidence, as has been demonstrated for autism (O’Roak et al., 2012a).

The goal of modern genetic studies of mental illness is to identify key pathways that, when disrupted, lead to disease, realizing that mutation of any of multiple genes may disrupt the same pathway (McClellan and King, 2010). Our goal for this project has been to use computational tools to define gene networks based on mutations in affected individuals, thereby providing empiric evidence that the networks are related to schizophrenia. The next step is to discover the biological pathways that underlie the computational networks (Konopka et al., 2012). The principal limitation in doing so is incomplete understanding of normal brain development. As pathways critical to neurodevelopment are revealed by mutations that disrupt them, it will become possible to carry out statistical testing for mutations leading to loss of integrity of a pathway rather than of a single gene.

Our results suggest that disruption of genes critical to neurogenesis of the prefrontal cortex are also critical to pathogenesis of schizophrenia. These findings have implications for the development of more biologically specific treatments of the illness (Insel, 2009). Given ethical, clinical, and neurobiological limitations to in utero interventions, treatment research inevitably focuses on compensatory strategies that address downstream effects on brain development. Medications targeting glutamate and T type calcium channel pathways are potential therapeutic mechanisms suggested by mutant genes in our patients. Large-scale collaborative efforts in genomics and neurobiology are needed to reveal additional therapeutic targets based on the genes and biological mechanisms underlying brain circuitry (Akil et al., 2010).

EXPERIMENTAL PROCEDURES

Exome Sequencing

From genomic DNA of the 399 participants, library construction, exome capture, and sequencing were carried out as previously described (Walsh et al., 2010) with some modifications. Genomic DNA libraries were captured by SeqCap EZ Exome v2 (Nimblegen) pools and then hybridized to bio-tynylated capture probes and sequenced to median depth of coverage of at least 100×. Paired-end sequence reads (2 × 101 bp) were collected, demultiplexed, filtered for quality, aligned to exome targets using BWA v0.6.1-r104, realigned with GATK v1.5-21, and then genotypes were called with MAQ v0.7.1, SAM tools v0.1.18, and GATK v1.5-21. For individual samples, on average, 93% of targeted coding regions were covered at ≥ 10× and 84% at ≥ 20×. For joint coverage of trios (Iossifov et al., 2012), on average, 89% of targeted coding regions were covered at ≥ 10× and 77% at ≥ 20× (Figure S1).

Artifacts were excluded by comparison with 800 exomes previously sequenced on the same instrument in our lab. Common variants, with MAF > 0.001 in dbSNP v137 or in the NHLBI Exome Sequencing Project (2012), were also excluded. To detect point mutations and small insertions and deletions (1–23 bp in length), variants were retained at sites with at least 30% of total reads representing the variant after adjusting by individual inspection for genomic complexities (e.g., multiple pseudogenes of CHEK2). No constraint was put on the proportion of variant reads in parental samples. Candidate de novo variants were tested by diagnostic PCR and Sanger sequencing. To detect CNVs, exome-based calls were made by CoNIFER (Krumm et al., 2012) and then validated by TaqMan in DNA from the proband or sibling and from the parents. For de novo CNVs not flanked by segmental duplications, exact genomic breakpoints were determined with diagnostic PCR primers and Sanger sequencing. De novo variants included in subsequent analyses were defined as variants of any type demonstrated by both exome sequencing and Sanger sequence validation to be present in a proband or a sibling and absent from both parents.

The predicted functional impact of each candidate de novo missense variant was assessed with in silico tools. Variants that met any of the following criteria were considered potentially damaging: frameshift; nonsense (unless the introduced nonsense was polymorphic or the natural stop in other species); missense mutation with Polyphen score ≥ 0.90 and/or SIFT p < 0.05 and/or Grantham score > 100 (Adzhubei et al., 2010; Kumar et al., 2009; Grantham, 1974); CNV disrupting a gene or completely deleting a gene or genes; or splice or enhancer alteration predicted by one or more algorithms and with GERP score > 5.0.

Proportions of probands and unaffected siblings with point mutations and with damaging mutations of all classes, were compared by chi-square tests using standard procedures. Attributable risk was estimated by p(OR – 1)/ [p(OR – 1) + 1] (Schlesselman, 1982), where p is the proportion of controls (unaffected siblings) with a de novo damaging mutations and OR is the odds ratio for harboring a de novo damaging mutation in cases versus controls.

Network Construction

Physical interactions from the Homo sapiens database were collected from GeneMANIA 2011-08-03 release of 150 studies (Mostafavi et al., 2008). There were 753,875 physical interactions in 12,010 proteins after removing duplicates. A physical interaction network was created for the schizophrenia genes and for each simulation gene set. Degree is defined as the number of neighbors of each node (Barabási and Oltvai, 2004; Watts and Strogatz, 1998).

Coexpression-based network analyses were conducted using existing data sets in the BrainSpan: Atlas of the Developing Human Brain. Normalized gene expression levels for 26 different brain tissues in 10 different developmental periods were obtained from the BrainSpan RNA-seq data set v3 of developing human brain. Tissue qualification, processing and dissection, and experimental and bioinformatics procedures are described in the technical white paper of the data set (http://help.brain-map.org/display/devhumanbrain/Documentation) (Supplemental Information). Network figures were created using Cytoscape 2.8.3 (Smoot et al., 2011).

To estimate the statistical significance of networks created from PPI and coexpression data, we generated a data set of de novo predicted damaging mutations among unaffected siblings by pooling results from unaffected siblings in this study with results from unaffected siblings reported in autism and schizophrenia exome studies (Krumm et al., 2012; Iossifov et al., 2012; O’Roak et al., 2012b; Sanders et al., 2012). We excluded variants observed in both probands and unaffected siblings, variants without a defined nucleotide change, and variants reported by the authors to have failed validation. From all unaffected siblings, 547 de novo events were included, for which functional consequences were predicted using the same criteria applied to probands in this study (Supplemental Information). A total of 276 events (2 frameshift, 21 nonsense, 248 missense, and 5 splice) in 267 genes were predicted to be damaging. Three genes not represented in either the GeneMania database or the BrainSpan Atlas were removed from analysis. Using the remaining 264 genes with damaging de novo events in the pooled group of unaffected siblings, we created 10,000 random gene sets that each contained 54 different genes, corresponding to the number of genes harboring predicted damaging events in probands.

PPI and coexpression networks were constructed for each gene set. The mean numbers and SD of nodes and edges in the simulated networks were used to generate normal distributions for PPI networks in each of the four anatomic regions and three developmental stages. Significance levels of network analyses and is not sufficient to prove causality of any one gene. Resequencing studies of candidate genes in very large cohorts may provide such statistical evidence, as has been demonstrated for autism (O’Roak et al., 2012a).

The goal of modern genetic studies of mental illness is to identify key pathways that, when disrupted, lead to disease, realizing that mutation of any of multiple genes may disrupt the same pathway (McClellan and King, 2010). Our goal for this project has been to use computational tools to define gene networks based on mutations in affected individuals, thereby providing empiric evidence that the networks are related to schizophrenia. The next step is to discover the biological pathways that underlie the computational networks (Konopka et al., 2012). The principal limitation in doing so is incomplete understanding of normal brain development. As pathways critical to neurodevelopment are revealed by mutations that disrupt them, it will become possible to carry out statistical testing for mutations leading to loss of integrity of a pathway rather than of a single gene.

Our results suggest that disruption of genes critical to neurogenesis of the prefrontal cortex are also critical to pathogenesis of schizophrenia. These findings have implications for the development of more biologically specific treatments of the illness (Insel, 2009). Given ethical, clinical, and neurobiological limitations to in utero interventions, treatment research inevitably focuses on compensatory strategies that address downstream effects on brain development. Medications targeting glutamate and T type calcium channel pathways are potential therapeutic mechanisms suggested by mutant genes in our patients. Large-scale collaborative efforts in genomics and neurobiology are needed to reveal additional therapeutic targets based on the genes and biological mechanisms underlying brain circuitry (Akil et al., 2010).
Genes in the merged network

Other genes with de novo damaging mutations

Tissue
- Dorsolateral prefrontal cortex
- Medial prefrontal cortex
- Orbital prefrontal cortex
- Ventrolateral prefrontal cortex

(legend on next page)
SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, two figures, and eight tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2013.06.049.

ACKNOWLEDGMENTS

We thank Jim Watson for advice and support throughout the project. This work was supported by a gift from Steve and Connie Lieber, by NARSAD grants to M.-C.K. and to T.W., and by NIMH grants R01MH083898, R01MH083849, R01MH083756, R01MH084071, and R03TW008696. Grants supporting the collection of clinical material for the NIMH repository are indicated in the Supplemental Information.

Received: April 3, 2013
Revised: June 3, 2013
Accepted: June 28, 2013
Published: August 1, 2013

REFERENCES


