Knockdown of the psychosis susceptibility gene ZNF804A alters expression of genes involved in cell adhesion

Matthew J. Hill1, Aaron R. Jeffries1, Richard J.B. Dobson2, Jack Price1 and Nicholas J. Bray1,∗

1Department of Neuroscience, Institute of Psychiatry, King’s College London, London, UK and 2Department of Biostatistics, NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, King’s College London, London, UK

Received October 6, 2011; Revised and Accepted November 7, 2011

Genome-wide association studies have convincingly implicated several novel genes in susceptibility to schizophrenia and bipolar disorder. The first genome-wide significant association with the broad phenotype of psychosis was with a polymorphism in the ZNF804A gene. However, the biological function(s) of ZNF804A have, to date, been entirely unknown. In this study, we manipulated the expression of ZNF804A in neural progenitor cells derived from human cortical neuroepithelium and assessed its effects on the cellular transcriptome. Gene ontology analysis of differentially expressed genes indicated a significant effect of ZNF804A knockdown on the expression of genes involved in cell adhesion, suggesting a role for ZNF804A in processes such as neural migration, neurite outgrowth and synapse formation. Several highly significant gene expression changes were confirmed in repeat cell culture experiments. Most consistent gene expression changes were seen for C2ORF80, a gene of as-yet-unknown function, and STMN3, a gene involved in neurite outgrowth and axonal and dendritic branching. These data, generated in a hypothesis-free manner, provide a basis for more targeted investigations of ZNF804A function.

INTRODUCTION

Genome-wide association studies (GWASs) have identified susceptibility loci for psychiatric disorders with unprecedented confidence (1,2). The first genome-wide significant finding for the broad phenotype of psychosis (encompassing both schizophrenia and bipolar disorder) was with the single-nucleotide polymorphism (SNP) rs1344706, located within an intron of the ZNF804A gene (3). Association between the T-allele of this SNP and schizophrenia is supported by several independent replication studies (4–6), and in a meta-analysis greatly exceeds accepted levels of genome-wide significance (\(P = 2.5 \times 10^{-11}\)), with an odds ratio of \(\sim 1.1\) (7). This allele also shows significant association with bipolar disorder alone (7). In addition, rare copy number variants affecting ZNF804A have been observed in individual cases of schizophrenia, bipolar disorder and anxiety disorder (5).

Although rs1344706 has been reported to show association with brain activity and structure (8,9), no information on the specific biological function(s) of ZNF804A has been published to date. ZNF804A is known to be expressed in the brain and is predicted to encode a protein with a C2H2 zinc finger domain, suggesting a role in the regulation of gene expression through DNA and/or RNA binding (10). However, any potential targets of ZNF804A have, to date, been unknown.

In order to shed light on the biological function(s) of ZNF804A, we performed hypothesis-free profiling of global gene expression in human neural progenitor cells following ZNF804A knockdown by RNA interference (RNAi). Gene ontology (GO) analysis was applied to genes for which significant expression changes were shared by two different ZNF804A small interfering RNA (siRNA) conditions. Repeat experiments showed reproducible changes in the expression of several selected genes following ZNF804A knockdown.
RESULTS

The primary experiment involved the following steps: knockdown of ZNF804A in human neural progenitor cells by separate use of two siRNAs targeting full-length ZNF804A with a negative control siRNA for comparison (n = 4 per condition), confirmation of ZNF804A RNA knockdown (relative to the negative control siRNA condition) by quantitative PCR (qPCR), microarray analysis of RNA samples using Illumina BeadChip arrays, identification of genes that were differentially expressed in both ZNF804A siRNA conditions relative to the negative control siRNA condition and analyses using the DAVID Bioinformatics Resource to assess potential enrichment of these genes in particular GO categories. We sought to confirm the most significant gene expression changes by qPCR, first using the same samples as were used for the microarray analysis, and then using independent samples derived from repeat ZNF804A knockdown experiments.

On the basis of uptake of Red Fluorescent oligonucleotide into neural cells (Supplementary Material, Figure S1), siRNA transfection efficiency in the primary experiment was estimated to be at least 80%. Cells were harvested 96 h after transfection, when mean RNA expression of ZNF804A was reduced by 35 and 36% relative to the negative control siRNA in the ZNF804A siRNA HSS150612 and HSS150613 conditions, respectively (both P < 0.0001; Fig. 1).

A total of 17111 microarray probes were detected at P > 0.95 in all 12 cellular RNA samples derived from the negative control siRNA and two ZNF804A siRNA conditions of the primary experiment. These probes correspond to transcripts for 13478 unique RefSeq genes, of which 10917 map to DAVID IDs. Expression at 899 microarray probes differed significantly (P < 0.05) between the ZNF804A HSS150612 and HSS150613 siRNA condition and the negative control siRNA condition, and expression at 1197 probes differed significantly between the ZNF804A HSS150613 siRNA condition and the negative control siRNA condition. Nominally significant (P < 0.05) gene expression changes were shared by the two ZNF804A siRNA conditions in the same direction, relative to the negative control, at 154 probes. Expression was reduced at 89 probes, and increased at 65 probes, in association with the two ZNF804A siRNA conditions. As an overlap of 154 probes is considerably more than would be expected by chance at the P < 0.05 threshold, it is likely that a large proportion of this list are genuine gene expression changes resulting from ZNF804A knockdown. However, we further tested this by comparing the P-values associated with the probes showing overlap with those not showing overlap for each ZNF804A siRNA comparison with the negative control. Consistent with an enrichment of ZNF804A-mediated expression changes in the overlapping gene list, the P-values associated with these changes were significantly smaller for the overlapping probes compared with the probes that were unique to each siRNA (Mann–Whitney P < 0.0005 for each comparison). The 154 probes correspond to transcripts for 151 unique RefSeq genes, of which 136 map to DAVID IDs. The full list of overlapping, differentially expressed genes is provided in Supplementary Material, Table S1.

Genes annotated as belonging to the primary biological process (GOTERM_BP_1) of ‘biological adhesion’ were statistically over-represented in the shared set of differentially expressed genes compared with the background list of all genes with detectable expression (P = 0.0022, Bonferroni corrected P = 0.044). The 12 differentially expressed genes belonging to this GO term are NCAM1, LAMA4, CCL2, NELL2, FBLN7, CTNND1, PCDH7, ANTXR1, SSPN, CD151, COL8A2 and PARI. All 12 genes also belong to the subsidiary GO term ‘cell adhesion’, which was associated with the smallest P-value (P = 0.0018) of all biological process terms in the GOTERM_BP_FAT category. GO terms in this category relating to regulation of cell growth/size were also significant, although, like the cell adhesion term, they did not survive Bonferroni correction for all terms in the category. Compared with the background list, genes annotated as belonging to the primary cellular compartment (GOTERM_CC_1) term of ‘extracellular region’ were also enriched in the shared set of differentially expressed genes shared by both ZNF804A siRNA conditions (P = 0.0036, Bonferroni corrected P = 0.039). Consistent with this, genes belonging to the terms ‘extracellular region’, ‘extracellular region part’, ‘plasma membrane’ and ‘cell surface’ in the GOTERM_CC_FAT category were also significantly over-represented, although their significance did not survive conservative Bonferroni correction for all terms in that category. All significant GO terms for differentially expressed genes are shown in Table 1. There was no significant enrichment of differentially expressed genes in any of the molecular function terms in the GOTERM_MF_1 or GOTERM_MF_FAT categories, nor in any KEGG pathway.

When the 89 down-regulated genes were analysed separately, ‘biological adhesion’ and ‘cell adhesion’ remained the most significant biological process GO terms (P = 0.0039) and ‘extracellular region’ the most significant cellular compartment term (P = 0.000015). In addition, ‘carbohydrate binding’ (P = 0.021) and ‘heparin binding’ (P = 0.039) were both significantly over-represented molecular function terms in the GOTERM_MF_FAT category. None of these GO terms was significantly over-represented in the 69
up-regulated genes. The most significant GO terms for up-regulated genes were ‘purine deoxyribonucleotide catabolic process’ (P = 0.020) in the biological process category and ‘metal ion binding’ in the molecular function category (P = 0.031).

Eight genes were identified as having particularly strong evidence for altered expression following ZNF804A knockdown on the basis of a false discovery rate (FDR) <0.05 in either of the ZNF804A siRNA conditions compared with the negative control siRNA condition. Genes that showed high-confidence expression changes in association with the HSS150612 siRNA were ZNF804A, C2ORF80, STMN3. Genes that showed high-confidence expression changes in association with the HSS150613 siRNA were SEZ6, FRZB, A2M, C2ORF80 and STMN3. Genes that showed high-confidence expression changes in association with the HSS150612 siRNA were SEZ6, FRZB, A2M, C2ORF80 and STMN3.

In order to test the reproducibility of individual gene expression changes, we performed a repeat experiment in which the effects of the two ZNF804A siRNAs were compared with a different negative control siRNA, again after 96 h. As with the primary experiment, transfection efficiency, as indexed by uptake of Red Fluorescent oligonucleotide, was estimated to be >80%. Knockdown of ZNF804A was slightly less
DISCUSSION

ZNF804A was identified as a psychosis susceptibility gene through genome-wide association scanning (3), an approach that does not rely on a priori hypotheses of disease etiology. Here, we provide the first clues to ZNF804A function through a similarly global, hypothesis-free design in which we assessed the effects of its knockdown on the cellular transcriptome. Genes belonging to the GO terms ‘biological adhesion’ and ‘extracellular region’ were enriched among those showing altered expression, the significance of both terms surviving correction for multiple testing. Repeat experiments demonstrated reproducible changes in the expression of specific genes as a result of ZNF804A knockdown.

The most specific, and significant, GO term in the biological process category was ‘cell adhesion’. This could suggest a role for ZNF804A in processes such as neural migration, neurite outgrowth and synapse formation, which are commonly hypothesized to be aberrant in schizophrenia. Interestingly, a recent pathway analysis of schizophrenia GWAS data found most significant enrichment of association signals in genes belonging to the KEGG pathway of ‘cell adhesion molecules’, which was also significantly associated in their bipolar disorder data set (11). Although we observed no enrichment of expression changes in any KEGG pathway, we note that only 33 of our 136 differentially expressed genes with DAVID IDs are represented in KEGG pathways, compared with 94 of these genes in the GOTERM_BP_FAT category.

We found replicable evidence for the down-regulation of several individual genes in response to ZNF804A knockdown, providing more specific insight into the molecular pathways by which ZNF804A operates. The genes showing most consistent expression differences across experiments were C2ORF80 and STMN3. C2ORF80 has yet to be functionally characterized, but STMN3 is known to encode a microtubule-stabilizing phosphoprotein (SCG10-like-protein/SCLIP), with roles in neurite outgrowth and axonal and dendritic branching (12–14). Although other genes showing high-confidence expression changes have also been implicated in these processes [e.g. A2M (15); CRYAB (16); SEZ6 (17)], these genes have various other functional roles, which could also be considered candidate mechanisms of ZNF804A function. CRYAB, for example, has been implicated in apoptosis (18) and glial pathology (19), and A2M is likely to have numerous cellular roles as a general proteinase inhibitor (20).

As our analyses were conducted at a single time point, they will provide only a snapshot of the transcriptional consequences of ZNF804A knockdown in our cells, and therefore some gene expression changes are likely to have been missed. Dynamic changes in gene expression, along with potential differences in the kinetics of individual siRNA, may explain some of the inconsistencies in repeat experiments.

Our data do not indicate whether altered expression of any given gene reflects direct actions of ZNF804A on that gene (through binding at the genomic locus or to its transcribed RNA) or secondary consequences. We therefore cannot draw any definitive conclusions as to whether ZNF804A serves as a positive or negative regulator of gene expression, although we note that the most significant gene expression changes were observed for genes that were down-regulated following ZNF804A knockdown. Methods based on chromatin or RNA immunoprecipitation will be crucial in determining any primary targets of ZNF804A. Like primary changes, secondary changes in gene expression could be part of the disease pathway or be aetiologically neutral, but may also reflect compensatory mechanisms. Genes that are part of the ZNF804A pathway—reflected in both primary and secondary gene expression changes—may be additional susceptibility loci for psychosis. As a possible example, we note that an SNP in CNNM2, a gene showing significantly reduced expression in both of our ZNF804A siRNA conditions (Supplementary Material, Table S1), shows genome-wide significant association with schizophrenia in the largest GWAS of the disorder to date (21).

Although ZNF804A contains coding variants with predicted effects on amino acid sequence, the association between this locus and schizophrenia centres on non-coding polymorphism; in particular, SNP rs1344706 within its second intron (7,22). Alleles of SNP rs1344706 have been found to differentially bind nuclear proteins from human neural cell lines (23), suggesting that susceptibility is mediated by effects on ZNF804A expression. Although the rs1344706 risk allele is reported to be associated with increased ZNF804A expression in adult dorsolateral prefrontal cortex (4) and lymphoblastoid cells (7), the results of the latter study suggest that this is not

<table>
<thead>
<tr>
<th>Gene</th>
<th>Fold change replication qPCR</th>
<th>P-value replication qPCR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HSS150612</td>
<td>HSS150613</td>
</tr>
<tr>
<td>A2M</td>
<td>0.84</td>
<td>0.70</td>
</tr>
<tr>
<td>ACTG2</td>
<td>0.84</td>
<td>1.15</td>
</tr>
<tr>
<td>C2ORF80</td>
<td>0.68</td>
<td>0.53</td>
</tr>
<tr>
<td>CRYAB</td>
<td>0.69</td>
<td>0.59</td>
</tr>
<tr>
<td>FRZB</td>
<td>0.97</td>
<td>0.66</td>
</tr>
<tr>
<td>SEZ6</td>
<td>1.04</td>
<td>1.25</td>
</tr>
<tr>
<td>STMN3</td>
<td>0.51</td>
<td>0.53</td>
</tr>
</tbody>
</table>

than was observed in the primary experiment, averaging 31% for the HSS150612 condition and 28% for the HSS150613 condition compared with the negative control siRNA, although these differences were again significant (both P < 0.005). Expression of the seven genes for which significant (P < 0.05) differences in expression between ZNF804A and negative control siRNA conditions had been confirmed in microarray samples was again assessed by qPCR. Average fold changes in target gene expression and associated P-values in the repeat experiment are shown in Table 3. Five of the seven genes showed significant (P < 0.05) expression differences in the same direction as in the microarray in association with at least one ZNF804A siRNA condition. These were A2M, C2ORF80, CRYAB, FRZB and STMN3. Most significant differences were seen for STMN3, which was reduced in expression by almost 50% in both repeat ZNF804A siRNA conditions relative to the negative control. Expression of CRYAB and C2ORF80 was also significantly reduced in association with both ZNF804A siRNAs in the repeat experiment.
directly the result of rs1344706. Indeed, using measures of ZNF804A allelic expression, we have found that the risk allele of rs1344706 has little effect in the adult brain, but significantly reduces ZNF804A expression in the fetal brain (M.J.H. and N.J.B., manuscript in preparation). Our manipulation of ZNF804A expression would therefore appear to be mechanistically inappropriate. In addition, the degree to which we experimentally reduced ZNF804A expression (∼35%) is consistent with the extent of common variability of cis-effects on ZNF804A expression, as observed in homogenized brain tissue at least [(7,24), M.J.H and N.J.B, manuscript in preparation].

We performed our experiments using clinical grade human neural progenitor cells with a normal karyotype (25). Although this cell line is likely to be more appropriate for our experiments than cells derived from lower species, other human tissues or neuroblastoma, we do not currently have information on the cell types that normally express ZNF804A, let alone those most relevant to its role in schizophrenia aetiology. However, we do know that our assayed cell line (weakly) expresses the nuclear protein that binds DNA sequence containing rs1344706 (23), and that it shows allelic imbalance of ZNF804A expression (unpublished data), indicating that variable cis-effects on ZNF804A expression are present in these cells at this developmental stage. In addition, our microarray data show that a large proportion of the human genome is expressed in these cells, and genes that we find differentially expressed as a result of ZNF804A knockdown have roles in cell types other than neural progenitors. This suggests that our findings are generally applicable to other cells of the human central nervous system. However, similar experiments in neural cells at later developmental stages and from different brain regions will help clarify the extent to which transcriptional effects of ZNF804A knockdown are cell-specific.

In summary, we provide the first clues to ZNF804A function through genome-wide expression profiling of human neural progenitor cells following ZNF804A knockdown. GO analyses indicate a significant enrichment of differentially expressed genes involved in biological, and more specifically cell, adhesion. Several genes showed reproducible expression changes in repeat experiments, providing molecular leads for more targeted investigations of ZNF804A function. Similar analyses of other genes showing robust association with psychoses may point to convergent molecular pathways and shared cellular mechanisms.

**MATERIALS AND METHODS**

**Cell culture**

All experiments were carried out using the CTX0E03 neural cell line obtained from ReNeuron Ltd (www.reneuron.com) under a Material Transfer Agreement. This clinical grade, karyotypically normal clonal cell line, described in detail elsewhere (25), was derived from 12-week human fetal cortical neuroepithelium. The CTX0E03 line has been conditionally immortalized by genomic incorporation of the c-mycER<sup>TAM</sup> transgene, to stimulate proliferation in the presence of the synthetic drug 4-hydroxy-tamoxifen (4-OHT). In their proliferative state, CTX0E03 cells stain positive for nestin, but are negative for both GFAP and β-III tubulin, consistent with neural progenitor cells (25).

CTX0E03 cells were grown in Nunclon<sup>Δ</sup>TM Surface T75 tissue culture flasks (Thermo Scientific, MA, USA) coated with Engelbreth–Holm–Swarm murine sarcoma basement membrane albumin (Sigma-Aldrich, Dorset, UK) at 1 μg/cm<sup>2</sup>. For expansion, cells were cultured in 10 ml of DMEM:F12, supplemented with 0.03% human serum albumin, 2 mM glutamine, human transferrin (100 μg/ml), putrescine dihydrochloride (16.2 μg/ml), human insulin (5 μg/ml), human progesterone (60 ng/ml), sodium selenite (40 ng/ml), 100 nm 4-OHT (all Sigma-Aldrich), and the growth factors bFGF (10 ng/ml) and EGF (20 ng/ml) (Peprotech, Inc., NJ, USA). Twenty-four hours before siRNA transfection, cells were divided equally between the T75 flasks that would be used for all siRNA conditions and cultured, on laminin, in 10 ml of DMEM:F12 containing all supplements listed above with the exception of 4-OHT. As CTX0E03 cells retain their nestin-positive, neural progenitor state in the presence of bFGF and EGF (25), 4-OHT was excluded from the media used for siRNA experiments so that proliferation was not artificially stimulated through c-myc over-expression.

**RNAi in cultured cells**

RNAi was performed using Stealth siRNA<sup>™</sup> (Invitrogen Ltd, Paisley, UK), which are chemically modified to reduce off-target effects. Two non-overlapping siRNAs (HSS150612 and HSS150613) targeting full-length ZNF804A mRNA (GI: 34740326) and a Stealth negative control siRNA with a similar guanine-cytosine (GC) content (Stealth RNAi<sup>™</sup> siRNA negative control low-GC duplex) were used for the primary experiments. HSS150612 targets exon 4, and HSS150613 targets exon 2, of full-length ZNF804A mRNA sequence (NM_194250). siRNA sequences are provided in Supplementary Material, Table S2. Each siRNA was combined with N-TERT<sup>™</sup> transfection reagent (Sigma-Aldrich), according to the manufacturer’s instructions, before adding to a T75 flask of seeded cells in 10 ml of the above media (minus 4-OHT) to yield a final siRNA concentration of 10 nm. For both the primary and repeat experiments, each siRNA was added to four separate flasks of seeded cells. In the repeat experiment, a different Stealth negative control siRNA (Stealth RNAi<sup>™</sup> siRNA negative control medium-GC duplex 2) was used for comparison, to control for any potential differences in gene expression arising from the original negative control siRNA rather than from the siRNA targeting ZNF804A. In order to estimate transfection efficiency, an additional T75 flask of seeded cells were transfected with 10 nm BLOCK-iTTM Alexa Fluor<sup>®</sup> Red Fluorescent oligonucleotide (Invitrogen), using the same N-TER reagent, and visualized after 24 h by fluorescence microscopy. Media (minus both 4-OHT and siRNA) was replaced at 48 h post-transfection. Cells from each T75 flask were separately harvested 96 h post-transfection by addition of Accutase (Millipore Ltd, Watford, UK) and centrifugation at 900g for 5 min at room temperature.
RNA preparation, cDNA synthesis and qPCR

Pelleted cells were separately homogenized in 1 ml of Tri-Reagent® (Applied Biosystems, Warrington, UK) and total RNA extracted according to the manufacturer’s instructions. Integrity of total RNA was assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, CA, USA). All samples had an RNA integrity number (RIN) > 9.5. Residual genomic DNA was removed by addition of 2 µl of Turbo DNA-free™ (Applied Biosystems) and incubation at 37°C for 30 min. cDNA was synthesized from 1 µg of total RNA from each extraction using random decamers and SuperScript® III (Invitrogen), according to the manufacturer’s protocol. qPCR was used to assess ZNF804A knockdown and expression of selected genes implicated by subsequent microarray analysis. Reactions were carried out in a total volume of 20 µl, containing diluted cDNA, 1 × HOT FIREPol® EvaGreen® qPCR Mix (Solis Biodyne, Tartu, Estonia) and primers at 200 nM, using an MJ Research Chromo 4 (Bio-Rad) and MJ Opticon Monitor analytic software (Bio-Rad). Primers were designed to target the same exons as the microarray probes (sequences provided in Supplementary Material, Table S3). Triplicate qPCR reactions were performed to measure each gene in each cDNA sample. The level of each gene was measured against a standard curve constructed by serial dilution of pooled cDNA from all assayed samples. A relative value was thus obtained for each of the three triplicate reactions for each cDNA sample. Mean measures of target genes were then normalized against those of an internal control gene (CFDP1) for each cDNA sample to yield a relative target gene expression value for all samples. CFDP1 was identified as a suitable internal control on the basis of whole-genome microarray data, where it showed the least variability (in terms of standard deviation) across all siRNA conditions used in this study. Normalized qPCR target gene expression values were compared between each ZNF804A siRNA condition and the negative control siRNA condition by individual t-tests (two-tailed).

Identification of differentially expressed genes and GO analysis

Microarray probes showing nominally significant (P < 0.05, uncorrected) differences between each ZNF804A siRNA and the negative control siRNA condition were first identified by individual t-tests (two-tailed) on normalized microarray data. To limit spurious results arising from low-expression genes, probes that were not detected in all 12 samples with a detection P-value > 0.95 were excluded. To refine the data set to those changes most likely to reflect ZNF804A knockdown rather than off-target effects of individual ZNF804A siRNA, we further identified significant gene expression changes associated with the two ZNF804A siRNA conditions that occurred in the same direction (i.e. up- or down-regulation) relative to the negative control siRNA condition. To test for enrichment of genuine gene expression changes resulting from ZNF804A knockdown in the overlapping gene set, we compared P-values between the overlapping and siRNA-specific gene set for each ZNF804A siRNA condition using the Mann–Whitney U test. We expected P-values to be significantly smaller in the overlapping gene set compared with the two siRNA-specific gene sets. The overlapping gene set was subjected to GO analysis through the DAVID Bioinformatics Resource 6.7 (http://david.abcc.ncifcrf.gov/) (27), using all gene probes that had a detection P-value > 0.95 in all samples as the background comparison.

Normalized microarray data were also analysed using the Significance Analysis of Microarrays (SAM) software (http://www-stat.stanford.edu/~tibs/SAM/) (28) in order to identify a smaller set of individual high-confidence gene expression changes for attempted confirmation and replication using qPCR. For this analysis, we set a delta value of 0.4 and excluded any gene probe that was not detected in all four samples in at least one of the three siRNA conditions. We sought to identify any gene expression differences between the negative control siRNA and either ZNF804A siRNA condition that had an FDR < 0.05 for subsequent qPCR analysis.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS

We are grateful to ReNeuron Ltd for providing the neural cell line used in this study and The Genome Centre, Barts and The London School of Medicine and Dentistry for microarray analysis.

Conflict of Interest statement. The authors declare no conflict of interest.

FUNDING

This work was supported by the Medical Research Council, UK (Grant ID G0802166). R.J.B.D. is funded by The NIHR Biomedical Research Centre for Mental Health at the South
REFERENCES


