Objective: rs12576775 was found to be associated with bipolar disorder (BD) in a genome-wide association study (GWAS). The GWAS signal implicates genes for the microRNAs miR-708 and miR-5579 and the first exon of the Odd Oz/ten-m homolog 4 gene (ODZ4). In the present study, miR-708, its surrounding region, and its targets were analyzed for potential BD-associated functional variants.

Methods: The miR-708 gene and surrounding regions were screened for variation using high-resolution melting (HRM) analysis in 1099 cases of BD, followed by genotyping of rare variants in an enlarged sample of 2078 subjects with BD, 1303 subjects with schizophrenia, and 1355 healthy controls. Whole-genome sequencing data from 99 subjects with BD were analyzed for variation in potential miR-708 binding sites. The minor allele frequencies (MAFs) of these variants were compared with those reported in reference individuals.

Results: Three variants detected by HRM were selected to be genotyped. rs754333774 was detected in three cases of BD, two cases of schizophrenia, and no controls. This variant is located 260 base pairs upstream from miR-708 and may play a role in controlling the expression of the miR. Four variants were identified in miR-708 targets binding sites. The MAFs of each of these variants were similar in BD and reference samples.

Conclusions: We report a single recurrent variant located near the miR-708 gene that may have a role in BD and schizophrenia susceptibility. These findings await replication in independent cohorts, as do functional analyses of the potential consequences of this variant.

KEYWORDS
bipolar disorder, microRNA, miR-708, sequencing, susceptibility, variation
miR-155-5p, miR-221-3p) was changed after 16 days of treatment. The addition to the findings with miR-137, several other microRNAs have been reported in SCZ, with 25 microRNAs impacted by rare CNVs in two or more unrelated subjects.18 Several studies have investigated polymorphic microRNA binding sites in target genes related to SCZ. These include rs3822674, which is located in the 3’ untranslated region (UTR) of the gene encoding complexin II (CPLX2) and is predicted to interfere with repression of CPLX2 expression by miR-49819; rs1130354 in the 3’UTR of dopamine receptor D2 gene (DRD2) was reported to interfere with miR-326-mediated repression of DRD2 expression20; and rs11122396 in the 3’ UTR of the disrupted-in-schizophrenia-1 (DISC1) gene disrupted miR-135b-5p-mediated control of DISC1 expression.21

The evidence that implicates miRNAs in the etiology of BD is not as strong as it is for SCZ. However, recent data have provided increasing support for the hypothesis that miRNAs also play a role in the etiology of BD. In the largest GWAS of BD to date, a SNP in an intergenic region flanking miR-2113 on chromosome 6q16.1 was the eighth strongest finding.11 A gene-based analysis of all known autosomal microRNAs using the same GWAS data found a significant association between nine microRNAs (including miR-708) and BD.22 In another study, bioinformatic analysis suggested that the BD-associated glutamate receptor-7 gene (GRM7) 3’ UTR variant, rs56173829, might modulate the binding of several microRNAs, including miR-4295, miR-130a-3p, and miR-130b-3p.23

Postmortem transcriptome studies have shown significant differences between miRNA expression levels in cortical brain tissue from people with BD and controls. While there is substantial heterogeneity in the miRNAs that have been reported to be associated with BD, several miRNAs have been reported in multiple studies.24 Two main studies have focused on miRNAs levels in the context of treatment with mood stabilizers in human tissue. One was conducted in lymphoblastoid cell lines, to assess the impact of lithium on miRNA expression.25 Of the 13 miRNAs previously suggested to respond to lithium and valproate, the regulation of four (miR-34a-5p, miR-152-3p, miR-155-5p, miR-221-3p) was changed after 16 days of treatment. The overall change in miRNA expression was small and no miRNA expression underwent a two-fold change. A second study focused on miR-134-5p. A decrease in miR-134-5p plasma levels was observed during manic phases, which correlated with symptom severity. This decrease was reversed upon successful treatment with various mood stabilizers.26 Thus, mood stabilizers used for the treatment of BD may influence psychopathology through the modification of miRNA biogenesis.

Bioinformatic analyses indicate that miR-708 is conserved in mammals within the ODZ4 intron, suggesting that miR-708 is co-expressed with ODZ4. The high level of conservation across mammalian species of the miR-708 precursor stem loop and the guide strand suggests that they are functionally important. It has been demonstrated that the expression of miR-708 correlates well with the expression of ODZ4 in adult mouse tissues. The observation of a significant accumulation of miR-708 and ODZ4 transcripts in the brain and eyes strongly suggests a physiological role for miR-708 in tissues in which ODZ4 is expressed.27 A study of postpartum psychosis (a disorder that is prevalent in up to 74% of women with an existing diagnosis of BD and a family history of puerperal psychosis28) suggested differential expression of miR-708 in the monocytes of affected patients compared with controls.29 miR-708 expression has been demonstrated to be upregulated significantly in mouse hippocampal neurons in an oxidative stress cell model.30

In the present study, miR-708, its surrounding region, and its targets were analyzed for possible functional variants associated with BD.

1.2 Detection and evaluation of new variants

High-resolution melting (HRM) variant screening was used to identify BD susceptibility variants 300 base pairs (bp) upstream and
downstream from the mir-708 gene (chr11:79112766-79113453, GRCh37/hg19).

This method of variant analysis allows cost-efficient detection of rare variations in large numbers of samples. HRM is particularly amenable to regions not efficiently targeted by established next generation sequencing selection panels such as introns and UTRs.

HRM was performed using three primer pairs in 1099 cases with BD. Reactions were carried out on a LightCycler 480 (Roche, Burgess Hill, UK). Primer sequences and reagents are shown in Table S1. Samples with abnormal HRM curves were then sequenced using the BigDye terminator v3.1 Cycle Sequencing kit (Applied Biosystems, Warrington, UK) on an ABI 3730xl DNA Analyzer (Applied Biosystems).

Sequencing data were analyzed using the Staden Package (http://staden.sourceforge.net/).32 The reference minor allele frequency (MAF) for the general population was evaluated in the data from the 1000 Genomes (1000G) project, 33 in the Exome Aggregation Consortium (ExAC), Cambridge, MA, USA (http://exac.broadinstitute.org [accessed January 2016]), from the Scripps Wellderly study (n = 534)34 and from two population cohorts in the UK10K study (UK10K ALSPAC [Avon Longitudinal Study of Parents and Children] and UK10K TWINS [TwinsUK n = 2432]).35 The latter two datasets were queried using the Reference Variant Store, (http://rvs.u.hpc.mssm.edu [accessed 5 August 2016]).

For the ExAC data, it is important to take into consideration the presence of BD and SCZ diagnoses in the database. This did not affect our study because the reported MAFs of the miR-708 variants selected did not exceed our exclusion threshold for genotyping. Bioinformatic analysis to determine the potential functional effect of SNPs was carried out using the University of California, Santa Cruz genome browser (http://genome.ucsc.edu/), Alibaba 2 (http://www.gene-regulation.com/pub/programs/alibaba2/index.html) and PROMO using version 8.3 of TRANSFACT (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3).

Only variants with a MAF lower than 0.01 in the general population were considered for further analysis.

### Genotyping

Genotyping of the selected SNPs was performed in an enlarged sample of 2078 cases with BD and 1355 ancestrally matched controls. It was performed in-house using the allele-specific polymerase chain reaction (PCR), using KASPar reagents (LGC Genomics, Hoddesdon, UK) on a LightCycler 480 real-time PCR machine. Allele-specific primers were designed for each of the SNPs using Primer Picker (KBiosciences, LGC Genomics, Hoddesdon, UK; primer sequences are listed in Table S2). The variants were also genotyped in 1305 SCZ samples. For all SNPs, 97% of samples were successfully genotyped. Genotyping for each heterozygote sample was repeated at least twice. All of these data were analyzed to confirm Hardy-Weinberg equilibrium. Allelic associations for SNPs were performed using Fisher’s exact test. Significance values shown for all analyses are uncorrected for multiple testing, and a cut-off significance value of P<.05 was used.

### Table 1

<table>
<thead>
<tr>
<th>SNP</th>
<th>Position in chr11</th>
<th>Variants</th>
<th>1000G MAF</th>
<th>Eur 1000G MAF</th>
<th>Exac MAF</th>
<th>Eur Exac MAF</th>
<th>Wellderly MAF</th>
<th>UK10K MAF</th>
<th>Position compared to miR-708</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs754333774</td>
<td>79113407</td>
<td>G&gt;A</td>
<td>nd</td>
<td>nd</td>
<td>nd</td>
<td>0.000 94</td>
<td>nd</td>
<td>nd</td>
<td>Upstream</td>
</tr>
<tr>
<td>rs768049399</td>
<td>79113040</td>
<td>A&gt;C</td>
<td>nd</td>
<td>nd</td>
<td>0.000 064</td>
<td>0.000 097</td>
<td>nd</td>
<td>nd</td>
<td>Downstream</td>
</tr>
<tr>
<td>rs56158925</td>
<td>79112878</td>
<td>C&gt;T</td>
<td>0.082</td>
<td>0.082</td>
<td>nd</td>
<td>0.069 29</td>
<td>0.063 54</td>
<td></td>
<td>Downstream</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Position on chr11</th>
<th>Position compared with miR-708</th>
<th>Change</th>
<th>N</th>
<th>Genotype counts</th>
<th>MAF</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs754333774</td>
<td>Upstream</td>
<td>G&gt;A</td>
<td>2041</td>
<td>BD</td>
<td>0.0007</td>
</tr>
<tr>
<td>79113407</td>
<td></td>
<td></td>
<td></td>
<td>0/3/2038</td>
<td></td>
</tr>
<tr>
<td>rs768049399</td>
<td>Downstream</td>
<td>A&gt;C</td>
<td>2001</td>
<td>BD</td>
<td>0.0002</td>
</tr>
<tr>
<td>79113040</td>
<td></td>
<td></td>
<td></td>
<td>0/1/2000</td>
<td></td>
</tr>
</tbody>
</table>

BD, bipolar disease; CTRL, control; MAF, minor allele frequency; N, total number; SCZ, schizophrenia. The genomic reference sequence used is GRCh37/hg19; change: the nucleotide change indicated is on the negative strand; genotype count: number of homozygotes for the minor allele/heterozygotes/homozygotes for the major allele.
Whole-genome sequencing (WGS) was performed on 99 subjects with BD-I selected from our BD cohort first on the basis of individuals with a strong positive family history of BD or bipolar spectrum disorder. Where the strength of the family history was tied, individuals with the earliest age at onset were selected. The mean age of onset for the cases selected for sequencing was 21.55 (standard deviation [SD] 8.90), and this was significantly lower than the total cohort (P = .0124). The genomic DNA was sequenced using 100 bp paired-end reads on a Hi-Seq 1000 (Illumina Inc., San Diego, CA, USA). Sequence data alignment to the National Center for Biotechnology Information human reference genome 37.1 (hg19) and variant calling was performed using the CASAVA 1.8.2 pipeline at Illumina (http://support.illumina.com/sequencing/sequencing_software/casava.html). The sequence data from these individuals was further analyzed and annotated using kGAP (Knome Inc., Boston, MA, USA). The BD WGS data was screened for variants in miR-708 binding sites (15 bp upstream and 1 bp downstream the microRNA seed) predicted by Targetscan 6.2: June 2012 (http://www.targetscan.org/).

## RESULTS

### 2.1 Variant selection and genotyping

Three single nucleotide variants were detected by HRM analysis across the region selected to be analyzed: rs56158925, rs754333774, and rs768049399 (Table 1). rs56158925 is located 200 bp downstream from mir-708. This variant has an overall MAF of 0.082 in the general population of the 1000G Project, and the same frequency in the data from the European subpopulation of the project and with a slightly lower frequency in the Wellderly and UK10K cohorts data (Table 1). The other two variants, rs754333774 and rs768049399, are located 260 bp upstream and 27 bp downstream, respectively, of the mir-708 gene. rs768049399 was annotated in the ExAC database, with an overall MAF of 6.4 × 10−5 and 9.7 × 10−5 in the European subpopulation of the 1000G Project, but not reported in the 1000G project or in the UK10K cohorts in the European subpopulation. rs768049399 was also not reported in the Wellderly and UK10K cohorts.

Genotyping assays were designed for the two SNPs with MAFs lower than 0.01 in the general population. Genotyping was conducted in the complete UCL case–control sample, including cases with BD and SCZ. The rs768049399 variant allele was detected in one case with BD and in one control (Table 2). The rs754333774 variant allele was not detected in cases or controls.

### 15. | Imputation and analysis of BD GWAS data

UCL BD and control GWAS data were included in the PGC-BD data set. These data have been subjected to a standardized quality control and imputation pipeline (https://sites.google.com/a/broadinstitute.org/ricopili/). Association testing was performed using PLINK2.38

<table>
<thead>
<tr>
<th>Gene</th>
<th>Seed position in 3' UTR</th>
<th>SNP</th>
<th>Position</th>
<th>Alleles</th>
<th>Ancestral allele</th>
<th>Variant position in the miR-708 binding site (5'-&gt;3')</th>
<th>BD WGS</th>
<th>1000G</th>
<th>Eur 1000G</th>
<th>Exac</th>
<th>Eur Exac</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOC3</td>
<td>41–47</td>
<td>rs5128</td>
<td>chr11:116703640</td>
<td>G &gt; C</td>
<td>C</td>
<td>..CCUAUCCACUGGacuccC..</td>
<td>0.915</td>
<td>0.77</td>
<td>0.88</td>
<td>0.84</td>
<td>0.9</td>
</tr>
<tr>
<td>GALNT13</td>
<td>764–770</td>
<td>rs707082</td>
<td>chr2:155307832</td>
<td>T &gt; C</td>
<td>C</td>
<td>..AAACUAAUAAACUGacuccC..</td>
<td>0.93</td>
<td>0.76</td>
<td>0.89</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>ST6GALNAC4</td>
<td>359–365</td>
<td>rs1043026</td>
<td>chr9:130670310</td>
<td>A &gt; G</td>
<td>G</td>
<td>..CCUCCGUCCCAACCGGCgacuccC..</td>
<td>0.995</td>
<td>0.93</td>
<td>1.00</td>
<td>nd</td>
<td>nd</td>
</tr>
<tr>
<td>SLC22A23</td>
<td>1812–1818</td>
<td>rs5873874</td>
<td>chr6:3271471-3271471</td>
<td>-&gt;GAT</td>
<td>GAT</td>
<td>..ACUGGCACGCGCgacuccC..</td>
<td>0.955</td>
<td>0.83</td>
<td>0.96</td>
<td>nd</td>
<td>nd</td>
</tr>
</tbody>
</table>

The germic reference sequence used was GRCh37/hg19. The variant position in the sequence is indicated with boldened letters. 1000G, 1000 Genome Project; APOC3, apolipoprotein C-III; Eur, European population; Exac, Exome Aggregation Consortium; GALNT13, polypeptide N-acetylgalactosaminyltransferase 13; nd, not detected; SNP, single-nucleotide polymorphism; UTR, untranslated region; SLC22A23, solute carrier family 22 member 23; ST6GALNAC4, ST6(alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4.
was detected in three cases with BD, two cases with SCZ, and no controls (Table 2).

Bioinformatic transcription factor binding analysis of the effect of rs754333774 did not identify consistent predictions for altered binding. By contrast, the variant allele of rs768049399 was predicted to destroy a binding site for the CCAAT/enhancer-binding protein beta transcription factor and to create a binding site for five transcriptional factors: hepatocyte nuclear factor (HNF)-3c; TATA box-binding protein; homeobox protein D8; HNF-1C; and HNF1-B.

2.2 | Variants in miR-708 binding sites

Targetscan 6.2 predicted 4377 transcripts with miR-708 binding sites (including many transcripts with overlapping 3′ UTRs), with a total of 381 conserved sites and 4999 poorly conserved sites (including many overlapping sites). These binding sites were analyzed in WGS from 99 subjects with BD for possible variants. Only four nonreference (GRCh37/hg19) allelic variants were identified. The variants were located in the genes encoding the following proteins: apolipoprotein C-III (APOC3); polypeptide N-acetylgalactosaminyltransferase 13 (GALNT13); ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 4 (ST6GALNAC4); and solute carrier family 22 member 23 (SLC22A23) (Table 3). All four variants are known common variants that have been annotated on The Single Nucleotide Polymorphism Database (dbSNP). rs5128 in APOC3 is 1 bp upstream from the miR-708 seed sequence, while the other three variants rs707082, rs1043026, rs5873874, located in GALNT13, ST6GALNAC4 and SLC22A23, respectively, are within the seed sequence of the miR-708 binding sites.

The allele frequencies of the four variants in the BD WGS samples were similar to those reported in both the entire sample and the European-only sample from the 1000G project (Table 3). Data were available only for rs5128 in the ExAC database, and these were similar to those of the BD WGS sample. The 3′ UTRs of the remaining genes were not included in the exome sequencing data in the ExAC database. Association testing of GWAS data for the first 491 UCL BD cases and 495 UCL controls, imputed using the 1000G project data as a reference panel, did not show evidence for involvement of rs5128 or rs707082 with BD (P= .5214 and .4286, respectively); rs1043026 and rs5873874 were not reliably imputed (Table S3).

3 | DISCUSSION

The potential role of microRNAs in psychiatric disorders has recently been highlighted by GWS association findings with SNPs within miR-137 and SCZ. Several reports have implicated microRNAs in BD, both genetically and biologically. The PGC-BD group performed a combined analysis of GWAS data that identified a GWS intronic variant in ODZ4 (TENM4), rs12576775. This association has been confirmed by two other independent studies.

The two variants in miR-708 with a MAF lower than 0.01 (rs768049399 and rs754333774) were genotyped in our BD and control cohorts. The strong prior findings with miRNA genes in SCZ led us to genotype the selected variants in our SCZ cohort, in addition to the subjects with BD and control subjects.

The rare allele of rs754333774 was found only in individuals with BD or SCZ; in our own data and, this finding was not statistically significant. This variant was detected in a single individual from the WELDERLY sample but was absent in a substantially larger number of individuals from the UK10K cohorts. Together the frequency of this variant in the combined UK10K cohorts and WELDERLY sample was lower than that in our BD and SCZ case cohorts. This finding is intriguing but requires validation in large independent samples of subjects with BD and or SCZ and healthy controls. No support for a role for rs768049399 in BD (and/or SCZ) was found.

In order for microRNAs to regulate gene expression, they need to bind to a target region, normally in the 3′ UTR of a gene. It is important that nucleotides 2–8 of the miRNA have a perfect match with their target, and this region is defined as a seed region. The microRNA binding site and, more specifically, the seed regions are well conserved; they are more likely to reside within those targets in the transcriptome with lower variant densities, especially target regions in which nucleotides have low mutation frequencies. Analysis of enrichment of GWS signals for miRNA genes and their putative target regions implicated miR-137 and its pathway in SCZ in different populations.

With the advent of WGS, it is therefore possible to implicate not only regions, but also specific variants, in disease. WGS data for 99 BD subjects were analyzed for possible variants in the miR-708 binding sites predicted by Targetscan. This analysis identified four miR-708 binding site variants, each of which was located in brain-expressed genes, including APOC3 and SLC22A23. Both of these genes have been reported to be associated with the response to antipsychotic drug treatment. Variants in APOC3 have been implicated in variations of cholesterol and triglyceride levels in SCZ patients treated with clozapine and olanzapine. Variants in SLC22A23 have been reported to be associated with QT prolongation in a GWAS of subjects with SCZ treated with quetiapine. However, comparison of the allele frequency data of the four miR-708 binding site variants between the 99 WGS subjects with BD and those in reference databases suggested that none of the miR708 3′UTR binding site variants were likely to be etiological. Indeed, analysis of imputed data from our own BD case–control sample did not show evidence for an association for the two variants that were reliably imputed. One of the challenges of studying miRNAs is the complexity of miRNA-target gene networks. The study of these networks requires systematic computational prediction of miRNA-target gene interactions. However, there is a certain degree of imprecision in the predictions made by current miRNA target gene algorithms. This imprecision can be observed in the inconsistent prediction scores that are obtained from the different algorithms. Most algorithms do not consider whether miRNA genes and their potential target are co-expressed and this could be an effective way of improving predictions. Furthermore both miRNAs and their targets are located in areas of the genome that are not often covered by whole
Exome sequencing. However, as larger scale WGS datasets become available, variation in 3'UTRs and intronic/intragenic regions will be better described, making it possible to begin to understand the functional role of these regions.

4 | CONCLUSIONS

In summary, we report a single recurrent variant located close to in the miR-708 gene that may have a role in susceptibility to BD and/or SCZ. This finding awaits replication in independent SCZ and BD cohorts, as does a functional analysis of the potential consequences of this variant.

ACKNOWLEDGEMENTS

The UCL clinical and control samples were collected with support from the Bipolar Organization, the Neuroscience Research Charitable Trust, the Central London NHS Blood Transfusion Service, the Camden and Islington NHS Foundation Trust and 10 other NHS Mental Health Trusts, the Stanley Foundation, the National Institute for Health Research Mental Health Research Network, and the NIHR-supported Primary Care Research Network. Genetic analysis of the UCL cohort has been supported by UK Medical Research Council project grants G9623693N, G0500791, G0701007, and G1000708. NLO/B was supported by a joint PhD studentship, funded by a UCL IMPACT award and Equilibrium, the Bipolar Foundation. The Stanley Foundation and the Stanley Psychiatric Research Center at the Broad Institute, Boston, MA, USA, funded the genome-wide association genotyping. The authors would also like to thank Leanne Graham, funded by a Wellcome Trust Biomedical Vacation Scholarship, for her contribution to the initial HRM screening of ODZ4.

DISCLOSURES

The authors of this paper do not have any commercial associations that might pose a conflict of interest in connection with this manuscript.

REFERENCES


SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Fiorentino, A., O'Brien, N. L., Sharp, S. L., Curtis, D., Bass, N. J. and McQuillin, A. (2016), Genetic variation in the miR-708 gene and its binding targets in bipolar disorder. Bipolar Disorders, 00: 1–7. doi: 10.1111/bdi.12448