Transcriptional targets of the schizophrenia risk gene MIR137

AL Collins, Y Kim, RJ Bloom, SN Kelada, P Sethupathy and PF Sullivan

INTRODUCTION

Schizophrenia is now known to be highly polygenic with a genetic architecture characterized by contributions from dozens and probably hundreds of different genetic loci. Although rare deleterious exonic and copy number variation (CNVs) have roles, numerous common variants of more subtle effect account for sizable fractions of the heritability of schizophrenia.1–4

A replicated and robust association with schizophrenia is in MIR137, the gene encoding the microRNA miR-137. This region was initially identified by the Psychiatric Genomics Consortium (P = 1.6 × 10−11)5. In the largest genome-wide association study (GWAS) yet conducted (by the Psychiatric Genomics Consortium, 36,989 cases), the second strongest finding was within MIR137 (P = 3.4 × 10−19),1 Pathway analyses have suggested that genes containing predicted miR-137 target sites are significantly enriched for smaller P-values for association with schizophrenia.5–7 This enrichment is not robustly identified in all studies,7 and one possibility is that miR-137 target predictions are not sufficiently accurate.

It is possible that a pathway regulated by miR-137 is important in the etiology of schizophrenia. microRNAs can target large numbers of mRNAs for degradation or translational inhibition, and a single microRNA can have major effects on cellular pathways.8–11 miR-137 has important roles in neural development12–15 and has been implicated in autism, intellectual disability, Rett syndrome, Alzheimer’s disease, Huntington’s disease and multiple cancers.14,16–28

Schizophrenia GWAS results, pathway analyses and the roles of miR-137 in development and disease are consistent with the broad hypothesis that genetic variation in miR-137 and its targets constitute a pathway whose dysfunction is involved in the etiology of schizophrenia. There are several major limitations in evaluating this hypothesis. A key issue is that miR-137 target genes are mostly based on computational algorithms of imperfect sensitivity and specificity. These algorithms predict only the direct targets of miR-137.12–20 Some direct targets of miR-137 have been verified12–14,18,25,26 including genes implicated in schizophrenia (CACNA1C, TCF4, CSMD1, C10orf26 and ZNF804A).33,34 However, many of these targets were identified in cancers, and their neuronal role is unclear. microRNA regulatory activity is complex and can include regulatory loops, post-translational modifications and non-canonical target sequences (for example, in coding regions and not 3′ untranslated regions).15,35,36 The ‘downstream’ effects of miR-137 (that is, its impact beyond a directly targeted transcript) have not been systematically evaluated.

We thus evaluated the effects of miR-137 overexpression in a human neuronal cell line. Our goal was to expand knowledge of the biology of miR-137 by identifying its empirical targets. We transduced neuronal stem cells with vectors expressing miR-137 or a control miRNA mimic and used high-throughput sequencing (RNA-seq) at 24 and 48 h after transduction to assess the effects of miR-137 overexpression. Overexpression of miR-137 had an impact on multiple cell cycle genes, consistent with other findings of effects on proliferation and downregulation in cancer. However, we also found that the indirect effects of miR-137 overexpression at 48 h had strong connections with GWAS findings for schizophrenia. These results provide functional support for a role of miR-137 in the etiology of schizophrenia.

MATERIALS AND METHODS

Cell line

ReNcell-V (Millipore, Billerica, MA, USA) is a human neural precursor cell line derived from the ventral mesencephalon.27 We chose this cell line because it was readily available, accessible to our methods and not derived from a cancer. Moreover, we genotyped two different clones with Illumina OminExpress arrays and confirmed the expected male sex, European

1Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA and 2Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. Correspondence: Dr PF Sullivan, Department of Genetics, University of North Carolina at Chapel Hill, CB#7264, 5097 Genomic Medicine, Chapel Hill, NC 27599-7264, USA. E-mail: pfsulliv@med.unc.edu

Received 15 November 2013; revised 27 April 2014; accepted 29 April 2014
ancestry, the absence of aneuploidy and the absence of CNV larger than >100 kb that affected a gene.

Infection of cells
ReNCell-VMs were grown on laminin-coated plates in ReNCell NSC Maintenance Media (Millipore) with 20 ng ml⁻¹ FGF-2 and 20 ng ml⁻¹ epidermal growth factor. We used pEZX-MR03 lentiviral vectors with miR-137 or a control packaged with an HIV-1-compatible system (GeneCopoeia, Rockville, MD, USA). Infection efficiency was assessed by microscopic evaluation of eGFP expression (contained in the pEZX-MR03 plasmid). To evaluate nonspecific effects because of overloading of the miRNA-processing machinery, we used a control lentiviral vector containing a miRNA mimic (that is, a random miR-like sequence of no similarity to any known miRNA). Flasks of cells were infected in parallel with the miR-137 or control lentivirus. Viral amounts were titrated to achieve ~100% infection rates without evidence of toxicity. For the 48-h time point, the cells were exposed to virus for 24 h and then replaced with fresh media. Cells were collected 24 or 48 h after infection, stored in RNA-stabilizing media (RNAlater, Qiagen, Venlo, The Netherlands) and frozen at −80 °C until analysis. Evaluation of two relatively early time points was performed on the same RNA samples as were used for RNA-seq except for performing qPCR that were carried out on replicate samples (using RNA extracted with the Norgen kit).

RNA extraction
Total RNA was extracted using either Maxwell 16 LEV Simply RNA kit (Madison, WI, USA; 24 h RNA-seq) or Norgen Biotek RNA/DNA/Protein Kit (Thorold, ON, Canada; 48 h RNA-seq and mature miR-137 assays). Although both methods yield total RNA, the Norgen kit extracts more mature miRNAs (data not shown) and is better suited for quantification of miR-137.

Quantitative PCR
We used TaqMan assays to measure expression of mature miR-137 (Applied Biosystems, assay ID 001129; Grand Island, NY, USA). Each sample was run in quadruplicate, and StepOne software (v2.1; SAS, Cary, NC, USA) was used to evaluate ΔCq, and fold change (RQ) values. The endogenous control was RNU6B. Statistical significance of the ΔCq values was assessed using nested analysis of variance (SAS v9.2, PROC GLM). All assays were performed on the same RNA samples as were used for RNA-seq except for the miR-137 assays at the 24-h time point, which were carried out on replicate samples (using RNA extracted with the Norgen kit).

RNA-seq
RNA-sequencing libraries were generated using Truseq RNA sample prep kits (Illumina, San Diego, CA, USA), which generate libraries from polyadenylated RNAs. Each library was barcoded, pooled and sequenced using Illumina HiSeq 2000 and single-end 100-bp reads. Single-end sequencing was chosen to maximize read depth and detection of small fold changes. There were three miR-137 overexpressions and three controls at 24 h, and five miR-137 overexpressions and six controls at 48 h. Read depth averaged ~55 million reads per replicate at 24 h and ~160 million reads per replicate at 48 h (see below for explanation of greater numbers of reads at 48 h). All procedures were randomized.

Data processing and analysis
We used the TopHat-Cufflinks-Cuffdiff pipeline to detect differential gene expression between neural stem cells with increased expression of miR-137 compared with controls. First, we used FastQC (see URLs) to evaluate sequencing quality. As one 48-h sample failed, we repeated RNA-seq on all of the 48-h samples. We verified sample identity across all experiments via the creation of a vector of genotypes based on high-confidence heterozygous exonic RNA-seq reads. RNA-seq reads were also aligned to the overexpression sequence (miR-137 or a miR-like mimic) to confirm experimental condition. Second, we mapped RNA-seq reads to the human genome reference (UCSC hg19) using TopHat2 and transcript annotations from the TopHat website (see URLs). Third, the mapped reads were assembled into transcripts using Cufflinks (v2.0.1; http://cufflinks.cbcb.umd.edu/). Fourth, we pooled the Cufflinks output files from multiple RNA-seq samples using Cuffcompare to assemble a comprehensive set of transcripts. Finally, we used Cuffdiff to evaluate differential gene expression of miR-137 overexpression versus the control vector separately at 24 and 48 h with correction for multiple comparisons using a false-discovery rate approach (q-value). We used the RNA-seq data to evaluate gene expression of transcripts whose protein products are known to characterize neuronal differentiation using a read count approach (Supplementary Methods). A few genes were reported more than once, and the transcript with the smallest P-value was retained.

Additional analyses
We compared the effects of miR-137 overexpression in neural progenitor cells to computational predictions of genes targeted by miR-137. We used Sylarray (see URLs) to test all known miRNA seed sequences in the RNA-seq gene level data. The analyses are performed at varying cutoffs to create a curve of log₂(P) values for each miRNA. We also compared differentially expressed genes from miR-137 overexpression to computationally predicted miR-137 targets (TargetScan v6.2; http://www.targetscan.org) with significance determined using permutation.

DAPPLE was used to evaluate whether the lists of genes had more protein–protein interactions than predicted by chance. We used DAVID to evaluate whether genes differentially expressed upon miR-137 overexpression tended to occur in specific biological pathways or processes including GO, KEGG and Biocarta (see URLs). DAVID generated enrichment scores for correlated pathways and Benjamini corrected P-values for individual pathways. PWM scan was used to predict high-confidence transcription factor-binding sites in gene promoters. We then identified transcription factor-binding site over-represented in the promoters of significantly differentially expressed genes by Monte Carlo simulation analysis.

We compared genes differentially expressed following miR-137 overexpression to an analogous experiment by Hill et al. and to a gene expression study in lymphoblastoid cell lines from schizophrenia cases and controls. We evaluated the relevance of genes differentially expressed following miR-137 overexpression to pathways previously implicated in the etiology of schizophrenia: expert-curated lists of genes found at synapses (1043 genes), genes encoding proteins found at the postsynaptic density (1457 genes). Genes in neuronal activity-regulated cytoskeleton-associated protein postsynaptic signaling complexes (28 genes), N-methyl-D-aspartate receptors (61 genes), L-voltage-gated calcium channel subunits (26 genes), genes whose proteins interact with calcium channel (Cav2, 207 genes) and genes whose miRNAs interact with FMRP (fragile X mental retardation protein, 842 genes).

Finally, we assessed the overlap of genes differentially expressed following miR-137 overexpression to schizophrenia GWAS results. This analysis is a relatively direct evaluation of the salience of sets of genes directly or indirectly regulated by miR-137 to genetic variation implicated in the etiology of schizophrenia. We used INRICH to test whether differentially expressed genes were enriched for smaller GWAS P-values in the largest published schizophrenia GWAS (Supplementary Methods).

RESULTS
We transduced human neural precursor cells (ReNCell-VM) with lentiviral vectors expressing either miR-137 or a scrambled miRNA-like control. We verified infection rates >99% via fluorescent imaging of the eGFP tag in the vector. Alignment of RNA-seq reads to the lentiviral sequence confirmed correct expression of the experimental vector in each sample.

Verification
We verified increased miR-137 expression in two ways. qPCR confirmed marked increases in the amounts of mature miR-137 at 24 h (111x increase) and 48 h (548x increase) in cells exposed to lentivirus expressing miR-137 compared with controls (Supplementary Figure S1A). We used RNA-seq to estimate expression of the primary polyadenylated MIR137 transcript from the miR-137 lentiviral vector and the MIR137 locus. Consistent with the qPCR results, we found 123x increase in the primary transcript at 24 h and 660x increase at 48 h (Supplementary Figure S1B). As miR-137 can effect neuronal proliferation and differentiation,12–14,29 we evaluated cell density, morphology and expression of differentiation markers in the same flask used for
RNA-seq. There were no significant differences in cell densities between miR-137 overexpression and control flasks at 24 (P = 0.53) or 48 h (P = 0.78), although the 48-h flasks had higher cell densities than at 24 h (P = 7.6 × 10^{−9}), as expected because of additional growth time. We noted no qualitative differences in cellular morphology between the experimental conditions at either time point. We used RNA-seq to compare expression of 12 genes whose protein products are commonly used to classify neuronal cell types (Supplementary Table S1). There were no significant differences between the miR-137 overexpression and control conditions. The overall pattern of results is consistent with that of a neuronal stem cell. Thus, large increases in expression of miR-137 did not lead to clear alterations in cell fate after 24 or 48 h.

We verified the downregulation of predicted targets of miR-137 in two ways. First, we applied Sylarray to the RNA-seq data to evaluate fold change differences associated with all known miRNA seed sequences. Genes with the miR-137 seed sequence (AGCAATTA) in the 3′-untranslated repeat evidenced the greatest downregulation of any seed sequence (Supplementary Figure S1C, minimum P = 1.1 × 10^{−22} in the 24-h condition and minimum P = 2.4 × 10^{−45} in the 48-h condition). As a complementary test, we used permutation to evaluate differential gene expression in 1115 genes with predicted miR-137 target sites (TargetScan v6.2). Genes with a predicted miR-137 target site were significantly more likely to be differentially expressed (Supplementary Figure S1D, P < 10^{−200} at both 24 and 48 h), and most were downregulated at 24 (82%) and 48 h (87%). These results are consistent with profound and relatively specific downregulation resulting from miR-137 overexpression.

General expression patterns
Supplementary Table S2 shows patterns of RNA-seq analyzability at 24 and 48 h. After removing transcripts that had insufficient sequencing depth, low expression levels or technical faults and resolving genes reported more than once, there were 14 682 transcripts analyzable at both 24 and 48 h. The correlations between miR-137 overexpression and control conditions were 0.999 at both 24 and 48 h, and the correlations between the 24 and 48 h conditions were 0.980–0.982 (Supplementary Figures S2a and S2b). We applied principal component analysis to the expression data matrix (Supplementary Figure S3). The first principal component accounted for 97% of the variance and completely distinguished the replicates from 24 and 48 h. These descriptive results suggest the comparability of the RNA-seq data within and between time points. The results also suggest that only a fraction of transcripts was altered by the miR-137 experimental manipulation.

Differential gene expression
Supplementary Figure S4 shows descriptive statistics for log_{2} fold changes, P-values and q-values for the 24 and 48 h analyses. These analyses compare the miR-137 overexpression to a control vector within each time point. At 24 h, 14 869 unique genes were analyzable and 202 (1.36%) had evidence of differential gene expression (q < 0.05, Supplementary Table S3). Of the differentially expressed genes, 130 genes (64.4%) had evidence of inhibition of gene expression following exposure to markedly elevated amounts of miR-137 and 72 genes (35.6%) showed enhancement. At 48 h, 428 (2.71%) of 15 780 analyzable and unique genes had evidence of differential expression (q < 0.05, Supplementary Table
48 h

| Time  | Direction of effect | Enrichment | Annotation cluster description |
|-------|---------------------|------------|--------------------------------|
| 24 h  | Downregulation (183 genes) | 3.04       | Extracellular matrix             |
|       |                      | 2.92       | Positive regulation of transcription |
|       |                      | 2.61       | Neuron differentiation/projection |
|       |                      | 2.55       | Extracellular signaling          |
|       |                      | 2.75       | Cell cycle                       |
|       | Upregulation (126 genes)   | 10.75      | Cell cycle                       |
|       | Downregulation (239 genes) | 7.16       | DNA replication                  |
|       |                      | 4.74       | Lumen/membrane enclosed          |
|       |                      | 4.42       | Cell cycle checkpoint            |
|       |                      | 3.31       | Chromosome segregation           |
|       |                      | 3.04       | Mitotic checkpoint/chromosome    |
|       |                      | 3.01       | Microtubule/spindle organization |
|       |                      | 2.79       | DNA binding                      |
| 48 h  | Upregulation (376 genes)   | 15.98      | Signaling                        |
|       |                      | 12.07      | Transmembrane                    |
|       |                      | 10.47      | Cell adhesion                    |
|       |                      | 5.87       | Plasma membrane                  |
|       |                      | 4.41       | EGF-like                         |
|       |                      | 4.18       | MHC class II/antigen processing  |
|       |                      | 3.89       | Ion transport/gated channels     |
|       |                      | 3.79       | Immunoglobulin-like              |
|       |                      | 3.63       | Calcium ion binding/ion binding  |
|       |                      | 3.21       | Angiogenesis                     |
|       |                      | 2.80       | Cell junction/synapse            |
|       |                      | 2.58       | Synaptic transmission            |
|       |                      | 2.55       | Matrix adhesion/integrins        |
|       |                      | 2.52       | Extracellular matrix             |

Abbreviations: EGF, epidermal growth factor; MHC, major histocompatibility complex. Pathway analysis from DAVID. As many annotated pathways contain overlapping genes, DAVID combines such pathways into clusters. Column 4 shows our labels for these clusters. Individual pathways are reported in Supplementary Tables S5a and S6b.

Of these, 157 genes (36.7%) showed inhibition after miR-137 overexpression and 271 genes (63.3%) showed enhancement.

We asserted above that the widely used, computationally derived TargetScan (v6.2) list of miR-137 targets is imperfect. The presence of a predicted TargetScan-binding site was associated with significant (q < 0.05) differential gene expression at 48 h (odds ratio 4.0, 95% confidence interval (CI) 3.2–5.1, P = 2.7 × 10^{-33}) and, as expected from the analyses above, this was mainly because of inhibitory effects of miR-137 overexpression (odds ratio 10.6, 95% CI 7.6–14.6, P = 6.4 × 10^{-69}) rather than enhancement (odds ratio 1.2, 95% CI 0.7–2.27, P = 0.52). However, these strong statistical associations were not robustly predictive as only 8.9% (87/1002) of analyzable genes with predicted miR-137 target sites were differentially expressed and only 20.3% (87/428) of differentially expressed transcripts had predicted miR-137 target sites. The direct and indirect effects of miR-137 thus appear to be complex and not robustly captured using TargetScan.

Pathway analyses

As the effects of miR-137 overexpression influenced multiple genes and never acted as a switch, we attempted to identify biological themes that could unify the effects of miR-137 across many different transcripts. We used a more liberal definition of differential expression in these analyses (q < 0.1 instead of 0.05). We used DAPPLE to assess whether the protein products of genes with significant differential expression had greater protein–protein interaction than expected by chance. We found significantly more interactions at 24 h (P = 0.01) and at 48 h (P = 0.001), and many of these genes formed a single network at each time point (Supplementary Figure S5). This result is consistent with the hypothesis that miR-137 influences multi-component functional networks rather than targeting of unrelated individual genes.

We applied pathway analysis to lists of differentially expressed genes using DAVID (Table 1). At 24 h, genes with inhibition after exposure to markedly increased amounts of miR-137 (presumably enriched for direct targets of miR-137) were enriched for pathways related to extracellular matrix, positive regulation of transcription, neuronal differentiation/projection and extracellular signaling (Supplementary Table S5a). Genes with increased expression at 24 h (presumptive indirect targets of miR-137) were enriched for cell cycle consistent with prior reports (Supplementary Table S5b). In the 48 h data, pathway analysis of differentially expressed genes yielded more changes. Interestingly, cell cycle pathways (upregulated at 24 h) were strongly downregulated at 48 h (Supplementary Table S6a). Pathways involved in DNA replication and cell cycle processes were also downregulated at 48 h. Upregulated pathways at 48 h...
analyses of pathways implicated by schizophrenia genome-wide studies

| Pathway                        | DE genes (q < 0.1) | 24 h | 48 h |
|-------------------------------|--------------------|------|------|
|                               | P                  | OR (95% CI) | P      | OR (95% CI) |
| Synaptic genes                | All                | 0.90  | 0.94 (0.57–1.56) | 2.8 x 10^-7 | 2.13 (1.63–2.78) |
| (1043 genes)                  | Downregulated      | 0.86  | 0.85 (0.41–1.74) | 1.0       | 0.95 (0.5–1.81)  |
|                               | Upregulated        | 0.70  | 1.11 (0.54–2.28) | 3.2 x 10^-8 | 2.54 (1.88–3.44) |
| Postsynaptic density          | All                | 0.78  | 0.88 (0.49–1.58) | 0.06      | 1.4 (0.99–1.99)  |
| (1457 genes)                  | Upregulated        | 1.0   | 0.98 (0.43–2.24) | 0.39      | 0.63 (0.28–1.44) |
| ARC                           | All                | 1.0   | N/A     | 0.0049   | 1.82 (1.23–2.71) |
| (28 genes)                    | Downregulated      | 1.0   | N/A     | 0.20      | 4.96 (0.61–40.5) |
|                               | Upregulated        | 1.0   | N/A     | 0.03      | 5.03 (1.4–18.09) |
| NMDA receptor                 | All                | 1.0   | 0.89 (0.12–6.45) | 0.17      | 1.94 (0.73–5.38) |
| (61 genes)                    | Upregulated        | 1.0   | N/A     | 0.24      | 2.24 (0.53–9.43) |
| FMRP interactors              | All                | 0.08  | 1.49 (0.96–2.32) | 4.7 x 10^-5 | 1.92 (1.43–2.57) |
| (842 genes)                   | Downregulated      | 0.07  | 1.61 (0.98–2.65) | 0.71      | 1.13 (0.57–2.21) |
|                               | Upregulated        | 1.0   | 0.84 (0.31–2.28) | 1.5 x 10^-4 | 1.98 (1.42–2.75) |
| Calcium channel subunits      | All                | 1.0   | N/A     | 0.0056   | 6.61 (2.19–19.97) |
|                               | Downregulated      | 1.0   | N/A     | 1.0       | N/A  |
| (26 genes)                    | Upregulated        | 1.0   | N/A     | 0.0098   | 5.68 (1.84–17.52) |
| Ca2 interactions              | All                | 0.08  | N/A     | 0.11      | 1.68 (0.9–3.1)   |
| (207 genes)                   | Downregulated      | 0.27  | N/A     | 0.73      | 0.47 (0.06–3.36) |
|                               | Upregulated        | 0.65  | N/A     | 0.04      | 2.09 (1.08–4.05) |

Abbreviations: ARC, neuronal activity-regulated cytoskeleton-associated protein postsynaptic signaling complexes; Cav2, voltage-gated calcium channel; CI, confidence interval; DE, differentially expressed; FMRP, fragile X mental retardation protein; NMDA, N-methyl-D-aspartate; OR, odds ratio. Boldface shows P < 0.01.
supports a role for miR-137 as an effector that indirectly regulates pathways critical to schizophrenia.

We compared these results with those of a similar experiment by Hill et al.\(^\text{51}\) (~1000-fold increase in miR-137 expression in the neuronal cell line CTXOE03, gene expression measured at 72 h using Illumina HT-12v4 arrays). For the genes present in both studies, there was a significant negative correlation in the log\(_2\) fold change values at 24 h (Spearman \(\rho = -0.33\), \(P = 5 \times 10^{-17}\)) and a significant positive correlation at 48 h (Spearman \(\rho = 0.25\), \(P = 3 \times 10^{-11}\)). Despite the use of a different vector, cell line and expression technology, these results are surprisingly comparable to these in our experiment. However, there was no significant overlap between our results and genes that were differentially expressing in lymphoblastoid cell lines from schizophrenia cases and controls (data not shown).\(^\text{52}\)

**DISCUSSION**

miR-137 is a particularly intriguing association for schizophrenia, given the strength of its association and the co-occurrence of additional associations in many of its predicted targets. These empirical findings from unbiased genomic screens are consistent with the hypothesis that a biological or developmental process orchestrated by miR-137 contributes to the etiology of schizophrenia. As we demonstrated, a key limitation of our current knowledge is that we have only an approximate idea of the genes influenced by miR-137 (computational prediction of direct targets using TargetScan and almost nothing on its indirect targets).

As an initial approach to understanding the reach of miR-137, we overexpressed miR-137 in human neuronal stem cells to detect genes whose transcripts showed differential expression upon exposure to large increases in amounts of miR-137. Notably, our results had approximate overlap with a similar study.\(^\text{51}\) Using a number of verification analyses, we demonstrated that we successfully increased the amounts of miR-137 and that these effects were relatively specific. Our approach allowed detection of direct targets influenced by RNA degradation plus secondary effects in the form of increased expression. Via identification of individual genes and larger-scale alterations in biological pathways, our findings provide a window into the cellular mechanisms regulated by miR-137 in neuronal stem cells. Moreover, the different patterns of findings at 24 and 48 h highlight the dynamic regulation of miR-137, and emphasize that findings at a single time point may not encompass all targets of a given miRNA. In aggregate, these admittedly circumstantial results increase support for a role of miR-137 in schizophrenia.

First, the direct and indirect targets of miR-137 that we identified had relatively small changes in expression, typically less than twofold. We identified no instances where miR-137 acted as an on- or off-switch. These findings are consistent with prior studies showing that microRNAs yield subtle changes in many genes.\(^\text{10,11}\) Moreover, many differentially expressed genes were not predicted targets of miR-137, and many predicted targets are not altered. We do not believe that this was due to off target effects or secondary effects on different microRNAs, given that Sylarray detected only enrichment in miR-137 seed sequences (Supplementary Figure S1C).

Second, miR-137 has been shown to affect proliferation and differentiation.\(^\text{14,25,26,59}\) We found that cell density, morphology and neuronal fate markers were not altered by miR-137 at 24 or 48 h. Given that cell cycle genes showed differential expression (Table 1), it is possible that longer periods of observation would have detected alterations in morphology or cell fate.

Third, our results were consistent with the coordinated influence of miR-137 on multiple sets of functionally interconnected genes (for example, via the significant enrichment of protein–protein interactions in the differentially expressed genes and the pathway analyses). These results support the hypothesis that miR-137 causes subtle differential regulation of multiple genes in interrelated pathways, leading to larger physiological changes to a cell or organism.

Fourth, as summarized above, highly significant schizophrenia GWAS findings near miR137 and enrichment of its predicted targets support a role of MIR137 in schizophrenia. However, the region of association includes other candidate genes (DPYD and MIR26822). Critically, genes influenced by miR-137 in a human neural stem cell line (particularly downregulated genes) were enriched for smaller GWAS \(p\)-values (Table 3). Moreover, pathway analyses provide supporting data. Genes upregulated at 48 h (presumptive indirect targets) were consistently enriched for genes and pathways related to schizophrenia (Tables 1 and 2). These included significant overlap with major histocompatibility complex class II genes (encompassed by the strongest schizophrenia finding to date), genes with TCF4 (a schizophrenia risk locus and known target of miR-137) promoter-binding sites and pathways previously implicated in schizophrenia (synaptic genes, genes whose mRNAs interact with FMRF and L-type voltage-gated calcium channels).

These results should be interpreted in the context of several limitations. First, we are cautious in our comparison of results between 24 and 48 h, given the change to a different method of RNA extraction. Second, we attempted to sequester miR-137 as an inhibition control but were unable to obtain technically satisfactory results. Third, the degree of overexpression of miR-137 was continual and far greater the naturally occurring levels of expression.

In summary, our results have identified genes and pathways regulated through miR-137, supporting a role of miR-137 in fine-tuning neurally important pathways including cell cycle regulation, L-type voltage-gated calcium channels, cell–cell interactions and synaptic function. These results provide empirical data supporting the hypothesis that direct and indirect miR-137 targets have roles in the etiology of schizophrenia.

**Schizophrenia GWAS results**

Previous work has shown significant enrichment of smaller schizophrenia GWAS \(P\)-values for predicted targets of miR-137.\(^\text{5,6}\) However, these analyses were based on TargetScan-predicted direct effects of miR-137 which, as described above, have incomplete overlap with the biological direct and indirect effects of miR-137 overexpression in human neural precursor cells. We thus compared the empirical findings from this report with results from the largest published schizophrenia GWAS (Table 3).\(^\text{7}\) There was no significant enrichment for genes differentially expressed at 24 h. However, there was significant enrichment of smaller GWAS \(P\)-values for downregulated targets of miR-137.

| Time | DE genes (\(q < 0.1\)) | Enrichment \(P\)-value |
|------|-----------------------|---------------------|
| 24 h | All                    | >0.1                 |
|      | Downregulation         | >0.1                 |
|      | Upregulation           | >0.1                 |
| 48 h | All                    | 0.014                |
|      | Downregulation         | 0.005                |
|      | Upregulation           | 0.011                |

Abbreviation: GWAS, genome-wide association study. \(P\)-values from INRICH evaluating the enrichment of lower GWAS \(P\)-values (\(P < 0.01\)) to differentially expressed genes (\(q < 0.1\)) from miR-137 overexpression. The extended major histocompatibility complex region (chr6:25–34 mb) was excluded due to its high gene density, extended linkage disequilibrium and strong GWAS signal in schizophrenia. Boldface shows \(P < 0.01\).
CONFLICT OF INTEREST
Dr Sullivan was on the SAB of Expression Analysis (Durham, NC, USA). The remaining authors declare no conflict of interest.

ACKNOWLEDGMENTS
We offer special thanks to Dr Wei Sun for statistical advice and direction, Peter S. Chines for assistance with the transcription factor-binding site enrichment analysis pipeline and to Dr David Rubinow for his support. The project was funded by the Foundation For Hope (Raleigh, NC, http://www.walkforhope.com).

AUTHOR CONTRIBUTION
All authors reviewed and approved the final version of the manuscript.

DISCLAIMER
The corresponding author had access to the full data set.

REFERENCES
1 Schizophrenia Working Group of the Psychiatric Genomics Consortium Biological Insights From 108 Schizophrenia-Associated Genetic Loci. Nature 2014.
2 Levinson DF, Duan J, Oh S, Wang K, Sanders AR, Shi J et al. Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VR1R2 duplications. Am J Psychiatry 2011; 168: 302–316.
3 Purcell SM, Moran JL, Fromer M, Ruderfer D, Solovieff N, Roussos P et al. A burden of ultra-rare disruptive mutations concentrated in synaptic gene networks increases risk of schizophrenia. Nature 2014; 506: 185–190.
4 Fromer M, Pocklington AJ, Kavanagh D, Williams H, Dwyer S, Gormley P et al. De novo mutations in schizophrenia identify pathogenic gene networks regulating synaptic strength and overlap with autism and intellectual disability. Nature 2014; 506: 179–184.
5 Ripeke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA et al. Genome-wide association study identifies five new schizophrenia loci. Nat Genet 2011; 43: 969–976.
6 Ripeke S, O’Dushlaine C, Chamberl K, Moran JL, Kahler AK, Akterin S et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nat Genet 2013; 45: 1150–1159.
7 Guella I, Sequeira A, Rollins B, Morgan L, Torri F, van Erp TG et al. Analysis of miR-137 expression and rs162579 in dorsolateral prefrontal cortex. J Psychiatr Res 2013; 47: 1215–1221.
8 Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004; 116: 281–297.
9 Carrington JC, Ambros V. Role of microRNAs in plant and animal development. Science 2003; 301: 336–338.
10 Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J et al. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 2005; 433: 769–773.
11 Selbach M, Schwanhauser B, Thierfelder N, Xia X, Shi Y, Guan W et al. Widespread changes in protein synthesis regulated by microRNAs. Nature 2008; 455: 58–63.
12 Tarantino C, Paolella G, Cazzulo L, Minopoli G, Pastore L, Parisi S et al. miRNA 34a, 100, and 137 modulate differentiation of mouse embryonic stem cells. FASEB J 2010; 24: 3255–3263.
13 Smrt RD, Szulwach KE, Pfeffer RL, Li X, Guo W, Pathania M et al. MicroRNA miR-137 regulates neuronal maturation by targeting ubiquitin ligase mind bomb-1. Stem Cells 2010; 28: 1060–1070.
14 Szulwach KE, Li X, Smrt RD, Li Y, Luo Y, Lin L et al. Cross-talk between microRNA and epigenetic regulation in adult neurogenesis. J Cell Biol 2010; 189: 127–141.
15 Sun G, Ye P, Murai K, Lang MF, Li S, Zhang H et al. miR-137 forms a regulatory loop with nuclear receptor TRXL and LSD1 in neural stem cells. Nat Commun 2011; 2: 529.
16 Devanna P, Vernes SC. A direct molecular link between the autism candidate gene OR2A and the schizophrenia candidate MIR137. Scit Rep 2014; 4: 3994.
17 Willemsen MH, Valles A, Kerkils LA, Mastebroek M, Olde Loohuis N, Kos A et al. Chromosome 1p21.3 microdeletions comprising DPDY and MIR137 are associated with intellectual disability. J Med Genet 2011; 48: 810–818.
18 Wu H, Tao J, Chen PJ, Shahab A, Ge W, Hart RP et al. Genome-wide analysis reveals methyl-CpG-binding protein 2-dependent regulation of microRNAs in a mouse model of Rett syndrome. Proc Natl Acad Sci USA 2010; 107: 18161–18166.

© 2014 Macmillan Publishers Limited
57 Darnell JC, Van Driesche SJ, Zhang C, Hung KY, Mele A, Fraser CE et al. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell 2011; 146: 247–261.

58 Lee PH, O’Dushlaine C, Thomas B, Purcell S. InRich: Interval-based enrichment analysis for genome-wide association studies. Bioinformatics 2012; 28: 1797–1799.

59 Silber J, Lim DA, Petritsch C, Persson AI, Maunakea AK, Yu M et al. mir-124 and mir-137 inhibit proliferation of glioblastoma multiforme cells and induce differentiation of brain tumor stem cells. BMC Med 2008; 6: 14.

60 Stefansson H, Ophoff RA, Steinberg S, Andreassen OA, Cichon S, Rujescu D et al. Common variants conferring risk of schizophrenia. Nature 2009; 460: 744–747.

61 Irish Schizophrenia Genomics Consortium, Wellcome Trust Case Control Consortium. Genome-wide association study implicates HLA-C*01:02 as a risk factor at the MHC locus in schizophrenia. Biol Psychiatry 2012; 72: 620–628.

62 Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 2006; 34: D108–D110.

63 Lips ES, Cornelisse LN, Toonen RF, Min JL, Hultman CM, Holmans PA et al. Functional gene group analysis identifies synaptic gene groups as risk factor for schizophrenia. Mol Psychiatry 2012; 17: 996–1006.

64 Ferreira MA, O’Donovan MC, Meng YA, Jones IR, Ruderfer DM, Jones L et al. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. Nat Genet 2008; 40: 1056–1058.