Title

Formation of Chimeric Genes by Copy-Number Variation as a Mutational Mechanism in Schizophrenia

Journal

American Journal of Human Genetics, 93(4)

ISSN

0002-9297

Authors

Rippey, Caitlin
Walsh, Tom
Gulsuner, Suleyman
et al.

Publication Date

2013-10-01

DOI

10.1016/j.ajhg.2013.09.004

Peer reviewed

Formation of Chimeric Genes by Copy-Number Variation as a Mutational Mechanism in Schizophrenia

Caitlin Rippey,[1,*] Tom Walsh,[1] Suleyman Gulsuner,[1] Matt Brodsky,[2] Alex S. Nord,[1,4] Molly Gasperini,[1] Sarah Pierce,[1] Cailyn Spurrell,[1] Bradley P. Coe,[1] Niklas Krumm,[1] Ming K. Lee,[1] Jonathan Sebat,[3] Jon M. McClellan,[2] and Mary-Claire King[1]

Chimeric genes can be caused by structural genomic rearrangements that fuse together portions of two different genes to create a novel gene. We hypothesize that brain-expressed chimeras may contribute to schizophrenia. Individuals with schizophrenia and control individuals were screened genome wide for copy-number variants (CNVs) that disrupted two genes on the same DNA strand. Candidate events were filtered for predicted brain expression and for frequency < 0.001 in an independent series of 20,000 controls. Four of 124 affected individuals and zero of 290 control individuals harbored such events ($p = 0.002$); a 47 kb duplication disrupted *MATK* and *ZFR2*, a 58 kb duplication disrupted *PLEKHD1* and *SLC39A9*, a 121 kb duplication disrupted *DNAJA2* and *NETO2*, and a 150 kb deletion disrupted *MAP3K3* and *DDX42*. Each fusion produced a stable protein when exogenously expressed in cultured cells. We examined whether these chimeras differed from their parent genes in localization, regulation, or function. Subcellular localizations of DNAJA2-NETO2 and MAP3K3-DDX42 differed from their parent genes. On the basis of the expression profile of the *MATK* promoter, *MATK-ZFR2* is likely to be far more highly expressed in the brain during development than the *ZFR2* parent gene. *MATK-ZFR2* includes a *ZFR2*-derived isoform that we demonstrate localizes preferentially to neuronal dendritic branch sites. These results suggest that the formation of chimeric genes is a mechanism by which CNVs contribute to schizophrenia and that, by interfering with parent gene function, chimeras may disrupt critical brain processes, including neurogenesis, neuronal differentiation, and dendritic arborization.

Introduction

Structural rearrangements arise continuously in the human genome. Although many are benign, a subset contributes to disease, particularly neurodevelopmental illnesses such as schizophrenia (MIM 181500).[1–4] The mechanisms by which rare structural variants lead to disease are as heterogeneous as the mutations themselves and likely include gene dosage effects (both overexpression[5] and haploinsufficiency[6]), unmasking of recessive alleles by a deletion on one chromosome,[7] and epistatic interactions between multiple events.[8] In addition, some rearrangements result in the formation of chimeric genes. Gene fusions are uniquely suited to play crucial roles in both evolution and disease because they can differ from parent genes in localization, regulation, and/or function.[9] Adaptive chimeric genes have been incorporated into genomes across many phyla, including a handful identified in the hominid lineage.[10,11] However, the majority of chimeras arising de novo in the genome are likely detrimental,[12] as is the case when oncogenic fusions such as BCR-ABL (MIM 151410)[13] arise somatically.

Less is known about the role of germline chimeras in human disease. Two prior reports of chimeric genes in individuals with schizophrenia suggest that chimerism may play a role in neuropsychiatric illness. We previously reported a rare deletion in an individual with juvenile onset schizophrenia that resulted in a chimera of *SKP2-SLC1A3* (MIM 601436 and 600111). This chimeric gene was predicted to alter the expression and function of *SLC1A3*, which codes for GLAST, a glial glutamate transporter known to regulate neurotransmitter concentration at excitatory synapses.[1] In addition, the well-documented t(1;11) translocation which segregates closely with psychiatric illness, including schizophrenia, in a large Scottish kindred disrupts *DISC1* (MIM 605210) on chromosome 1 and the noncoding RNA *DISC1FP1* (also called *Boymaw*) on chromosome 11.[4] Chimeric transcripts fusing these two genes are present in carrier lymphoblasts,[14] and a recent study demonstrated that the resulting fusion proteins have altered localization and lead to severe mitochondrial dysfunction.[15] A systematic survey of copy-number variation in persons with autism (MIM 209850) revealed potential candidate chimeric events but no difference in frequency between affected and control individuals.[16]

In this study, we seek to systematically characterize the role of chimeric genes in schizophrenia by identifying rare, germline chimeras in individuals with schizophrenia and control individuals. We hypothesize that brain-expressed chimeric genes in these individuals contribute to schizophrenia by disrupting critical neuronal pathways involving the parent genes.

[1]Departments of Medicine and of Genome Sciences, University of Washington, Seattle, WA 98195, USA; [2]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA 98195, USA; [3]Departments of Psychiatry, Department of Cellular and Molecular Medicine, Institute for Genomic Medicine, University of California, La Jolla, CA 92093, USA
[4]Present address: Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
*Correspondence: cfields@uw.edu
http://dx.doi.org/10.1016/j.ajhg.2013.09.004.

Subjects and Methods

Subjects

Our case series comprised 124 individuals with schizophrenia or schizoaffective disorder whose diagnostic status was confirmed by diagnostic interview with the Structured Clinical Interview for DSM-IV disorders (SCID) for adults or the KID-SCID for youths. The average age of onset of schizophrenia was 19 years (SD = 5.0 years, range = 9–35 years). When possible, we enrolled family members of affected individuals, but, for most affected individuals, family member contact was not possible because of institutionalization. DNA was extracted from whole blood for array comparative genomic hybridization (aCGH), and lymphoblastoid cell lines were also generated in our lab for each individual. These cell lines were used for RNA and protein analyses. The control group comprised 122 individuals without mental illness from National Institute of Mental Health Genomics Research and Repository Distribution 5 and 168 individuals without mental illness who agreed that their DNA may be used as controls for projects in our laboratory. All control individuals were unrelated to each other, older than 35 years, and matched to affected individuals by self-identified race. The project was approved by the appropriate University of Washington and Washington state ethics committees, and informed consent was obtained for every participant.

Detection of Copy-Number Variants

DNA from all 124 individuals with schizophrenia and 122 of the control individuals was hybridized to NimbleGen HD2 microarray slides (average probe density = 1 probe per kb) against a well-characterized reference subject (SKN1). aCGH data were normalized, and copy-number variants (CNVs) were called with a sliding-window algorithm with a ten-probe minimum size threshold,[17] which allowed the detection of events greater than 40 kb. Candidate CNVs were deletions and duplications disrupting two genes transcribed on the same strand. DNA from 168 additional control individuals was screened by exome sequencing. Library construction, exome capture, and sequencing were carried out as previously described.[18] CNVs were derived with CoNIFER.[19] It was previously determined that exome sequencing on our laboratory platform is as sensitive as NimbleGen HD2 aCGH in detecting CNVs that would be candidates for chimeric events.[19] CNVs detected by either platform were defined as rare if no event with greater than 60% overlap was present on the Database of Genomic Variants[20] at a frequency greater than 0.001, as previously described.[1] Furthermore, candidate chimeric events were screened against CNVs detected in 19,585 individuals (Cooper et al., 2011,[21] and data not shown), and events present at a frequency 0.001 or greater were excluded. To determine genomic breakpoints, we performed long-range PCR amplification on genomic DNA with TaKaRa LA Taq as described in the manufacturer's protocol. PCR products were Sanger sequenced with breakpoint primers and additional sequencing primers as necessary in order to reach breakpoints. All primer sequences are included in Table S1 (available online).

Bioinformatic Analysis of Gene Expression in Human Brain

For analysis of *MATK* and *ZFR2* expression across developmental time points in the human brain, we used RNA sequencing (RNA-seq) data from the BrainSpan Atlas of the Developing Human Brain. Normalized gene expression levels for 26 different brain tissues in ten different developmental periods were obtained from the BrainSpan RNA-seq data set version 3. Tissue qualification, processing and dissection, and experimental and bioinformatics procedures are described in the technical white paper of the data set. Normalized values of expression of parent genes were obtained from 149 frontal cortex samples at different developmental ages. Developmental ages were grouped into ten developmental periods: early fetal (8–12 weeks gestation), early-mid fetal (13–18 weeks gestation), late-mid fetal (19–24 weeks gestation), late fetal (25–38 weeks gestation), early infancy (birth–5 months), late infancy (6–18 months), early childhood (19 months–5 years), late childhood (6–11 years), adolescence (12–19 years), and adulthood (20–60+ years). Median expression levels for each gene in each period were calculated and plotted.

Transcript Analysis and Expression Constructs

First strand cDNA synthesis of subject lymphoblast cell line RNA or of total RNA from different human tissues (Clontech Laboratories) was performed with SuperScript III (Life Technologies) primed with random hexamers and/or gene-specific primers. Chimeric transcripts were amplified with PCR primers targeted to flanking exons and Sanger sequenced.

To generate expression constructs, we amplified *MAP3K3*, *DDX42*, *DNAJA2*, *MAP3K3-DDX42* transcripts 1 and 2, *PLEKHD1-SLC39A9* transcripts 1 and 2, and *DNAJA2-NETO2* transcripts 1 and 2 from subject lymphoblast cDNA, *PLEKHD1* and *ZFR2* from human brain cDNA (Clontech), *SLC39A9* from a commercial cDNA clone (Origene), and *MATK* fragment from subject genomic DNA. *MAP3K3-DDX42* and *PLEKHD1-SLC39A9* chimeras and parents were cloned into TOPO-TA pcDNA3.1 C-terminal V5- and His-tagged mammalian expression vector (Life Technologies). *DNAJA2-NETO2* chimeras and parents were cloned into an N-terminal V5-tagged pcDNA3 expression vector (a gift of Wendy Roeb) with the Gateway system (Life Technologies). *ZFR2* and *MATK-ZFR2* constructs were assembled with the Gibson method.[22] All inserts were fully Sanger sequenced and free of nonsynonymous PCR errors.

Cell Culture and Transfections

Human embryonic kidney (HEK) 293 cells (ATCC) were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin and streptomycin. Cortical neurons from P0 C57BL/6 mice were plated at a density of ~25,000–50,000 per well onto a bed of confluent astrocytes on poly-D-lysine- and collagen-treated coverslips. Neurons were cultured in DMEM with B27 and 10% horse serum. Mitotic inhibitor (5 μM 5-fluoro-2′-deoxyuridine and 12.5 μM uridine) was added at 2 days in vitro (DIV). Experiments were performed according to the guidelines for the care and use of animals approved by the Institutional Animal Care and Use Committee at the University of Washington. Cells were transfected with Lipofectamine 2000 (Life Technologies) according to the manufacturer's protocol. Cortical cells were cotransfected at 9 or 10 DIV with GFP pCMV (0.6 μg) and experimental construct or empty vector (2.4 μg). Media was changed after 4 hr, and transfections proceeded for 24 hr.

Western Blotting and Immunoprecipitation

HEK 293 cells were lysed in radio-immunoprecipitation assay buffer (50 mM Tris-HCL [pH 7.4], 1% NP-40, 0.25% sodium

deoxycholate, 150 mM NaCl, and 1 mM EDTA) with complete mini protease inhibitor cocktail tablets (Roche) and Halt Protease and Phosphatase Inhibitor Single-Use Cocktail (Thermo Fisher Scientific) and quantitated with the Bradford method. Western blots were imaged with an Odyssey Infrared Imager system (LI-COR Biosciences). For analysis of MAPK activation, blots were probed with phosphospecific antibodies and stripped and reprobed with antibody to total JNK, p38, or actin. Primary antibodies used were mouse anti-V5 (1:5,000, Life Technologies), rabbit anti-β-actin (1:1,000, Santa Cruz Biotechnology), mouse anti-β-actin (1:10,000, Sigma-Aldrich), rabbit anti-MEK5 (1:10,000, Abcam), rabbit anti-phospho-MEK5 (1:500, Santa Cruz Biotechnology), rabbit anti-p-ERK1/2 (1:1,000, Cell Signaling Technology), mouse anti-phospho-JNK (1:500, Santa Cruz Biotechnology), rabbit anti-phospho-p38 MAPK (1:1,000, Cell Signaling), rabbit anti-JNK (1:1,000, Cell Signaling), and rabbit anti-p38α (1:500, Santa Cruz Biotechnology). Experiments were performed in triplicate, and paired two-tailed t tests were used for statistical analysis. For immunoprecipitation, transfected HEK 293 cells were lysed with NP-40 buffer (1% NP-40, 0.15 M NaCl, and 0.01 M Sodium phosphate) with same protease and phosphatase inhibitors and quantitated as above. Lysates were immunoprecipitated with Protein G Dynabeads (Life Technologies) conjugated to mouse anti-V5 (1:100, Life Technologies) according to the manufacturer's protocol. Two independent replicates were performed.

Immunocytochemistry

HEK 293 cells or cultured neurons were fixed with 4% paraformaldehyde for 20 min at room temperature. Cells were permeabilized with 0.5% Triton X-100 in PBS for 10 min at room temperature and blocked in 5% normal goat serum (0.25% Triton X-100 in PBS for 1 hr at 4°C). Cultured cells were incubated overnight at 4°C with mouse monoclonal V5 primary antibody (1:200, Life Technologies), rinsed four times in PBS, and incubated in Alexa 568-conjugated goat secondary antibody (1:400, Life Technologies) for 1–2 hr at room temperature. Then, cells were washed three times in PBS and mounted on slides with ProLong Gold Antifade media containing DAPI (Life Technologies).

Microscopy

Immunofluorescence images were acquired on a Zeiss 510 Meta Confocal Microscope with 40× or 63× oil immersion objectives. Gain and offset were adjusted for each image in order to optimize dynamic range. For localization in HEK 293 cells, we measured the fluorescence intensity of DAPI and transfected protein along a 30 μm line drawn through the center of each transfected cell in a 40× image with Plot Profile function in ImageJ. Dendritic localization was analyzed in images taken with a 63× objective and 3×–4× zoom. Images were acquired in z series (0.4 μm steps), rendered with maximum intensity projections, and analyzed in ImageJ. Profiles of GFP and Alexa 568 fluorescence were made with the Plot Profile function for a rectangular selection around the longest straight segment of a dendrite. Branching sites were indicated manually on the basis of GFP imaging of dendrites. To correct for variation in background staining intensity and transfection efficiency, we normalized fluorescence values, and mean intensities at and between branching sites were calculated. To correct for increased fluorescence at branch sites due to the widening of dendrites, we divided normalized mean Alexa 568 fluorescence intensity by normalized mean GFP fluorescence intensity. Data presented are from three independent experiments.

Results

Discovery of Chimeric Genes

We screened DNA from 124 affected individuals with schizophrenia with aCGH then scanned genome wide for CNVs longer than 30 kb. CNVs in DNA of 290 matched controls were detected by either the same aCGH platform (n = 122) or exome sequencing (n = 168; CNVs derived with CoNIFER[19]). We eliminated CNVs present in the Database of Genomic Variants.[20] From the remaining events, we selected those predicted to delete or duplicate both the 5′ end of one gene and the 3′ end of another. Breakpoints of all candidate chimeric events were obtained by PCR with diagnostic primers followed by Sanger sequencing. To detect chimeras formed by inverted tandem duplications, we designed breakpoint primers in order to test for an inverted orientation when a duplication disrupted two genes on opposite strands. This approach did not yield additional putative chimeras.

Several filters were applied to each validated genomic chimera: (1) the 5′ parent gene was expressed in brain, as determined by RT-PCR or by published report, (2) the event had a frequency <0.001 in CNVs from 19,585 controls, and (3) the predicted protein product included >5% of the protein coding sequence of at least one gene.

Four candidate chimeric events passed through all the filters, all of which were in affected individuals (Figure 1). These are fusions of Mitogen-activated protein kinase kinase kinase 3 (*MAP3K3* [MIM 602539]) to DEAD box protein 42 (*DDX42* [MIM 613369]; *MAP3K3-DDX42*), DNAJ homolog subfamily A member 2 (*DNAJA2* [MIM 611322]) to Neuropilin and tolloid-like 2 (*NETO2* [MIM 607974]; *DNAJA2-NETO2*), Pleckstrin homology domain-containing family D member 1 (*PLEKHD1*) to Solute carrier family 39 member 9 (*SLC39A9*), and Megakaryocyte-associated tyrosine kinase (*MATK* [MIM 600038]) to Zinc finger RNA-binding protein 2 (*ZFR2*) (*MATK-ZFR2*). In total, we detected four chimeric events meeting our criteria in 124 affected individuals and zero in 290 control individuals (p = 0.002 by two-tailed Fisher's exact test). The characteristics of these candidate chimeric genes are outlined in Table 1, and excluded events are described in Table S2.

Evaluation of Transcripts of Parent and Chimeric Gene mRNA

Next, we obtained lymphoblast RNA from each individual and designed primers in order to detect the predicted chimeric mRNA. We detected *MAP3K3-DDX42*, *DNAJA2-NETO2*, and *PLEKHD1-SLC39A9* chimeric transcripts (Figure 2A). *MAP3K3-DDX42* had two stable transcripts. *MAP3K3-DDX42* transcript 2 is spliced at noncanonical sites in the middle of exons, the donor site being located in *MAP3K3* exon 4 and the acceptor site in *DDX42* exon 17. Both *MAP3K3-DDX42* chimeras are in frame.

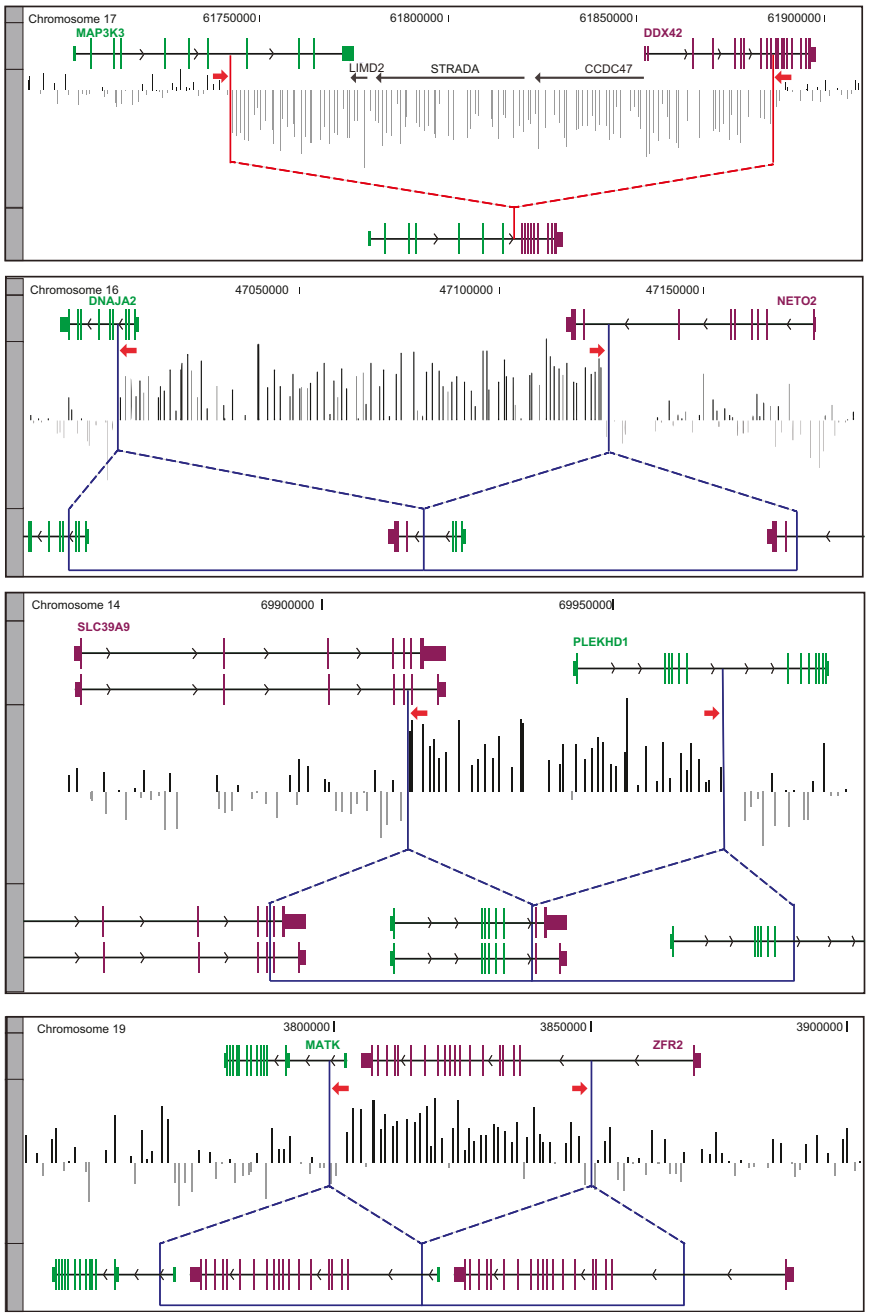

Figure 1. CNVs Resulting in the Formation of Chimeric Genes in Individuals with Schizophrenia

Deletion (red dashed lines) and duplications (blue dashed lines) result in the formation of *MAP3K3-DDX42*, *DNAJA2-NETO2*, *PLEKHD1-SLC39A9*, and *MATK-ZFR2* chimeric genes. For each event, we indicate hg19 genomic coordinates, impacted genes with direction of transcription (5′ parent genes are green, 3′ parent genes are purple, and deleted genes are black), regions targeted by PCR primers for breakpoint validation (red arrows), histograms of aCGH hybridization *Z* scores, and resulting chimeric gene.

Two *DNAJA2-NETO2* transcripts resulted from alternate splicing of *NETO2* exon 8. Both are out of frame, adding 11 (transcript 1) or 8 (transcript 2) frame-shifted codons before a premature stop. *PLEKHD1-SLC39A9* also had two transcripts, resulting from the inclusion of alternate 3′ UTRs of *SLC39A9*.

The chimeric *MATK-ZFR2* transcript could not be amplified from lymphoblasts of the duplication carrier, consistent with the expression profiles of the two parent genes, both of which were expressed only in the adult and fetal brain (Figure 2B). Evaluation of the parent gene *ZFR2* revealed a second transcript characterized by the deletion of *ZFR2* exon 15 (95 base pairs) and, therefore, an altered reading frame. Both *ZFR2* transcripts were consistently detected in all brain regions (Figure 2C) and, therefore, may be included in brain-expressed *MATK-ZFR2* chimeras.

Functional Evaluation of Fusion Proteins

All four chimeric genes produced stable proteins when transiently expressed in mammalian cells.

MAP3K3-DDX42

The *MAP3K3-DDX42* chimera is the only chimera resulting from a deletion, so the individual's genome includes only a single functional copy of each parent gene. Furthermore, the 150 kb event fully deletes three interstitial genes (*LIMD2*, *STRADA* [MIM 608626], and *CCDC47*), one of which—*STRADA*—is known to regulate dendritic and axonal outgrowth.[23] In addition, the *MAP3K3-DDX42* chimera contains substantial coding sequence from both constituent genes; therefore, the fusion protein could impact two different pathways. For the present study, we focused on the function of fusion proteins in the well-known ERK5 pathway (encoded by *MAPK7* [MIM 602521), which includes the 5′ gene, *MAP3K3*.[24]

The two MAP3K3-DDX42 fusion proteins were 75.1 and 12.9 kD (Figures 3A and 3B). In transfected HEK 293 cells, localizations of both fusion isoforms differed from those of their parent proteins (Figure 3C). MAP3K3 was predominantly cytoplasmic, and DDX42 was predominantly nuclear, consistent with previously reported localizations.[25; 26] In contrast, the localization of transfected MAP3K3-DDX42 isoforms 1 and 2 was highly variable: preferentially or exclusively in the nucleus in some cells, in the cytoplasm in others, and, in still others, diffused equally throughout the cell (Figure 3D).

Table 1. Chimeric Genes in Persons with Schizophrenia

Case ID	CNV Coordinates (hg19)	Gain/ Loss	CNV Size (kb)	5′ GENE (size in aa)	3′ GENE (size in aa)	Same frame[b]	Length of predicted fusion proteins[c]	Brain Expression[d]		Freq.[e]	Age at onset
								5′ gene	3′ gene		
SZ113	chr17:61737993-61887860[a]	Loss	150	*MAP3K3* (657)	*DDX42* (938)	Y	128 + 554	Y	Y	0	16
SZ61	Chr16:47004235-47124878	Gain	121	*DNAJA2* (412)	*NETO2* (525)	N	120 + 11	Y	Y	0	26
SJW19	chr14:69921835-69979453	Gain	57.6	*PLEKHD1* (506)	*SLC39A9* (307)	Y	186 + 121	Y	Y	0.00075	16
SZ25	Chr19:3800538-3847280	Gain	46.7	*MATK* (508)	*ZFR2* (939)	Y	0 (*5′-UTR*) + 839	Y	Y	0.00072	20

The following abbreviations are used: CNV, copy-number variant; kb, kilobases; aa, amino acids.
[a]*MAP3K3-DDX42* breakpoints sequenced to nearby *alu*.
[b]Indicates whether the second gene is predicted to be in the same frame as the first one when fused.
[c]Amino acids contributed by 5′ gene and amino acids contributed by 3′ gene.
[d]Determined by RT-PCR of human brain cDNA.
[e]Frequency of CNV in 19,585 control individuals.

We observed a similar pattern in transfected mouse cortical neurons (Figure S1).

The canonical role of MAP3K3 is the activation of mitogen-activated protein kinase ERK5.[24] In response to extracellular signals, MAP3K3 and its downstream kinase MEK5, encoded by *MAP2K5* (MIM 602520), are activated by heterodimerization and phosphorylation and, in turn, phosphorylate ERK5.[27; 28] MAP3K3-MEK5 dimerization occurs at N-terminal PB1 (Phox Bem 1p) domains[29] present in both proteins, whereas phosphorylation is mediated by C-terminal kinase domains.[24] The expression of a truncated MAP3K3 PB1 domain or a kinase-dead MAP3K3 results in the dominant-negative inhibition of ERK5 activation.[27,29] Similar to these engineered dominant-negative proteins, MAP3K3-DDX42 isoform 1 includes an intact PB1 domain but has the DDX42 helicase domain in place of the kinase domain (Figure 3A). We examined whether this fusion protein binds to MEK5. Immunoprecipitation of transfected HEK 293 cells indicated that MAP3K3 and MAP3K3-DDX42 isoform 1 bind endogenous phosphorylated MEK5, whereas DDX42 and MAP3K3-DDX42 isoform 2 do not (Figure 4A). This suggests that MAP3K3-DDX42 isoform 1 likely acts as a dominant-negative inhibitor of ERK5 signaling.

MAP3K3 also has a noncanonical role in the activation of three other MAPK cascades: p38 (encoded by *MAPK14* [MIM 600289]),[30] JNK (encoded by *MAPK8* [MIM 601158]),[30] and ERK1 and ERK2 (encoded by *MAPK1* [MIM 176948] and *MAPK3* [MIM 601795].[31] MAP3K3 activity in these pathways is not thought to involve heterodimerization at the PB1 domain.[29] We compared the activation of each of these pathways in cells transfected with *MAP3K3* or *MAP3K3-DDX42* chimeras alone or those cotransfected with *MAP3K3* and the *MAP3K3-DDX42* chimeras (Figures 4B and 4C). MAP3K3 activated ERK1 ($p = 0.006$), ERK2 ($p = 0.009$), and p38 ($p = 0.011$), but MAP3K3-DDX42 fusions did not, as predicted by the lack of a kinase domain in fusion proteins. Our data also suggest that both fusions may slightly suppress the activation of ERK1 and p38 (but it was not statistically significant). JNK was not activated by MAP3K3 or the fusion proteins, but there was evidence that JNK activation was suppressed by the fusion proteins (but it was not statistically significant).

DNAJA2-NETO2

DNAJA2-NETO2 fusion protein isoforms 1 and 2 were 14.7 and 14.3 kD (Figures 5A and 5B). The parent protein DNAJA2 is 45.7 kD. Protein levels of isoform 1 were lower than protein levels of isoform 2 and full-length DNAJA2, suggesting nonsense-mediated decay of *DNAJA2-NETO2* transcript variant 1 as its premature termination codon arises more than 100 nucleotides 5′ of the nearest exon boundary.[32] Furthermore, DNAJA2-NETO2 localization was altered in HEK 293 cells. The parent protein DNAJA2 was excluded from the nucleus, as previously reported,[33] whereas the fusion proteins were localized throughout the nucleus and cytoplasm (Figures 5C and 5D).

PLEKHD1-SLC39A9

PLEKHD1-SLC39A9 fusion protein isoforms 1 and 2 are 34.0 and 26.4 kD (Figures 6A and 6B). PLEKHD1 is 59.2 kD. SLC39A9 is predicted to have a molecular weight of 32.4 kD but appears much smaller according to western blot (~15 kD), suggesting that the protein product may be cleaved.

MATK-ZFR2

ZFR2 contributes the entire coding sequence of the MATK-ZFR2 fusion protein. Little is known about *ZFR2*. We detected two *ZFR2* transcript variants in the human brain and predicted *MATK-ZFR2* chimeras for both variants. *ZFR2* transcript variant 2 is characterized by the omission of exon 15, which leads to a shift of the reading frame. This frameshift does not result in premature truncation but, instead, codes for a 946aa protein that differs from the 938aa ZFR2 isoform 1 in its C-terminal 198 residues (Figure S2). We generated expression constructs

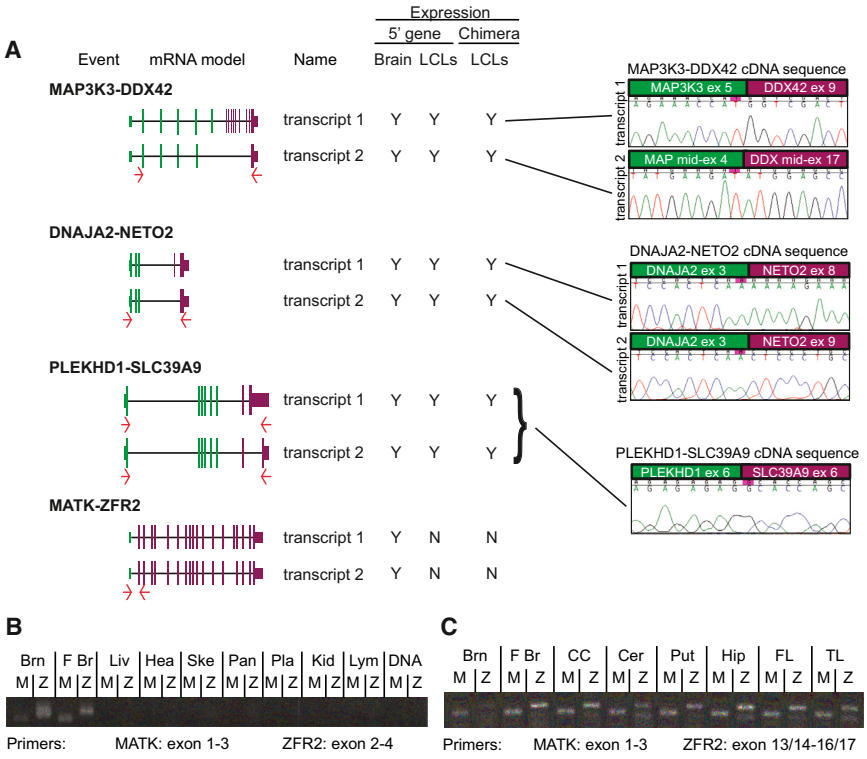

Figure 2. Characterization of Chimeric Transcripts

(A) *MAP3K3-DDX42*, *DNAJA2-NETO2*, and *PLEKHD1-SLC39A9* chimeric transcripts were detected by RT-PCR of RNA from lymphoblast cell lines (LCLs). Sanger sequence traces show chimeric junctions. *MATK-ZFR2* chimeric transcripts were not detected in lymphoblasts. Red arrows indicate exons targeted by forward and reverse primers. The expression of each 5′ parent gene, tested by RT-PCR, is indicated for human brain and LCLs.

(B) Parent genes *MATK* (M) and *ZFR2* (Z) were expressed only in brain and fetal brain, as determined by RT-PCR of cDNA from indicated tissues. Tissue legend: Brn, adult brain; F Br, fetal brain; Liv, liver; Hea, heart; Ske, skeletal muscle; Pan, pancreas; Pla, placenta; Kid, kidney; Lym, lymphoblasts; and DNA, genomic DNA (negative control).

(C) *MATK* and *ZFR2* were expressed in every brain region tested, and two *ZFR2* transcripts were present, distinguished by the alternate splicing of exon 15. Tissue legend: Brn, adult brain; F Br, fetal brain; CC, cerebral cortex; Cer, cerebellum; Put, putamen; Hip, hippocampus; FL, frontal lobe; and TL, temporal lobe.

for both *ZFR2* transcripts and their respective chimeras with *MATK* (Figure 7A). The chimeras include a 5′ UTR of *MATK* and all but the first exon of *ZFR2*, and as a result, lack the wild-type start codon. Observed molecular weights of the fusion proteins (predicted to be 97.3 and 97.8 kD) (Figure 7B) were consistent with translation initiation occurring at the first Methionine codon in *ZFR2* exon 2, Met41, which is in the correct reading frame and has a strong Kozak sequence (GGG**ATG**G).[34]

One potential pathogenic mechanism of the *MATK-ZFR2* chimera could be the altered expression of *ZFR2* as a result of fusion to promoter sequences upstream of or within the 5′ UTR of *MATK*. We compared temporal expression patterns of *MATK* and *ZFR2* with RNA-seq data from human prefrontal cortex from the BrainSpan Atlas of the Developing Human Brain (Figure 7C). *MATK* is upregulated between the late fetal period and early infancy, whereas *ZFR2* expression remains at a relatively constant low level. We also compared the expression patterns of parent genes across brain subregions and in whole adult and fetal brain with publicly available gene expression data from 128 human tissues curated from 1,068 published microarray experiments[35] (Figure S3). *MATK* and *ZFR2* differ in expression most dramatically in a subset of brain structures, including the cingulate cortex, pons, caudate nucleus, and globus pallidus. Fusion of *MATK* and its upstream regulatory elements to most of the coding sequence of *ZFR2* may lead to the aberrant overexpression of ZFR2 isoforms 1 and 2 at critical developmental times and in specific brain regions.

Because *MATK* and *ZFR2* are expressed only in the brain, we evaluated localization in transfected mouse cortical neurons. The localization of MATK-ZFR2 fusion proteins did not appear to differ from their full-length ZFR2 counterparts, but the two isoforms of ZFR2 had distinct localization patterns. ZFR2 isoform 1 and MATK-ZFR2 isoform 1 were predominantly in the nucleus but were also diffusely expressed in the cell soma and processes (Figures 8A and 8B). In contrast, ZFR2 isoform 2 and MATK-ZFR2 isoform 2 were largely excluded from the nucleus with a granular staining pattern (Figures 8C and 8D). Within the dendrites, we found that ZFR2 isoform 2 and its chimera localized preferentially to branch sites (Figures 8E and 8F). Nuclear localization of ZFR2 isoform 1 may be driven by a putative nuclear localization signal at amino acids 913–933[36] (Figure 7A) within the C-terminal region absent from ZFR2 isoform 2, potentially accounting for the difference in localization between the two isoforms.

Discussion

We found an increased burden of rare, brain-expressed chimeric genes in individuals with schizophrenia, and there were four events in 124 affected individuals compared to zero events in 290 control individuals. Because each event arises in only a single affected individual in our series, we cannot establish causality for any one chimera. However, the results support the theory that rare chimeric genes as a group contribute to schizophrenia

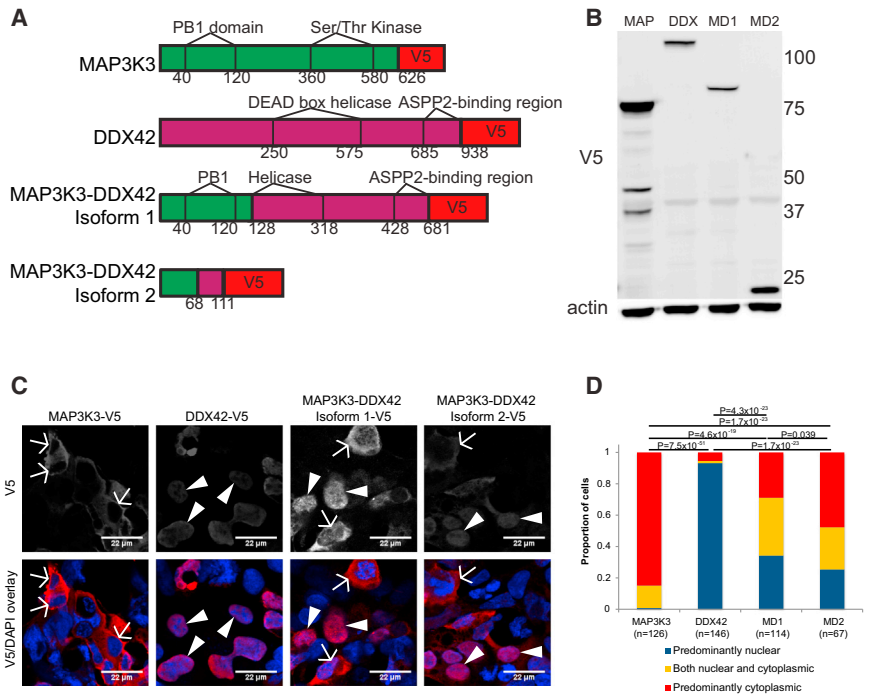

Figure 3. Synthesis and Altered Localization of MAP3K3-DDX42 Fusion Proteins
(A) The expression constructs for *MAP3K3-DDX42* chimeras and parent genes were transfected into HEK 293 cells. Known functional domains and placement of the V5 epitope tag are indicated. The figure is not to scale.
(B) Western blot shows protein levels for transfected genes and loading control. MAP and DDX are parent genes, and MD1 and MD2 are MAP3K3-DDX42 isoforms 1 and 2. Molecular weights of protein ladder are indicated in kD.
(C) Representative micrographs of transfected cells show cytoplasmic localization (arrows) of MAP3K3 and nuclear localization (triangles) of DDX42. The localization of MAP3K3-DDX42 fusion proteins ranges from nuclear (triangles) to cytoplasmic (arrows).
(D) Proportions of cells with predominantly nuclear, both nuclear and cytoplasmic, or predominantly cytoplasmic localization are indicated for each construct. Results are from the indicated number of cells across three experiments. p values show chi-square comparisons of distributions.

and are consistent with multiple lines of evidence suggesting that mutations leading to chimeric genes are likely to be deleterious.

MAP3K3-DDX42

Our results suggest that MAP3K3-DDX42 is likely to act as a dominant-negative inhibitor of ERK5 signaling by binding to activated MEK5. The ERK5 signaling cascade is critically important in neuronal differentiation and proliferation, neuroprotection, and adult neurogenesis[37–39] and may be misregulated in neuropsychiatric illnesses including autism and major depressive disorder (MIM 608516).[40,41] Our findings also suggest that MAP3K3-DDX42 may have slight inhibitory effects on JNK, p38, and ERK1 signaling, all of which function in brain processes, including synaptic plasticity and inflammation.[42–44] These effects may be more pronounced against a genetic background haploid at this locus, as in this affected individual.

DDX42, the 3′ parent of *MAP3K3-DDX42*, codes for a brain-expressed RNA helicase thought to mediate response to viral infection of the CNS.[45] Recently, DDX42 was shown to modulate the effect of the proapoptotic protein ASPP2, which is encoded by *TP53BP2* (MIM 602143).[26] DDX42 interferes with the induction of apoptosis by ASPP2 and excludes ASPP2 from the nucleus. Intriguingly, a brain-specific function for ASPP2 has recently emerged: the control of polarity and proliferation of neural progenitor cells during CNS development.[46] ASPP2 targeting to the apical junctional complex in these neural progenitors is thought to be critical for these functions.[46] MAP3K3-DDX42 isoform 1 includes the DDX42 residues that bind ASPP2, but the protein is mislocalized. DDX42 is predominantly nuclear, whereas the subcellular localization of fusion isoform 1 is highly variable. Aberrant cytoplasmic localization of the fusion protein, including the C terminus of DDX42, may allow it to access ASPP2 at the apical junctional complexes, potentially interfering with its critical role in neural progenitor proliferation.

DNAJA2-NETO2

DNAJA2-NETO2 is an out-of-frame chimera, and as such *NETO2* does not contribute any protein-coding sequence to the fusion protein. Nonetheless, the contribution of a naturally occurring 3′ end to the mRNA may improve the ability of RNA processing machinery to recognize and maintain this aberrant mRNA in comparison to genomic deletions or duplications that result in a simple truncation. 3′ UTRs have been shown to have profound, sequence-dependent regulatory effects on their respective genes via microRNA binding, polyadenylation, translational control, and mRNA localization (for review, see Jia et al., 2013[47]). *NETO2* is involved in glutamate signaling in the brain,[48] and its 3′ end should stabilize rather than destabilize the chimeric mRNA in brain tissue.

DNAJA2 is a brain-expressed member of the J protein family, which consists of cochaperones of heat shock protein 70 (HSP70) proteins.[33] DNAJA2 promotes degradation of the cardiac hERG potassium channel[49] and enhances the HSP70-mediated refolding of neural G proteins.[50] These roles are likely mediated by two distinct but interrelated functions of DNAJA proteins. First, they activate folding by HSP70 via ATP hydrolysis.[51] Second, they

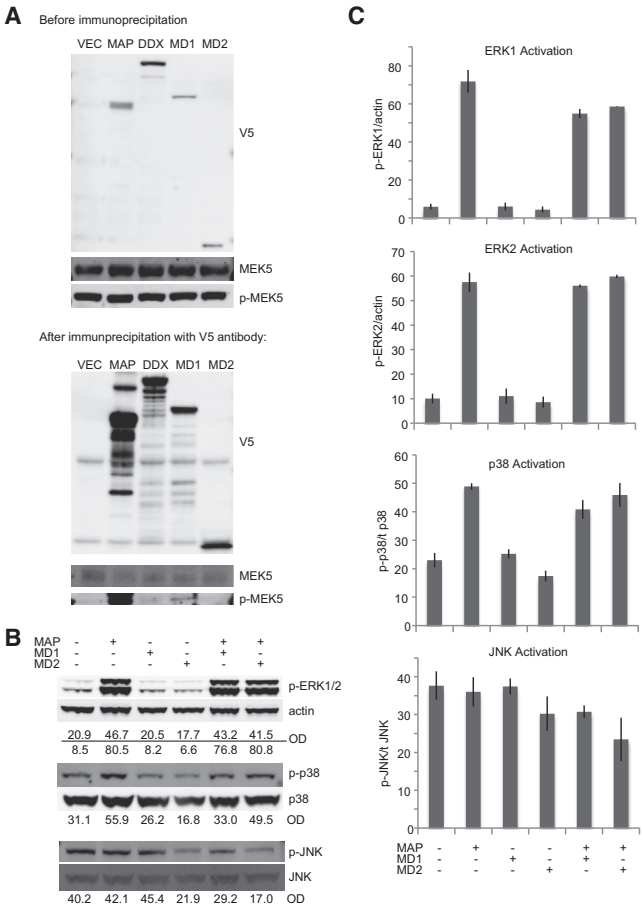

Figure 4. MAP3K3-DDX42 Fusion Protein Interacts with MAPK Signaling Pathways

(A) HEK 293 cells were transfected with empty vector (VEC), *MAP3K3-DDX42* parent genes (MAP and DDX), or chimeras (MD1 and MD2), and lysates were immunoprecipitated with a V5 antibody. Western blots of lysates before and after immunoprecipitation were stained with antibodies against V5, MEK5, and phospho-MEK5 and show binding of phospho-MEK5 by MAP3K3 and MAP3K3-DDX42 isoform 1.

(B) We evaluated the activation of MAPK pathways in HEK 293 cells transfected with the indicated combinations of *MAP3K3* and *MAP3K3-DDX42* constructs with phospho-specific antibodies against ERK1 and ERK2, p38, and JNK. Representative blots and optical density (OD) of each band normalized to an appropriate loading control are shown.

(C) Mean normalized densitometry values from three replicate experiments are shown for ERK1, ERK2, p38, and JNK. Equivalent amounts of each construct were transfected for all conditions, and the total amount of transfected DNA was kept constant by the addition of empty vector as necessary. Error bars represent SEM.

homodimerize, forming a clamp-like structure that binds target proteins for delivery to HSP70.[52] The first function is shared by all J proteins and is mediated by the N-terminal J domain. The latter function is more specific to DNAJA proteins and is mediated by central and C-terminal domains.

DNAJA2-NETO2 fusion proteins include only the J domain. Similar mutations in other J domain family proteins have been studied. One pertinent study found that the expression of an isolated J domain in mammalian cells completely blocks luciferase refolding by endogenous HSP70 and impairs cell growth.[53] This dominant-negative effect is thought to result from inefficient ATPase activity of the isolated J domain, which, therefore, prevents binding and activation by full-length J proteins and other chaperones. DNAJA2-NETO2 fusion proteins may have an analogous dominant-negative function, preventing the proper folding of key substrates. Furthermore, the observed mislocalization of DNAJA2-NETO2 proteins to the cytoplasm may contribute to a deleterious effect via a gain-of-function mechanism.

PLEKHD1-SLC39A9

Both parent genes of the *PLEKHD1-SLC39A9* chimera are broadly expressed, including in the fetal and adult brain, but their role in the CNS has not been studied. *PLEKHD1* is an uncharacterized gene that has an N-terminal Pleckstrin homology (PH) domain. PH domains exist in a wide range of proteins and are generally involved in intracellular signaling by binding to phosphatidylinositol lipids and recruitment of target proteins to the cell membrane.[54] *SLC39A9* codes for the zinc transporter ZIP9, which was recently shown to enhance AKT1 (encoded by *AKT1* [MIM 164730]) and ERK1 and ERK2 phosphorylation in immune cells by regulating intracellular zinc concentration.[55] Signaling through AKT1, ERK1, and ERK2 is critical in the brain. As mentioned above, ERK1 and ERK2 are also activated by MAP3K3 as well as DISC1,[56] and are involved in processes including neuroprotection[39] and synaptic plasticity.[44] AKT1 is a protein kinase central to multiple signaling cascades in the brain and has been identified as a point of convergence of many known schizophrenia risk factors.[56,57] The fusion proteins contain both the PH domain of PLEKHD1 and the zinc transporter domain of ZIP9 and potentially interfere with functions of both genes.

MATK-ZFR2

ZFR2 is a previously uncharacterized gene that is similar by sequence homology and predicted protein structure to *ZFR*, a member of the double-stranded RNA-binding family.[58] ZFR mediates nucleocytoplasmic shuttling and the transport of ribonuclear particles into neurites.[59] ZFR binds the brain-expressed RNA-binding protein STAU2 (MIM 605920), and ZFR2 likely has a different binding partner. *C. elegans* have a single gene, Y95B8A.7, orthologous to *ZFR* and *ZFR2*, which, when targeted by RNAi, leads to defects in axon guidance.[60]

The localization of ZFR2 isoform 1 is similar to that of ZFR in neuronal cells, which is present at high levels in the nucleus and lower levels in the cytoplasm. By contrast, ZFR2 isoform 2 is excluded from the nucleus and shows preferential localization to dendritic branch sites. This staining pattern is similar to that of other neuronal RNA-binding proteins, such as Staufen and FMRP in *Drosophila*

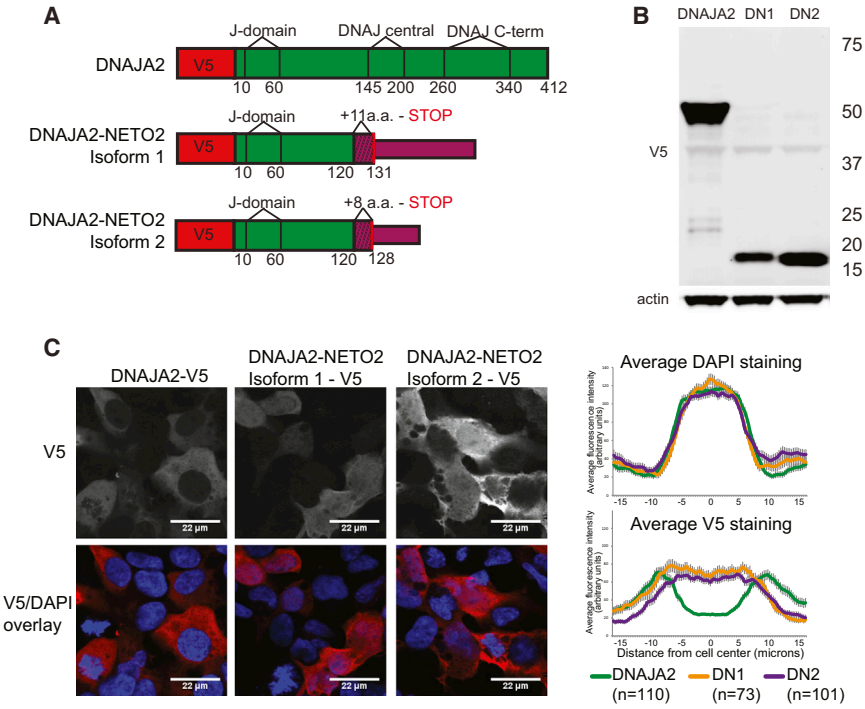

Figure 5. Synthesis and Altered Localization of DNAJA2-NETO2 Fusion Proteins
(A) Expression constructs for *DNAJA2-NETO2* chimeras and full-length *DNAJA2* were transfected into HEK 293 cells. Known functional domains and placement of V5 epitope tag are indicated. The figure is not to scale.
(B) Western blot shows protein levels for transfected genes and loading control. DN1 and DN2 are DNAJA2-NETO2 isoforms 1 and 2. Molecular weights of protein ladder are indicated in kD.
(C) Representative micrographs of transfected cells show the nuclear exclusion of DNAJA2 in comparison to the increased nuclear localization of both DNAJA2-NETO2 chimeras. Average fluorescence intensity of DAPI and V5 stain are shown along a 30 µm line drawn centered on the nucleus in the indicated number of cells for each transfection condition. Results are from three independent experiments. Error bars show SEM.

neurons,[61] and zipcode-binding protein 1 (ZBP1), which regulates the formation of dendritic arbors in rat hippocampal neurons.[62] Intriguingly, both knockdown and the overexpression of *ZBP1* have been shown to impair dendritic arborization.[62] We postulate that ZFR2 isoforms 1 and 2 have distinct functions within the neuronal RNA binding pathway. On the basis of evidence from subcellular localization, isoform 1 may be involved in the nucleocytoplasmic shuttling of ribonucleoparticles, whereas isoform 2 may regulate dendritic branching by transporting specific ribonucleoparticles into neuronal processes. The fusion of *ZFR2* to the 5′ UTR and putative regulatory regions of *MATK* is predicted to result in overexpression of both isoforms of ZFR2 in the brain at critical developmental time points, potentially disrupting these functions and leading to aberrant dendritic arborization. Additional functional work is necessary in order to evaluate the functions of these proteins and determine the likely outcome of overexpression.

The exonic deletion and consequent frameshift differentiating ZFR2 isoforms 1 and 2 may have acted as an evolutionary mechanism for the emergence of a distinct protein involved in dendritic branching. Notably, three stop codons present in the frame-shifted sequence in mouse and rat are absent from the chimpanzee and human sequence, suggesting that this protein is primate specific. Frame shifting has been proposed and demonstrated to be a source of proteins that can be targeted by selection but generally in the evolutionarily permissive context of a prior gene duplication.[63] In the case of *ZFR2*, a new protein instead emerges from alternate transcription and translation of two extensively overlapping open reading frames encoded by a single nonduplicated gene.

Features of Chimerism

Chimerism is an important mutational mechanism that potentially impacts evolution and disease. Newly

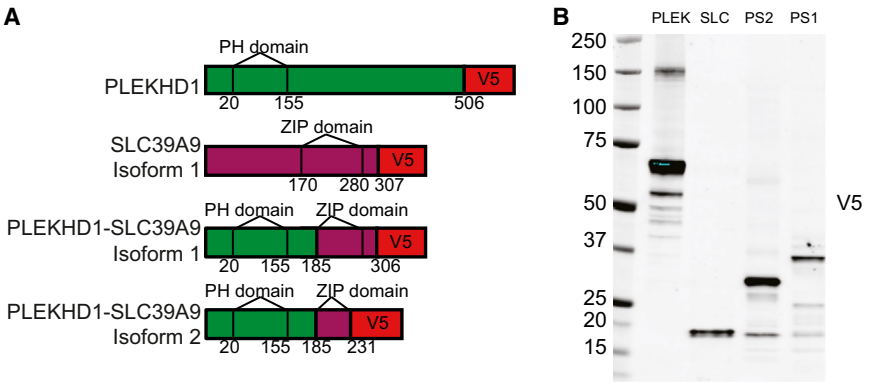

Figure 6. Stable Synthesis of PLEKHD1-SLC39A9 Fusion Proteins
(A) Expression constructs for *PLEKHD1-SLC39A9* chimeras and parent genes were transfected into HEK 293 cells. Known functional domains and placement of V5 epitope tag are indicated. The figure is not to scale.
(B) Western blot shows protein levels for transfected genes and loading control. PLEK and SLC are parent genes, and PS1 and PS2 are PLEKHD1-SLC39A9 isoforms 1 and 2. Molecular weights of protein ladder are indicated in kD.

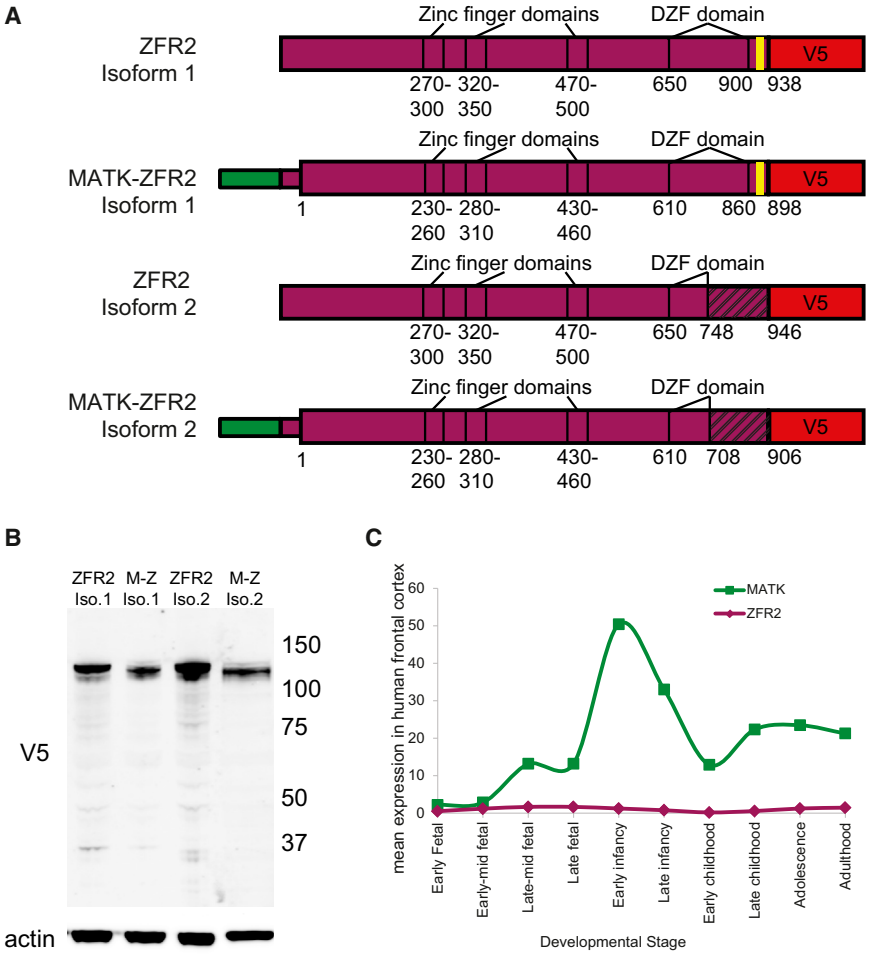

Figure 7. Synthesis of MATK-ZFR2 Fusion Proteins and Developmental Regulation of Parent Genes

(A) Expression constructs for two isoforms of parent gene *ZFR2* and respective *MATK-ZFR2* chimeras were transfected into HEK 293 cells. Known functional domains and placement of V5 epitope tag are indicated. Yellow bars are putative nuclear localization signals, and black diagonal lines indicate the region altered by exon skipping and frameshift. The figure is not to scale.

(B) Western blot shows protein levels for transfected genes and loading control. Molecular weights of protein ladder are indicated in kD.

(C) Normalized expression values for *MATK* and *ZFR2* in the human prefrontal cortex plotted at different developmental ages. Data were obtained from the Allen Brain Atlas BrainSpan RNA-seq data set and represent 149 human frontal cortex samples.

emerging chimeric genes have the potential to create proteins with new functions. For example, a recent study compared the phenotypic impacts of different classes of engineered mutations in the eleven genes of the yeast mating pathway.[64] The recombination of functional domains into chimeric genes led to a striking diversification in mating response, whereas whole-gene duplications, single-domain duplications, and the coexpression of distinct functional domains did not. This pattern is also evident in our results. The recombination of functional domains in the *MAP3K3-DDX42* or *PLEKHD1-SLC39A9* chimeras produces proteins with features of both parents and the potential to interact with multiple pathways. Such proteins could not result from simple truncation, deletion, or duplication of parent genes.

Even when a chimera includes protein-coding sequence from only one gene, the recombination of regulatory regions may also result in functional changes. A survey of 14 conserved chimeric genes in *Drosophila melanogaster* revealed dramatic regulatory differences between parent genes and chimeras at the level of both subcellular localization and tissue-specific expression.[65] We also observed striking differences in regulatory patterns between parent genes *MATK* and *ZFR2*, and both DNAJA2-NETO2 and MAP3K3-DDX42 fusions had altered subcellular localization. Altered localization potentially influences the mechanism of pathogenesis. When present in the same cellular compartment as a parent gene, the fusion protein can carry out predicted dominant-negative functions, such as the inhibition of MAPK pathway signaling by MAP3K3-DDX42 fusions. When present in different locations relative to parent proteins, the same chimera may have potential gain-of-function effects, such as aberrantly binding to ASPP2.

Evidence for Causality

For private or very rare mutations, evidence of causality is best obtained by functional studies demonstrating biological relevance of the gene and biological consequences of the mutation. For the chimeric genes in our case series, the differences in localization, regulation, or function between parent and fusion genes support roles for these events in schizophrenia. Many of the parent genes play important roles in the brain, including, potentially, the gene *ZFR2*, for which we provide evidence for a recently evolved role in dendritic branching. The neurodevelopmental pathways within which these genes operate have been previously implicated in schizophrenia, including neural precursor proliferation, neuronal migration and integration, glutamate signaling, and neurite outgrowth (for review, see Brandon and Sawa, 2011[66]).

Our results also provide additional evidence that schizophrenia is characterized by vast genetic heterogeneity. In our cohort, only four cases are potentially explained by

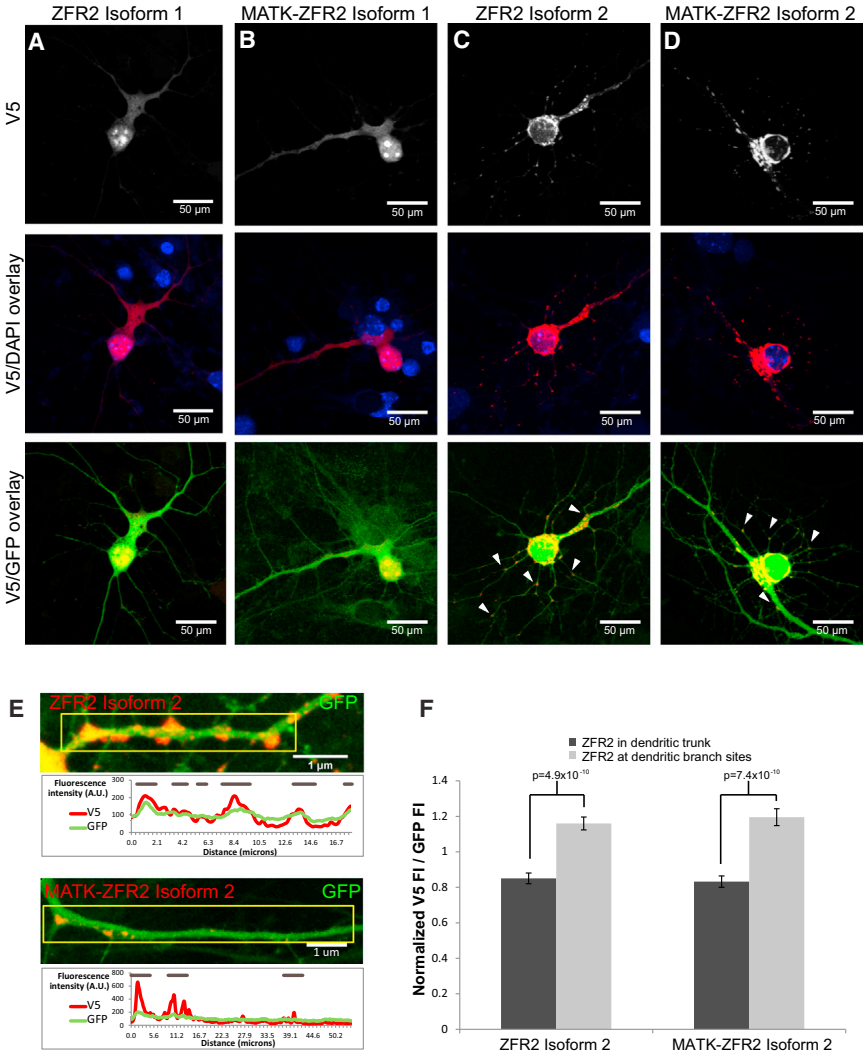

Figure 8. Localization of ZFR2 Isoforms 1 and 2 and MATK-ZFR2 Fusion Proteins

(A–D) Representative micrographs of transfected mouse cultured cortical neurons show diffuse localization of ZFR2 isoform 1 (A) and MATK-ZFR2 isoform 1 (B) throughout the nucleus and cytoplasm, and bright puncta in the nucleus not colocalized with nucleoli. ZFR2 isoform 2 (C) and MATK-ZFR2 isoform 2 (D) are largely excluded from the nucleus and show a granular staining pattern with many granules present at branch sites (triangles).

(E) Representative micrographs of single dendrites and corresponding fluorescence profiles from mouse cortical cells transfected with ZFR2 isoform 2 or MATK-ZFR2 isoform 2. Dendritic morphology was visualized by cotransfected GFP. Horizontal gray bars indicate manually annotated branch sites.

(F) Mean fluorescence intensity (FI) of V5 in the dendritic trunk versus dendritic branch sites normalized to GFP FI in order to correct for increased dendrite diameter at branch sites. Data are from three independent replicates. p values represent significance by two-tailed t test. Error bars represent SEM.

chimeras. Each fusion gene is rare, perhaps even unique to one affected individual. Furthermore, each chimera differs dramatically in regards to the architecture of the fusion, the functions of the genes involved, and the predicted mode of action, including the disruption of regulation and localization.

In the modern era of gene discovery, it is now possible to solve one case at a time. The disruption of a gene important to neurodevelopment, by any number of different mutational mechanisms (including chimerism), may contribute to neuropsychiatric disease. Given that the majority of human genes are expressed in the brain, and because each gene may be disrupted by a number of different mutations, the number of potential genetic causes of schizophrenia is vast. Although any given damaging mutation impacting a gene or genes likely explains only a small number of cases, the sum total of individually rare severe events disrupting brain development may explain a substantial portion of illness.

The enrichment of rare brain-expressed chimeras in persons with schizophrenia, and the analyses suggesting a functional impact of these fusion genes, support the theory that chimeric genes play a role in the illness, at least in a small number of affected individuals. Exome and whole-genome sequencing data in large case series can be readily exploited in order to detect chimeric genes resulting from CNVs as well as inversions and translocations. Future studies will reveal more about the overall contribution of this class of mutation to schizophrenia and other neuropsychiatric illnesses.

Supplemental Data

Supplemental Data include Supplemental Experimental Procedure, three figures, and two tables and can be found with this article online at http://www.cell.com/AJHG.

Acknowledgments

This project was supported by National Institutes of Health grants R01 MH083989, F30 MH085467-03, and T32 GM07266 and a fellowship from the Achievement Rewards for College Scientists Foundation. This study makes use of data generated by the Wellcome Trust Case Control Consortium.

Received: June 30, 2013
Revised: August 15, 2013
Accepted: September 10, 2013
Published: October 3, 2013

Web Resources

The URLs for data presented herein are as follows:

Allen Brain Atlas BrainSpan technical white sheet, http://help.brain-map.org/display/devhumanbrain/Documentation

BrainSpan: Atlas of the Developing Human Brain, http://www.brainspan.org

NCBI GenBank, http://www.ncbi.nlm.nih.gov/genbank/

National Institute of Mental Health Genomics Research and Repository Distribution 5, https://www.nimhgenetics.org/available_data/controls/detailed_sample_information.php

Online Mendelian Inheritance in Man (OMIM), www.omim.org

Accession Numbers

The chimeric mRNA sequences reported in this paper were deposited into the NCBI GenBank and are available at the following accession numbers: MAP3K3-DDX42 transcript variant 1, KF612910; MAP3K3-DDX42 transcript variant 2, KF612911; DNAJA2-NETO2 transcript variant 1, KF612912; DNAJA2-NETO2 transcript variant 2, KF612913; PLEKHD1-SLC39A9 transcript variant 1, KF612914; and PLEKHD1-SLC39A9 transcript variant 2, KF612915.

References

1. Walsh, T., McClellan, J.M., McCarthy, S.E., Addington, A.M., Pierce, S.B., Cooper, G.M., Nord, A.S., Kusenda, M., Malhotra, D., Bhandari, A., et al. (2008). Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science *320*, 539–543.
2. Stefansson, H., Rujescu, D., Cichon, S., Pietiläinen, O.P., Ingason, A., Steinberg, S., Fossdal, R., Sigurdsson, E., Sigmundsson, T., Buizer-Voskamp, J.E., et al.; GROUP. (2008). Large recurrent microdeletions associated with schizophrenia. Nature *455*, 232–236.
3. Kirov, G., Grozeva, D., Norton, N., Ivanov, D., Mantripragada, K.K., Holmans, P., Craddock, N., Owen, M.J., and O'Donovan, M.C.; International Schizophrenia Consortium; Wellcome Trust Case Control Consortium. (2009). Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. Hum. Mol. Genet. *18*, 1497–1503.
4. Millar, J.K., Wilson-Annan, J.C., Anderson, S., Christie, S., Taylor, M.S., Semple, C.A., Devon, R.S., St Clair, D.M., Muir, W.J., Blackwood, D.H., and Porteous, D.J. (2000). Disruption of two novel genes by a translocation co-segregating with schizophrenia. Hum. Mol. Genet. *9*, 1415–1423.
5. Vacic, V., McCarthy, S., Malhotra, D., Murray, F., Chou, H.H., Peoples, A., Makarov, V., Yoon, S., Bhandari, A., Corominas, R., et al. (2011). Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. Nature *471*, 499–503.
6. Friedman, J.I., Vrijenhoek, T., Markx, S., Janssen, I.M., van der Vliet, W.A., Faas, B.H., Knoers, N.V., Cahn, W., Kahn, R.S., Edelmann, L., et al. (2008). CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. Mol. Psychiatry *13*, 261–266.
7. Williams, N.M. (2011). Molecular mechanisms in 22q11 deletion syndrome. Schizophr. Bull. *37*, 882–889.
8. Williams, H.J., Monks, S., Murphy, K.C., Kirov, G., O'Donovan, M.C., and Owen, M.J. (2013). Schizophrenia two-hit hypothesis in velo-cardio facial syndrome. Am. J. Med. Genet. B. Neuropsychiatr. Genet. *162B*, 177–182.
9. Ranz, J.M., and Parsch, J. (2012). Newly evolved genes: moving from comparative genomics to functional studies in model systems. How important is genetic novelty for species adaptation and diversification? Bioessays *34*, 477–483.
10. Courseaux, A., and Nahon, J.L. (2001). Birth of two chimeric genes in the Hominidae lineage. Science *291*, 1293–1297.
11. Thomson, T.M., Lozano, J.J., Loukili, N., Carrió, R., Serras, F., Cormand, B., Valeri, M., Díaz, V.M., Abril, J., Burset, M., et al. (2000). Fusion of the human gene for the polyubiquitination coeffector UEV1 with Kua, a newly identified gene. Genome Res. *10*, 1743–1756.
12. Rogers, R.L., Bedford, T., and Hartl, D.L. (2009). Formation and longevity of chimeric and duplicate genes in Drosophila melanogaster. Genetics *181*, 313–322.
13. Heisterkamp, N., Stam, K., Groffen, J., de Klein, A., and Grosveld, G. (1985). Structural organization of the bcr gene and its role in the Ph' translocation. Nature *315*, 758–761.
14. Zhou, X., Chen, Q., Schaukowitch, K., Kelsoe, J.R., and Geyer, M.A. (2010). Insoluble DISC1-Boymaw fusion proteins generated by DISC1 translocation. Mol. Psychiatry *15*, 669–672.
15. Eykelenboom, J.E., Briggs, G.J., Bradshaw, N.J., Soares, D.C., Ogawa, F., Christie, S., Malavasi, E.L., Makedonopoulou, P., Mackie, S., Malloy, M.P., et al. (2012). A t(1;11) translocation linked to schizophrenia and affective disorders gives rise to aberrant chimeric DISC1 transcripts that encode structurally altered, deleterious mitochondrial proteins. Hum. Mol. Genet. *21*, 3374–3386.
16. Holt, R., Sykes, N.H., Conceição, I.C., Cazier, J.B., Anney, R.J., Oliveira, G., Gallagher, L., Vicente, A., Monaco, A.P., and Pagnamenta, A.T. (2012). CNVs leading to fusion transcripts in individuals with autism spectrum disorder. Eur. J. Hum. Genet. *20*, 1141–1147.
17. Nord, A.S., Roeb, W., Dickel, D.E., Walsh, T., Kusenda, M., O'Connor, K.L., Malhotra, D., McCarthy, S.E., Stray, S.M., Taylor, S.M., et al.; STAART Psychopharmacology Network. (2011). Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. Eur. J. Hum. Genet. *19*, 727–731.
18. Gulsuner, S., Walsh, T., Watts, A.C., Lee, M.K., Thornton, A.M., Casadei, S., Rippey, C., Shahin, H., Nimgaonkar, V.L., Go, R.C., et al.; Consortium on the Genetics of Schizophrenia (COGS); PAARTNERS Study Group. (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. Cell *154*, 518–529.
19. Krumm, N., Sudmant, P.H., Ko, A., O'Roak, B.J., Malig, M., Coe, B.P., Quinlan, A.R., Nickerson, D.A., Eichler, E.E., and Eichler, E.E.; NHLBI Exome Sequencing Project. (2012). Copy number variation detection and genotyping from exome sequence data. Genome Res. *22*, 1525–1532.
20. Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W., and Lee, C. (2004). Detection of large-scale variation in the human genome. Nat. Genet. *36*, 949–951.
21. Cooper, G.M., Coe, B.P., Girirajan, S., Rosenfeld, J.A., Vu, T.H., Baker, C., Williams, C., Stalker, H., Hamid, R., Hannig, V., et al. (2011). A copy number variation morbidity map of developmental delay. Nat. Genet. *43*, 838–846.
22. Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly

of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.
23. Orlova, K.A., Parker, W.E., Heuer, G.G., Tsai, V., Yoon, J., Baybis, M., Fenning, R.S., Strauss, K., and Crino, P.B. (2010). STRADalpha deficiency results in aberrant mTORC1 signaling during corticogenesis in humans and mice. J. Clin. Invest. *120*, 1591–1602.
24. Chao, T.H., Hayashi, M., Tapping, R.I., Kato, Y., and Lee, J.D. (1999). MEKK3 directly regulates MEK5 activity as part of the big mitogen-activated protein kinase 1 (BMK1) signaling pathway. J. Biol. Chem. *274*, 36035–36038.
25. Uhlik, M.T., Abell, A.N., Johnson, N.L., Sun, W., Cuevas, B.D., Lobel-Rice, K.E., Horne, E.A., Dell'Acqua, M.L., and Johnson, G.L. (2003). Rac-MEKK3-MKK3 scaffolding for p38 MAPK activation during hyperosmotic shock. Nat. Cell Biol. *5*, 1104–1110.
26. Uhlmann-Schiffler, H., Kiermayer, S., and Stahl, H. (2009). The DEAD box protein Ddx42p modulates the function of ASPP2, a stimulator of apoptosis. Oncogene *28*, 2065–2073.
27. Xu, B.E., Stippec, S., Lenertz, L., Lee, B.H., Zhang, W., Lee, Y.K., and Cobb, M.H. (2004). WNK1 activates ERK5 by an MEKK2/3-dependent mechanism. J. Biol. Chem. *279*, 7826–7831.
28. Nakamura, K., Uhlik, M.T., Johnson, N.L., Hahn, K.M., and Johnson, G.L. (2006). PB1 domain-dependent signaling complex is required for extracellular signal-regulated kinase 5 activation. Mol. Cell. Biol. *26*, 2065–2079.
29. Nakamura, K., and Johnson, G.L. (2003). PB1 domains of MEKK2 and MEKK3 interact with the MEK5 PB1 domain for activation of the ERK5 pathway. J. Biol. Chem. *278*, 36989–36992.
30. Huang, Q., Yang, J., Lin, Y., Walker, C., Cheng, J., Liu, Z.G., and Su, B. (2004). Differential regulation of interleukin 1 receptor and Toll-like receptor signaling by MEKK3. Nat. Immunol. *5*, 98–103.
31. Wang, X., Zhang, F., Chen, F., Liu, D., Zheng, Y., Zhang, Y., Dong, C., and Su, B. (2011). MEKK3 regulates IFN-gamma production in T cells through the Rac1/2-dependent MAPK cascades. J. Immunol. *186*, 5791–5800.
32. Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev. *21*, 1833–1856.
33. Terada, K., and Mori, M. (2000). Human DnaJ homologs dj2 and dj3, and bag-1 are positive cochaperones of hsc70. J. Biol. Chem. *275*, 24728–24734.
34. Kozak, M. (1987). An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. *15*, 8125–8148.
35. Kupershmidt, I., Su, Q.J., Grewal, A., Sundaresh, S., Halperin, I., Flynn, J., Shekar, M., Wang, H., Park, J., Cui, W., et al. (2010). Ontology-based meta-analysis of global collections of high-throughput public data. PLoS ONE *5*, 5.
36. Dingwall, C., Robbins, J., Dilworth, S.M., Roberts, B., and Richardson, W.D. (1988). The nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. J. Cell Biol. *107*, 841–849.
37. Pan, Y.W., Zou, J., Wang, W., Sakagami, H., Garelick, M.G., Abel, G., Kuo, C.T., Storm, D.R., and Xia, Z. (2012). Inducible and conditional deletion of extracellular signal-regulated kinase 5 disrupts adult hippocampal neurogenesis. J. Biol. Chem. *287*, 23306–23317.
38. Li, T., Pan, Y.W., Wang, W., Abel, G., Zou, J., Xu, L., Storm, D.R., and Xia, Z. (2013). Targeted deletion of the ERK5 MAP kinase impairs neuronal differentiation, migration, and survival during adult neurogenesis in the olfactory bulb. PLoS ONE *8*, e61948.
39. Cavanaugh, J.E., Jaumotte, J.D., Lakoski, J.M., and Zigmond, M.J. (2006). Neuroprotective role of ERK1/2 and ERK5 in a dopaminergic cell line under basal conditions and in response to oxidative stress. J. Neurosci. Res. *84*, 1367–1375.
40. Dwivedi, Y., Rizavi, H.S., Teppen, T., Sasaki, N., Chen, H., Zhang, H., Roberts, R.C., Conley, R.R., and Pandey, G.N. (2007). Aberrant extracellular signal-regulated kinase (ERK) 5 signaling in hippocampus of suicide subjects. Neuropsychopharmacology *32*, 2338–2350.
41. Yang, K., Sheikh, A.M., Malik, M., Wen, G., Zou, H., Brown, W.T., and Li, X. (2011). Upregulation of Ras/Raf/ERK1/2 signaling and ERK5 in the brain of autistic subjects. Genes Brain Behav. *10*, 834–843.
42. Mingorance-Le Meur, A. (2006). JNK gives axons a second chance. J. Neurosci. *26*, 12104–12105.
43. Tong, L., Prieto, G.A., Kramár, E.A., Smith, E.D., Cribbs, D.H., Lynch, G., and Cotman, C.W. (2012). Brain-derived neurotrophic factor-dependent synaptic plasticity is suppressed by interleukin-1β via p38 mitogen-activated protein kinase. J. Neurosci. *32*, 17714–17724.
44. Sweatt, J.D. (2001). The neuronal MAP kinase cascade: a biochemical signal integration system subserving synaptic plasticity and memory. J. Neurochem. *76*, 1–10.
45. Lin, C.W., Cheng, C.W., Yang, T.C., Li, S.W., Cheng, M.H., Wan, L., Lin, Y.J., Lai, C.H., Lin, W.Y., and Kao, M.C. (2008). Interferon antagonist function of Japanese encephalitis virus NS4A and its interaction with DEAD-box RNA helicase DDX42. Virus Res. *137*, 49–55.
46. Sottocornola, R., Royer, C., Vives, V., Tordella, L., Zhong, S., Wang, Y., Ratnayaka, I., Shipman, M., Cheung, A., Gaston-Massuet, C., et al. (2010). ASPP2 binds Par-3 and controls the polarity and proliferation of neural progenitors during CNS development. Dev. Cell *19*, 126–137.
47. Jia, J., Yao, P., Arif, A., and Fox, P.L. (2013). Regulation and dysregulation of 3′UTR-mediated translational control. Curr. Opin. Genet. Dev. *23*, 29–34.
48. Tang, M., Ivakine, E., Mahadevan, V., Salter, M.W., and McInnes, R.R. (2012). Neto2 interacts with the scaffolding protein GRIP and regulates synaptic abundance of kainate receptors. PLoS ONE *7*, e51433.
49. Walker, V.E., Wong, M.J., Atanasiu, R., Hantouche, C., Young, J.C., and Shrier, A. (2010). Hsp40 chaperones promote degradation of the HERG potassium channel. J. Biol. Chem. *285*, 3319–3329.
50. Rosales-Hernandez, A., Beck, K.E., Zhao, X., Braun, A.P., and Braun, J.E. (2009). RDJ2 (DNAJA2) chaperones neural G protein signaling pathways. Cell Stress Chaperones *14*, 71–82.
51. Horne, B.E., Li, T., Genevaux, P., Georgopoulos, C., and Landry, S.J. (2010). The Hsp40 J-domain stimulates Hsp70 when tethered by the client to the ATPase domain. J. Biol. Chem. *285*, 21679–21688.
52. Stirling, P.C., Bakhoum, S.F., Feigl, A.B., and Leroux, M.R. (2006). Convergent evolution of clamp-like binding sites in diverse chaperones. Nat. Struct. Mol. Biol. *13*, 865–870.
53. Michels, A.A., Kanon, B., Bensaude, O., and Kampinga, H.H. (1999). Heat shock protein (Hsp) 40 mutants inhibit Hsp70 in mammalian cells. J. Biol. Chem. *274*, 36757–36763.
54. Meuillet, E.J. (2011). Novel inhibitors of AKT: assessment of a different approach targeting the pleckstrin homology domain. Curr. Med. Chem. *18*, 2727–2742.

55. Taniguchi, M., Fukunaka, A., Hagihara, M., Watanabe, K., Kamino, S., Kambe, T., Enomoto, S., and Hiromura, M. (2013). Essential role of the zinc transporter ZIP9/SLC39A9 in regulating the activations of Akt and Erk in B-cell receptor signaling pathway in DT40 cells. PLoS ONE *8*, e58022.
56. Wu, Q., Li, Y., and Xiao, B. (2013). DISC1-related signaling pathways in adult neurogenesis of the hippocampus. Gene *518*, 223–230.
57. Zheng, W., Wang, H., Zeng, Z., Lin, J., Little, P.J., Srivastava, L.K., and Quirion, R. (2012). The possible role of the Akt signaling pathway in schizophrenia. Brain Res. *1470*, 145–158.
58. Kleines, M., Gärtner, A., Ritter, K., and Schaade, L. (2001). Cloning and expression of the human single copy homologue of the mouse zinc finger protein zfr. Gene *275*, 157–162.
59. Elvira, G., Massie, B., and DesGroseillers, L. (2006). The zinc-finger protein ZFR is critical for Staufen 2 isoform specific nucleocytoplasmic shuttling in neurons. J. Neurochem. *96*, 105–117.
60. Schmitz, C., Kinge, P., and Hutter, H. (2007). Axon guidance genes identified in a large-scale RNAi screen using the RNAi-hypersensitive Caenorhabditis elegans strain nre-1(hd20) lin-15b(hd126). Proc. Natl. Acad. Sci. USA *104*, 834–839.
61. Barbee, S.A., Estes, P.S., Cziko, A.M., Hillebrand, J., Luedeman, R.A., Coller, J.M., Johnson, N., Howlett, I.C., Geng, C., Ueda, R., et al. (2006). Staufen- and FMRP-containing neuronal RNPs are structurally and functionally related to somatic P bodies. Neuron *52*, 997–1009.
62. Perycz, M., Urbanska, A.S., Krawczyk, P.S., Parobczak, K., and Jaworski, J. (2011). Zipcode binding protein 1 regulates the development of dendritic arbors in hippocampal neurons. J. Neurosci. *31*, 5271–5285.
63. Raes, J., and Van de Peer, Y. (2005). Functional divergence of proteins through frameshift mutations. Trends Genet. *21*, 428–431.
64. Peisajovich, S.G., Garbarino, J.E., Wei, P., and Lim, W.A. (2010). Rapid diversification of cell signaling phenotypes by modular domain recombination. Science *328*, 368–372.
65. Rogers, R.L., and Hartl, D.L. (2012). Chimeric genes as a source of rapid evolution in Drosophila melanogaster. Mol. Biol. Evol. *29*, 517–529.
66. Brandon, N.J., and Sawa, A. (2011). Linking neurodevelopmental and synaptic theories of mental illness through DISC1. Nat. Rev. Neurosci. *12*, 707–722.