Human Molecular Genetics, 2012, Vol. 21, No. 19 ***4171–4186***
doi:10.1093/hmg/dds240
Advance Access published on June 23, 2012

RBFOX1 regulates both splicing and transcriptional networks in human neuronal development

Brent L. Fogel[1,*], Eric Wexler[2], Amanda Wahnich[1], Tara Friedrich[1], Chandran Vijayendran[1], Fuying Gao[1], Neelroop Parikshak[1], Genevieve Konopka[1,†] and Daniel H. Geschwind[1,2,3,*]

[1]Program in Neurogenetics, Department of Neurology, [2]Department of Psychiatry and [3]Department of Human Genetics, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA 90095, USA

Received March 19, 2012; Revised May 21, 2012; Accepted June 14, 2012

RNA splicing plays a critical role in the programming of neuronal differentiation and, consequently, normal human neurodevelopment, and its disruption may underlie neurodevelopmental and neuropsychiatric disorders. The RNA-binding protein, fox-1 homolog (*RBFOX1*; also termed *A2BP1* or *FOX1*), is a neuron-specific splicing factor predicted to regulate neuronal splicing networks clinically implicated in neurodevelopmental disease, including autism spectrum disorder (ASD), but only a few targets have been experimentally identified. We used RNA sequencing to identify the *RBFOX1* splicing network at a genome-wide level in primary human neural stem cells during differentiation. We observe that *RBFOX1* regulates a wide range of alternative splicing events implicated in neuronal development and maturation, including transcription factors, other splicing factors and synaptic proteins. Downstream alterations in gene expression define an additional transcriptional network regulated by *RBFOX1* involved in neurodevelopmental pathways remarkably parallel to those affected by splicing. Several of these differentially expressed genes are further implicated in ASD and related neurodevelopmental diseases. Weighted gene co-expression network analysis demonstrates a high degree of connectivity among these disease-related genes, highlighting *RBFOX1* as a key factor coordinating the regulation of both neurodevelopmentally important alternative splicing events and clinically relevant neuronal transcriptional programs in the development of human neurons.

INTRODUCTION

The genetic programs that regulate neuronal differentiation and development are of fundamental importance. In addition to transcriptional levels of control, RNA splicing plays a critical role in establishing these genome-wide programs in a tissue-specific manner (1–3). Neuron-specific RNA splicing programs are therefore of considerable interest in understanding cellular differentiation and neurodevelopment (4). Examples include the polypyrimidine tract-binding protein (*PTBP1*), which plays a key role in regulating the switch between splicing programs leading to neuronal and non-neuronal cell differentiation (5), and the *NOVA1* splicing factor, which directly regulates hundreds of neuron-specific splicing events involving such critical developmental processes as synaptogenesis and neuronal migration (6,7). As virtually all multi-exon genes undergo alternative splicing, with the majority occurring in a tissue-specific manner (2,8), a more detailed understanding of the regulation of such events is essential if we are to better understand neuronal differentiation.

RBFOX1 (also known as *A2BP1* or *FOX1*) is a neuron-specific splicing factor that exerts both positive and negative regulatory effects on alternative splicing (9) and has been previously implicated in several neurodevelopmental and neuropsychiatric disorders including autism spectrum disorder (ASD), mental retardation and epilepsy (10–15), attention-deficit hyperactivity disorder (16), bipolar disorder, schizoaffective disorder and schizophrenia (17–19). Previous work using bioinformatic predictions has suggested that *RBFOX1* affects a large tissue-specific splicing regulatory network (2,4,20,21) and this has been further supported by RNA-protein-binding studies of its homolog *RBFOX2* (*RBM9* or *FOX2*) in human embryonic stem cells (22), and studies in knockout mice which show altered synaptic

*To whom correspondence should be addressed at: University of California at Los Angeles, 695 Charles Young Drive South, Gonda Room 2506, Los Angeles, CA 90095, USA. Email: bfogel@ucla.edu (B.L.F.); dhg@mednet.ucla.edu (D.H.G.).
†Present address: Department of Neuroscience, University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390, USA.

transmission, increased membrane excitability and a predisposition to seizures (23). Together, these studies suggest that *RBFOX1* is a high-level regulatory factor in the neuronal development cascade acting through the control of RNA splicing patterns and subsequent alterations in gene expression.

To address whether *RBFOX1* regulates early genetic programs important for proper neuronal development and integrative coordinate functioning of neurons, we sought to identify the specific splicing events regulated by *RBFOX1* by performing a genome-wide analysis using RNA sequencing in differentiated primary human neural progenitor (PHNP) cells. Here we identify the splicing regulatory network downstream of *RBFOX1* and show that it regulates the alternative splicing of a large set of genes important for neuronal development, maintenance and proliferation. We further demonstrate that disruption of *RBFOX1* function also leads to widespread transcriptional changes involving additional genes important to these pathways. Furthermore, a subset of these transcriptionally altered genes show coordinate expression and are implicated in the molecular pathogenesis of ASD and other neurodevelopmental conditions, suggesting that *RBFOX1* function is also important for the correct regulation of a network of clinically relevant neuronal transcriptional programs in human neurodevelopment.

RESULTS

Characterization of *RBFOX1* expression in PHNP cells

To study the role of *RBFOX1* in human brain development, we used cultures of fetal PHNP cells (24–26) to avoid evolutionary or species-specific differences in *RBFOX1* splicing regulatory programs (23). As *RBFOX1* is itself highly alternatively spliced (9) (Fig. 1A), we performed quantitative real time-polymerase chain reaction (qRT-PCR) using primers directed to the first two constitutive exons of all *RBFOX1* isoforms (Fig. 1B) to measure its expression in progenitors and differentiating neurons. We determined that undifferentiated PHNP cells do not produce significant amounts of *RBFOX1*. However, upon differentiation into neurons, there was a substantial increase in gene expression, including isoforms with an active RNA-binding domain (27) (Fig. 1B). Consistent with this differentiation-induced increase in expression, *in situ* hybridization of *RBFOX1* in a 19-week human fetal brain shows the majority of expression to be in regions harboring post-mitotic neurons, with a reduced signal in the germinal zones relative to the basal ganglia and cortical plate (Fig. 1C). This is similar to what is seen in developmental mapping within the mouse brain, suggestive of expression in post-mitotic projection neurons and interneurons (28), implicating *RBFOX1* in the development and maturation of human neurons.

To further validate the use of the PHNPs, we assessed whether the *RBFOX1* isoform profile was comparable with that of the fetal or adult human brain (Fig. 1D). Similar isoform expression, including both nuclear and cytoplasmic forms (29), was noted among the human fetal brain, adult brain and the undifferentiated and differentiated PHNPs, indicating that *RBFOX1* isoform expression in this primary cell culture system is similar to that of the human brain *in vivo*. Furthermore, PHNPs recapitulate brain-specific inclusion of exon 16 and exclusion of muscle-specific exon 17 (30) (Fig. 1D).

RNA interference-mediated knockdown of *RBFOX1* in differentiated PHNP cells

To better study its role in neuronal development, we reduced all *RBFOX1* isoforms in PHNPs through RNA interference by introducing a short hairpin RNA (shRNA) against exon 9 (sh*RBFOX1*) (Supplementary Material, Fig. S1A). PHNPs were then differentiated over a period of 4 weeks as *RBFOX1* mRNA levels appear to be maximal at this point (Fig. 1B), suggesting a functionally important time point. Transduction into PHNP cells resulted in approximately a 50% reduction of mRNA levels (data not shown) and a 60–70% reduction of *RBFOX1* protein after 4 weeks of differentiation compared with the non-targeting hairpin (Supplementary Material, Fig. S1B). This system recapitulates the peripheral blood levels of *RBFOX1* found in clinically haploinsufficient patients with autism and developmental delay (10,12) and the reduction observed in brains from patients with ASD (15).

Knockdown of *RBFOX1* changes the alternative splicing pattern of a network of neuronal genes involved in cell differentiation and proliferation

To assess genome-wide effects of *RBFOX1* knockdown on maturing human neurons, RNAs from five biological replicates transduced either with sh*RBFOX1* or shGFP were analyzed using next-generation sequencing. Analysis of alternative splicing was performed according to published methods (8,31). We identified 996 significant alternative splicing events involving 603 unique genes (Supplementary Material, File S1).

Defining direct and indirect alternative splicing events mediated by RBFOX1

As our methods will detect alternative splicing of genes regulated by *RBFOX1* both directly and indirectly, we examined the intronic sequence flanking the most significant alternatively spliced exons for the presence of the *RBFOX1*-binding sequence (U)GCAUG (4,9,32,33) (Fig. 2B, Supplementary Material, File S2). Overall, the canonical *RBFOX1* site was detected in 56% of the 996 total significant alternative splicing events, suggesting that other splicing factors may contribute to the observed alternative splicing changes or that additional novel *RBFOX1*-binding sites may exist. The presence of the *RBFOX1* site did not correlate with the magnitude of the splicing effect (Supplementary Material, File S2). We also found that the *RBFOX1*-binding site was enriched above baseline only in the downstream introns (Fig. 2B, Supplementary Material, File S2), similar to what has previously been seen with genes showing altered alternative splicing in autistic brains with reduced *RBFOX1* levels (15).

We next examined whether the locations of the observed *RBFOX1* sites were consistent with the alternative splicing changes detected by RNA sequencing (Fig. 2B, Supplementary Material, File S2). After excluding alternative exons that possessed *RBFOX1* sites both upstream and downstream, only 51% of the remaining sites were positioned consistently

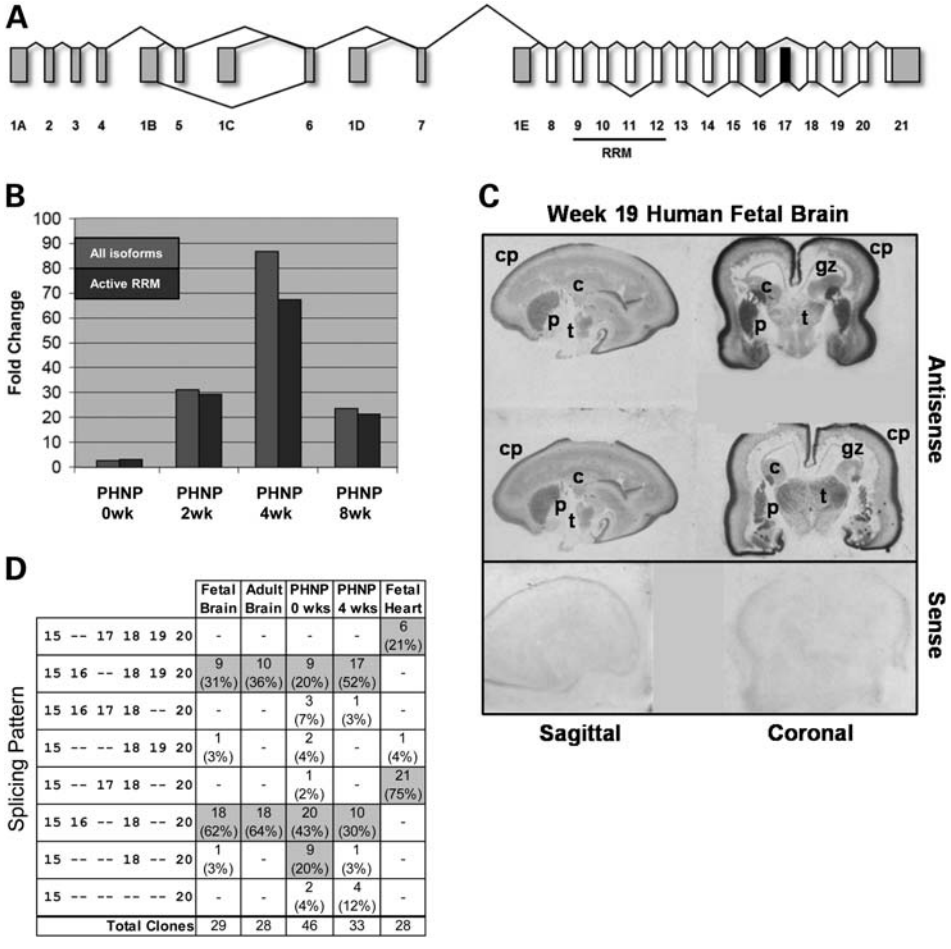
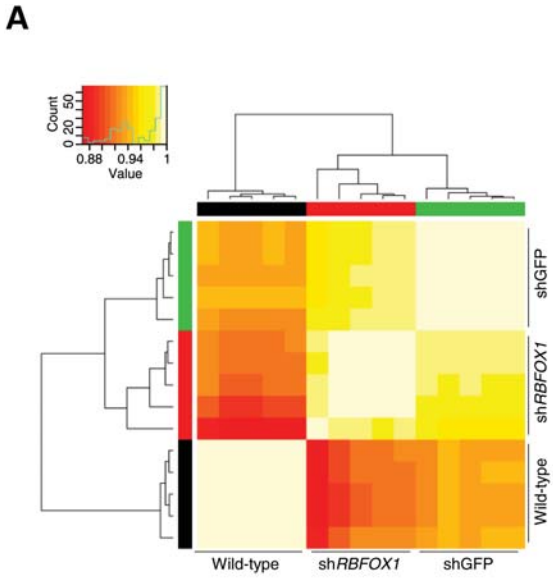


Figure 1. Characterization of *RBFOX1* expression in human brain and fetal-derived neural progenitor cells. (**A**) A schematic illustration of the *RBFOX1* genomic organization is shown. Figure adapted from Underwood *et al*. (9). Untranslated exons are shown in light gray, translated exons are in white. The brain-specific exon 16 is shown in dark gray, whereas the muscle-specific exon 17 is shown in black. The location of the RNA-binding domain (RRM) is indicated. (**B**) The expression level of *RBFOX1* was assessed by qRT-PCR, with mRNA from PHNP cells differentiated for the indicated times. Primers were directed against exons 8–9 to detect all *RBFOX1* isoforms (light gray). Autoregulatory alternative splicing of exon 11 eliminates RNA binding, so primers against exons 9–11 were utilized to detect isoforms with an active RNA-binding domain (dark gray). (**C**) *In situ* hybridization was performed with a human fetal brain, age of 19 weeks, using an S^{35}-labeled antisense riboprobe directed against exons 8–13 of *RBFOX1*. Two representative coronal and sagittal sections are shown. The sense probe is used as a control (lower panels). (**D**) To quantitate the pattern of *RBFOX1* isoforms expressed in the indicated tissues and cell lines, RT-PCR was performed using primers to amplify exons 15–20, which represent the largest region of alternative splicing diversity in the gene. Amplified products were subcloned and sequenced. Total clones are indicated with the representative counts and percentages of the various alternative spliced isoforms. The most highly expressed patterns are highlighted in gray. c, caudate; cp, cortical plate; gz, germinal zone; p, putamen; t, thalamus.

with the observed splicing change, suggesting this canonical model (9,33) may not be universal to all alternative exons. Interestingly, 63% of all correctly positioned *RBFOX1* sites were found downstream of the alternative exon (Fig. 2B, Supplementary Material, File S2), suggesting that the majority of downstream *RBFOX1* sites likely do act as enhancers, whereas upstream sites may act as silencers or enhancers depending on local context.

In mice, *Rbfox1* and *Nova1* coordinate alternative splicing during neurodevelopment (6), so we examined the intronic sequences surrounding all the alternatively spliced exons in our data set for the enrichment of the *NOVA1*-binding site, YCAY (4,6). We found a bias against the site in the upstream introns and no significant enrichment in the downstream introns where *RBFOX1* sites are enriched (Fig. 2B). In contrast, binding sites for the unrelated splicing factor, *PTBP1* (4), were found to be enriched in both upstream and downstream introns (Fig. 2B), likely consistent with its role in the repression of neuronal exons in non-neuronal tissues (5). On more detailed inspection, we did find upstream enrichment of alternate sequences that *NOVA1* has been reported to bind (34) (Supplementary Material, File S2), but *in vivo* cross-linking and immunoprecipitation experiments do not suggest these to be of functional significance (35). Therefore, these data suggest the splicing events observed here are predominantly regulated independent of the *NOVA1* pathway. To test this prediction, we chose a list of *Nova1* splicing targets derived from the mouse brain (6), where an estimated 15% may be under coordinate regulation with *Rbfox1/2* (6), and found high concordance (Fig. 2D, Supplementary Material, File S3) indicating that, despite a lack of enrichment for the *NOVA1* core binding site (Fig. 2B), many of these genes are

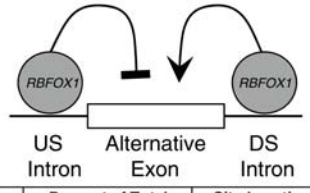


B

RBFOX1 (U)GCAUG	Observed	Predicted	Enrichment
Upstream	311	310 ±14	ns
Downstream	341	311 ±16	↑ p < 0.05

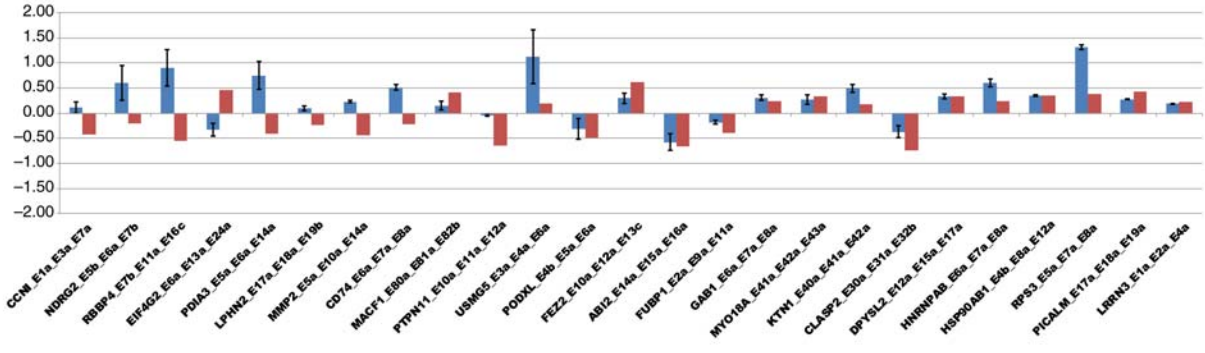
NOVA1 YCAY	Observed	Predicted	Enrichment
Upstream	4476	4554 ±34	↓ p < 0.05
Downstream	4493	4524 ±32	ns

PTBP1 CTCTCT TCTT	Observed	Predicted	Enrichment
Upstream	1086	1054 ±16	↑ p < 0.05
Downstream	1046	1009 ±16	↑ p < 0.05



RBFOX1 Site Location	Splicing Events	Percent of Total Splicing Events (996)	Site Location Predicts Splicing	Percent of Splicing Events
Upstream	213	21%	77	36%
Downstream	243	24%	154	63%
Both	98	10%	n/a	n/a
Total	554	56%	231 - 329	51 - 59%

C



D

List	RBFOX1-Dependent Genes	Autistic Brain-Associated Genes	RBFOX1 Bioinformatic	RBFOX2 Targets	Nova Targets	Post Synaptic Density Genes	Mitochondrial Genes	ER Genes	Ataxia Genes	Epilepsy Genes	Autism-Associated
Reference	Current Study	Voineagu et al 2011	Zhang et al 2008	Yeo et al 2009	Zhang et al 2010	Online	Online	Online	Supplemental	Supplemental	Supplemental
Total Events	996	212	1706	1876	417	1080	319	204	52	46	304
Unique Genes	603	194	1103	1681	303	1080	319	204	51	46	247
Overlap	-	41	107	167	46	136	11	15	4	2	9
Percent of List	-	21%	10%	10%	15%	13%	3%	7%	8%	4%	4%
Percent of RBFOX1 Genes	-	7%	18%	28%	8%	23%	2%	2%	1%	0%	1%
P-Value	-	8.58×10^{-25}	4.26×10^{-31}	2.43×10^{-51}	2.27×10^{-21}	4.36×10^{-53}	9.25×10^{-2}	3.62×10^{-4}	n/a	n/a	n/a

Figure 2. RNA sequencing detects altered alternative splicing patterns in PHNP cells with *RBFOX1* knockdown. (**A**) Heat map showing clustering of gene expression among five biological replicates representing three experimental conditions; wild-type (black), *RBFOX1* knockdown (sh*RBFOX1*, red) and non-targeting RNA interference (shGFP, green). Analysis is of the top 250 most significant genes, using the Bayes method with a Spearman correction. (**B**) Analysis of the intronic regions 400 nucleotides upstream and downstream of the alternative exons whose splicing was most significantly affected by *RBFOX1* knockdown for the presence of the binding sites for *RBFOX1*, *NOVA1* or *PTBP1*. Observed sites are shown as well as the number predicted by iterative analysis of an equivalent number of random introns culled from all human genes. Enrichment of the various sites is indicated by gray boxes and arrows. Significance is based on the normal distribution. ns, not significant. A schematic illustration of the predicted effects on alternative splicing based on the location of the *RBFOX1*-binding site is shown, with downstream sites enhancing and upstream sites repressing exon inclusion. The correlation between *RBFOX1*-binding site location and splicing changes identified by RNA sequencing in this study is shown. (**C**) Validation of splicing changes detected by RNA sequencing. Exons are labeled using a sequential annotation based on location within the gene. Genomic coordinates can be found in Supplementary Material, File S1. qRT-PCR or semi-qRT-PCR was used to calculate the ratio of exon inclusion in the *RBFOX1* knockdown cells lines when compared with the shGFP control line. A selection of 25 genes is shown with the differential fold change (log base 2) in exon inclusion detected by RNA sequencing shown in red and the observed fold change by RT-PCR shown in blue. Standard error of the mean is indicated by black bars. (**D**) Comparison of the *RBFOX1* gene set with published gene lists. The number of overlapping genes is indicated along with the percentage they represent from each list. Lists were derived from the references indicated and are also shown in Supplementary Material, File S3. Online sources for gene lists include the Genes to Cognition (G2C) database (http://www.genes2cognition.org/), the Organelle DB (http://organelledb.lsi.umich.edu/), the Online Mendelian Inheritance in Man database (http://www.omim.org/), the GeneTests database (http://www.ncbi.nlm.nih.gov/sites/GeneTests/) and the Simons Foundation Autism Research Initiative database (https://sfari.org/). Lists referenced as supplemental are composites of multiple lists derived from the above sources. *P*-values were determined based on hypergeometric probability. ER, endoplasmic reticulum.

regulated by *NOVA1*, likely through splicing events distinct from those observed here. This observation highlights the independent yet coordinative effects of *RBFOX1* and *NOVA1* on human neuronal development.

Because we did not observe *RBFOX1*- or *NOVA1*-binding sites near many of the genes we found to be alternatively spliced, we explored the hypothesis that *RBFOX1* might regulate the alternative splicing or differential expression of other splicing factors that could in turn affect downstream splicing events. We found that three splicing/RNA-processing factors (34) were alternatively spliced in our data set of predicted direct *RBFOX1* targets: heterogeneous nuclear ribonucleoprotein D (*HNRNPD*), known to regulate gene expression in the developing brain (36), and *HNRNPH1* and *HNRNPA1*, proteins that can modulate splice site selection both independently and in collaboration (37) (Supplementary Material, File S3). In our data set, *HNRNPD* sites were found to be enriched downstream of 29% of the alternative exons (285 of 996 events), whereas *HNRNPH1* was enriched upstream of 21% (205 of 996 events) (Supplementary Material, File S2). We also found that the splicing factor *ELAVL2*, thought to play a role in neuronal differentiation (38), was differentially expressed. *ELAVL2*-binding sequences were enriched both upstream and downstream of alternatively spliced exons in our data set, with bindings sites present in 36% (355) and 31% (305) of introns, respectively (Supplementary Material, File S2). Together, binding sites for either differentially spliced or differentially expressed splicing factors were found in the flanking introns of 67% of alternative exons (666 of 996) with 480 upstream events (48%) and 411 downstream events (41%) (Supplementary Material, File S2). More than half of these (56%, 374 of 666) also had a flanking *RBFOX1* site but 29% overall (292 of 996) did not, indicating that although the majority of events are likely direct, regulatory factors downstream of *RBFOX1* likely contribute to some of the splicing changes seen following *RBFOX1* knockdown.

Validation of observed RBFOX1*-dependent alternative splicing events and classification as a neurodevelopmental splicing network*

To validate these *RBFOX1*-dependent splicing results, we performed quantitative RT-PCR (qRT-PCR) and obtained a supportive confirmation rate of 68% for a set of 25 splicing events (Fig. 2C). The presence of residual *RBFOX1* in these cell lines (Supplementary Material, Fig. S1B) likely influenced these results, as many of the observed RNA sequencing changes were less than 2-fold (Fig. 2C) and this level has been previously noted to reduce splicing validation by this method (39). Therefore, to independently validate our results, we compared the *RBFOX1*-dependent gene set with various relevant published gene lists, including genes bioinformatically predicted to be regulated by *RBFOX1* (20) or regulated by its homolog *RBFOX2* (22), as well as genes showing altered alternative splicing in autistic brains with reduced *RBFOX1* levels (15). Our set showed highly significant concordance rates with all these lists, and no enrichment when compared with control lists of genes from cellular organelles (Fig. 2D, Supplementary Material, File S3).

We hypothesized that the genes within this splicing network were likely involved in key neuronal pathways which might be regulated as part of the cellular role of *RBFOX1*. Gene ontology analysis revealed that genes with increased exon inclusion are involved in biological processes related to gene expression, such as cellular protein metabolism, post-transcriptional modification, translation and RNA processing, including RNA splicing (Table 1, Supplementary Material, File S4). Gene ontology analysis of the corresponding subset of genes with increased exon skipping revealed involvement in neuronal development and cytoskeletal organization (Table 1, Supplementary Material, File S4), processes related to cell differentiation. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed similar results (Supplementary Material, File S4), as did restricting the gene ontology analysis to those genes most likely to be direct targets of *RBFOX1* (Supplementary Material, File S4).

Since previous analysis of gene expression in ASD brains identified *RBFOX1* as a hub in a gene co-expression module enriched for genes involved in synaptic transmission (15), synaptic transmission is altered in *Rbfox1* knockout mice (23) and a number of autism-related genes are involved in synaptic function (40) (SFARI database, http://genes.sfari.org), we examined whether genes whose alternative splicing was affected by *RBFOX1* knockdown were also enriched in the human post-synaptic density (PSD). We compared proteomic analysis of the murine PSD (41) with the *RBFOX1* splicing targets identified here and found that nearly a quarter of the *RBFOX1* target set is enriched in PSD proteins (136 genes, 23% of total targets) (Fig. 2D; Supplementary Material, File S3), indicating a role of *RBFOX1* in the splicing regulation of genes involved in synaptic maturation and function (see also Supplementary Material, Analysis).

Knockdown of *RBFOX1* results in the differential expression of a distinct network of genes, including a subset related to neurodevelopmental disease

Given that altered splicing in the *RBFOX1*-depleted PHNP cells affects factors related to gene expression (Table 1, Supplementary Material, File S4) and evidence from autistic brains with reduced *RBFOX1* levels suggests that transcriptional programs play an important role in neurodevelopmental disease (15), we next examined how gene expression, irrespective of splicing, was affected by *RBFOX1* knockdown. We identified 981 significant differentially expressed genes in the *RBFOX1* knockdown cell line (Supplementary Material, File S5). Validation was performed using qRT-PCR and a confirmation rate of 86% was obtained for a sample of 44 differentially expressed genes (Fig. 3A).

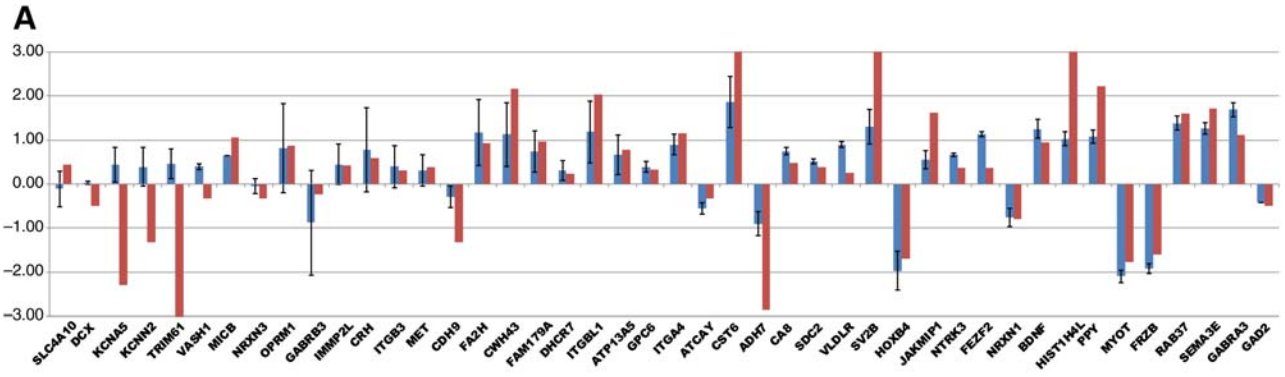
Differentially expressed genes downstream of RBFOX1 *comprise a transcriptional network distinct from the* RBFOX1 *splicing network*

Given that RNA transcription and processing are coupled and alterations in the splicing of a transcript can influence its expression (42,43), we compared the *RBFOX1*-dependent transcriptional and splicing data sets and found that the differentially expressed genes regulated downstream of *RBFOX1* were not alternatively spliced (Fig. 3B, Supplementary Material, File S6). Gene ontology analysis revealed that genes with increased expression upon *RBFOX1* knockdown are related to

Table 1. Gene ontologies of *RBFOX1*-dependent differentially spliced genes

GO term	*P*-value	Gene members
GENES WITH INCREASED EXON INCLUSION (308 GENES)		
GO:0044267—cellular protein metabolic process	1.39×10^{-5}	*HSP90AB1*, *MORF4L1*, *RPL17*, *NRBP1*, *PDIA6*, *RPS2*, *CANX*, *RPS3*, *OS9*, *APP*, *GAB1*, *PSMD2*, *RPN2*, *PPP2R1A*, *TNIK*, *PTPRG*, *DDB1*, *WNK1*, *PTPRS*, *CCT6A*, *DGUOK*, *MYH9*, *LRPAP1*, *PRKD1*, *PJA2*, *ST13*, *PSMA2*, *EIF4G2*, *PJA1*, *PPM1G*, *MAP4K4*, *EIF2AK1*, *DCUN1D2*, *PFDN5*, *CD81*, *RPS10*, *USP22*, *PAM*, *USP8*, *PTPLAD1*, *NNAT*, *ADH5*, *ABI1*, *ABCA1*, *EPHB2*, *PTK2*, *EIF3G*, *PSMB1*, *EIF3E*, *RPL3*, *EIF3L*, *C19ORF62*, *RPL4*, *RPS20*, *HSPA8*, *DCLK1*, *HSPA9*, *EEF1A1*, *RPSA*, *PTPRZ1*, *TRIM28*, *LRRN3*, *PTPRA*, *EIF4B*, *WSB1*, *CCT7*, *HSP90B1*, *PSMC5*, *RPL23*, *CUL4A*, *UBA1*, *CDC42BPA*, *AHSA1*
GO:0051246—regulation of protein metabolic process	1.25×10^{-4}	*HSP90AB1*, *A2M*, *PTPLAD1*, *KITLG*, *IGF2BP2*, *APP*, *SET*, *PSMB1*, *EIF3E*, *GAB1*, *PSMD2*, *QKI*, *PUM2*, *SAMD4A*, *PPP2R1A*, *CCDC88A*, *NDFIP1*, *FLNA*, *PSMA2*, *EIF4B*, *ATXN2*, *EIF4G2*, *PSMC5*, *EIF2AK1*, *CD81*
GO:0006414—translational elongation	1.33×10^{-4}	*RPSA*, *RPL17*, *EEF1A1*, *RPL23*, *RPL3*, *RPS10*, *RPL4*, *RPS20*, *RPS2*, *RPS3*
GO:0010608—post-transcriptional regulation of gene expression	2.11×10^{-4}	*DHX9*, *SYNCRIP*, *IGF2BP2*, *FLNA*, *EIF4B*, *ATXN2*, *EIF4G2*, *APP*, *EIF2AK1*, *EIF3E*, *QKI*, *PABPC1*, *PUM2*, *SAMD4A*
GO:0008380—RNA splicing	3.71×10^{-4}	*FUS*, *DHX9*, *PPP2R1A*, *POLR2E*, *SYNCRIP*, *HNRNPA1*, *SF3B2*, *NONO*, *HNRNPA3*, *HNRNPH3*, *SNRNP200*, *QKI*, *HNRNPC*, *PABPC1*, *HNRNPH1*, *PUF60*
GO:0016070—RNA metabolic process	4.65×10^{-4}	*RNASEN*, *FUS*, *HSD17B10*, *POLR2E*, *MYEF2*, *SYNCRIP*, *SF3B2*, *NONO*, *HNRNPA3*, *DDX17*, *APP*, *HSF2*, *EIF3E*, *DDX24*, *QKI*, *PABPC1*, *HNRNPC*, *KHDRBS1*, *PPP2R1A*, *DHX9*, *SMG7*, *HNRNPA1*, *HNRPDL*, *EIF4G2*, *ATXN2*, *HNRNPH3*, *PSMC5*, *GTF2I*, *SNRNP200*, *SUPT16H*, *NFE2L1*, *HNRNPH1*, *NCOR1*, *PUF60*
GO:0000902—cell morphogenesis	1.36×10^{-3}	*NRP2*, *PTPRZ1*, *TRIM28*, *CLU*, *MYH9*, *PCM1*, *EPHB2*, *CTNNB1*, *SS18*, *PTK2*, *APP*, *MACF1*, *RHOA*, *FEZ2*, *DCLK1*, *ADD1*, *HNRNPAB*
GO:0006900—membrane budding	3.21×10^{-3}	*PICALM*, *ARF1*, *COPB1*, *ARCN1*
GO:0010605—negative regulation of macromolecule metabolic process	3.59×10^{-3}	*RNASEN*, *HSP90AB1*, *A2M*, *CBX3*, *IGF2BP2*, *CTNNB1*, *RPS3*, *CBX5*, *SET*, *PSMB1*, *EIF3E*, *NPM1*, *PSMD2*, *HNRNPAB*, *ENO1*, *KHDRBS1*, *PPP2R1A*, *TRIM28*, *NDFIP1*, *ILF3*, *FLNA*, *PSMA2*, *PSMC5*, *EIF2AK1*, *PHB2*, *NCOR1*
GO:0048193—Golgi vesicle transport	3.65×10^{-3}	*NRBP1*, *ARF1*, *COPB1*, *ARCN1*, *SORT1*, *SAR1A*, *DTNBP1*, *HSPA8*, *DNM2*
GENES WITH INCREASED EXON SKIPPING (400 GENES)		
GO:0044267—cellular protein metabolic process	3.26×10^{-8}	*MORF4L1*, *CUL3*, *APP*, *PLOD2*, *TLK2*, *RPN2*, *DDOST*, *ADAM9*, *WNK1*, *PTPRS*, *EEF2*, *SKP1*, *MYH9*, *PPP1CC*, *MARK2*, *MAP4K4*, *SPAG9*, *CPE*, *RPS17*, *PSMA4*, *MRPL47*, *RPS10*, *EP400*, *FGFR1*, *PAM*, *ABCA1*, *PSMA7*, *CALR*, *KARS*, *CD74*, *EPHB2*, *PSMB5*, *PSMB4*, *EIF3A*, *RPS27*, *PSMB7*, *RPL7*, *RPL8*, *RPL3*, *DYRK4*, *EIF3L*, *RPL4*, *DNAJA3*, *RPSA*, *EEF1A1*, *PTPN13*, *SUGT1*, *RPS6*, *RPS8*, *PTPN11*, *EIF4B*, *PSMC5*, *PPIB*, *VCP*, *UBA1*, *SPTBN1*, *ABL1*, *UCHL1*, *CANX*, *RPS3*, *PRMT1*, *MGRN1*, *USP11*, *PAK1*, *PSMD7*, *PRKCA*, *PPP2R1A*, *HSP90AA1*, *PIGX*, *SARS*, *DDB1*, *LRPAP1*, *EIF4G2*, *DDR1*, *HIPK2*, *CD81*, *USP22*, *FKBP8*, *PML*, *RPS15A*, *ABI2*, *PTPMT1*, *MTMR3*, *GALNT10*, *SQSTM1*, *DCLK2*, *NEDD4L*, *DCLK1*, *CSNK1A1*, *TCP1*, *RRBP1*, *PTPRZ1*, *TRIM28*, *PTPRA*, *WSB1*, *HSP90B1*, *CSNK1D*, *RPL13A*, *PSMD10*
GO:0006414—translational elongation	1.83×10^{-7}	*RPSA*, *EEF1A1*, *RPS15A*, *EEF2*, *RPS6*, *RPS8*, *RPS3*, *RPS27*, *RPL7*, *RPS17*, *RPL13A*, *RPL8*, *RPL3*, *RPS10*, *RPL4*
GO:0000902—cell morphogenesis	8.45×10^{-7}	*UCHL1*, *CLU*, *SOX9*, *CTNNB1*, *EPHB2*, *CUL3*, *APP*, *SLC1A3*, *ROBO2*, *CLASP2*, *TOP2B*, *DCLK1*, *DLG1*, *PRKCA*, *ACTB*, *PTPRZ1*, *TRIM28*, *CELSR2*, *DOCK7*, *MYH9*, *PCM1*, *PTPN11*, *MARK2*, *NOTCH1*, *DST*, *ADD1*, *FEZ1*
GO:0022008—neurogenesis	2.24×10^{-6}	*FGFR1*, *CLU*, *XRCC6*, *UCHL1*, *PEX5*, *ABI2*, *CALR*, *TIMP2*, *TTC3*, *EPHB2*, *CTNNB1*, *APP*, *SLC1A3*, *ATXN10*, *MTCH1*, *ROBO2*, *AGRN*, *TOP2B*, *DCLK1*, *PRKCA*, *ACTB*, *MYO6*, *PTPRZ1*, *LDB1*, *DOCK7*, *CELSR2*, *PTPN11*, *NOTCH3*, *NOTCH1*, *BAX*, *PBX1*, *CUX1*, *DBN1*, *DST*, *SMARCA4*, *FEZ1*
GO:0031175—neuron projection development	2.10×10^{-5}	*PRKCA*, *ACTB*, *FGFR1*, *MYO6*, *PTPRZ1*, *CLU*, *UCHL1*, *ABI2*, *DOCK7*, *CELSR2*, *PTPN11*, *EPHB2*, *NOTCH1*, *APP*, *ATXN10*, *ROBO2*, *TOP2B*, *DST*, *DCLK1*, *FEZ1*
GO:0051246—regulation of protein metabolic process	3.79×10^{-5}	*A2M*, *CPEB2*, *PML*, *CALR*, *PSMA7*, *PSMB5*, *PSMB4*, *APP*, *PSMB7*, *QKI*, *PUM2*, *NEDD4L*, *PAK1*, *PSMD7*, *DNAJA3*, *ADAM9*, *PRKCA*, *PPP2R1A*, *NRD1*, *DOCK7*, *SKP1*, *FLNA*, *EIF4B*, *EIF4G2*, *PSMC5*, *PSMD10*, *BAX*, *PSMA4*, *PRKAR1A*, *CD81*, *PPP2R4*
GO:0048666—neuron development	3.82×10^{-5}	*PRKCA*, *ACTB*, *FGFR1*, *MYO6*, *PTPRZ1*, *CLU*, *UCHL1*, *ABI2*, *DOCK7*, *CELSR2*, *PTPN11*, *EPHB2*, *NOTCH1*, *APP*, *SLC1A3*, *ATXN10*, *MTCH1*, *ROBO2*, *AGRN*, *TOP2B*, *DST*, *DCLK1*, *FEZ1*
GO:0031325—positive regulation of cellular metabolic process	6.61×10^{-5}	*ARNT2*, *XRCC6*, *PML*, *ABCA1*, *CALR*, *SOX9*, *PSMA7*, *TCF7L2*, *CTNNB1*, *PSMB5*, *PSMB4*, *PSMB7*, *APP*, *MEIS2*, *SQSTM1*, *AGRN*, *PSMD7*, *DNAJA3*, *ADAM9*, *PRKCA*, *HSP90AA1*, *MYO6*, *MTA2*, *TRIM28*, *NRD1*, *TEAD1*, *DOCK7*, *SKP1*, *DDX5*, *NOTCH1*, *PSMC5*, *BPTF*, *PSMD10*, *PSMA4*, *HIPK2*, *SMARCC2*, *CD81*, *PBX1*, *PPP2R4*, *USP22*, *NFIB*, *SMARCA4*
GO:0007015—actin filament organization	6.84×10^{-5}	*ACTN4*, *CALD1*, *ELN*, *FSCN1*, *ABI2*, *DBN1*, *PLS3*, *FLNA*, *ADD1*, *DLG1*
GO:0032990—cell part morphogenesis	7.01×10^{-5}	*PRKCA*, *ACTB*, *OPA1*, *PTPRZ1*, *CLU*, *UCHL1*, *DOCK7*, *CELSR2*, *PCM1*, *PTPN11*, *EPHB2*, *NOTCH1*, *APP*, *BAX*, *ROBO2*, *TOP2B*, *DST*, *DCLK1*, *FEZ1*

Each represented gene contains at least one exon which was differentially spliced in the *RBFOX1* knockdown cell line. A total of 603 unique genes were analyzed; however, 105 of the genes contain at least one exon which showed increased inclusion as well as at least one other exon which showed increased skipping. Gene ontology (GO) analysis was performed using the indicated gene groups with the publically available DAVID bioinformatic software package (http://david.abcc.ncifcrf.gov/; NIAID, NIH). Gene subsets were chosen for analysis to better illustrate potential points of differential regulation. The top 10 parent GO terms from the level 5 biological process category are shown based on the *P*-values calculated by the DAVID algorithm. To maximize diversity of the GO terms, child terms belonging to a parent term already present on the list were removed. Gene members for each GO term are shown. The complete list of GO terms can be found in Supplementary Material, File S4.



B

List	*RBFOX1*-Dependent Genes	Autistic Brain-Associated Genes	*RBFOX1* Bioinformatic	*RBFOX2* Targets	*Nova* Targets	Post Synaptic Density Genes	Mitochondrial Genes	ER Genes	Ataxia Genes	Epilepsy Genes	Autism-Associated
Reference	Current Study	Voineagu et al 2011	Zhang et al 2008	Yeo et al 2009	Zhang et al 2010	Online	Online	Online	Supplemental	Supplemental	Supplemental
Total Events	996	212	1706	1876	417	1080	319	204	52	46	304
Unique Genes	603	194	1103	1681	303	1080	319	204	51	46	247
Overlap	24	7	67	66	27	74	1	8	5	7	27
Percent of List	4%	4%	6%	4%	9%	7%	0%	4%	10%	15%	11%
Percent of Differentially Expressed Genes	2%	1%	7%	7%	3%	8%	0%	1%	1%	1%	3%
P-Value	7.05×10^{-2}	1.28×10^{-1}	2.07×10^{-3}	2.76×10^{-2}	2.94×10^{-4}	6.45×10^{-6}	6.50×10^{-6}	1.33×10^{-1}	n/a	n/a	n/a

Figure 3. Characterization of differential gene expression in *RBFOX1* knockdown cells. (**A**) A selection of 44 genes is shown with the differential fold change (log base 2) in gene expression detected by RNA sequencing shown in red and the observed fold change by qRT-PCR shown in blue. Standard error of the mean is indicated by black bars. (**B**) Comparison of the *RBFOX1* differentially expressed gene set with published gene lists. The number of overlapping genes is indicated along with the percentage they represent from each list. Lists were derived from the references indicated and are also shown in Supplementary Material, File S6. Online sources for gene lists are as described for Figure 2. Lists referenced as supplemental are composites of multiple lists derived from the above sources. *P*-values were determined based on hypergeometric probability. ER, endoplasmic reticulum.

cell migration, signaling and proliferation (Table 2, Supplementary Material, File S7), whereas genes with decreased expression are involved in neuronal development, synapse function and cell differentiation (Table 2, Supplementary Material, File S7), remarkably similar pathways to those seen for the genes whose alternative splicing was dependent on *RBFOX1*. KEGG pathway enrichment was also similar (Supplementary Material, File S7). Despite the overlap in functional pathways, the transcriptionally regulated gene set did not overlap with *RBFOX1*-dependant splicing targets identified here or elsewhere (Fig. 3B, Supplementary Material, File S6). Together this implies that similar functional pathways and programs are being regulated by these two processes, splicing and transcription, and that although these *RBFOX1*-regulated transcriptional and splicing networks are distinct from one another, they converge on common biological processes important to neuronal differentiation.

A subset of genes in the RBFOX1 *transcriptional network are responsible for neurodevelopmental disorders*

As *RBFOX1* has been implicated in several neurodevelopmental and neuropsychiatric disorders (10–19), we hypothesized the mechanism could be via the missplicing of other disease genes, as has been previously suggested (15). We compared our differential splicing list with lists of genes associated with spinocerebellar ataxia, epilepsy and autism but found these represented no more than 1–2% of the genes within the set (Fig. 2D, Supplementary Material, File S3). Despite haploinsufficiency of *RBFOX1* previously being shown to cause autism and ASD (10,12), we notably did not find a significant enrichment for alternatively spliced genes currently predicted to be related to autism (40) (SFARI database, http://genes.sfari.org) (Fig. 2D, Supplementary Material, File S3). This lack of strong overlap indicates that the relationship between *RBFOX1* function and neurodevelopmental disease is unlikely to occur primarily due to the direct *RBFOX1*-controlled missplicing of known disease genes, as was suggested by previous analysis of the ASD brain (15).

In contrast to the differentially spliced genes, those with differential expression were 3-fold enriched in known disease genes (Figs 2D and 3B, Supplementary Material, Files S3 and S6), suggesting altered transcriptional regulation could play a role in *RBFOX1*'s association with neuropsychiatric disease. Among these differentially expressed genes were several well-supported ASD candidates including *DHCR7*, *GABRB3*, *ITGB3*, *NRXN1* and *MET* (40) (Supplementary Material, File S6). This subset of transcriptionally regulated genes implicated in ASD suggests an important role for *RBFOX1* in the regulation of critical transcriptional programs required for proper human neurodevelopment that can increase risk for common neurodevelopmental disorders when disrupted.

Weighted gene co-expression network analysis demonstrates connectivity among differentially expressed genes related to ASD

To better understand the systems level organization of the transcriptional network altered by *RBFOX1* knockdown, we performed weighted gene co-expression network analysis (WGCNA) (15,44–46). To identify the modules most

Table 2. Gene ontologies of *RBFOX1*-related differentially expressed genes

GO term	*P*-value	Gene members
GENES WITH INCREASED EXPRESSION (625 GENES)		
GO:0030334—regulation of cell migration	1.12×10^{-10}	*IL6ST*, *ENPP2*, *ITGB3*, *ADORA1*, *TPM1*, *PRR5*, *TEK*, *THBS1*, *FGF2*, *ADAM9*, *EGFR*, *COL18A1*, *PARD6B*, *PLD1*, *SMAD3*, *ITGA2*, *PTPRU*, *NEXN*, *KDR*, *CDH13*, *LAMA4*, *LAMA3*, *CXCL16*, *PDGFRB*, *HBEGF*, *TGFBR3*, *IGFBP3*
GO:0001568—blood vessel development	1.27×10^{-9}	*CAV1*, *HTATIP2*, *CD44*, *CTGF*, *DHCR7*, *ITGAV*, *RHOB*, *ZC3H12A*, *ANGPT1*, *ADRA2B*, *LOX*, *THBS1*, *PLXND1*, *FGF2*, *C1GALT1*, *CYR61*, *PLAT*, *COL18A1*, *TGFBR2*, *ITGA4*, *SLIT2*, *ANXA2P2*, *ANXA2*, *KDR*, *CDH13*, *LAMA4*, *ITGA7*, *TGFBR3*, *COL1A1*, *TNFAIP2*, *PLAU*
GO:0007584—response to nutrient	5.91×10^{-8}	*MUC1*, *A2M*, *CAV1*, *SLC8A1*, *RARG*, *CCL2*, *CFB*, *IL6ST*, *IGFBP7*, *TGFBR2*, *ITGA2*, *VDR*, *CCND1*, *CDKN2B*, *CD44*, *ANGPT1*, *COL1A1*, *LRP2*, *IGFBP2*, *LCT*, *VLDLR*
GO:0031960—response to corticosteroid stimulus	1.63×10^{-7}	*KCNMA1*, *A2M*, *CAV1*, *CCL2*, *IGFBP7*, *PTPRU*, *CDO1*, *ABCB4*, *CCND1*, *ADM*, *FAS*, *COL1A1*, *IGFBP2*, *FOSL1*, *NEFL*, *ADAM9*
GO:0001525—angiogenesis	6.95×10^{-7}	*COL18A1*, *HTATIP2*, *TGFBR2*, *SLIT2*, *ANXA2P2*, *ANXA2*, *KDR*, *CDH13*, *CTGF*, *ZC3H12A*, *RHOB*, *ANGPT1*, *ADRA2B*, *PLXND1*, *THBS1*, *FGF2*, *TNFAIP2*, *C1GALT1*, *PLAU*, *CYR61*
GO:0007167—enzyme-linked receptor protein signaling pathway	2.24×10^{-6}	*TWSG1*, *CCL2*, *LTBP2*, *IL6ST*, *GDF6*, *RPE65*, *FSTL1*, *LIF*, *CTGF*, *LEFTY2*, *TEK*, *PTN*, *ANGPT1*, *NRG1*, *FGF2*, *ADAM9*, *PLAT*, *EGFR*, *PTPRE*, *ARID5B*, *TGFBR2*, *MET*, *AXL*, *SMAD3*, *PTPRU*, *KDR*, *NTRK3*, *PDGFRB*, *HBEGF*, *TGFBR3*, *GRB7*
GO:0008285—negative regulation of cell proliferation	2.38×10^{-6}	*CAV2*, *CAV1*, *IGFBP7*, *PAWR*, *POU1F1*, *ADORA1*, *CD9*, *VDR*, *BDNF*, *GPC3*, *CDKN2B*, *GPNMB*, *THBS1*, *FGF2*, *FOSL1*, *DHCR24*, *COL18A1*, *RARG*, *KAT2B*, *TGFBR2*, *S100A11*, *SMAD3*, *PTPRU*, *CD164*, *NOTCH2*, *CDH13*, *ADAMTS8*, *ADM*, *TGFBR3*, *PMP22*, *EMP3*, *IGFBP3*
GO:0009888—tissue development	4.05×10^{-6}	*CAV2*, *TWSG1*, *CAV1*, *TNNC1*, *ELN*, *CYTL1*, *PAX2*, *TPM1*, *LIF*, *VDR*, *GPC3*, *CD44*, *PLOD1*, *CTGF*, *LHX2*, *DMD*, *ALDH1A3*, *SERPINE1*, *PTN*, *AHNAK2*, *NRG1*, *GPNMB*, *COL11A1*, *DHCR24*, *ADAM9*, *COL18A1*, *F11R*, *SATB2*, *SCUBE1*, *UGCG*, *ANXA1*, *SMAD3*, *RCAN1*, *COL5A3*, *COL5A2*, *SLIT2*, *KDR*, *FZD6*, *NOTCH2*, *LAMA3*, *TGFBR3*, *PDGFRB*, *ID3*, *LAMC1*, *COL1A1*, *PLAU*, *EMP1*
GO:0045785—positive regulation of cell adhesion	5.04×10^{-6}	*CD47*, *CDH13*, *CD36*, *FBLN2*, *EMID2*, *SMAD3*, *ITGA2*, *NRG1*, *THBS1*, *TPM1*, *ADAM9*, *CYR61*
GO:0051272—positive regulation of cell motion	5.82×10^{-6}	*EGFR*, *COL18A1*, *PLD1*, *IL6ST*, *SMAD3*, *ITGA2*, *KDR*, *CDH13*, *PRR5*, *CXCL16*, *HBEGF*, *PDGFRB*, *THBS1*, *FGF2*, *ADAM9*
GENES WITH DECREASED EXPRESSION (356 GENES)		
GO:0022008—neurogenesis	3.37×10^{-9}	*GPRIN1*, *NRTN*, *CDK5R1*, *HELT*, *NNAT*, *ONECUT2*, *TH*, *L1CAM*, *BRSK1*, *KIT*, *KCNIP2*, *GPC2*, *BCL11B*, *MAPT*, *BAI1*, *ROBO2*, *DCX*, *DNMT3B*, *TUBB3*, *DSCAM*, *DTX1*, *NRXN3*, *STMN2*, *SOX11*, *CELSR3*, *ARTN*, *NRXN1*, *DLX2*, *DLX1*, *SEMA6C*, *DLX5*, *PDGFRA*, *CNTN2*, *CHRNB2*, *KALRN*
GO:0048666—neuron development	1.41×10^{-5}	*GPRIN1*, *CDK5R1*, *NRTN*, *NRXN3*, *ONECUT2*, *TH*, *CELSR3*, *L1CAM*, *NRXN1*, *KCNIP2*, *SEMA6C*, *DLX5*, *BCL11B*, *BAI1*, *CNTN2*, *ROBO2*, *CHRNB2*, *DCX*, *DSCAM*, *KALRN*
GO:0032990—cell part morphogenesis	7.01×10^{-5}	*CDK5R1*, *NRXN3*, *ONECUT2*, *CELSR3*, *L1CAM*, *NRXN1*, *SEMA6C*, *DLX5*, *BCL11B*, *BAI1*, *CNTN2*, *CHRNB2*, *ROBO2*, *DCX*, *DSCAM*, *KALRN*
GO:0007268—synaptic transmission	1.13×10^{-4}	*KCNMB4*, *GABRB3*, *NRXN3*, *SLC12A5*, *TH*, *BSN*, *NRXN1*, *KCNIP2*, *GAD2*, *GRIA2*, *SYN1*, *DLG4*, *CHRNB2*, *CORT*, *KCNQ2*, *GAD1*, *CACNA1B*
GO:0000902—cell morphogenesis	2.82×10^{-4}	*ARHGEF2*, *CDK5R1*, *NRXN3*, *ONECUT2*, *CELSR3*, *L1CAM*, *NRXN1*, *BRSK1*, *SEMA6C*, *DLX5*, *BCL11B*, *BAI1*, *CNTN2*, *ROBO2*, *CHRNB2*, *DCX*, *DSCAM*, *KALRN*
GO:0009887—organ morphogenesis	6.33×10^{-4}	*FGF6*, *MAFB*, *FGF9*, *ONECUT2*, *TH*, *TLE2*, *T*, *DLX2*, *COL9A1*, *HOXB4*, *ACVR2B*, *DLX1*, *HOXB2*, *EYA2*, *CHD7*, *AES*, *FREM2*, *HOXB5*, *DLX5*, *HOXB6*, *PDGFRA*, *DSCAM*, *PITX2*
GO:0007417—central nervous system development	7.83×10^{-4}	*CDK5R1*, *HELT*, *MAFB*, *SOX11*, *SLC6A3*, *NNAT*, *BCAN*, *CYP26A1*, *ZBTB16*, *TAGLN3*, *DLX2*, *DLX1*, *HOXB2*, *CHD7*, *SBK1*, *BCL11B*, *ROBO2*, *CHRNB2*, *DCX*
GO:0009952—anterior/posterior pattern formation	1.05×10^{-3}	*T*, *HOXB4*, *ACVR2B*, *HOXB2*, *HOXB5*, *HOXB6*, *CYP26A1*, *ZBTB16*, *AXIN2*, *MESP2*
GO:0048568—embryonic organ development	1.21×10^{-3}	*T*, *HOXB4*, *DLX2*, *HOXB2*, *CHD7*, *FGF9*, *MAFB*, *DLX5*, *HOXB5*, *TH*, *HOXB6*
GO:0048706—embryonic skeletal system development	2.77×10^{-3}	*HOXB4*, *DLX2*, *DLX1*, *HOXB2*, *FGF9*, *HOXB5*, *HOXB6*

A total of 981 unique genes were analyzed. Gene ontology (GO) analysis was performed using the indicated gene groups with the publically available DAVID bioinformatic software package (http://david.abcc.ncifcrf.gov/; NIAID, NIH). Gene subsets were chosen for analysis to better illustrate potential points of differential regulation. The top 10 parent GO terms from the level 5 biological process category are shown based on the *P*-values calculated by the DAVID algorithm. To maximize diversity of the GO terms, child terms belonging to a parent term already present on the list were removed. Gene members for each GO term are shown. The complete list of GO terms can be found in Supplementary Material, File S7.

related to *RBFOX1* function, we assessed the module eigengene relationship with *RBFOX1* knockdown status (47). Two modules were identified that showed the largest difference between *RBFOX1* knockdown and control, designated as the blue and yellow modules, respectively (Fig. 4, Supplementary Material, File S8).

Figure 4. WGCNA in the *RBFOX1* knockdown cell line reflects pathways important to neurodevelopment and to autism. For clarity, only the most highly connected module members are shown. Genes with the highest connectivity (i.e. hubs) are indicated in red. (**A**) Blue module. (**B**). Yellow module. Differentially expressed ASD genes are in purple.

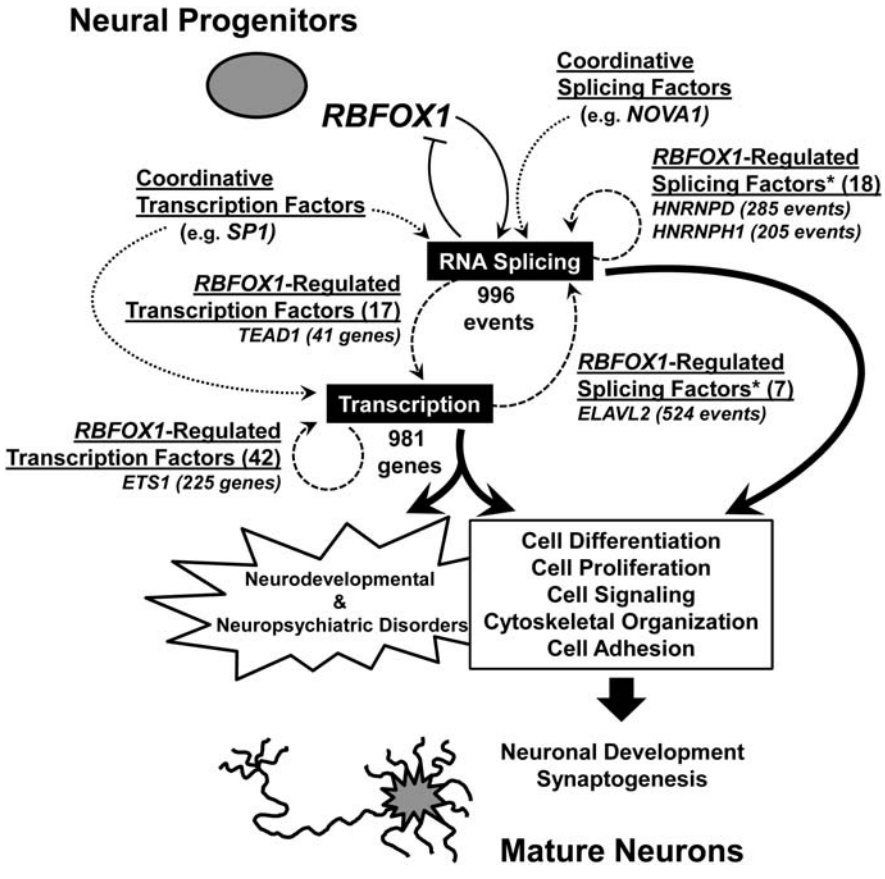


Figure 5. A model of *RBFOX1* function in PHNPs. During neuronal differentiation, *RBFOX1* is induced and directly modulates (solid arrows) an RNA splicing network (black box) which in turns coordinately regulates a transcriptional network of additional genes (black box). Downstream genes present within these networks can further modulate either RNA splicing or transcription to generate additional layers of control (dashed arrows). *RBFOX1* can further alter its own splicing and downregulate the activity of these networks as indicated. Additional regulatory factors and/or programs can also contribute to the coordinate regulation of these networks (dotted arrows). The number of factors identified in this study at each of the steps is indicated with selected proteins of interest as mediators noted with their predicted regulatory contribution indicated, see text for complete details. The genes within these networks affect a number of key cellular developmental processes that promote neural development and synaptogenesis, leading to the formation of mature neurons. Disruption of these pathways, particularly genes in the transcriptional network, can lead to neurodevelopmental and/or neuropsychiatric phenotypes in humans. The asterisk indicates factors involved in splicing and/or other aspects of RNA-processing.

Enrichment of ASD genes in the WCGNA modules

Since the differentially expressed data set was enriched in genes with established relationships to neurodevelopmental disease, we assessed the top *RBFOX1* modules for the enrichment of ASD genes. The blue module consisted of 737 genes (Fig. 4A, Supplementary Material, File S8) and its members include the transcription factor *FOXP2*, which functions in the regulation of genes important to CNS development and language (45), as well as several other genes linked to ASD (*CADPS2*, *DMD*, *HLA-DRB1*, *INPP1*, *ITGB3*, *LRFN5*, *MDGA2*, *NTRK3*, *OPRM1*, *PARK2* and *SDC2*) (48).

The yellow module contained 312 genes (Fig. 4B, Supplementary Material, File S8) and its hub genes include *DLG1*, a protein important in modulating cytoskeletal interactions to establish cell polarity (49), and implicated as a potential candidate gene in both ASD and schizophrenia (50,51). The yellow module contained several other ASD candidate genes including *CNTN3*, *DPYD*, *IMMP2L*, *MACROD2*, *MET*, *RAPGEF4*, *REEP3*, *SLC4A10* and *TPH2* (48). Of these ASD genes, six are differentially expressed following *RBFOX1* knockdown (*DLG1*, *MET*, *IMMP2L*, *REEP3*, *SLC4A10* and *TPH2*) and represent a key subset of the yellow module with a statistically significant ($P \leq 1 \times 10^{-4}$) greater average connectivity and shorter mean path length relative to all other module members (Supplementary Material, File S8). These observations show that *RBFOX1* functionally contributes to the transcriptional regulation of important neurodevelopmental cellular programs that are clinically relevant to autism. This confirms previous indirect studies in human brain, suggesting a role for *RBFOX1* in regulating ASD-related transcription (15).

DISCUSSION

Here we demonstrate that the neuronal splicing factor *RBFOX1* regulates the alternative splicing of an extensive network of genes involved in neuronal differentiation and maintenance. We examined these alternatively spliced genes and observed enrichment for pathways involved in cellular proliferation, cell signaling, cytoskeletal organization and cell adhesion, gene expression and regulation and neuronal development (Table 1, Supplementary Material, File S4),

suggesting that *RBFOX1* may represent a molecular switch in neuronal progenitors which promotes correct differentiation into functional neurons (Fig. 5). In humans, *RBFOX1* haploinsufficiency (similar to the level of knockdown achieved here) (Supplementary Material, Fig. S1B) causes neurodevelopmental disease such as ASD (10,12). Based on previous observations in autistic brains (15), we initially had predicted that genes undergoing differential alternative splicing would be related to ASD or similar clinical disorders. However, this was not observed. Instead, it was the network of genes that were differentially expressed in cells lacking *RBFOX1* that were associated with neurodevelopmental disease. Interestingly, these differentially expressed genes were also related to similar biological pathways as those that were alternative spliced (Tables 1 and 2, Supplementary Material, Files S4 and S7). A subset of these differentially expressed genes were also highly connected in a module of co-expressed genes that were enriched in ASD candidates (48) (Fig. 4B, Supplementary Material, Files S6 and S8). The *RBFOX1* protein is therefore involved in regulating both neuronal alternative splicing and coordinative transcriptional programs, both of which contribute to normal neurodevelopment (Fig. 5).

It is interesting to compare these data with recent studies, using microarray analysis in the brains of adult knockout mice lacking *Rbfox1*, identifying differentially spliced genes which impacted synaptic transmission and membrane excitability, leading to an increased susceptibility for seizure events (23). Similar to the mouse, a number of the *RBFOX1*-dependent alternatively spliced genes in our data set appear to be involved in synapse formation and function (Fig. 2D, Supplementary Material, File S3), areas of particular interest in the study of the pathogenesis of neurodevelopmental disease and autism (52,53). Despite this, we did not observe significant differences in the *RBFOX1* knockout cell line with regard to basic cellular morphology (data not shown). We attribute this to our cells not being fully devoid of *RBFOX1* (Supplementary Material, Fig. S1B), with residual activity potentially preventing observation of the more extreme cellular phenotypes, although still reflecting the underlying molecular pathogenesis affecting splicing and transcription. In this sense, the current model approximates human haploinsufficiency observed in patients, as suggested by the highly significant overlap between our alternative splicing gene data set and the genes with altered splicing identified in autistic brains with reduced *RBFOX1* (15) (Fig. 2D, Supplementary Material, File S3). To extend this comparison further, we evaluated our list of *RBFOX1*-dependent splicing events for genes related to epilepsy and found altered splicing of *FLNA*, a cytoskeletal gene involved in neural migration and heterotopia (54), and *SLC1A3*, a glutamate transport protein associated with episodic ataxia and epilepsy (55) (Fig. 2D, Supplementary Material, File S3). Additional genes associated with epilepsy were detected in the differentially expressed *RBFOX1*-dependent gene set including *DCX*, *GABRB3*, *GAD2*, *KCNQ2*, *SLC12A5*, *SV2B* and *SYN1* (Fig. 3B, Supplementary Material, File S6). This observation is similar to the increase in ASD genes in this transcriptional data set and consistent with the seizures observed in the knockout mice (23), as well as the clinical finding of seizure disorders in patients with *RBFOX1* mutations (10,12).

Correlations between *RBFOX1*-dependent genes and neurodevelopmental disorders, including ASD

In addition to *HNRNPD*, *HNRNPH1* and *HNRNPA1*, 15 additional RNA-binding proteins (56) were identified in the alternatively spliced data set (Supplementary Material, File S3), which may contribute to the neurodevelopmental phenotypes associated with *RBFOX1*. These include *DDX17*, a developmentally essential factor involved in miRNA processing and transcriptional co-regulation in cell proliferation (57,58), and *NPM1*, implicated in neuronal cell proliferation (59). Several of these genes also show altered expression in neurodevelopmental disorders, including *NONO* and *FUBP1* (60), *RPL3* (61) and *RPL7* and *RPS17* (62). In addition to *ELAVL2*, another six RNA-binding proteins (56) were also differentially expressed (Supplementary Material, File S6).

We also found 17 transcription factors that were alternatively spliced, including *TEAD1*, whose binding site was enriched in the promoters of the differentially expressed gene set (Supplementary Material, Analysis and File S9). Several of these transcription factors are known to be involved in neurodevelopment, including *ARID1B*, which causes an ASD-associated syndrome when haploinsufficient (63,64); *ARNT2*, a gene critical to hypothalamic development and associated with autistic traits (65); *RPL7*, differentially expressed in mothers of ASD children (62); *GTF2I*, involved in the Williams–Beuren neurodevelopmental syndrome (66); *MEIS2*, involved in forebrain development (67); and *NOTCH3*, causative for the inherited leukoencephalopathy CADASIL (68) (Supplementary Material, File S3). Forty-two transcription factors (69) were also differentially expressed (Supplementary Material, File S6), including *ETS1*, whose binding site was enriched in genes found in the blue module (Supplementary Material, Analysis and File S9). Binding sites for two transcription factors with some circumstantial evidence for involvement in ASD (70,71), *SP1* and *RORA*, were enriched among the differentially expressed genes (Supplementary Material, File S9), and transcriptional regulation may also occur among the differentially spliced genes (Supplementary Material, Analysis).

One ASD candidate gene which may be particularly relevant is the cell polarity gene *DLG1* (49–51), which is differentially expressed with *RBFOX1* knockdown (Supplementary Material, File S5) and is a hub in the yellow WGCNA module that is enriched for ASD genes (Fig. 4B, Supplementary Material, File S8). *DLG1* is also differentially spliced in the *RBFOX1* knockdown line (Supplementary Material, File S3) and in autistic brains (15) (Supplementary Material, File S3), and is also a target of *Nova1* in the mouse brain (6) (Supplementary Material, File S3). Further study of *DLG1* may provide valuable insights into the downstream effects of disrupting the *RBFOX1* networks. *NLGN1*, another interesting ASD candidate gene (72), was also differentially expressed.

RBFOX1 and cerebellar ataxia

Because *RBFOX1* was originally identified as an ataxin-2-binding protein (73,74), its role in cerebellar ataxia is another potential area of interest. Here we observed differential alternative splicing of the ataxia genes *ATXN2*, *ATXN10*, *FMR1* and *SLC1A3* (Fig. 2D, Supplementary Material, File

S3). Furthermore, we identified *RBFOX1*-dependent differential expression of several other ataxia genes (Fig. 3B, Supplementary Material, File S6), including *CA8* and *VLDLR*, both of which cause a developmental syndrome characterized by cerebellar hypoplasia, congenital ataxia and mental retardation (75,76). *CA8* further influences cell signaling via the function of the inositol triphosphate receptor, *ITPR1*. Mutations in *ITPR1* also cause spinocerebellar ataxia, type 16 (76,77), and *ITPR1* was also found to be differentially expressed in the *RBFOX1*-deficient cells (Fig. 3B, Supplementary Material, File S6). Interestingly, *CA8* was found in the blue WGCNA module, and *ITPR1* was in the yellow module (Fig. 4, Supplementary Material, File S8), indicating a close relationship between *RBFOX1* and these cerebellar ataxia genes. Whether certain variants in *RBFOX1* could lead to clinical ataxic syndromes independent of neurodevelopmental phenotypes has yet to be determined.

Regulation of *RBFOX1* downstream splicing and transcription events

Our data reflect both direct and indirect regulatory events resulting from the loss of *RBFOX1* protein. The finding that several RNA-processing factors and transcription factors are indeed alternatively spliced (Supplementary Material, File S3), and that other factor-binding motifs are enriched in the introns and promoters of the differentially spliced and expressed genes, respectively (Supplementary Material, Files S2 and S9), points to a complex regulatory system, likely involving multiple *cis*- and *trans*-acting elements, that is disrupted by the loss of *RBFOX1* (Fig. 5). One key question will be whether novel intronic sequences found to be enriched in the differentially spliced data set (data not shown) represent specific *RBFOX1*-associated regulatory elements. The observation of enriched motifs in the 3′-untranslated regions and promoter regions of the *RBFOX1* splicing network including some, like *SP1*, shared with the *RBFOX1* transcriptional network (Supplementary Material, File S9) suggests that certain regulatory factors and/or signaling pathways may converge to coordinately regulate both networks in the parallel execution of the *RBFOX1*-mediated developmental program (Fig. 5). Elucidation of the factors mediating these processes in both cell lines and in the human brain (15), and how the loss of *RBFOX1* impacts their function, will have important relevance to our understanding of neurodevelopment and the pathogenesis of neurodevelopmental diseases.

MATERIALS AND METHODS

Plasmid construction

The *RBFOX1* short-RNA hairpin targeting exon 9 was originally derived from clone ID TRCN0000073260 (Open Biosystems, Thermo Fisher Scientific, Rockford, IL, USA) expressed in the pLKO.1 vector. As use of this vector in our PHNP cultures can lead to vector-related cell death regardless of the RNA hairpin expressed, this hairpin was redesigned in the shRNAmir-30 context and subcloned in the pGIPZ vector (Open Biosystems, Thermo Fisher Scientific) at the *Bam*HI and *Xho*I sites as per the manufacturer's recommendations. For viral production, the hairpin was transferred to the pLCR vector (78) at the *Eco*RI and *Xho*I sites. The *RBFOX1* overexpression vector was generated by shuttling *RBFOX1* transcript 4 (NM_018723.3) from pENTR-*A2BP1* (Addgene, Cambridge, MA, USA) to pPELTR5.5.2-GA, using Gateway cloning (Invitrogen, Life Technologies, Carlsbad, CA, USA) to create pPELTR5.5.2-GA-*RBFOX1*. The pPELTR5.5.2-GA vector is a second generation lentiviral vector where the expression of the insert is under the control of an EF1-HTLV fusion promoter and includes the addition of an N-terminal HA epitope tag. A red fluorescent reporter protein is also expressed in *cis* with the insert with an intervening T2A sequence to allow for co-translational cleavage of the two proteins.

Human tissues

Normal human tissues were obtained from the NICHD Brain and Tissue Bank for Developmental Disorders at the University of Maryland School of Medicine (Baltimore, MD, USA; NICHD contract numbers N01-HD-4-3368 and N01-HD-4-3383) (fetal heart age 19 weeks and fetal brains age 18 and 19 weeks). The role of the NICHD Brain and Tissue Bank is to distribute tissue and therefore cannot endorse the studies performed or the interpretation of results. Additional tissues were obtained from the UCLA Center for AIDS Research Gene and Cellular Therapy Core (fetal brain age 17 weeks) and the UCLA Alzheimer's Disease Center Neuropathology Core (adult brain age 97 years).

Cell culture, transfection and transduction

PHNP cells were cultured from a 17-week human fetal brain, proliferated and differentiated as previously described (26). Genotyping was performed to ensure cells did not harbor major structural aberrations (24). Biological replicates (7–10 per condition) were grown on individual 10 cm plates and all transductions and other sample preparations were handled independently. For shRNA transduction, high-titer virus production was performed by the UCLA Jonsson Comprehensive Cancer Center Vector Core (Los Angeles, CA, USA), utilizing the specified lentiviral vectors, and viral transduction was performed using standard protocols.

HeLa cells (ATCC, Manassas, VA, USA) were grown under recommended conditions. Cells were transfected with the specified vectors using the FuGene 6 transfection reagent (Roche, Indianapolis, IN, USA) according to the manufacturer's recommended protocol. After 48–72 h, cells were harvested and protein was extracted for western blotting.

Preparation of RNA/protein from cultured cells and tissues

Cells were harvested and RNA was extracted using the miRNeasy Mini Kit (Qiagen, Valencia, CA, USA) as per the manufacturer's recommended instructions. For PHNP cells, protein was isolated during the RNA extraction procedure according to the supplemental protocols available at the manufacturer's website (www.qiagen.com). For HeLa cells, whole-cell protein lysates were prepared by directly lysing cells in 0.5% Nonidet P-40 buffer containing protease and

phosphatase inhibitors. For tissues, RNA and protein were prepared from 30 mg tissue using the AllPrep DNA/RNA/Protein Mini Kit (Qiagen) according to the manufacturer's instructions. All protein concentrations were determined using the Bradford (Bio-Rad, Hercules, CA, USA) or BCA (Pierce, Thermo Fisher Scientific) protein assays.

QRT-PCR

RNA was initially treated with amplification-grade DNase I (Invitrogen, Life Technologies) and then converted to cDNA, using the SuperScript III First-Strand Synthesis System Kit (Invitrogen, Life Technologies). The resulting cDNA (25–50 ng/reaction) was utilized for qRT-PCR, using Bio-Rad iTaq SYBR Green Supermix with ROX (Bio-Rad) according to the manufacturer's recommendations. Reactions (10 μl volume) were performed in triplicate using either an ABI 7900HT (Applied Biosystems, Life Technologies, Carlsbad, CA, USA) or a LightCycler 480 (Roche) qPCR machine. Data analysis was performed using the comparative C_t method normalized to β-actin. Reactions where the replicates deviated by more than 0.2 C_t from the mean were excluded from further analysis. Each data point represents a minimum of two individual reactions. For confirmation of differential gene expression and splicing from transduced PHNP cell lines, an equal mixture of RNA from five biological replicates was used as a template to minimize bias from an individual sample. All primers were designed to cross an exon–exon junction and, for the detection of splicing events, this would span the junction between the alternative exon and one of its neighboring exons.

For semi-qRT-PCR analysis, 10 ng of cDNA was amplified within the linear range, using Accuprime *Pfx* Supermix (Invitrogen, Life Technologies) with primer pairs designed to detect products with and without the alternatively spliced exon. The program used was 94°C for 15 s, 55°C for 15 s and 72°C for 15 s for an average of 37 cycles. Products were separated on a 3.5% agarose gel, stained with ethidium bromide, photographed and analyzed by densitometry, using the publically available software ImageJ (http://imagej.nih.gov/ij/).

RBFOX1 gene nomenclature and isoform analysis

RBFOX1 genomic organization nomenclature used in this report is as published in Underwood *et al* (9). For isoform analysis, RNA was harvested from the indicated cultured cell lines and human tissues, converted to cDNA as above and RT-PCR was performed using PCR Supermix (Invitrogen, Life Technologies) with primers spanning exons 15–20, the region of highest alternative splicing diversity. The resulting products were subcloned using the TOPO TA cloning kit (Invitrogen) and 30–50 individual inserts were sequenced.

In situ hybridization

Sagittal and coronal sections of human fetal brain, age 19 weeks, were probed using an S^{35}-labeled antisense riboprobe directed against exons 8–13 of *RBFOX1* as previously described (79). The sense riboprobe was used as the control.

Immunoblotting

Protein (50–100 μg) was loaded onto 10–12% SDS–polyacrylamide gels, separated by electrophoresis, transferred onto Immun-Blot PVDF membranes (Bio-Rad), blocked in 5% milk, probed with primary antibody diluted into TBST for 1 h to overnight at 4°C and then probed with a secondary antibody diluted into TBS for 1 h to overnight at 4°C. Detection was performed using SuperSignal West Pico chemiluminescent substrate (Pierce, Thermo Fisher Scientific) according to the manufacturer's instructions and membranes were exposed to film. Band quantitation was performed from scanned images, using publically available densitometry software and all samples were normalized to β-actin levels.

Antibodies

Primary antibodies were against HA-tag (1:2000, Covance, Princeton, NJ, USA), *RBFOX1* (rabbit polyclonal, 1:500, Aviva Systems Biology, San Diego, CA, USA) or β-actin (mouse monoclonal, 1:250 000, Sigma-Aldrich, St Louis, MO, USA). Secondary antibodies used were goat anti-rabbit horseradish peroxidase (1:2000, Cell Signaling Technology, Danvers, MA, USA) and goat anti-mouse horseradish peroxidase (1:4000, Millipore, Billerica, MA, USA).

RNA sequencing and data analysis

RNA sequencing at a 75 bp single-end read scale was performed on an Illumina Platform G2 sequencer (Yale Center for Genome Analysis, West Haven, CT, USA), using polyA-enriched RNA from five biological replicates of the shGFP and sh*RBFOX1* PHNP cell lines differentiated for 4 weeks. Unique sequence reads were filtered for repetitive and ribosomal RNA sequences and then aligned to the human genome (hg19, NCBI 37), using the Burrows-Wheeler alignment tool (BWA) (80), allowing a maximum of three mismatches to generate reads mapping to exons ($54\,779\,644 \pm 1\,258\,444$ for shGFP and $55\,146\,141 \pm 2\,027\,402$ for sh*RBFOX1*) and exon junctions as defined by Refseq. Biological replicate counts were pooled for both gene expression analysis and alternative splicing analysis. The annotation used for mapping potential alternatively spliced reads consisted of a previously published genome-wide alternative gene expression database source (39). To be counted as a junction, a read had to overlap two adjacent exons by a minimum of 5 nucleotides and total counts had to equal or exceed 20. Splicing events were defined as C1–A1A2–C2, where the alternative exon is A1A2 (A1, 5′ 70 nucleotides; A2 3′ 70 nucleotides), and C1 and C2 represent the adjacent 70 nucleotides of the upstream and downstream exons, respectively. The corresponding junctions were therefore defined as C1A1 and A2C2 for the alternative isoform and C1C2 for the constitutive isoform. Events were excluded if the counts for either [C1A1] or [A2C2] exceeded the other by >2-fold. Percent inclusion was calculated as (([C1A1] + [A2C2])/2)/([C1C2] + ([C1A1] + [A2C2])/2). Statistical comparisons for both differential splicing and differential gene expression were made using a negative binomial distribution model with the DESeq software package (81). A cutoff

threshold of $P \leq 0.05$ was used for significance. Examination of the most variant gene expression indicated a reproducible difference between the sh*RBFOX1* and the shGFP lines, but also between transduced and wild-type cells as well (Fig. 2A). To correct for variation due to the transduction process, only the shGFP and sh*RBFOX1* lines were used for further comparisons. Gene ontologies and KEGG pathways were determined using the DAVID Bioinformatic resources (NIAID, NIH). Statistical comparisons to other gene lists were made using the hypergeometric probability. WGCNA was performed as previously described (45,47). Connectivity and shortest path calculation were performed using the igraph software package (82).

All RNA sequencing data are deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/) (accession number GSE36710).

Transcription, splicing and RNA-processing factor binding site analysis

Intronic sequences flanking each significant alternative exon were interrogated for 400 nucleotides in either direction for the presence of splicing and RNA-processing factor binding sites. To calculate enrichment for sites assessed individually, an equivalent number of random intronic sequences were assessed from a pool of all genes in the human genome and this process was permuted 100 times to generate an average and standard deviation. Significance was calculated using the normal distribution. In addition, we also utilized a database of experimentally derived target RNA sequences (34) to generate position-weight matrices that were used to scan the intronic sequence using the Clover algorithm (83). For these searches, enrichment was calculated based on three different background data sets comprising all introns in the human genome, human CpG islands, and human chromosome 20 sequence.

For transcription-factor binding site analysis, 1000 base pairs upstream of the transcription start site of each gene were examined using experimentally defined position-weight matrices from the JASPAR database (84) with the Clover algorithm (83). The MEME algorithm (85) was employed for predicting site occurrence. Enrichment was calculated based on three different background data sets comprising all human gene promoters, human CpG islands, and human chromosome 20 sequence.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

The authors wish to thank Irina Voineagu, Benjamin Blencowe and Xinchen Wang for helpful comments regarding the splicing analysis; Daning Lu for technical assistance with cell culture; Hongmei Dong for technical assistance with *in situ* hybridization; and Emily Batton, Sonia Kantak and Emily Matthews for additional technical assistance.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the National Institute of Mental Health (grants K08MH86297 to B.L.F., R37MH060233 to D.H.G., R01MH081754 to D.H.G., K08MH074362 to E.W., R00MH090238 to G.K. and T32MH073526 to N.P.); the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant p30HD004612 to F.G. and D.H.G.); the UCLA Caltech Medical Scientist Training Program (to N.P.); the Dr Miriam and Sheldon G. Adelson Medical Research Foundation (to C.V.); and the John Douglas French Alzheimer's Research Foundation (to E.W.).

REFERENCES

1. Calarco, J.A., Zhen, M. and Blencowe, B.J. (2011) Networking in a global world: establishing functional connections between neural splicing regulators and their target transcripts. *RNA*, **17**, 775–791.
2. Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P. and Burge, C.B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**, 470–476.
3. Blencowe, B.J., Ahmad, S. and Lee, L.J. (2009) Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes. *Genes Dev.*, **23**, 1379–1386.
4. Yeo, G.W., Nostrand, E.L. and Liang, T.Y. (2007) Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. *PLoS Genet.*, **3**, e85.
5. Boutz, P.L., Stoilov, P., Li, Q., Lin, C.H., Chawla, G., Ostrow, K., Shiue, L., Ares, M. Jr. and Black, D.L. (2007) A post-transcriptional regulatory switch in polypyrimidine tract-binding proteins reprograms alternative splicing in developing neurons. *Genes Dev.*, **21**, 1636–1652.
6. Zhang, C., Frias, M.A., Mele, A., Ruggiu, M., Eom, T., Marney, C.B., Wang, H., Licatalosi, D.D., Fak, J.J. and Darnell, R.B. (2010) Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science*, **329**, 439–443.
7. Yano, M., Hayakawa-Yano, Y., Mele, A. and Darnell, R.B. (2010) Nova2 regulates neuronal migration through an RNA switch in disabled-1 signaling. *Neuron*, **66**, 848–858.
8. Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.*, **40**, 1413–1415.
9. Underwood, J.G., Boutz, P.L., Dougherty, J.D., Stoilov, P. and Black, D.L. (2005) Homologues of the *Caenorhabditis elegans* Fox-1 protein are neuronal splicing regulators in mammals. *Mol. Cell. Biol.*, **25**, 10005–10016.
10. Martin, C.L., Duvall, J.A., Ilkin, Y., Simon, J.S., Arreaza, M.G., Wilkes, K., Alvarez-Retuerto, A., Whichello, A., Powell, C.M., Rao, K. *et al.* (2007) Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **144**, 869–876.
11. Wang, K., Zhang, H., Ma, D., Bucan, M., Glessner, J.T., Abrahams, B.S., Salyakina, D., Imielinski, M., Bradfield, J.P., Sleiman, P.M. *et al.* (2009) Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*, **459**, 528–533.
12. Bhalla, K., Phillips, H.A., Crawford, J., McKenzie, O.L., Mulley, J.C., Eyre, H., Gardner, A.E., Kremmidiotis, G. and Callen, D.F. (2004) The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. *J. Hum. Genet.*, **49**, 308–311.
13. Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J. *et al.* (2007) Strong association of de novo copy number mutations with autism. *Science*, **316**, 445–449.
14. Wintle, R.F., Lionel, A.C., Hu, P., Ginsberg, S.D., Pinto, D., Thiruvahindrapduram, B., Wei, J., Marshall, C.R., Pickett, J., Cook, E.H. *et al.* (2011) A genotype resource for postmortem brain samples from the Autism Tissue Program. *Autism Res.*, **4**, 89–97.

15. Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R.M., Blencowe, B.J. and Geschwind, D.H. (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*, **474**, 380–384.
16. Elia, J., Gai, X., Xie, H.M., Perin, J.C., Geiger, E., Glessner, J.T., D'Arcy, M., de Berardinis, R., Frackelton, E., Kim, C. *et al.* (2010) Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol. Psychiatry*, **15**, 637–646.
17. Xu, B., Roos, J.L., Levy, S., van Rensburg, E.J., Gogos, J.A. and Karayiorgou, M. (2008) Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat. Genet.*, **40**, 880–885.
18. Le-Niculescu, H., Patel, S.D., Bhat, M., Kuczenski, R., Faraone, S.V., Tsuang, M.T., McMahon, F.J., Schork, N.J., Nurnberger, J.I. Jr. and Niculescu, A.B. III (2009) Convergent functional genomics of genome-wide association data for bipolar disorder: comprehensive identification of candidate genes, pathways and mechanisms. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **150B**, 155–181.
19. Hamshere, M.L., Green, E.K., Jones, I.R., Jones, L., Moskvina, V., Kirov, G., Grozeva, D., Nikolov, I., Vukcevic, D., Caesar, S. *et al.* (2009) Genetic utility of broadly defined bipolar schizoaffective disorder as a diagnostic concept. *Br. J. Psychiatry*, **195**, 23–29.
20. Zhang, C., Zhang, Z., Castle, J., Sun, S., Johnson, J., Krainer, A.R. and Zhang, M.Q. (2008) Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. *Genes Dev.*, **22**, 2550–2563.
21. Yeo, G.W., Xu, X., Liang, T.Y., Muotri, A.R., Carson, C.T., Coufal, N.G. and Gage, F.H. (2007) Alternative splicing events identified in human embryonic stem cells and neural progenitors. *PLoS Comput. Biol.*, **3**, 1951–1967.
22. Yeo, G.W., Coufal, N.G., Liang, T.Y., Peng, G.E., Fu, X.D. and Gage, F.H. (2009) An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat. Struct. Mol. Biol.*, **16**, 130–137.
23. Gehman, L.T., Stoilov, P., Maguire, J., Damianov, A., Lin, C.H., Shiue, L., Ares, M. Jr., Mody, I. and Black, D.L. (2011) The splicing regulator Rbfox1 (A2BP1) controls neuronal excitation in the mammalian brain. *Nat. Genet.*, **43**, 706–711.
24. Konopka, G., Wexler, E., Rosen, E., Mukamel, Z., Osborn, G.E., Chen, L., Lu, D., Gao, F., Gao, K., Lowe, J.K. *et al.* (2012) Modeling the functional genomics of autism using human neurons. *Mol. Psychiatry*, **17**, 202–214.
25. Rosen, E.Y., Wexler, E.M., Versano, R., Coppola, G., Gao, F., Winden, K.D., Oldham, M.C., Martens, L.H., Zhou, P., Farese, R.V. Jr. *et al.* (2011) Functional genomic analyses identify pathways dysregulated by progranulin deficiency, implicating wnt signaling. *Neuron*, **71**, 1030–1042.
26. Wexler, E.M., Paucer, A., Kornblum, H.I., Palmer, T.D. and Geschwind, D.H. (2009) Endogenous Wnt signaling maintains neural progenitor cell potency. *Stem Cells*, **27**, 1130–1141.
27. Damianov, A. and Black, D.L. (2010) Autoregulation of Fox protein expression to produce dominant negative splicing factors. *RNA*, **16**, 405–416.
28. Hammock, E.A. and Levitt, P. (2011) Developmental expression mapping of a gene implicated in multiple neurodevelopmental disorders, a2bp1 (fox1). *Dev. Neurosci.*, **33**, 64–74.
29. Lee, J.A., Tang, Z.Z. and Black, D.L. (2009) An inducible change in Fox-1/A2BP1 splicing modulates the alternative splicing of downstream neuronal target exons. *Genes Dev.*, **23**, 2284–2293.
30. Nakahata, S. and Kawamoto, S. (2005) Tissue-dependent isoforms of mammalian Fox-1 homologs are associated with tissue-specific splicing activities. *Nucleic Acids Res.*, **33**, 2078–2089.
31. Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M. and Misteli, T. (2010) Regulation of alternative splicing by histone modifications. *Science*, **327**, 996–1000.
32. Jin, Y., Suzuki, H., Maegawa, S., Endo, H., Sugano, S., Hashimoto, K., Yasuda, K. and Inoue, K. (2003) A vertebrate RNA-binding protein Fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *EMBO J.*, **22**, 905–912.
33. Kuroyanagi, H. (2009) Fox-1 family of RNA-binding proteins. *Cell. Mol. Life Sci.*, **66**, 3895–3907.
34. Piva, F., Giulietti, M., Burini, A.B. and Principato, G. (2012) SpliceAid 2: a database of human splicing factors expression data and RNA target motifs. *Hum. Mutat.*, **33**, 81–85.
35. Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X. *et al.* (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*, **456**, 464–469.
36. Dobi, A., Szemes, M., Lee, C., Palkovits, M., Lim, F., Gyorgy, A., Mahan, M.A. and Agoston, D.V. (2006) AUF1 is expressed in the developing brain, binds to AT-rich double-stranded DNA, and regulates enkephalin gene expression. *J. Biol. Chem.*, **281**, 28889–28900.
37. Fisette, J.F., Toutant, J., Dugre-Brisson, S., Desgroseillers, L. and Chabot, B. (2010) hnRNP A1 and hnRNP H can collaborate to modulate 5′ splice site selection. *RNA*, **16**, 228–238.
38. Hambardzumyan, D., Sergent-Tanguy, S., Thinard, R., Bonnamain, V., Masip, M., Fabre, A., Boudin, H., Neveu, I. and Naveilhan, P. (2009) AUF1 and Hu proteins in the developing rat brain: implication in the proliferation and differentiation of neural progenitors. *J. Neurosci. Res.*, **87**, 1296–1309.
39. Griffith, M., Griffith, O.L., Mwenifumbo, J., Goya, R., Morrissy, A.S., Morin, R.D., Corbett, R., Tang, M.J., Hou, Y.C., Pugh, T.J. *et al.* (2010) Alternative expression analysis by RNA sequencing. *Nat. Methods*, **7**, 843–847.
40. Abrahams, B.S. and Geschwind, D.H. (2008) Advances in autism genetics: on the threshold of a new neurobiology. *Nat. Rev. Genet.*, **9**, 341–355.
41. Collins, M.O., Husi, H., Yu, L., Brandon, J.M., Anderson, C.N., Blackstock, W.P., Choudhary, J.S. and Grant, S.G. (2006) Molecular characterization and comparison of the components and multiprotein complexes in the postsynaptic proteome. *J. Neurochem.*, **97** (Suppl. 1), 16–23.
42. Perales, R. and Bentley, D. (2009) 'Cotranscriptionality': the transcription elongation complex as a nexus for nuclear transactions. *Mol. Cell*, **36**, 178–191.
43. Pandit, S., Wang, D. and Fu, X.D. (2008) Functional integration of transcriptional and RNA processing machineries. *Curr. Opin. Cell Biol.*, **20**, 260–265.
44. Winden, K.D., Oldham, M.C., Mirnics, K., Ebert, P.J., Swan, C.H., Levitt, P., Rubenstein, J.L., Horvath, S. and Geschwind, D.H. (2009) The organization of the transcriptional network in specific neuronal classes. *Mol. Syst. Biol.*, **5**, 291.
45. Konopka, G., Bomar, J.M., Winden, K., Coppola, G., Jonsson, Z.O., Gao, F., Peng, S., Preuss, T.M., Wohlschlegel, J.A. and Geschwind, D.H. (2009) Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature*, **462**, 213–217.
46. Zhang, B. and Horvath, S. (2005) A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.*, **4**, Article17.
47. Oldham, M.C., Konopka, G., Iwamoto, K., Langfelder, P., Kato, T., Horvath, S. and Geschwind, D.H. (2008) Functional organization of the transcriptome in human brain. *Nat. Neurosci.*, **11**, 1271–1282.
48. Basu, S.N., Kollu, R. and Banerjee-Basu, S. (2009) AutDB: a gene reference resource for autism research. *Nucleic Acids Res.*, **37**, D832–D836.
49. Manneville, J.B., Jehanno, M. and Etienne-Manneville, S. (2010) Dlg1 binds GKAP to control dynein association with microtubules, centrosome positioning, and cell polarity. *J. Cell Biol.*, **191**, 585–598.
50. Mulle, J.G., Dodd, A.F., McGrath, J.A., Wolyniec, P.S., Mitchell, A.A., Shetty, A.C., Sobreira, N.L., Valle, D., Rudd, M.K., Satten, G. *et al.* (2010) Microdeletions of 3q29 confer high risk for schizophrenia. *Am. J. Hum. Genet.*, **87**, 229–236.
51. Willatt, L., Cox, J., Barber, J., Cabanas, E.D., Collins, A., Donnai, D., FitzPatrick, D.R., Maher, E., Martin, H., Parnau, J. *et al.* (2005) 3q29 microdeletion syndrome: clinical and molecular characterization of a new syndrome. *Am. J. Hum. Genet.*, **77**, 154–160.
52. van Spronsen, M. and Hoogenraad, C.C. (2010) Synapse pathology in psychiatric and neurologic disease. *Curr. Neurol. Neurosci. Rep.*, **10**, 207–214.
53. Boda, B., Dubos, A. and Muller, D. (2010) Signaling mechanisms regulating synapse formation and function in mental retardation. *Curr. Opin. Neurobiol.*, **20**, 519–527.
54. Spalice, A., Parisi, P., Nicita, F., Pizzardi, G., Del Balzo, F. and Iannetti, P. (2009) Neuronal migration disorders: clinical, neuroradiologic and genetics aspects. *Acta Paediatr.*, **98**, 421–433.

55. Jen, J.C., Wan, J., Palos, T.P., Howard, B.D. and Baloh, R.W. (2005) Mutation in the glutamate transporter EAAT1 causes episodic ataxia, hemiplegia, and seizures. *Neurology*, **65**, 529–534.
56. Galante, P.A., Sandhu, D., de Sousa Abreu, R., Gradassi, M., Slager, N., Vogel, C., de Souza, S.J. and Penalva, L.O. (2009) A comprehensive in silico expression analysis of RNA binding proteins in normal and tumor tissue: identification of potential players in tumor formation. *RNA Biol.*, **6**, 426–433.
57. Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., Mihara, M., Naitou, M., Endoh, H., Nakamura, T. *et al.* (2007) DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. *Nat. Cell Biol.*, **9**, 604–611.
58. Wortham, N.C., Ahamed, E., Nicol, S.M., Thomas, R.S., Periyasamy, M., Jiang, J., Ochocka, A.M., Shousha, S., Huson, L., Bray, S.E. *et al.* (2009) The DEAD-box protein p72 regulates ERalpha-/oestrogen-dependent transcription and cell growth, and is associated with improved survival in ERalpha-positive breast cancer. *Oncogene*, **28**, 4053–4064.
59. Qing, Y., Yingmao, G., Lujun, B. and Shaoling, L. (2008) Role of Npm1 in proliferation, apoptosis and differentiation of neural stem cells. *J. Neurol. Sci.*, **266**, 131–137.
60. Martins-de-Souza, D., Gattaz, W.F., Schmitt, A., Rewerts, C., Maccarrone, G., Dias-Neto, E. and Turck, C.W. (2009) Prefrontal cortex shotgun proteome analysis reveals altered calcium homeostasis and immune system imbalance in schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.*, **259**, 151–163.
61. Bowden, N.A., Scott, R.J. and Tooney, P.A. (2008) Altered gene expression in the superior temporal gyrus in schizophrenia. *BMC Genomics*, **9**, 199.
62. Kuwano, Y., Kamio, Y., Kawai, T., Katsuura, S., Inada, N., Takaki, A. and Rokutan, K. (2011) Autism-associated gene expression in peripheral leucocytes commonly observed between subjects with autism and healthy women having autistic children. *PLoS ONE*, **6**, e24723.
63. 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.* (2011) Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. *Eur. J. Hum. Genet.*, **19**, 727–731.
64. Halgren, C., Kjaergaard, S., Bak, M., Hansen, C., El-Schich, Z., Anderson, C., Henriksen, K., Hjalgrim, H., Kirchhoff, M., Bijlsma, E. *et al.* (2011) Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. *Clin. Genet*, [Epub ahead of print].
65. Chakrabarti, B., Dudbridge, F., Kent, L., Wheelwright, S., Hill-Cawthorne, G., Allison, C., Banerjee-Basu, S. and Baron-Cohen, S. (2009) Genes related to sex steroids, neural growth, and social-emotional behavior are associated with autistic traits, empathy, and Asperger syndrome. *Autism Res.*, **2**, 157–177.
66. Merla, G., Brunetti-Pierri, N., Micale, L. and Fusco, C. (2010) Copy number variants at Williams-Beuren syndrome 7q11.23 region. *Hum. Genet.*, **128**, 3–26.
67. Larsen, K.B., Lutterodt, M.C., Laursen, H., Graem, N., Pakkenberg, B., Mollgard, K. and Moller, M. (2010) Spatiotemporal distribution of PAX6 and MEIS2 expression and total cell numbers in the ganglionic eminence in the early developing human forebrain. *Dev. Neurosci.*, **32**, 149–162.
68. Joutel, A., Corpechot, C., Ducros, A., Vahedi, K., Chabriat, H., Mouton, P., Alamowitch, S., Domenga, V., Cecillion, M., Marechal, E. *et al.* (1996) Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature*, **383**, 707–710.
69. Fulton, D.L., Sundararajan, S., Badis, G., Hughes, T.R., Wasserman, W.W., Roach, J.C. and Sladek, R. (2009) TFCat: the curated catalog of mouse and human transcription factors. *Genome Biol.*, **10**, R29.
70. Thanseem, I., Anitha, A., Nakamura, K., Suda, S., Iwata, K., Matsuzaki, H., Ohtsubo, M., Ueki, T., Katayama, T., Iwata, Y. *et al.* (2012) Elevated transcription factor specificity protein 1 in autistic brains alters the expression of autism candidate genes. *Biol. Psychiatry*, **71**, 410–418.
71. Sarachana, T., Xu, M., Wu, R.C. and Hu, V.W. (2011) Sex hormones in autism: androgens and estrogens differentially and reciprocally regulate RORA, a novel candidate gene for autism. *PLoS ONE*, **6**, e17116.
72. Glessner, J.T., Wang, K., Cai, G., Korvatska, O., Kim, C.E., Wood, S., Zhang, H., Estes, A., Brune, C.W., Bradfield, J.P. *et al.* (2009) Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature*, **459**, 569–573.
73. Lim, J., Hao, T., Shaw, C., Patel, A.J., Szabo, G., Rual, J.F., Fisk, C.J., Li, N., Smolyar, A., Hill, D.E. *et al.* (2006) A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell*, **125**, 801–814.
74. Shibata, H., Huynh, D.P. and Pulst, S.M. (2000) A novel protein with RNA-binding motifs interacts with ataxin-2. *Hum. Mol. Genet.*, **9**, 1303–1313.
75. Ozcelik, T., Akarsu, N., Uz, E., Caglayan, S., Gulsuner, S., Onat, O.E., Tan, M. and Tan, U. (2008) Mutations in the very low-density lipoprotein receptor VLDLR cause cerebellar hypoplasia and quadrupedal locomotion in humans. *Proc. Natl Acad. Sci. USA*, **105**, 4232–4236.
76. Turkmen, S., Guo, G., Garshasbi, M., Hoffmann, K., Alshalah, A.J., Mischung, C., Kuss, A., Humphrey, N., Mundlos, S. and Robinson, P.N. (2009) CA8 mutations cause a novel syndrome characterized by ataxia and mild mental retardation with predisposition to quadrupedal gait. *PLoS Genet.*, **5**, e1000487.
77. Iwaki, A., Kawano, Y., Miura, S., Shibata, H., Matsuse, D., Li, W., Furuya, H., Ohyagi, Y., Taniwaki, T., Kira, J. *et al.* (2008) Heterozygous deletion of ITPR1, but not SUMF1, in spinocerebellar ataxia type 16. *J. Med. Genet.*, **45**, 32–35.
78. Wexler, E.M., Rosen, E., Lu, D., Osborn, G.E., Martin, E., Raybould, H. and Geschwind, D.H. (2011) Genome-wide analysis of a Wnt1-regulated transcriptional network implicates neurodegenerative pathways. *Sci. Signal.*, **4**, ra65.
79. Abu-Khalil, A., Fu, L., Grove, E.A., Zecevic, N. and Geschwind, D.H. (2004) Wnt genes define distinct boundaries in the developing human brain: implications for human forebrain patterning. *J. Comp. Neurol.*, **474**, 276–288.
80. Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
81. Anders, S. and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol.*, **11**, R106.
82. Csardi, G. and Nepusz, T. (2006) The igraph software package for complex network research. *Int. J. Complex Systems*, 1695.
83. Frith, M.C., Fu, Y., Yu, L., Chen, J.F., Hansen, U. and Weng, Z. (2004) Detection of functional DNA motifs via statistical over-representation. *Nucleic Acids Res.*, **32**, 1372–1381.
84. Portales-Casamar, E., Thongjuea, S., Kwon, A.T., Arenillas, D., Zhao, X., Valen, E., Yusuf, D., Lenhard, B., Wasserman, W.W. and Sandelin, A. (2010) JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Res.*, **38**, D105–D110.
85. Bailey, T.L. and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, **2**, 28–36.